Les protéines modulaires : un socle architectural pour la multifonctionnalité. Cas des protéines du complément

Véronique Rossi

To cite this version:

Véronique Rossi. Les protéines modulaires : un socle architectural pour la multifonctionnalité. Cas des protéines du complément. Biologie structurale [q-bio.BM]. Université Joseph Fourier, 2015. tel-01261973

HAL Id: tel-01261973
https://hal.univ-grenoble-alpes.fr/tel-01261973
Submitted on 26 Jan 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE GRENOBLE

Institut de Biologie Structurale

Mémoire

Pour obtenir le diplôme d'

HABILITATION A DIRIGER LES RECHERCHES

Spécialité : **Biologie**

Présenté par

Véronique Rossi

préparé à l'Institut de Biologie Structurale, Grenoble

Les protéines modulaires : un socle architectural pour la multifonctionnalité

Cas des protéines du complément

Soutenance le 18 septembre 2015,

Membres du Jury :

Jean-François HERNANDEZ

Directeur de Recherche CNRS, Université de Montpellier (Rapporteur)

David HULMES

Directeur de Recherche CNRS, Université de Lyon (Rapporteur)

Stéphane DEDIEU

Professeur, Université de Reims (Rapporteur)

Franck FIESCHI

Professeur, Université de Grenoble Alpes (Examinateur)

Nicole THIELENS

Directeur de Recherche INSERM, Grenoble (Examinateur)

Université Joseph Fourier / Université Pierre Mendès France / Université Stendhal / Université de Savoie / Grenoble INP

Messieurs les Directeurs de Recherche au CNRS Jean-François Hernandez et David Hulmes, et monsieur le Professeur Stéphane Dedieu ont accepté d'être rapporteurs de ce travail, j'en suis très honorée.

Je suis reconnaissante à monsieur le Professeur Franck Fieschi et à Nicole Thielens, Directrice de Recherche à l'INSERM de participer à ce jury.

Je tiens à exprimer toute ma gratitude à Gérard ARLAUD qui m'a accueillie dans son laboratoire pour commencer cette aventure dans le domaine du complément. Bien sûr sa rigueur scientifique et son aptitude toute particulière pour poser le parafilm m'ont été d'un grand secours pendant toutes ces années passées au LEM. Il y a des choses qui restent…et bien d'autres encore…

Cette HDR, mes collègues et ma famille en ont entendu parler depuis longtemps et je les remercie de tout cœur de leurs encouragements. Merci à Nicole Thielens qui n'a cessé de me soutenir dans cette démarche, merci également à Franck Fieschi, qui à chaque rencontre abordait la question de « quand ? ».

Merci aux autres HDR-istes féminines confirmées, enseignantes-chercheuses de l'IBS, Eve de Rosny et Claire Durmort. Et surtout, j'ai une gratitude éternelle pour l'Université Joseph Fourier, dans laquelle je suis enseignante depuis 15 ans, et qui maintenant, me donne l'opportunité de redevenir étudiante pour quelques mois.

C'est toujours avec un grand plaisir que je vais au laboratoire et c'est grâce à mes collègues et amis de longtemps, Isabelle avec laquelle on a fini par se comprendre sans ne rien dire et qui m'apporte depuis toujours un soutien inestimable. Evelyne, la souriante, qui partage mon bureau depuis … et son chocolat ! Pascale et ses éclats de rire. Christine, un remède pour les discussions scientifiques. Philippe, mon collègue de fac et d'efferocytose. Une pensée pour Romain et Rabia et les agréables moments du matin.

Merci à mes nouveaux collègues et aventuriers de LRP1, Catherine, Jean-Philippe, Jean-Baptiste, Anne et Dominique.
Merci à tous les protagonistes de mes travaux de recherches et à tous les étudiants passés et futurs qui me transmettent toujours leur jeune énergie.

Table des matières

Table des matières 7

Table des illustrations........ 9

Chapitre I Curriculum vitae 11

Diplômes........ 12

Expérience scientifique 12

Prix scientifiques et financements 14

Responsabilités collectives et animation scientifique........ 14

Co-auteurs et collaborations 14

Encadrement d'étudiants 16

Liste des publications 17

Liste des communications 20

Activité d'enseignement........ 23

Expérience en recherche 25

Chapitre II Présentation générale de l'activité de recherche........ 27

2.1 - Préambule 27

2.2 - Introduction générale........ 32

2.2.1 - Le complément, un système majeur de l'immunité innée 32

2.2.2 - Les complexes activateurs des voies classique et lectine 33

2.2.3 - L'organisation modulaire des protéines du complément 35

Chapitre III Protéases à sérine des complexes activateurs du complément 39

3.1 - Relations structure/fonction des protéases C1s et C1r 41

3.1.1 - Brève description du complexe C1 et des protéases associées. 41

3.1.2 - Le domaine protéase à sérine (SerPr) – catalyse enzymatique et sites de reconnaissance. 42

3.1.3 - Elaboration d'un modèle tridimensionnel de la région catalytique (CCP1-CCP2-SerPr) de C1s par réticulation chimique. 45

3.1.4 - Structure du fragment CCP2-SerPr de C1s par cristallographie aux rayons X et identification d'une famille de protéases à motif structural homologue. 47

3.1.5 - Mise en évidence de deux mécanismes distincts de reconnaissance des substrats C2 et C4 par C1s. 50

3.1.6 - Cartographie fonctionnelle de la région catalytique de C1s : Identification des déterminants moléculaires responsables de la spécificité de C1s pour ses deux substrats, C2 et C4. 50

3.1.7 - Régulation de l'activité enzymatique de C1s par C1-inhibiteur – une interaction atypique avec l'héparine. 54

3.2 - Les protéases MASP de la voie lectine - Un système ancestral majeur de défense innée contre l'infection 57
3.2.1 - Caractérisation structurale et fonctionnelle des protéases MASP-1 et MASP-2 58
3.2.2 - Identification des déterminants structuraux responsables de l'efficacité accrue de MASP-2 vis-à-vis de ses substrats C2 et C4 – Caractérisation de molécules recombinantes chimères de C1s et MASP-2 60
3.2.3 - L'énigme MASP-3 61
Chapitre IV Le complexe C1 – Modèle revisité 63
4.1 - Identification des sites d'assemblage du complexe C1 – Le modèle de C1 revisité. 65
Chapitre V Aparté sur C9 et AIL – Stage post-doctoral 67
5.1 - Implication des ponts disulfures de C9 dans la toxicité du MAC. 69
5.2 - Topographie de l'insertion membranaire de C9 - utilisation de mutants de glycosylation. 70
5.3 - Etude de la protéine membranaire AIL - Construction d'un système membranaire artificiel 71
Chapitre VI Protéines de l'immunité innée à l'interface hôte-pathogène – Etude du récepteur CR1 et de la thrombospondine 1 73
6.1 - Etude du récepteur CR1 76
6.2 - La thrombospondine 1 : une plateforme d'ancrage des collagènes de défense 79
Chapitre VII Projets de recherche 81
7.1 - Projet 1 : Inhibiteurs bactériens de protéases à sérines : nouvelles fonctions et potentiel thérapeutique contre les maladies inflammatoires de l'intestin – ANR SerpinGUTARGET 83
7.1.1 - Introduction sur le fonctionnement d'une serpine 83
7.1.2 - Demande de financement ANR SerpinGUTarget 85
7.2 - Projet 2 : Rôle et mécanisme d'action du récepteur CD91/LRP-1 dans l'efferocytose 89
7.2.1 - Description du projet de groupe : Les collagènes de défense dans l'efferocytose. 89
7.2.2 - Demande de financement AGIR (Université Joseph Fourier) 94
7.2.3 - Description détaillée de la structure multimodulaire de LRP1 et de son mode d'interaction calcium-dépendant. 98
7.2.4 - Résultats préliminaires 102
7.2.5 - Perspectives 106
Références bibliographiques 109
Annexes 117
Publications du Chapitre III 119
Publication du Chapitre IV 121
Publication du Chapitre V 123
Publication du Chapitre VI 125

Table des illustrations

Figure 1 : Les voies classique et lectine d'activation du complément et les différents sujets étudiés 30

Figure 2 : Arbre phylogénique simplifié de l'évolution des principaux groupes du monde animal et de leur système immunitaire. 33

Figure 3 : Représentation schématique des complexes activateurs de la voie classique et lectine. 35

Figure 4 : Structure modulaire des protéines du complément. 37

Figure 5 : Structure modulaire des protéases à sérine des complexes activateurs des voies classique et lectine. 42

Figure 6 : Mécanisme de la catalyse enzymatique des protéases à sérine. 43

Figure 7 : Localisation des éléments de reconnaissance de la trypsine communs à toutes les protéases de la famille. 44

Figure 8 : Modèle et structure de la région catalytique de C1s. 46

Figure 9 : Structure modulaire des protéases de la famille CCP-SP. 48

Figure 10 : Illustration du rôle potentiel du bras rigide CCP2 de C1s dans le pivotement du domaine catalytique SP pour l'activation de la cascade du complément. 49

Figure 11 : Mutations réalisées sur C1s pour l'étude de la cartographie fonctionnelle. 53

Figure 12 : Modèle tridimensionnel de la région catalytique de C1s. 54

Figure 13 : Fragments serpine de C1-inhibiteur produits dans les cellules d'insectes. 56

Figure 14 : Chimères de C1s et MASP-2 exprimées dans le système baculovirus/cellules d'insectes. 60

Figure 15 : Modèle 3-D de l'assemblage du complexe C1 66

Figure 16 : Représentations schématiques du complexe C1. 66

Figure 17: Représentation schématique de la structure modulaire de C9. 70

Figure 18 : Topographie de l'interaction de C9 avec la membrane par des mutants de glycosylation. 71

Figure 19 : Les collagènes de défense C1q, MBL et ficolines, senseurs de l'immunité innée. 75

Figure 20 : Représentation structurale des modules CCP24-25 de CR1. 78

Figure 21 : Structure et mécanisme d'inhibition des serpines. 84

Figure 22 : Les différents états conformationnels des serpines. 84

Figure 23 : Vue simplifiée du réseau d'interaction de l'efferocytose médiée par C1q (ANR C1qEffero 2015). 92

Figure 24 : Représentation schématique de la structure modulaire de LRP1/CD91 99

Figure 25 : Modèle commun d'interaction calcium-dépendante entre les récepteurs à CR et leur ligand et entre les modules CUB de MASP et la MBL. 100

Figure 26 : Production des clusters solubles de LRP1 et caractérisation de l'interaction avec C1q 103

Figure 27 : Minirécepteurs de LRP1 - Outils cellulaires 105

Chapitre I

Curriculum Vitae

__Etat Civil :__

Véronique ROSSI

Née le 27 mars 1969 à Grenoble

Française, mariée (épouse ECHINARD)

2 enfants (2004 et 2012)

__Activité professionnelle :__

Maître de conférences (6$^{\text{ème}}$ échelon)

Université Joseph Fourier - Grenoble Alpes

CNU 64 – Biochimie et Biologie moléculaire

*Adresse **professionnelle** :*

Groupe **IRPAS**
Réponse Immunitaire aux Pathogènes et au Soi Altéré
Institut de Biologie Structurale IBS2
(CNRS-CEA-UJF)
71 avenue des Martyrs
CS 10090
38044 GRENOBLE CEDEX 9

☎ : +33 (0)4 57 42 85 38
webpages : http://www.ibs.fr
E-mail : veronique.rossi@ibs.fr

Diplômes

1997 : Doctorat de Biologie

Thèse de doctorat de Biologie de l'Université Joseph Fourier soutenue le 10 Janvier 1997 à Grenoble,
Mention : Très honorable avec félicitations du jury.

1993 : Diplôme d'Etudes Approfondies (DEA)

DEA de **Biologie Cellulaire et Moléculaire** de l'Université Joseph Fourier, Grenoble.
Mention : Bien.

1992 : Maîtrise

Maîtrise de Biochimie de l’Université Joseph Fourier – ERASMUS à Glasgow.
Mention : Bien.

1991 : Licence

Licence de Biochimie de l’Université Joseph Fourier, Grenoble.
Mention : Très Bien.

1990 : DEUG

DEUG de Chimie Biochimie de l’Université Joseph Fourier, Grenoble.
Mention : Bien.

Expérience scientifique

Depuis septembre 2000 :

Maître de conférences de Biochimie (section 64), UFR de Chimie/Biologie, Université Joseph Fourier, Grenoble.
Rattachée au Laboratoire d’Enzymologie Moléculaire (LEM) puis au groupe IRPAS depuis 2010, Institut de Biologie Structurale (IBS).

Octobre 1998 - 2000 :

Attachée Temporaire d’Enseignement et de Recherche en Biochimie **(ATER)**, Université Joseph Fourier, Grenoble. Travail de recherche effectué au LEM, IBS. Directeur de laboratoire : Dr. Gérard Arlaud. ***Sujet d’étude*** : Etude du mécanisme d’activation de la voie lectine du complément : caractérisation structurale et fonctionnelle des protéases associées à la « Mannan-Binding Lectin »

Mars 1997 - 0ctobre 1998 :

Séjour post-doctoral, "**Research associate**", School of Biological Sciences (SBS), University of Missouri Kansas City (UMKC), Directeur de laboratoire: Professeur Alfred Esser.
Sujet d'étude : Etude de l'interaction de la protéine C9 du complexe d'attaque membranaire du complément avec les membranes.
Etude de la protéine AIL de Yersinia enterocolitica, un agent de résistance de la bactérie à son élimination par le complément.

1993-1997 :

Travaux de thèse

Sujet : Structure et fonction d'une protéase modulaire : étude de la région catalytique de la sous-unité C1s du complexe C1 du complément humain.
***Directeur* de thèse** : Dr. Gérard Arlaud, LEM, IBS, Grenoble.
Financement : CEA (Contrat de Formation par la Recherche).

1992-1993 :

Diplôme d'Etudes Approfondies **(DEA)** de **Biologie Cellulaire et Moléculaire** à l'Université Joseph Fourier, Grenoble.

Sujet : Etude structurale du domaine catalytique de la protéase C1s du complément humain : réticulation chimique et modélisation moléculaire tridimensionnelle.
Directeur : Dr. Gérard Arlaud, LEM, IBS, Grenoble.

1991 :

Stage d'été (1 mois) Laboratoire d'immunochimie du Professeur M.G Colomb - INSERM U238 - CEA Grenoble

Sujet : Techniques d'études du facteur C3 du complément humain.
Responsable de stage : Dr. Serge Chesne.

Prix scientifiques et financements

Prix Amersham du meilleur étudiant en travaux pratiques - Glasgow University (1992)
Bourse de DEA (bourse d'excellence) – Université Joseph Fourier (1993)
Bourse de Thèse CEA – 1993-1997
Bourse pour la participation au congrés *XVIth International Complement Workshop, Boston* (1996)
Financement NIH – School of Biological Sciences (SBS), University of Missouri Kansas City (UMKC) (1997-1998).

Prime d'excellence scientifique (PES) de l'université Joseph Fourier depuis 2013
Financement AGIR pour la recherche (20 000 euros /2 ans) – Université Joseph Fourier (2014-2015)
Financement ANR SerpinGUTarget en tant que partenaire – (45 000 euros/3ans) (2015-2018)

Responsabilités collectives et animation scientifique

- **2009-2011 :** membre élu du conseil d'unité de l'Institut de Biologie Structurale
- **2011 :** membre nommé commission de spécialiste - poste MCU de biologie structurale N°1335 à l'UVHCI
- **2007 et 2008 :** membre élu de la commission de spécialiste de biologie (sections 64-69) – assesseur du bureau de la commission (rang B)
- **2004-2006 :** membre élu du conseil d'UFR de biologie
- Animation de la fête de la science IBS **2010, 2014**
- **2013 -** Membre rapporteur du jury de Thèse EPHE de Rémi Baroso
- Membre de deux comités de thèse.

  Rida Awad : début de thèse septembre **2010** : Equipe « Réponse immunitaire aux pathogènes et au soi altéré » (IBS), Caractérisation des régulateurs d'ELMO. Identification des ligands du domaine polyproline d'ELMO

  Alexandre Appolaire : début septembre **2010** : Equipe « Extremophile and Large Molecular Assemblies » (IBS) Caractérisation de complexes protéolytiques TET issus d'Archées hyperthermophiles

Co-auteurs et collaborations

En gras les co-auteurs de mes publications

Collaborations internes au laboratoire LEM-IBS

Directeur du laboratoire : **Arlaud G.J.**
Techniciens/Ingénieurs biochimistes : **Bally I., Lacroix M.B., Lunardi T.**
Chercheurs : **Thielens N.M.,** chercheur senior du laboratoire spécialiste des molécules du complément. **Hernandez J.F**, chimiste spécialiste de la synthèse chimique de peptides.
Post doctorants : **Chevallier S.**, **Cseh S.**
Etudiants : **Teillet F.**, étudiante en M2 ; **Ancelet S.,** étudiante en BTS ; **Jacquet M.**, étudiant en thèse

Collaborations internes à l'IBS - Grenoble

Laboratoire de cristallographie des protéines	Directeur : Juan **Fontecilla-Camps**
Christine Gaboriaud, chercheur, collaboratrice principale chargée de la résolution structurale par cristallographie aux rayons X des protéines du complément étudiées au sein de notre groupe	

Laboratoire de spectrométrie de masse	Directeur : Eric Forest
Pétillot Y., Jacquinod M, et Ulrich J. - Analyse des protéines en spectrométrie de masse electrospray	

Laboratoire de séquençage et analyse d'acides aminés	Directeur : Jean Gagnon
Gagnon J., Analyse d'acides aminés et séquençage	

Laboratoire de résonance magnétique nucléaire	Directeur : Dominique Marion
Gans P. et Blackledge M.- Analyse RMN d'un peptide de C1s	

Laboratoire des glucosaminoglycanes	Directeur : Hugues Lortat-Jacob
Vives R. et Sadir R. - Purification des polymères de glucosaminoglycanes	

Collaborations internes au laboratoire de biologie cellulaire et biophysique UMKC Kansas City

Directeur du laboratoire : **Esser A.F.**, Post-doctorant : **Wang Y.**

Collaborations externes au laboratoire de recherche

Collaborations internationales

Department of medical microbiology and immunology **University of Aahrus, Denmark**	
Jensenius J.C., Thiel, S. Jensen L., Moller-Kristensen M., Vorup Jensen T., collaboration dans le cadre des travaux sur la voie lectine du complément	

Department of Biochemistry and Molecular Biology **Monash University, Australia**	
Pike R.N., Kerr F.K., Thomas A.R., Wijeywickrema L.C., Whisstock J.C., Kaiserman D., Matthews A.Y., Bird P.I., collaboration dans le cadre d'une étude sur la spécificité enzymatique de MASP-2	

University at of Alabama **Birmingham, USA**	
Xu Y., Volanakis J.E., collaboration dans l'étude de la caractérisation de l'inhibition de la protéase C1r-like	

Collaborations nationales

Hopital Européen Georges Pompidou **Paris, France**	
Fremeaux-Bacchi V., ADNc de C1 inhibiteur	

Université de Reims **France**	
Dedieu S., collaboration dans le cadre du projet impliquant le récepteur LRP1 depuis 2014	

INRA MICALIS, Jouy en Josas **France**	
Maguin E., collaboration ANR SerpinGUTARGET – 3 ans de 2015 - 2018	

Encadrement d'étudiants

Moniteurs CIES

2004 - 2007 - Nicolas COQUELLE
2009 - 2012 - Mélanie VERNERET
2010 - 2013 - Delphine DAYDE

Stagiaires de BTS ANABIOTECH (2 ans en alternance)

2005-2007 - Marina LAMAIRIA - Recrutée au CHU, Grenoble.
2007-2009 - Sarah ANCELET - CDD assistant-ingénieur CNRS (ANR) dans le laboratoire (3 ans) puis CHU Grenoble.

Stagiaires de licence (étudiants de l'UJF - 1 mois)

L1
2006 - Céline CRETTALAZ, poursuite en thèse
2008 - Jennifer FRANCES
2010 - Emi CITTY
2013 - Benjamin FOUET, stage d'excellence UJF
L2
2001 - Céline LAFAYE, poursuite à l'INSA puis en thèse
2013 - François HAMBERT
L3
2008 - Eva GLEIZE, poursuite en école d'ingénieur de chimie de Rennes
2011 - Morgane BERTRAND (L3 Biologie) et Adeline BESSON (L3 Chimie/Biologie)
2014 - Julie LOPES (L3 Chimie/Biologie)

Stagiaires de M1

2008 et 2009 - Neda HOGHOUGHI, poursuite en thèse à l'IAB

Stagiaires de M2

2001-2002 - Florence TEILLET - M2 Biologie Structurale et Fonctionnelle – poursuite en thèse avec Nicole Thielens (LEM)
Sujet : *Ingénierie et caractérisation fonctionnelle d'une molécule chimère des protéases modulaires C1s et MASP-2 du complément.*

2008-2009 - Ludovic LENCLUME – M2 Biologie Structurale et Fonctionnelle
Sujet : *Caractérisation de l'activité enzymatique de la protéase à sérine MASP-3*

2009-2010 - Isma KERMIA – M2 Biologie cellulaire intégrative
Sujet : *Expression et purification du fragment N-terminal de la trombospondine-1 humaine et de la calréticuline de Trypanosoma cruzi.*

2009-2010 - David RODGERS - durée 2 mois et ½ dans le cadre du Wellcome Trust – University of GLASGOW. Subject : *Study of the interaction properties of the N-terminal fragment of human thrombospondin-1.*

Thèse d'université

2009 -2012 – Mickael JACQUET : Thèse Irtelis (CEA) – Encadrement à 80 % - thèse soutenue le 24 septembre 2012
Sujet : *Reconnaissance des signaux de danger et activation des mécanismes effecteurs de l'immunité innée.*

Liste des publications

22 réferences dans Pubmed avec un facteur H de 17, *IF (facteur d'impact 2013 sur une durée de 5 ans)*

Publication originales dans des journaux internationaux

P1 - **Rossi V.**, Gaboriaud C., Lacroix M., Ulrich J., Fontecilla-Camps J.C., Gagnon J. & Arlaud G.J. **(1995)** Structure of the catalytic region of human complement protease C1s: study by chemical cross-linking and three-dimensional homology modeling. ***Biochemistry*** **34**, 7311-7321. *(IF 3,17)*

P2 - Pétillot Y., Thibault P., Thielens N.M., **Rossi V.**, Lacroix M., Coddeville B., Spik G., Schumaker V.N., Gagnon J. & Arlaud G.J. **(1995)** Analysis of the N-Linked oligosaccharides of human C1s using electrospray ionisation mass spectrometry. ***FEBS Lett.*** **358**, 323-328. *(IF 3,48)*

P3 - Lacroix M., **Rossi V.**, Gaboriaud C., Chevallier S., Jacquinod M., Thielens N.M., Gagnon J. & Arlaud G.J. **(1997)** Structure and assembly of the catalytic region of human complement protease C1r: a three-dimensional model based on chemical cross-linking and homology modeling. ***Biochemistry*** **36**, 6270-6282. *(IF 3,17)*

P4 - **Rossi V.**, Bally I., Thielens N.M., Esser A.F. & Arlaud G.J. **(1998)** Baculovirus-mediated expression of truncated modular fragments from the catalytic region of human complement serine protease C1s. Evidence for the involvement of both complement control protein modules in the recognition of the C4 protein substrate. ***J. Biol. Chem***. **273**, 1232-1239. *(IF 5,02)*

P5 - Gans P., **Rossi V.**, Gaboriaud C., Bally I., Hernandez J.-F., Blackledge M. J. & Arlaud G. J. **(1998)** NMR structures of the C-terminal end of human complement serine protease C1s. ***Cell. Mol. Life Sci.*** **54,2:** 171-178. *(IF 6,36)*

P6 - Gaboriaud C., **Rossi V.**, Fontecilla-Camps J. C. & Arlaud G. J. **(1998)** Evolutionary conserved rigid module - domain interactions can be detected at the sequence level: the examples of complement and coagulation proteases. ***J. Mol. Biol.*** **282**, 459-470. *(IF 3,88)*

P7 - Gaboriaud C., **Rossi V.**, Bally I., Arlaud G.J. & Fontecilla-Camps J.C. **(2000)** Crystal structure of the catalytic domain of complement C1s : a serine protease with a handle. ***EMBO J.*** **19**, 1755-1765. *(IF 9,6)*

P8 - Thielens N.M., Cseh S., Thiel S., Vorup-JensenT., **Rossi V.**, Jensenius J.C. & Arlaud G.J. **(2001)** Interaction properties of human mannan-binding lectin (MBL)-associated serine proteases-1 and –2, MBL-associated protein 19, and MBL. ***J. Immunol***. **166**, 5068-5077. *(IF 5,67)*

P9 - **Rossi V.**, Cseh S., Bally I., Thielens N.M., Jensenius J.C. & Arlaud G.J. **(2001)**. Substrate specificities of recombinant mannan-binding lectin-associated serine proteases-1 and -2. ***J Biol Chem.*** **276**, 40880-40887. *(IF 5,02)*

P10 - Moller-Kristensen M., Jensenius J.C., Jensen L., Thielens N., **Rossi V.**, Arlaud G. & Thiel S. **(2003)** Levels of mannan-binding lectin associated serine protease-2in healthy individuals. ***J. Immunol. Methods***. **282**, 159-167. *(IF 2,45)*

P11 - Bally I., **Rossi V.**, Thielens N.M., Gaboriaud C. & Arlaud G.J. **(2005)** Functional Role of the Linker between the Complement Control Protein Modules of Complement Protease C1s. ***J. Immunol.*** **175**, 4536-4542 *(IF 5,67)*

P12 - **Rossi V.**, Teillet F., Thielens N.M., Bally I., Arlaud G.J. **(2005)** Functional characterization of complement proteases C1s/MASP-2 chimeras reveals the higher C4 recognition efficacy of the MASP-2 complement control protein modules. ***J. Biol. Chem.*** **280**, 41811-41818. *(IF 5,02)*

P13 - Kerr F.K., Thomas A.R., Wijeyewickrema L.C., Whisstock J.C., Boyd S.E., Kaiserman D., Matthews A.Y., Bird P.I., Thielens N.M., **Rossi V*.** & Pike R.N.* **(2008)** Elucidation of the substrate specificity of the MASP-2 protease of the lectin complement pathway and identification of the enzyme as a major physiological target of the serpin, C1-inhibitor. ***Mol. Immunol.*** **45**, 670-677. (* last co-authors). *(IF 2,92)*

P14 - Bally I., **Rossi V.**, Lunardi T., Thielens N.M., Gaboriaud C. & Arlaud G.J. **(2009)** Identification of the C1q Binding Sites of Human C1r and C1s. A refined Three-dimensional Model of the C1 Complex of Complement. ***J. Biol. Chem.*** **284**, 19340-19348. *(IF 5,02)*

P15 - **Rossi V.**, Wang Y. & Esser A.F. **(2010)** Topology of the membrane-bound form of complement protein C9 probed by glycosylation mapping, anti-peptide antibody binding, and disulfide modification. ***Mol. Immunol.*** **47**, 1553-1560. *(IF 2,92)*

P16 - **Rossi V.**, Bally I., Ancelet S., Xu Y., Frémeaux-Bacchi V., Vivès R.R., Sadir R., Thielens N. & Arlaud G.J.**(2010)** Functional Characterization of the Recombinant Human C1 inhibitor Serpin Domain: Insights into Heparin Binding. ***J. Immunol.*** **184**, 4982-4989. *(IF 5,67)*

P17 – Jacquet M., Lacroix M., Ancelet S., Gaboriaud C, Thielens N.M. & **Rossi V.**, **(2013)**, Deciphering complement receptor type 1 (CR1) interactions with recognition proteins of the lectin complement pathway***. J. Immunol*** **190**, 3721-31. *(IF 5,67)*

P18 – Gaboriaud C., Gupta R., Martin L., Lacroix M., Serre L., Teillet F., Arlaud G., **Rossi V.** & Thielens N.M. **(2013)**, Enzymatic properties of the serine protease domain of MASP-3 and its crystal structure in complex with ecotin. ***PLoS One***. Jul 4;8(7) e67960. *(IF 4,24)*

Articles de revues

O1 - Arlaud G.J., Volanakis J.E., Thielens N.M., Narayana S.V.L., **Rossi V.** & Xu Y. **(1998)** The atypical serine proteases of the complement system. ***Adv. Immunol***. **69**, 249-307. *(IF 8,02)*

O2 - Arlaud G.J., **Rossi V.**, Thielens N.M., Gaboriaud C., Bersch B., & Hernandez J.-F. **(1998)** Structural and functional studies on C1r and C1s: New insights into the mechanisms involved in C1 activity and assembly. In: *C1 and the Collectins in Health and Disease* (Colomb M.G., Loos M. & Reid K.B.M., Eds.) ***Immunobiology*** **199**, 303-316. *(IF 3,18)*

O3 - Arlaud G.J., Gaboriaud C., Thielens N.M., **Rossi V.**, Bersch B., Hernandez J.F. & Fontecilla-Camps J.C. **(2001)** Structural biology of C1: dissection of a complex molecular machinery. ***Immunol. Rev***. **180**, 136-145. *(IF 11,89)*

O4 - Arlaud G.J., Gaboriaud C., Thielens N.M., Budayova-Spano M., **Rossi V.** & Fontecilla-Camps J.C. **(2002)** Structural biology of the C1 complex of complement unveils the mechanisms of its activation and proteolytic activity. ***Mol. Immunol.*** **39**, 383-394. *(IF 2,92)*

O5 - Arlaud G.J., Gaboriaud C., Thielens N.M. & **Rossi V.** **(2002)** Structural biology of C1. ***Biochem. Soc. Trans***.(30) : 1001-1006. *(IF 3.21)*

O6 - Gaboriaud C., Thielens N.M., Gregory L.A., **Rossi V.**, Fontecilla-Camps J.C., & Arlaud G.J. **(2004)** Structure and activation of the C1 complex of complement: unraveling the puzzle. ***Trends Immunol.*** **25**, 368-373. *(IF 9,71)*

O7 - Gaboriaud C., Ling W-L, Thielens N.M., Bally I., & **Rossi V.** **(2014)** Deciphering the fine details of C1 assembly and activation mechanisms:'mission impossible'? ***Frontiers Immunol*** *(DOI : 10.3389/fimmu*.2014.00565) 5 : 565 (p1-7).

Chapitres dans des ouvrages

O8- Arlaud G.J., Thielens N.M., Rossi V., Gaboriaud C., Fontecilla-Camps J.C. **(2001)** Protéines de reconnaissance et protéases impliquées dans l'immunité innée. ***Regard sur la biochimie* n°3**. 15-20.

O9- Arlaud G.J., Thielens N.M., **Rossi V.**, Gaboriaud C., **(2004)** Complement component C1s. *Handbook of Proteolytic Enzymes 2nd Edn.* ***Academic Press*** *vol2, pp 1620-1622 (Barrett A.J., Rawlings N.D & Woessner J.F., Eds)*.

O10- Arlaud G.J., Thielens N.M., **Rossi V.**, Gaboriaud C., **(2011)** Complement component C1s. *Handbook of Proteolytic Enzymes 3rd Edn.* ***Academic Press*** *vol 3, chapter 632, pp 2853-2857 (Rawlings N.D & Salvesen., Eds)*.

O11 – **Rossi V.,** Bally I., Lacroix M., Arlaud G.J., Thielens N.M. **(2014)** Classical complement pathway components C1r and C1s – Purification from human serum and in recombinant form and functional characterization. Complement system methods and protocols. ***Methods Mol Biol.*** **1100**:43-60.

Liste des communications

C1 - **Rossi V.**, Gaboriaud C., Lacroix M., Ulrich J., Fontecilla-Camps J.C., Gagnon J. & Arlaud G.J. (1994) Structure de la région catalytique de la protéase C1s du complément. Etude biochimique et modélisation moléculaire. R*éunion IMABIO, La Grande Motte, France.*

C2 - **Rossi V.**, Gaboriaud C., Lacroix M.B., Ulrich J., Fontecilla-Camps J. C., Gagnon J. & Arlaud G.J. (1994) Structure of the catalytic region of the human complement protease C1s: biochemical and three-dimensional modeling studies. International workshop on: *Extracellular Modules: Sequence, Structure, Function and Evolution*, Margretetorp, Sweden.

C3 - Pétillot Y., Thibault P., Thielens N.M., **Rossi V.**, Lacroix, M., Coddeville B., Spik G., Schumaker V.N., Gagnon J. & Arlaud G.J. (1995) Analysis of the N-linked oligosaccharides of human C1s using ESI-MS. *43rd ASMS Conference on Mass Spectrometry and Allied Topics, Atlanta, USA.*

C4 - **Rossi V.**, Gaboriaud C., Lacroix M., Ulrich J., Fontecilla-Camps J.C., Gagnon J. & Arlaud G.J. (1995) Structure de la région catalytique de la protéase C1s du complément. Réticulation chimique et modélisation moléculaire. *XXIIème Forum des Jeunes Chercheurs de la SFBBM, Grenoble, France.*

C5 - Lacroix M.B., **Rossi V.**, Gaboriaud C., Pétillot Y., Thielens N.M., Gagnon J. & Arlaud G.J. (1996) Structure of the catalytic regions of human C1r. Study by chemical cross-linking and 3-D homology modeling. *Mol. Immunol.* **33**, 19.*XVIth International Complement Workshop, Boston, USA.*

C6 - Thielens N.M., **Rossi V.**, Bally I., Esser A. F. & Arlaud G.J. (1996) Production de protéines recombinantes dans le système baculovirus/cellules d'insectes à l'aide du kit BAC-TO-BAC™ expression de fragments des protéines C1r et C1s du complément. *48 h de Biologie structurale – Autrans, France.*

C7 - Gaboriaud C., **Rossi V.**, Fontecilla J. C. & Arlaud G. J. (1997) Conservation through evolution of the close interaction observed between the last CCP module and the serine protease domain of the C1r and C1s complement proteases. *IInd International Conference on Molecular Structural Biology,Vienna, Austria.*

C8 - Lacroix M., **Rossi V.**, Gaboriaud C., Chevallier S., Jaquinod M., Thielens N. M., Gagnon J. & Arlaud G. J. (1997) A three-dimensional model of the catalytic region of human C1r. *IVth International Workshop on C1 and the Collectins, Mainz, Germany.*

C9 - Gaboriaud C., **Rossi V.**, Fontecilla J. C. & Arlaud G. J. (1997) Conservation through evolution of the close interaction observed between the last CCP module and the serine protease domain of C1r and C1s. *IVth International Workshop on C1 and the Collectins, Mainz, Germany.*

C10 - **Rossi V.**, Bally I., Thielens N. M., Esser A. F. & Arlaud G. J. (1997) Baculovirus-mediated expression of truncated fragments from the catalytic region of human C1s: evidence for the involvement of both CCP modules in C4 recognition. *IVth International Workshop on C1 and the Collectins, Mainz, Germany.*

C11 - **Rossi V.**, Xia C. & Esser A.F. (1998) Effect of disulfide-mutagenesis on folding and functional activity of human C9. *Mol. Immunol.* ***35**, 6-7. XVIIth International Complement Workshop, Rhodes, Greece.*

C12 - **Rossi V.**, Tatar L.D., Beer K.B., Miller V.L. & Esser A.F. (1998) Structural studies of Ail-mediated resistance of *E.coli* to killing by complement. *Mol. Immunol.* ***35**, 6-7. XVIIth International Complement Workshop, Rhodes, Greece.*

C13 - **Rossi V**., Serban D., Wang Y., Shivanna B. & Esser A.F. (1999) Membrane insertion and structure of complement protein C9. *Meeting on Lipid protein Interplay: Mechanisms and Implication for cell Function, Jena, Germany.*

C14 - **Rossi V**., Bally I., Thielens N.M. & Arlaud G.J. (1999) Functional topology of the catalytic region of human C1s: screening using site-directed mutagenesis. *7th European meeting on Complement in Human Disease, Helsinki, Finland.*

C15 - Gaboriaud C., **Rossi V.**, Bally I., Arlaud G. & Fontecilla J.C. (1999) Structure du domaine catalytique de C1s humain, initiateur de la cascade du complément. *Hourtin, France.*

C16 - **Rossi V.**, Gaboriaud C., Bally I., Fontecilla J.C. & Arlaud G. J. (1999) Topologie fonctionnelle de la région catalytique de C1s. *48 heures de Biologie Structurale, Autrans, France.*

C17 - **Rossi V.**, Gaboriaud C., Bally I., Fontecilla J.C. & Arlaud G. J. (2000) Functional topology of the catalytic region of human C1s: insights into the molecular determinants of C1s specificity. *XVIIIth International Complement Workshop, Salt Lake City, USA.*

C18 - Thielens N., Cseh S., Thiel S., **Rossi V.**, Jensenius J.C. & Arlaud G. J. (2000) Expression and characterization of the Map19 component of the MBL pathway of complement. *XVIIIth International Complement Workshop, Salt Lake City, USA.*

C19 - Gaboriaud C., **Rossi V.**, Bally I., Arlaud G. J. & Fontecilla J.C. (2000) X-Ray Structure of the human C1s catalytic domain: a protease with a CCP module handle. *XVIIIth International Complement Workshop, Salt Lake City, USA.*

C20 - **Rossi V**., Cseh S., Thielens N., Bally I., Jensenius J. & Arlaud G. (2000) Expression and proteolytic activities of the human MBL-associated serine proteases MASP-1 and MASP-2. *Irst Innate Immunity Workshop, Santorini, Greece.*

C21 - **Rossi V.,** Bally I., Thielens N. & Arlaud G.J. (2001) Topologie fonctionnelle de la region catalytique de C1s Humain. *Journées DBMS/IBS, Grenoble, France.*

C22- Thielens N., Cseh S., Thiel S., Vorup-Jensen T., **Rossi V.**, Jensenius J. & Arlaud G.J. (2001) Propriétés d'interaction des protéases à sérine MASP-1 et MASP-2, de la protéine Map19, et de la Mannan-Binding Lectin (MBL). *Journées DBMS/IBS, Grenoble, France.*

C23- **Rossi V.**, Cseh S., Bally I., Thielens N.M., Jensenius J.C. & Arlaud G.J. (2001) Substrate specificities of recombinant mannan-binding lectin-associated serine proteases-1 and –2. *Vth international workshop on C1 and the collectins, Frankfurt, Germany.*

C24 - Bally I., **Rossi V.**, Thielens N.M. & Arlaud G.J. (2002) Residue GLN339, at the interface between the CCP1 and CCP2 modules of C1s, is a key element of C4 recognition. *XIXth International Complement Workshop, Palerm, Italia.*

C25- **Rossi V.**, Teillet F., Bally I. & Arlaud G.J. (2002) C1s/MASP-2 Chimeras: tools to determine the relative contributions of the CCP modules and serine protease domain of MASP-2 to its higher C4 cleaving activity. *XIXth International Complement Workshop, Palermo, Italia.*

C26 - **Rossi V.**, Teillet F., Bally I. & Arlaud G.J. (2002) CHIMERES des protéases modulaires C1s et MASP-2 : des outils pour déterminer la contribution relative des modules CCP et du domaine protéase à sérine de MASP-2 dans sa spécificité pour son substrat C4. *Forum des jeunes chercheurs – Dijon*, France.

C27 - Arlaud G.J., Gaboriaud C., Thielens N.M., Juanhuix J., Gregory L., **Rossi V.**, Budayova-Spano M. & Fontecilla-Camps J.C. (2003) Structural biology of the C1 complex of complement sheds light on its enzymatic and recognition properties. *2nd workshop on complement associated diseases, animal models, and therapeutics. Myconos, Greece.*

C28 - **Rossi V.**, Teillet F., Bally I. & Arlaud G.J. (2003) The CCP modules of MASP-2 recognize C4 more efficiently, but the serine protease domain of C1s is catalytic more active. *9th European meeting on complement in human disease. Trieste, Italia.*

C29 - Bally I., **Rossi V.**, Thielens N.M., Gaboriaud C. & Arlaud .G.J. (2008) Identification of the C1q binding sites of C1r and C1s. A refined 3-D model of the C1 complex. *XXIIth International Complement Workshop, Bale, Suisse.*

C30 - **Rossi V.**, Bally I., Xu Y., Frémeaux-Bacchi V., Vivès R., Ancelet S. & Arlaud G.J. (2010) Expression and functional characterization of the serpin domain of human C1 inhibitor. *12th European meeting on complement and Human Diseases (EMCHD), Visegrad, Hongary.*

C31 – Jacquet M.C., Lacroix M., Gaboriaud C., Thielens N. & **Rossi V.** (2011) Deciphering complement receptor 1 (CR1) interaction with defense collagens. *13h European meeting on complement and Human Diseases (EMCHD), Leiden, Holland.*

C32 - Bally I., Gout E., Wicker-Plancquart C., Kleman J.P., Chouquet A., Thielens N.M. and **Rossi V.** (2015) Role of defense collagens in efferocytosis mediated by LRP1. Development of molecular and cellular tools for the dissection of LRP1 interaction mechanisms. *15h International European meeting on complement and Human Diseases (EMCHD), Uppsala, Sweden.*

Activité d'enseignement

Depuis septembre 2000 : Maître de conférences en Biologie, Université Joseph Fourier
(192 heures, équivalent TD par an)
50% en premier cycle (L1, L2 et L3) et 50% en deuxième cycle (M1)

- **Participation active** à la mise en place de la nouvelle habilitation de l'Université Joseph Fourier pour l'harmonisation européenne des diplômes (système LMD) en 2003. Conception des modules de biochimie Bio 121 de L1 et Bio 241 de L2. Implication dans l'élaboration des travaux pratiques dans chacun de ces deux modules à Grenoble.

- **Responsable** de l'installation de la plateforme de travaux pratiques de biologie de L1, L2 et L3 dans l'antenne de l'Université Joseph Fourier de Valence.

- **Encadrement** de **2008 à 2011** de groupes d'élèves de première et terminale avec leur professeur pour des TP de biochimie dans le cadre de l'ouverture au public des Lycées de la plateforme d'enseignement pratique de biologie de l'université Joseph Fourier (CUBE).

2014- **Développement de supports pédagogiques** pour les premiers cycles – (i) mini films d'apprentissage des techniques pratiques de laboratoire, (ii) application pour tablettes numériques d'apprentissage des structures de base des macromolécules biologiques.

2010 - **Responsable du parcours de la Licence internationale de Biologie** ouvert en 2011. Mise en place du parcours. Cette licence a une double vocation: (i) former les étudiants Français à l'apprentissage des sciences en anglais scientifique et (ii) ouvrir notre parcours de licence de biologie à un public d'étudiants étrangers anglophones.

https://dlst.ujf-grenoble.fr/index.php?module=classique&url=international/BIOI_ang.html

>Practical informations

>Practical informations

>Why Grenoble

>Why Grenoble ?

To be in Grenoble is a great opportunity to get immersed in a multicultural community studying arts and science. Grenoble's universities host 60 000 students with more than 9000 foreigners each year. The city is the leading research center in France outside Paris area and is often described as the "French silicone valley" with booming high tech businesses.
You will leave this university with a fluent command of French and a new insight in French history, literature and way of life.

Having fun in Grenoble:
Skiing
Mountain hiking
(fauna, flora of the Alps)
An outstanding art museum
Concert halls, theatres
The heart of Europe: 3h TGV to Paris, 3h

TGV to the Mediterranean Sea, Driving distance from Spain, Italy, Germany, Switzerland…

Enjoy the great French way of life – You may become addicted!

Londres
Berlin
Varsovie
Kiev
Francfort
Luxembourg
Prague
Paris
Vienne
Budapest
Genève
Lyon
Milan
GRENOBLE
Madrid
Rome

Contact

Programme coordinator: veronique.rossi@ibs.fr
Secretary: marie-pierre.bassani@ujf-grenoble.fr

Postal address : Licence de Biologie Internationale
DLST – 480, avenue CENTRALE
DOMAINE UNIVERSITAIRE
BP53
38041 Grenoble Cedex 9

www-biologie.ujf-grenoble.fr

Conception graphique : service com UJF / crédits photos : SHUTTERSTOCK service com UJF, EH/service com UJF

Université
Joseph Fourier
GRENOBLE

Bachelor of science in Biology

Grenoble
FRANCE

>Bachelor of sciences in Biology

>Bachelor of sciences in Biology

University Joseph Fourier (UJF-Grenoble), one of the best French universities according to international ranking, welcomes the students of any nationality for a full programme of undergraduate biology studies in English. Grenoble's laboratories are among the best in Europe and our teaching programmes are closely linked to their research themes. Our University offers a great opportunity for biology students to get intense experimental training using the practical work platform (CUBE)* and summer internship in our research institutes.

BIODIVERSITY

STRUCTURAL BIOLOGY

CELL BIOLOGY

PHYSIOLOGY

The programmes are open to groups of students coming for a semester from our partner universities around the world under a student exchange agreement. The foreign students will be mixed with French students.

Any course of the programme can be taught in English and French courses for foreigners (any level) are offered at any semester.

For additional information or customized arrangements, please contact relations.internationales@ujf-grenoble.fr

CUBE
http://www-biologie.ujf-grenoble.fr/CUBE/

>International Programme of Biology

>International Programme of Biology

SEMESTER 1
- Introduction to Cell Biology
- General Chemistry
- Algebra and Elementary Analysis
- Physics for Life Sciences
- Multimedia Informatics

SEMESTER 2
- Molecular Cell Components
- Animal and Plant Biology
- Organic and Inorganic Chemistry
- Dynamical Systems and Modelling
- French courses

SEMESTER 3
- Cell Biology
- Genetics
- Thermodynamics and Kinetics
- Statistics for Life sciences
- French courses

SEMESTER 4
- Biochemistry1
- Physiology
- Cell Imaging or Ecology
- Supervised Experimental Project
- French Courses

SEMESTER 5
- Experimental Training in Molecular and Cell Biology 1 and 2
- Biochemistry 2
- Molecular basis of Gene Expression
- French Courses

SEMESTER 6
- Developmental Biology
- Cellular Neurobiology
- Bioenergetics
- Structural Biology
- Ecology

experimental biology

The curriculum could be adapted to fit specific requirement

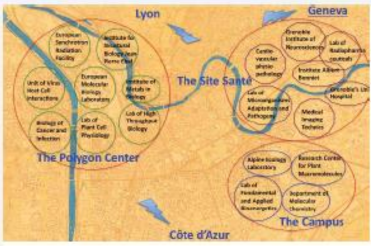
The courses content can be consulted at http://dlst.ujf-grenoble.fr/international/BIOl_ang.html

>Biology research in Grenoble

>Biology research in Grenoble

Grenoble is a fertile environment, famous for research and technology involving 20% of its active population. Science is the driving force of the town and our Biology laboratories provide a vast playground for a strong experimental training encompassing the main fields of modern biology. Our research laboratories are geographically grouped on three specialised campuses.

>Grenoble Institutes for Research in Biology


Lyon
Geneva
European Synchrotron Radiation Facility
Institute for Structural Biology Jean Pierre Ebel
Unit of Virus Host Cell Interactions
European Molecular Biology Laboratory
Institute of Metals in Biology
Biology of Cancer and Infection
Lab of Plant Cell Physiology
Lab of High Throughput Biology
The Polygon Center
The Site Santé
Grenoble Institute of Neurosciences
Lab of Radiopharmaceuticals
Cardio-vascular physio pathology
Institute Albert Bonniot
Grenoble's University Hospital
Lab of Microorganisms Adaptation and Pathogeny
Medical Imaging Technics
Alpine Ecology Laboratory
Research Center for Plant Macromolecules
Lab of Fundamental and Applied Bioenergetics
Department of Molecular Chemistry
The Campus
Côte d'Azur

>Research Laboratories

THE POLYGON CENTER:
8 research institutes including 2 prestigious European research centres (European Synchrotron Research Facility (ESRF) and the European Molecular Biology Laboratory (EMBL). Research areas include Structural Biology, Plant Biology and Biology of viruses and cancer.

THE SITE "SANTÉ":
7 health research centres comprising medical imaging, neurosciences, cancer, and hospital research.

THE CAMPUS:
4 research laboratories in the field of Molecular Chemistry, Ecology and Bioenergetics.

http://www.ujf-grenoble.fr/recherche/chimie-biologie-sante/laboratoires/

Expérience en recherche

COMPETENCES : d'une manière générale, étude des **relations structure-fonction** des protéines, biochimie, biologie moléculaire, enzymologie et biologie structurale.

Techniques biochimiques :

- Purification de protéines et peptides sur colonnes par HPLC et FPLC, (chromatographie d'affinité, filtration sur gel, chromatographie en phase inverse, échange d'ions et chromatofocalisation)
- Modifications chimiques des protéines (réticulation, clivages chimiques, protéolyse ménagée)
- Analyse de résultats de séquençage protéique et de spectrométrie de masse
- Western blot, tests ELISA
- Marquage radioactif des protéines (tritium, iode 125, Carbone 14)
- Cinétique enzymatique
- Interaction en phase solide et SPR

Modélisation moléculaire par homologie.

Biologie moléculaire :

- PCR, clonage, séquençage, mutagenèse dirigée
- Production de protéines recombinantes dans les systèmes eucaryotes (baculovirus/cellules d'insecte, levures, cellules de mammifères) et dans les systèmes bactériens

Langues étrangères:

- Anglais : lu, parlé, écrit, couramment; année de maîtrise à l'université de Glasgow (programme ERASMUS), stage post-doctoral de 18 mois aux Etats-Unis à l'université de Kansas City (chercheur associé).
- Italien : lu, parlé, couramment.

Chapitre II

Présentation générale de l'activité de recherche

2.1 - Préambule

En tant que biochimiste de formation, mon centre d'intérêt dès mes débuts a porté sur les **relations structure-fonction** de plusieurs **protéines du complément humain.** L'ensemble de ces protéines est soluble et d'organisation structurale complexe, souvent modulaire, ce qui rend les études structurales et fonctionnelles difficiles à aborder sur l'ensemble de la molécule. Pour cette raison, nous avons développé une stratégie d'étude basée sur de la **dissection moléculaire,** dans laquelle des portions de molécules sont produites dans des systèmes hétérologues, principalement des systèmes d'expression eucaryotes, et caractérisées du point de vue fonctionnel et structural.

Le **complément** est un système plasmatique qui représente **l'élément majeur de la réponse innée.** Son rôle dans l'élimination des pathogènes (bactéries, virus) est bien connu depuis longtemps et est induit par la reconnaissance de PAMPs (Pathogen-Associated Molecular Patterns). En revanche, la mise en évidence du rôle du complément dans l'élimination des éléments du soi altéré ACAMPs (Apoptotic Cell-Associated Molecular Patterns) est beaucoup plus récente.

Dans le cas de la lutte contre les pathogènes, le complément génère trois types d'effets biologiques synergiques : l'inflammation, la phagocytose et la lyse cellulaire. Dans le cas de l'élimination des cellules apoptotiques ou autres cellules du soi modifiées, il induit une élimination « silencieuse » anti-inflammatoire qui contribue au maintien de la tolérance immune. L'ensemble des mécanismes effecteurs liés à l'élimination des corps étrangers est générée par l'activation en cascade de protéases et est finement régulé.

On distingue trois voies d'activation du complément en réponse à différents éléments déclencheurs, la **voie alterne, la voie classique** et **la voie lectine** qui a été découverte plus récemment, il y a une vingtaine d'années. Le disfonctionnement du système du complément aboutit à des maladies inflammatoires et/ou auto-immunes c'est pourquoi la connaissance de ses mécanismes d'action présente à la fois un intérêt fondamental et un intérêt thérapeutique dans diverses situations pathologiques.

Dans ce contexte biologique, mes travaux de recherche ont porté d'une part sur les complexes activateurs des voies **classique et lectine** qui **déclenchent** l'action du complément, et d'autre part sur ceux qui exercent des effets biologiques aboutissant à leur **destruction** (Figure 1).

Mon travail de thèse a été consacré à l'étude des **relations structure-fonction de la protéase à sérine C1s** de la **voie classique** d'activation du complément humain. Ce projet a nécessité la mise en place et la pratique de collaborations scientifiques avec des spécialistes de disciplines complémentaires au sein de l'Institut de Biologie Structurale dont la principalecollaboratrice en biologie structurale, Christine Gaboriaud du Laboratoire de Cristallogénèse et Cristallographie des Protéines (LCCP).

Mes travaux de recherches se sont poursuivis par un **stage post-doctoral aux Etats-Unis** qui m'a conduit à l'étude les interactions protéine/membrane dans deux cas : (i) la **protéine cytotoxique C9** qui contribue à l'élimination des micro-organismes par **lyse cellulaire** et (ii) l'insertion membranaire de la **protéine AIL de *Yersinia enterocolitica*** qui est responsable de la résistance de cette bactérie pathogène à l'attaque par le système du complément.

Jusqu'en 2009, mes travaux de recherche ont porté principalement sur l'étude des déterminants moléculaires impliqués dans l'activité enzymatique (reconnaissance spécifique des substrats et catalyse) des protéases C1s et MASP-2 qui sont responsables du déclenchement des voies classique et lectine du complément. Ces travaux ont été abordés sous trois angles différents : (i) **Cartographie fonctionnelle** par mutagenèse dirigée de la région catalytique de C1s**,** (ii) caractérisation fonctionnelle de la **protéase à sérine MASP-2** de la **voie lectine** et (iii) dissection du mécanisme d'inhibition de C1s par son inhibiteur physiologique **C1 inhibiteur** et de sa potentialisation par l'héparine. Ces travaux ont été conduits dans le **laboratoire d'Enzymologie Moléculaire (LEM)** dirigé par Gérard Arlaud.

Depuis 2010, Nicole Thielens a pris la direction d'un nouveau groupe, issu du regroupement de 3 équipes de l'IBS, nommé **IRPAS (Réponse immunitaire aux pathogènes et au soi altéré)**. Je fais partie de l'équipe de Nicole Thielens qui est l'équipe de « biochimistes » du groupe, spécialiste depuis de nombreuses années des relations structure-fonction des protéines du complément. Cette équipe comprend en tout 6 personnes permanentes, 3 chercheurs, Christine Gaboriaud biologiste structural, Nicole Thielens et moi-même, et 3 ingénieurs, Isabelle Bally (CEA), Evelyne GOUT (CNRS) et Jean Pierre Andrieu (CEA, plateforme de séquençage protéique). Dans cette nouvelle équipe, j'ai eu l'opportunité de réorienter ma thématique de recherche sur les récepteurs des collagènes de défense dont le rôle d'opsonine conduit à la phagocytose des éléments que l'organisme doit éliminer. J'ai encadré Mickael Jacquet en thèse, dans le cadre d'un sujet sur le récepteur CR1 qui est impliqué dans la phagocytose des complexes immuns reconnus par les collagènes de défense. J'ai également abordé le rôle de la thrombospondine 1 en tant que plateforme de reconnaissance des collagènes de défense.

Depuis 2014, je suis porteur d'un projet ayant pour objectif de caractériser le rôle du récepteur CD91/ LRP1 dans la phagocytose des cellules apoptotiques **(efferocytose),** un projet qui est central dans la nouvelle thématique de l'ensemble du groupe et que je décrirai dans la première partie de mes projets de recherche.

Je consacrerai une deuxième partie de mes projets de recherche à la **caractérisation enzymatique d'inhibiteurs procaryotes de protéases à sérine, les Serpines**, dans le cadre d'une ANR SerpinGUTARGET obtenue en tant que partenaire de l'équipe d'Emmanuelle Maguin de l'INRA MICALIS de Jouy en Josas.

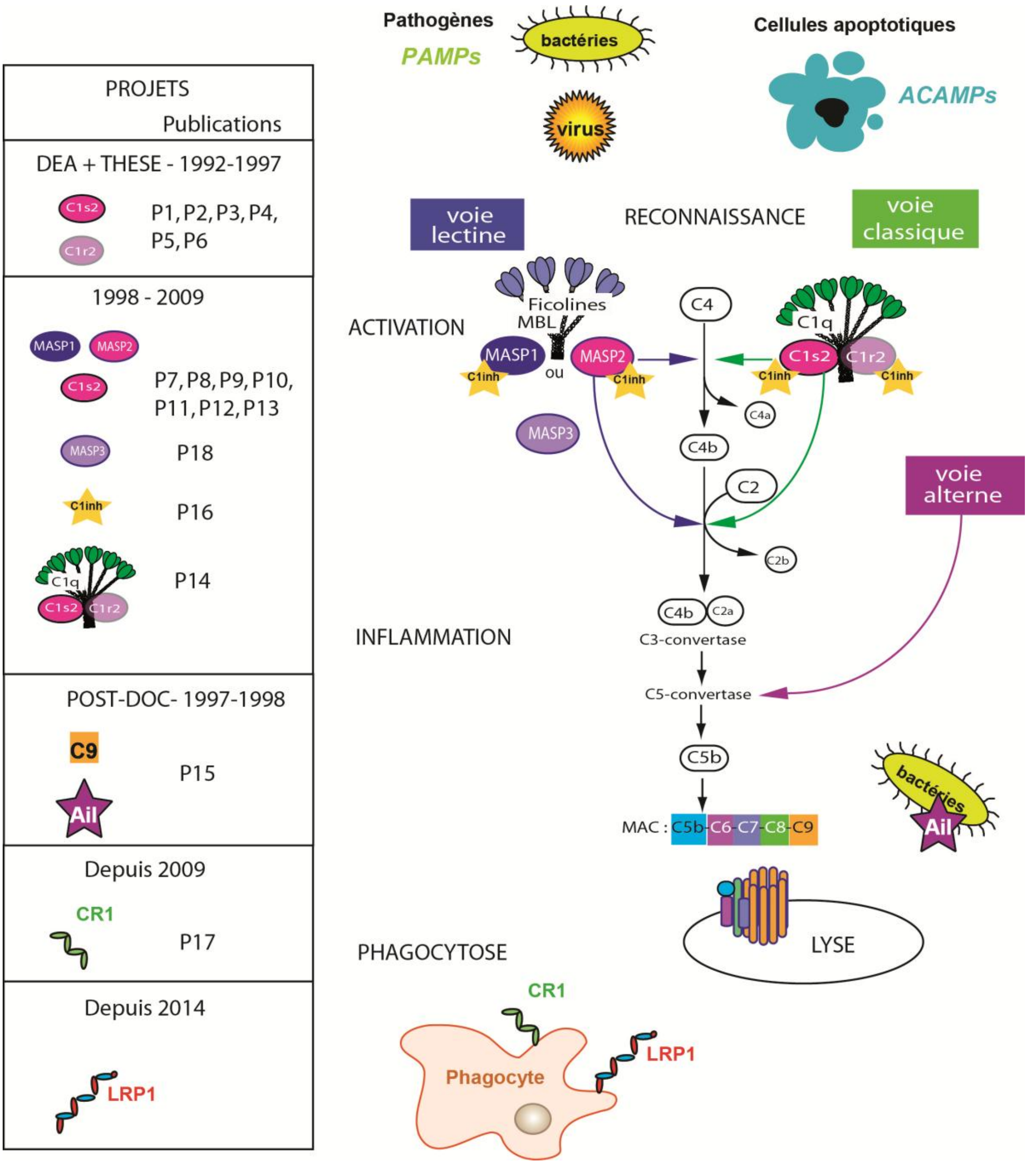
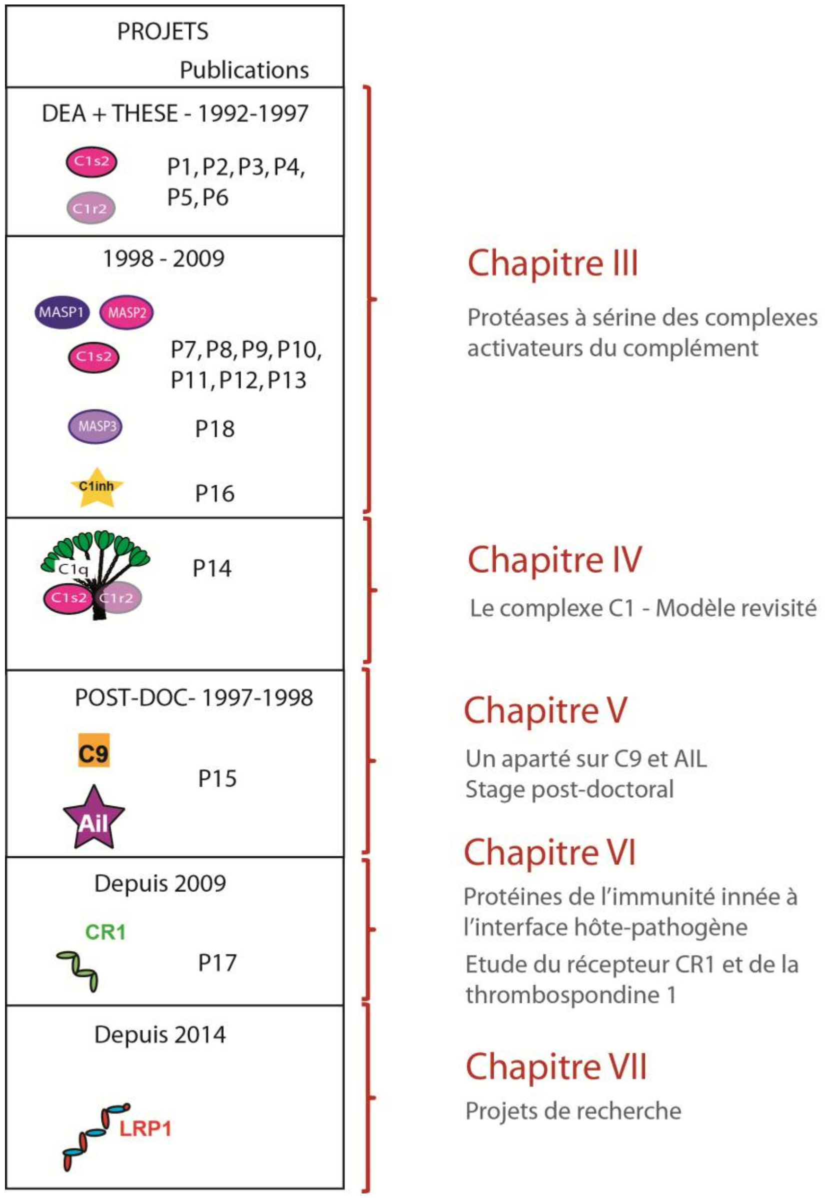

PROJETS
Publications
DEA + THESE - 1992-1997
C1s2
C1r2
P1, P2, P3, P4, P5, P6
1998 - 2009
MASP1
MASP2
C1s2
P7, P8, P9, P10, P11, P12, P13
MASP3
P18
C1inh
P16
C1q
C1s2
C1r2
P14
POST-DOC- 1997-1998
C9
Ail
P15
Depuis 2009
CR1
P17
Depuis 2014
LRP1
Pathogènes
PAMPs
bactéries
virus
Cellules apoptotiques
ACAMPs
voie lectine
RECONNAISSANCE
voie classique
ACTIVATION
Ficolines
MBL
MASP1
ou
MASP2
C1inh
C1inh
MASP3
C4
C4a
C4b
C1q
C1s2
C1r2
C1inh
C1inh
C2
C2b
voie alterne
INFLAMMATION
C4b
C2a
C3-convertase
C5-convertase
C5b
MAC : C5b-C6-C7-C8-C9
bactéries
Ail
LYSE
PHAGOCYTOSE
CR1
LRP1
Phagocyte

Figure 1 : Les voies classique et lectine d'activation du complément et les différents sujets étudiés

La voie classique est déclenchée par la reconnaissance directe, ou par l'intermédiaire d'un anticorps, d'un agent pathogène (PAMPs) ou d'un élément du soi altéré (ACAMPs) par C1q. La voie lectine fait intervenir soit la MBL soit une ficoline pour la reconnaissance spécifique de motifs oligosaccharidiques de surface. C1q est associé à un tétramère de deux protéases à sérine C1r et C1s tandis que la MBL ou les ficolines sont assocciées à un dimère soit de MASP-1 soit de MASP-2. Il existe une protéase homologue MASP-3 dont on ne connait encore pas le rôle précis. C1s et MASP-2 déclenchent la cascade protéolytique du complément par coupure de C2 et C4 qui génère trois types d'effets biologiques : (i) L'inflammation, sous l'effet des polypeptides C3a et C5a, (ii) la phagocytose causée par l'opsonisation des microorganismes par C3b et (iii) la lyse, par formation d'un complexe lytique (MAC, Complexe d'Attaque Membranaire) constitué des protéines C5b à C9 à la surface du microorganisme. L'inhibiteur physiologique de C1s, C1-inhibiteur est une serpine (SERine Protease INhibitor) et inhibe également C1r et MASP-1 et 2. Les collagènes de défense, C1q, MBL ou ficolines ont également un rôle d'opsonines et participent à l'élimination des agents opsonisés par interaction directe avec des récepteurs sur les phagocytes comme CR1 et LRP1. Les publications issus de ces projets sont indiquées dans le tableau de gauche.

Dans ce manuscrit, chaque chapitre comporte une description résumée des résultats principaux. Le détail peut être consulté dans les publications originales ajoutées en annexe que je présente dans leur intégralité pour les plus marquantes et avec la première page pour les autres.


PROJETS
Publications
DEA + THESE - 1992-1997
C1s2
C1r2
P1, P2, P3, P4,
P5, P6
1998 - 2009
MASP1
MASP2
C1s2
P7, P8, P9, P10,
P11, P12, P13
MASP3
P18
C1inh
P16
C1q
C1s2
C1r2
P14
POST-DOC- 1997-1998
C9
P15
Ail
Depuis 2009
CR1
P17
Depuis 2014
LRP1
Chapitre III
Protéases à sérine des complexes activateurs du complément
Chapitre IV
Le complexe C1 - Modèle revisité
Chapitre V
Un aparté sur C9 et AIL
Stage post-doctoral
Chapitre VI
Protéines de l'immunité innée à l'interface hôte-pathogène
Etude du récepteur CR1 et de la thrombospondine 1
Chapitre VII
Projets de recherche

2.2 - Introduction générale

Les généralités concernant mes sujets de recherche sont décrites dans cette partie, les points précis seront détaillés dans chaque partie traitée.

2.2.1 - Le complément, un système majeur de l'immunité innée

L'organisme humain est quotidiennement en contact avec des éléments potentiellement dangereux. Il doit lutter contre des agents infectieux externes (du NON-SOI) mais se prémunir également contre l'accumulation nocive d'éléments internes modifiés (du SOI). La protection de l'individu est assurée par un système immunitaire très spécialisé qui orchestre deux ensembles en synergie : l'immunité innée et l'immunité acquise. Lors d'une primo-infection c'est le système immunitaire inné qui entre immédiatement en action faisant intervenir une cohorte de protéines humorales et de cellules phagocytaires circulantes qui assurent la première barrière de lutte contre les agents à éliminer. Le système adaptatif entre ensuite en jeu pour produire une défense « mémoire » qui nécessite la production de protéines spécialisées à grande spécificité, les anticorps. L'immunité innée est un système très ancien qui constitue l'élément de lutte unique chez les invertébrés et qui au cours de l'évolution a été conservé chez les vertébrés pour assurer une protection optimisée en relation avec le système adaptatif.

L'immunité innée représente donc une première barrière de protection de l'organisme assurée par la veille permanente de molécules solubles ou des récepteurs membranaires constitutifs, les « pattern recognition molecules » ou PRMs. Ces protéines agissent immédiatement lorsqu'elles reconnaissent des motifs moléculaires spécifiques des pathogènes (PAMPs "Pathogen-associated molecular patterns") ou des éléments du SOI modifié (ACAMPs "apoptotic cell-associated molecular patterns").
Le complément humain est un acteur majeur du système immunitaire inné découvert à la fin du XIX siècle par Paul Ehrlich (1890) et Jules Bordet (1898). Initialement décrit comme un système intervenant en « complément » du système adaptatif, il a été montré depuis que ce terme « complément » s'est avéré par la suite sous-estimé par rapport à ses multiples effets. Le complément orchestre des processus inflammatoires et immunologiques qui jouent bien d'avantage que le simple rôle d'élimination d'éléments de « danger ». En effet, au cours de ces dix dernières années il a été montré que le complément excerce le rôle d'un système complexe de surveillance immune, capable de faire la discrimination entre les cellules saines de l'hôte, les débris cellulaires et les agents infectieux externes et d'adapter son action en conséquence. En plus de son implication dans l'élimination des microbes, le complément est maintenant reconnu comme participant à des

processus divers tels que le contrôle de l'homéostasie cellulaire du soi, la maturation des synapses, l'angiogenèse, la mobilisation des cellules hématopoiétiques.. (Ricklin et al., 2010). Cette versatilité n'est pas étonnante puisque le complément représente l'un des plus anciens systèmes de l'immunité innée (800 millions d'années) et a évolué en parallèle d'autres systèmes plus récents, tel que le système adaptatif (450 millions d'années) (Figure 2) (Sunyer et al., 1998).

Je me limiterai à décrire ici le fonctionnement du complément dans son rôle d'élimination des pathogènes et des éléments du soi altéré.

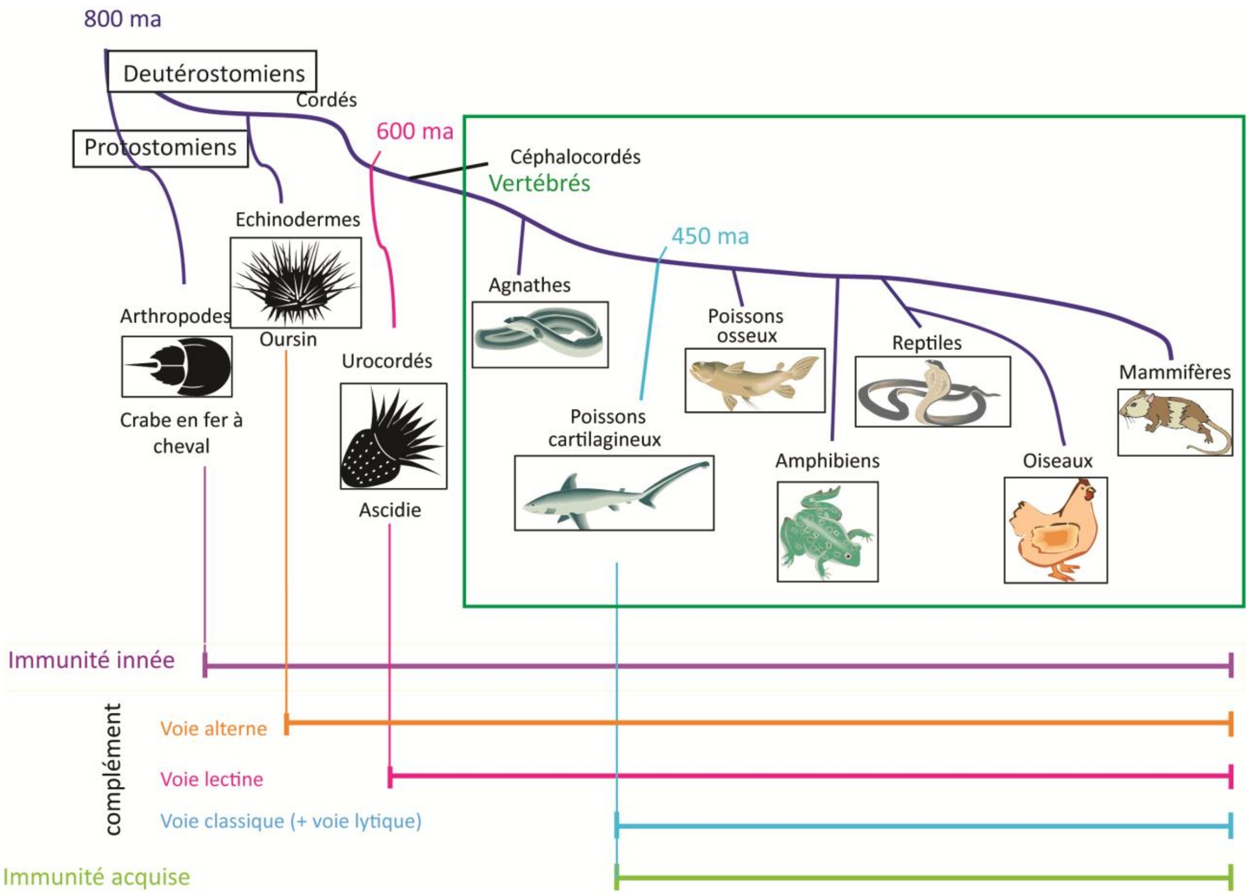

800 ma
Deutérostomiens
Cordés
Protostomiens
600 ma
Céphalocordés
Vertébrés
Echinodermes
450 ma
Agnathes
Arthropodes
Oursin
Poissons osseux
Reptiles
Mammifères
Urocordés
Crabe en fer à cheval
Poissons cartilagineux
Amphibiens
Oiseaux
Ascidie
Immunité innée
complément
Voie alterne
Voie lectine
Voie classique (+ voie lytique)
Immunité acquise

Figure 2 : Arbre phylogénique simplifié de l'évolution des principaux groupes du monde animal et de leur système immunitaire.
Adapté d'après Sunyer et al.,1998. ma, millions d'années.

2.2.2 - Les complexes activateurs des voies classique et lectine

Le complément est composé d'une trentaine de protéines qui circulent dans le plasma sous forme de précurseurs inactifs et de protéines exprimées à la surface de cellules endogènes. L'activation du complément résulte de la reconnaissance de signaux à la surface des éléments à éliminer par des molécules de reconnaissance (PRMs) et déclenche une réaction d'activation en cascade qui conduit à l'élimination des éléments ciblés. Trois voies d'activation distinctes sont connues à ce jour, la voie classique, la voie lectine et la voie alterne. Toutes convergent pour générer les effets biologiques induisant l'élimination des PAMPS ou ACAMPS (inflammation, la phagocytose et lyse cellulaire) (Figure 1).

Les voies classique et lectine ne diffèrent que par la nature de leurs complexes activateurs. Ces complexes sont relativement homologues puisqu'ils sont composés d'une protéine de reconnaissance de la famille des collagènes de défense (C1q ou MBL/ficoline) associé à un complexe de protéases à sérine modulaires (C1s/C1r ou MASPs (MBL-Associated Serine Proteases)(Figure 3). Les collagènes de défense du complément sont des molécules oligomériques solubles composées de tiges collagènes prolongées par des domaines globulaires. Cette organisation structurale leur confère un rôle bivalent. Leurs régions globulaires à l'extrémité C-terminale de la protéine ciblent spécifiquement les signaux de dangers à la surface des pathogènes (PAMPS) et des cellules du SOI altéré (ACAMPS). Les tiges collagènes, à l'autre extrémité, interagissent avec un ensemble d'effecteurs immuns conduisant à l'activation du complément par les protéases à sérine associées (C1s, C1r, et MASPs).

Les éléments à éliminer vont activer différentiellement la voie classique ou la voie lectine en fonction des motifs de reconnaissance présents à leur surface. La MBL ou les ficolines (voie lectine) apparues plus tôt dans l'évolution (600 millions d'années) sont des protéines de type lectine qui ciblent spécifiquement des motifs oligosaccharidiques tandis que C1q (voie classique) apparu plus récemment (450 millions d'années), à la même période que le système adaptatif, a un spectre de reconnaissance plus spécialisé et reconnait les complexes immuns ainsi que d'autres cibles ne nécessitant pas de synergie avec le système adaptatif, telles que l'ADN ou les cellules apoptotiques (Walport, 2001a, Walport, 2001b; Arlaud et al., 2001) (Figure 2).

S'ils présentent certaines similitudes, les complexes activateurs sont différents par le nombre de protéases associées aux collagènes de défense. Dans le complexe C1 de la voie classique, C1q est associé à un tétramère calcium-dépendant de protéases à sérine composé de deux molécules de C1s et deux molécules de C1r. La voie lectine fait quant à elle intervenir plusieurs types de collagènes de défense de type lectine, la MBL ou les ficolines M,L ou H associées à un dimère de protéases à sérine MASP-1 ou -2 sous forme d'homo-complexes.

L'activation de ces complexes enzymatiques nécessite donc une première phase de reconnaissance par les régions globulaires des collagènes de défense, puis l'activation des protéases associées. Dans le cas du complexe C1, la fixation de C1q sur sa cible engendre un changement conformationnel qui provoque l'autoactivation de C1r et l'activation consécutive de C1s par C1r. Dans le cas des complexes activateurs de la voie lectine, la fixation induit l'autoactivation du dimère de MASPs associées.

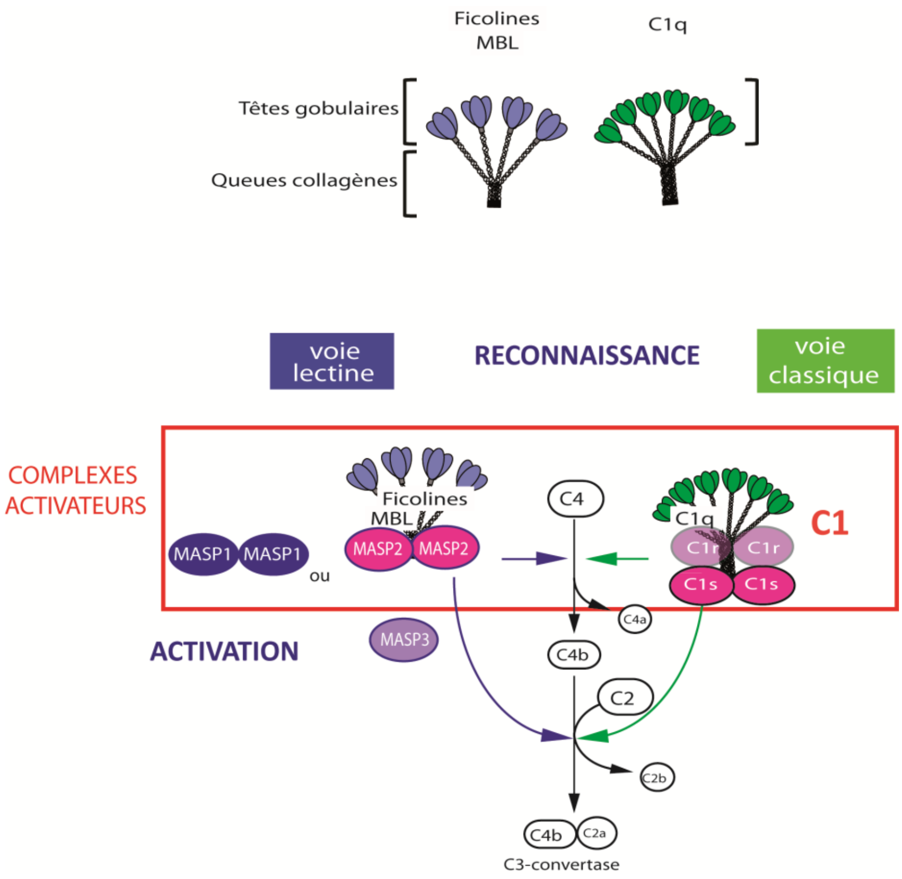

Ficolines
MBL
C1q
Têtes gobulaires
Queues collagènes
voie lectine
RECONNAISSANCE
voie classique
COMPLEXES ACTIVATEURS
MASP1
MASP1
ou
Ficolines
MBL
MASP2
MASP2
C4
C1q
C1r
C1r
C1s
C1s
C1
C4a
ACTIVATION
MASP3
C4b
C2
C2b
C4b
C2a
C3-convertase

Figure 3 : Représentation schématique des complexes activateurs de la voie classique et lectine.

2.2.3 - L'organisation modulaire des protéines du complément

Les protéines du complément sont constituées pour la plupart par l'assemblage d'un répertoire limité de motifs structuraux à repliement autonome, les modules (Figure 4). On recense plus de 60 familles différentes de modules dont la composition varie de 30 à 350 acides aminés (Bork et al., 1996) et dont certaines (par exemple les modules de type EGF) comptent plusieurs centaines d'exemplaires. Les modules d'une même famille dérivent vraisemblablement d'un même gène ancestral, et sont caractérisés par des repliements similaires. Sur ce squelette commun se greffent des motifs variables qui déterminent la spécificité fonctionnelle d'un module particulier. Ils jouent un rôle essentiel dans les processus de reconnaissance moléculaire qui sont à la base des phénomènes biologiques. Les protéines du complément peuvent contenir jusqu'à plusieurs dizaines de modules. Selon leur structure, on distingue différentes familles dont les membres pourraient provenir d'un ancêtre commun, chaque module remplissant un rôle précis d'interaction ou d'activité protéolytique (Figure 4). Ainsi, les protéines régulatrices du complément, CR1, CR2, MCP, DAF, facteur H et C4bp sont toutes constituées de modules CCP (Complement Control Protein repeat). Les protéines C6 à C9 du complexe terminal qui forment le pore transmembranaire responsable de la lyse cellulaire contiennent toutes un module (MACP (Transmembranaire), un module de type EGF et un motif LM de type LDL récepteur. Les protéases C1r, C1s, MASP-1 à 3, C2 et B contiennent quant à elles toutes un domaine protéase à sérine associé directement ou par l'intermédiaire d'autres modules à des modules CCP (Forneris et al., 2012). Bien que la structure d'ensemble du domaine protéase à sérine des enzymes du complément soit homologue des protéases à sérine pancréatiques, leur structure modulaire ajoute des caractéristiques qui leur confèrent des fonctions hautement spécialisées. Le rôle fonctionnel de chaque membre d'une famille (interaction, insertion dans la membrane, activité enzymatique) est donc défini par la présence des différents modules, leur spécificité propre étant vraisemblablement fonction de la variabilité au sein même des modules protéiques.

L'organisation modulaire de ces protéines leur permet d'intervenir de façon efficace dans le mécanisme d'activation en cascade du complément et présente une grande similitude avec les protéines impliquées dans la coagulation sanguine, qui est également un système d'activation en cascade. Au sein d'une même molécule, en plus d'être le socle d'une fonction souvent multiple, l'agencement modulaire apporte la flexibilité nécessaire aux mécanismes d'actions séquentielles mis en jeu (interaction, activation, protéolyse).

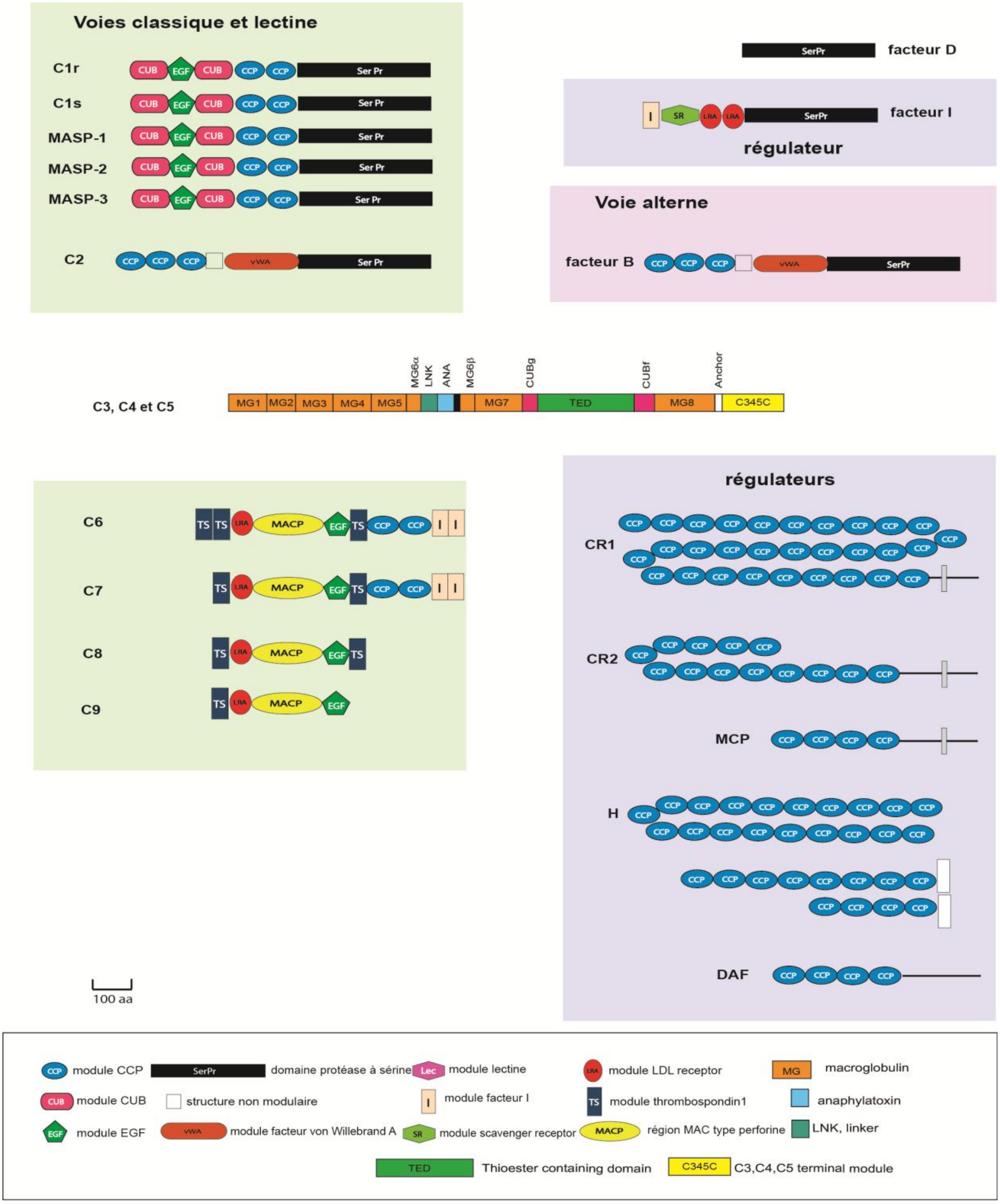

Voies classique et lectine
C1r
C1s
MASP-1
MASP-2
MASP-3
C2
CUB
EGF
CCP
Ser Pr
vWA
facteur D
SerPr
facteur I
I
SR
LRA
régulateur
Voie alterne
facteur B
C3, C4 et C5
MG1
MG2
MG3
MG4
MG5
MG6α
LNK
ANA
MG6β
MG7
CUBg
TED
CUBf
MG8
Anchor
C345C
C6
C7
C8
C9
TS
MACP
régulateurs
CR1
CR2
MCP
H
DAF
100 aa
module CCP
domaine protéase à sérine
Lec
module lectine
module LDL receptor
MG
macroglobulin
module CUB
structure non modulaire
module facteur I
module thrombospondin1
anaphylatoxin
module EGF
module facteur von Willebrand A
module scavenger receptor
région MAC type perforine
LNK, linker
Thioester containing domain
C3,C4,C5 terminal module

Figure 4 : Structure modulaire des protéines du complément.
Les protéines d'une même famille sont représentées dans les encadrés, ceux qui sont extérieurs sont homologues mais n'appartiennent pas au système complémentaire. D'après Boch et Bairoch 1995.

Chapitre III

Protéases à sérine des complexes activateurs du complément

Ce chapitre décrit la partie principale de mon activité de recherche qui commence en DEA en 1992, lorsque j'ai rejoint le laboratoire d'enzymologie moléculaire de l'IBS sous la direction de Gérard Arlaud. Dans le cadre de mon DEA, puis de ma thèse, jusqu'en janvier 1997, mes travaux ont porté sur les protéases à sérine C1s et C1r qui forment avec C1q le complexe activateur C1 du complément.

J'ai ensuite effectué pendant 18 mois un stage post-doctoral aux états Unis de 1997 à 1998. Ayant réintégré en 1998 le laboratoire dans le cadre d'un contrat d'attaché temporaire à l'enseignement et à la recherche (ATER), je me suis intéressée aux protéases MASP-1 et MASP-2 de la voie lectine d'activation du complément. Cette voie d'activation venait d'être découverte et avait un intérêt tout particulier du point de vue phylogénique, car apparue dans l'évolution antérieurement à la voie classique. Elle a suscité un grand nombre de questionnements puisqu'elle agit de façon similaire à la voie classique, par un déclenchement analogue par l'intermédiaire de protéases ayant une spécificité très homologue de celle de C1r et C1s.

Au début des années 1990, les connaissances sur les aspects moléculaires des protéines du complément étaient réduites à des caractérisations biochimiques mises en relation avec des modèles structuraux très schématiques. Mon arrivée dans le laboratoire de Gérard Arlaud coïncide avec l'inauguration de l'Institut de Biologie Structurale qui regroupait des équipes de biologistes et des équipes spécialisées dans les méthodes de biologie structurale (cristallographie aux rayons X, RMN, spectrométrie de masse). Ce rassemblement m'a offert l'opportunité de travailler en collaboration avec Christine Gaboriaud, spécialiste de cristallographie aux rayons X, afin d'apporter à la caractérisation des protéines du complément une dimension structurale plus précise que celle dont nous disposions.

L'introduction de la biologie moléculaire dans le laboratoire en 1996 à la fin de mon travail de thèse a été un tremplin vers le décryptage moléculaire des fonctions associées aux protéines du complément et a permis l'obtention de structures tridimensionnelles par cristallographie aux rayons X. Jusqu'alors, les essais de cristallographie avaient été vains pour deux raisons principales qui empêchaient la cristallogenèse: (i) la grande hétérogénéité de glycosylations au sein d'une même molécule, (ii); la structure modulaire très flexible de molécules ayant une taille supérieure à 80 kDa.
C'est la stratégie de dissection, envisageable uniquement avec le clonage de fragments restreints qui nous a permis d'élargir un champ d'investigation qui avait alors atteint ses limites.

3.1 - Relations structure/fonction des protéases C1s et C1r

3.1.1 - Brève description du complexe C1 et des protéases associées.

Le complexe C1, responsable de l'activation de voie classique du complément, comprend une protéine de reconnaissance C1q (collagène de défense) ainsi qu'un tétramère protéasique C1s-C1r-C1r-C1s comportant deux protéases à sérine homologues, C1r et C1s (Figure 1, 3 et 16). La Figure 5 représente l'organisation modulaire des protéases homologues que l'on retrouve dans le complexe C1 (C1s et C1r) et les complexes activateurs de la voie lectine (MASP-1 et MASP-2). MASP-3 est une protéase à sérine homologue des autres MASPs mais bien qu'elle s'associe aux molécules de reconnaissance de la voie lectine, elle n'a pas de rôle connu dans l'activation du complément. Chaque protéase contient un domaine protéase à sérine (SerPr) homologue de la trypsine auquel s'ajoutent en N-terminal 5 modules n'ayant pas d'activité enzymatique. Les protéines de cette famille contiennent toutes deux régions fonctionnelles communes : (i) une région d'interaction, composée des trois modules N-terminaux CUB1/EGF/CUB2 impliquée dans l'association des protéases entre elles et avec la protéine de reconnaissance des complexes activateurs (ii) une région « catalytique » composée des deux modules CCP1et CCP2 (Complement Control Protein repeat) et du domaine protéase à sérine (SerPr).

Sous leur forme proenzyme ces protéases ne sont constituées que d'une seule chaîne polypeptidique. Leur activation résulte de la coupure d'une liaison Arg-Ile (indiquée par la flèche Figure 5) soit par un mécanisme d'autoactivation dans le cas de C1r, soit par coupure par une autre protéase à sérine du complexe protéasique, c'est le cas de C1s qui est clivée par C1r. Dans le cas de MASP-1 et MASP-2, on sait depuis peu que le mécanisme d'activation est similaire, MASP-1 s'autoactive et active MASP-2 (Heja et al., 2012.Degn et al., 2012, Megyeri et al., 2013, Degn et al., 2014).

Le mécanisme d'activation du complexe C1 fait donc intervenir des changements conformationnels très importants sur ces protéases pour de telle sorte que dans un premier temps s'associer à la protéine de reconnaissance C1q puis ensuite s'activer lorsque cette dernière s'est fixée sur une cible à éliminer, pour finalement dans le cas de C1s aller cliver les substrats C2 et C4. Dans les conditions physiologiques, C1-inhibiteur contrôle l'action du complexe C1 en étant associé à C1s et C1r. Les changements de conformations seront nécessaires à la dissociation de C1-inhibiteur pour permettre l'activation de la cascade protéolytique.

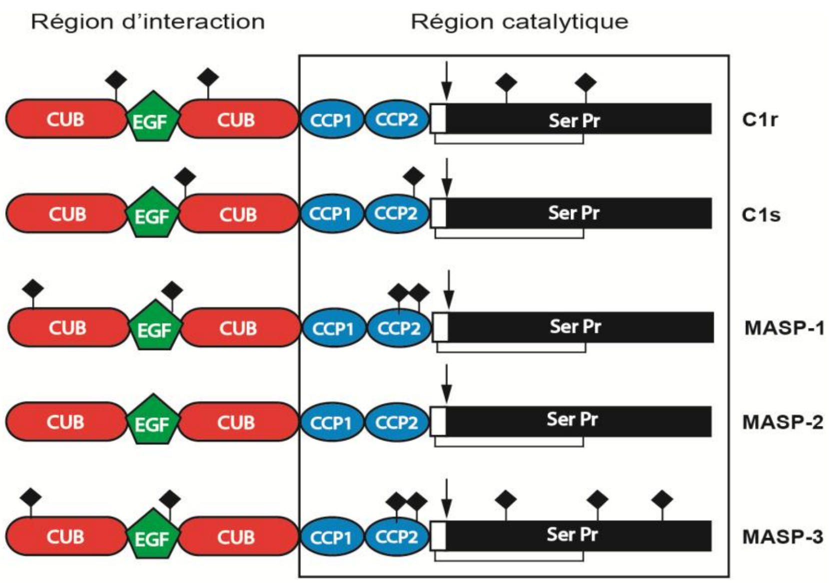


Figure 5 : Structure modulaire des protéases à sérine des complexes activateurs des voies classique et lectine.
La nomenclature et les symboles utilisés pour les modules protéiques sont définis par Bork et Bairoch (1995). CUB, module présent dans diverses protéines impliquées dans le développement, EGF, module homologue du facteur de croissance épidermique. CCP, Complement Control Protein repeat, module présent dans les protéines régulatrices du complément. Ser Pr, domaine protéase à sérine. (◆), oligosaccharides N-liés. Le segment précédant le domaine protéase à sérine est homologue du peptide d'activation du chymotrypsinogène. La flèche indique le site d'activation des protéases. L'unique pont disulfure représenté est celui reliant le peptide d'activation au domaine protéase à sérine.

3.1.2 - Le domaine protéase à sérine (SerPr) – catalyse enzymatique et sites de reconnaissance.

La région catalytique des protéases que j'ai étudiées comporte un domaine protéase à sérine homologue de la trypsine. La structure d'un grand nombre de protéases à sérine de cette famille a été déterminée par cristallographie aux rayons X. Ces protéses ont toutes une topologie globale identique qui comporte une hélice alpha C-terminale ainsi que deux domaines globulaires d'environ 120 acides aminés chacun, de type tonneau β-antiparallèle (du même type que les immunoglobulines). Ces deux régions ménagent une cavité qui contient la triade catalytique du site actif Ser 195, His 97 et Asp 102 (Figure 6 et 7) (Perona and Craik, 1997, Hedstrom, 2002). S'ajoutent aux résidus de cette triade, 4 résidus conservés (Asp 189, Gly 193, Ser 214 et Gly 216) dans le site de reconnaissance principal appelé site S1, ils permettent le positionnement du substrat comme schématisé sur la Figure 6. Dans ce site S1, la sérine 195 de la triade catalytique peut, par attaque nucléophile, hydrolyser la liaison peptidique qui relie un résidu P1 Arg ou Lys (dans le cas de la trypsine et des protéases des voies classiques et lectines) au résidu P'1. Le positionnement de substrats protéiques dans les protéases à sérine fait intervenir également d'autres résidus proche du site de coupure appelés sous-sites et plus éloignés, en surface de l'enzyme, dans des boucles qui sont spécifiques (longueur et composition) d'une protéase à sérine donnée. (Figure 7). C'est l'ensemble de ces déterminants structuraux qui détermine la spécificité d'une protéase à sérine pour ses substrats.

Dans le cas de C1s, un de mes questionnements principaux a été de définir le rôle précis des modules CCP dans sa spécificité enzymatique très restreinte. C1s ne clive en effet que trois substrats physiologiques au sein du complément, C2, C4 et C1-inhibiteur contrairement à la trypsine ou d'autres protéases digestives qui ne contiennent que le domaine protéase à serine et ont une spécificité plus large.

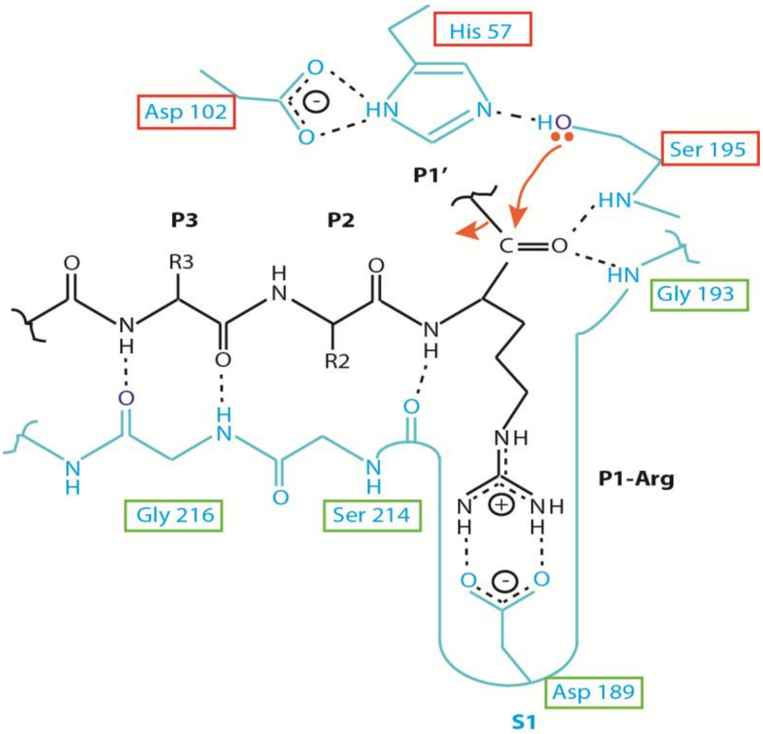


Figure 6 : Mécanisme de la catalyse enzymatique des protéases à sérine.
Les groupements des acides aminés de la protéase impliqués dans la catalyse sont montrés en bleu en interaction avec un substrat fixé au site P1. Les liaisons hydrogènes qui permettent le positionnement du substrat sont indiquées en pointillés. La nomenclature utilisée pour les acides aminés du substrat est Pn, …, P2, P1, P1', P2', ….Pn', où P1-P1' indique la liaison peptidique qui est hydrolysée par l'enzyme. Sur l'enzyme, les acides aminés impliqués dans l'interaction sont nommés Sn, …S2, S1-S1', S2', …Sn'. Adapté de Perona and Craik, 1997.

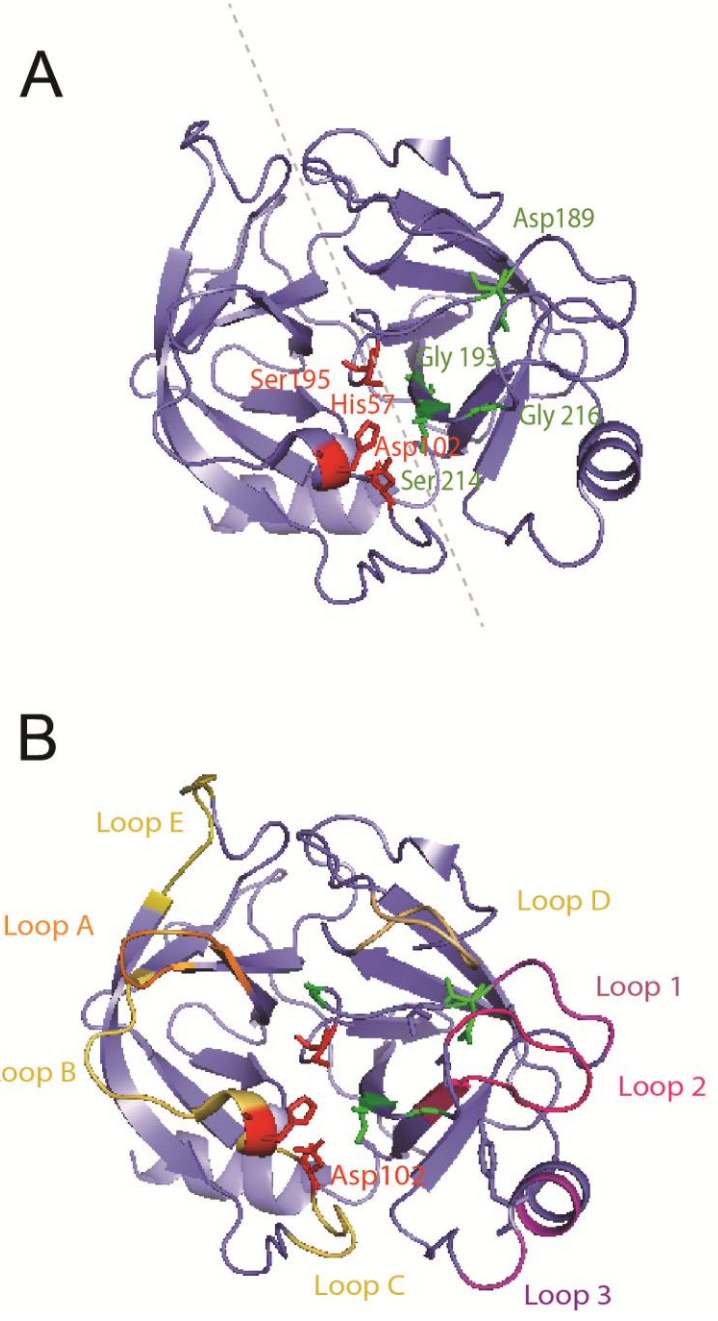

A
Asp189
Gly 193
Ser195
His57
Gly 216
Asp102
Ser 214
B
Loop E
Loop D
Loop A
Loop 1
Loop B
Loop 2
Asp102
Loop C
Loop 3

***Figure 7 : Localisation des éléments de reconnaissance de la trypsine communs à toutes les protéases de la famille**.*

(A) Ossature typique des protéases à sérine de type trypsique sous forme de deux tonneaux béta (séparés par les pointillés). La triade catalytique His 57, Asp102, Ser 195 est représentée en rouge. Les résidus qui positionnent le substrat dans le site S1 de l'enzyme sont représentés en vert (cf Figure 6)

(B) Localisation des huit boucles de surface qui déterminent la spécificité des sites S1 et des sous-sites de reconnaissance. Loop A (résidus 34-41), Loop B (résidus 56-64), Loop C (résidus 97-103), Loop D (résidus 143-149), Loop E (résidus 74-80), Loop 1 (résidus 185-188), Loop 2 (résidus 217-225), Loop 3 (résidus 169-174). Adapté de Perona et Craik, 1997.

3.1.3 - Elaboration d'un modèle tridimensionnel de la région catalytique (CCP1-CCP2-SerPr) de C1s par réticulation chimique.

Publications : **P1** Rossi et al., *Biochemistry* (1995), **P5** Gans et al., *Cell.Mol.Life Sci*. (1998), **P2** Pétillot et al., *FEBS Lett.* (1995) **P3** Lacroix et al., *Biochemistry* (1997)

Communications : C1, C2, C3, C4, C5, C8.

Projet : Les travaux antérieurs du laboratoire avaient défini l'organisation modulaire de C1s (Figure 5) mais la structure tridimensionnelle n'était pas connue. Mon travail a permis d'élaborer un modèle tridimensionnel de sa région catalytique, constituée des deux modules CCP1 et CCP2 et du domaine protéase à sérine SerPr. Ce modèle a contribué à la détermination des relations structure-fonction de C1s, plus particulièrement en ce qui concerne son rôle au sein du complexe C1 et sa spécificité vis-à-vis de ses substrats C4 et C2.

Résultats : Dans un premier temps, nous avons construit un modèle relativement précis de la région catalytique de C1s grâce aux données de réticulation chimique. Le traitement du fragment CCP1-CCP2-SerPr par l'agent réticulant EDC nous a permis d'identifier deux liaisons ioniques qui maintiennent de façon rigide le module CCP2 et le domaine SerPr et qui font intervenir respectivement : (i) la lysine K405 du module CCP2 avec l'acide glutamique E672 de l'extrémité C-terminale du domaine SerPr et (ii) l'acide glutamique E418 à l'extrémité du peptide d'activation libéré lors de l'activation de C1s par C1r avec la lysine K608 du domaine SerPr. (Figure 8A) (Rossi et al., 1995). La même étude sur la protéase homologue C1r, a conduit à mettre en évidence une organisation de la région CCP2-SerPr en une structure compacte du même type que celle de la région homologue de C1s (Lacroix et al., 1997). Le modèle de C1s a été construit par homologie avec des structures connues de protéases à sérine et de modules CCP et affiné par la caractérisation structurale de l'extrémité C-terminale (20 acides aminés) du domaine SerPr par RMN. La structure tridimensionnelle en forme d'hélice alpha a été caractérisée par Pierre Gans du laboratoire de RMN de l'IBS et a été ajoutée au modèle (Gans et al., 1998).

Le modèle obtenu est assez précis et a constitué, à défaut d'une structure cristallographique, une base de réflexion plus précise que celle dont nous disposions.

Pendant les 3 années de mon travail de thèse, de nombreux essais de cristallisation de ce fragment purifié à partir du plasma humain ont été réalisés dans le laboratoire de Juan Fontecilla par Christine Gaboriaud (LCCP, IBS). Le fragment CCP1-CCP2-SerPr plasmatique cristallise sous la forme d'aiguilles qui ne sont malheureusement pas utilisables pour la diffraction aux rayons X probablement à cause de son hétérogénéité due à la présence de plusieurs isoformes du sucre N-lié au module CCP2.

Pour améliorer l'homogénéité de ce fragment nous avons caractérisé le sucre par spectrométrie de masse (électrospray). Nous avons identifié trois populations comportant respectivement un sucre N-lié sous une forme biantennaire ou deux formes triantennaires différentes dont l'une comporte un fucose (Petillot et al., 1995). Les espèces bi- et tri-antennaires ont ensuite été séparées et des essais de cristallisation ont été tentés, en vain.

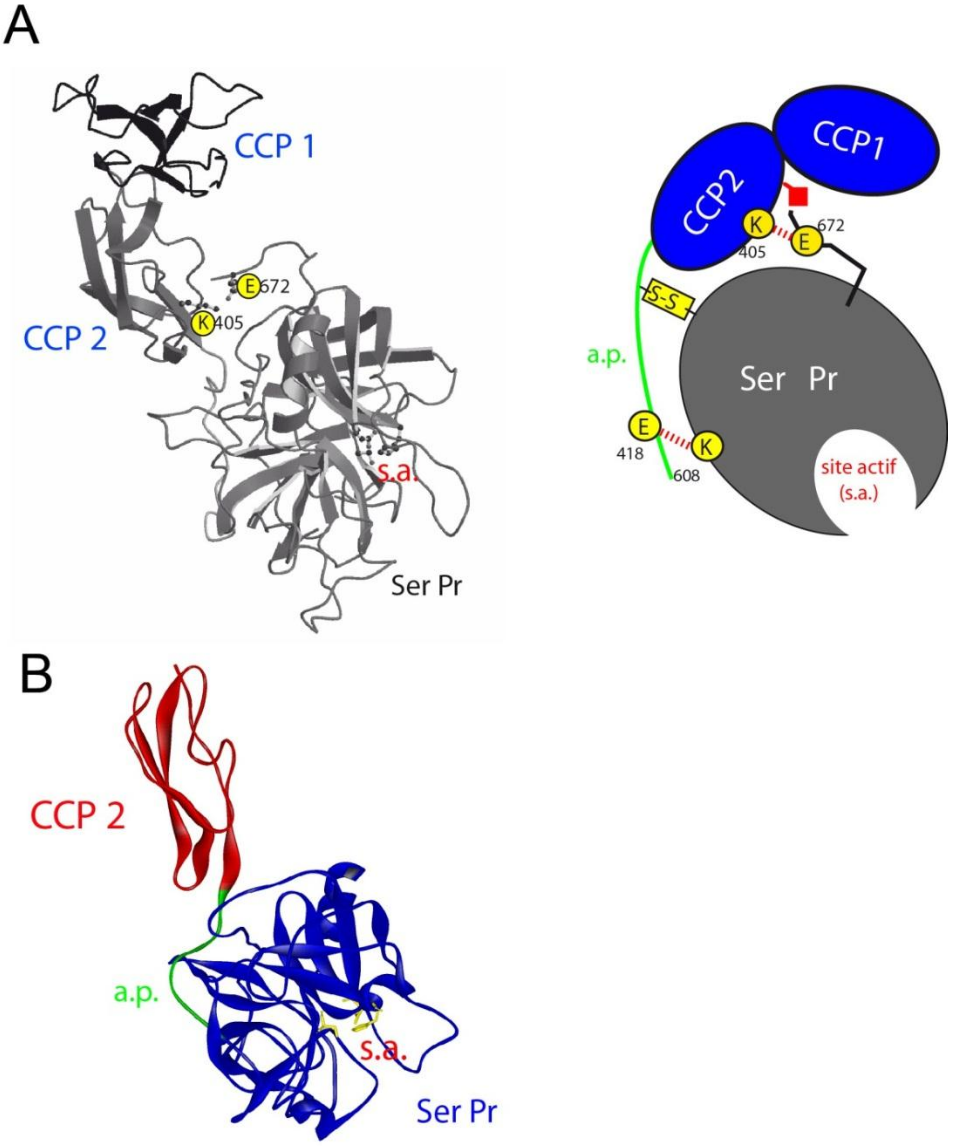


Figure 8 : Modèle et structure de la région catalytique de C1s.
(A) : Modèle en 3-D et schéma de l'assemblage de la région catalytique CCP1-CCP2-SerPr de C1s obtenu par des études de réticulation chimique et modélisation moléculaire tridimensionnelle *(Rossi et al., Biochemistry 1995)*
Le module CCP2 est assemblé au domaine protéase à sérine (Ser Pr) par l'intermédiaire du peptide d'activation (a.p en vert). Les trois points d'ancrage (en jaune) ayant servi pour cet assemblage correspondent aux deux réticulations entre K608-E418 et K405-E672 et au pont disulfure. Le module CCP1 est positionné par rapport au module CCP2 dans la même orientation que les modules H15 et H16 du facteur H du complément ayant servi de référence. s.a. site actif.
(B) : Structure 3-D de la région catalytique CCP2-SerPr de C1s obtenue par cristallographie aux rayons X à 1,7 Å de résolution *(Gaboriaud et al., EMBO J. 2000)*

3.1.4 - Structure du fragment CCP2-SerPr de C1s par cristallographie aux rayons X et identification d'une famille de protéases à motif structural homologue.

Publications : **P6** Gaboriaud et al., *J. Mol. Biol.* (1998), **P7** Gaboriaud et al., *EMBO J.* (2000).

Communications : **C7, C9, C15, C27**

Projet : Les techniques de purification à partir du plasma n'ayant pas permis d'obtenir de fragments des protéines du complément assez homogènes pour être cristallisés et utilisés en diffraction aux rayons X, nous avons changé de stratégie en 1996 et mis en place, grâce à la biologie moléculaire, l'expression de protéines recombinantes du complément dans des systèmes d'expression hétérologues.

Résultats : C'est finalement grâce à la production de fragments de C1s dans les cellules d'insectes que la structure tridimensionnelle du domaine catalytique recombinant, comportant le module CCP2 et le domaine protéase à sérine SerPr a été déterminée à 1,7 Å de résolution (Rossi et al., 1998, Gaboriaud et al., 2000) (Figure 8B). Ce résultat représente une avancée considérable dans nos connaissances car il a conduit à résoudre la première structure tridimensionnelle d'une protéine modulaire du complément.

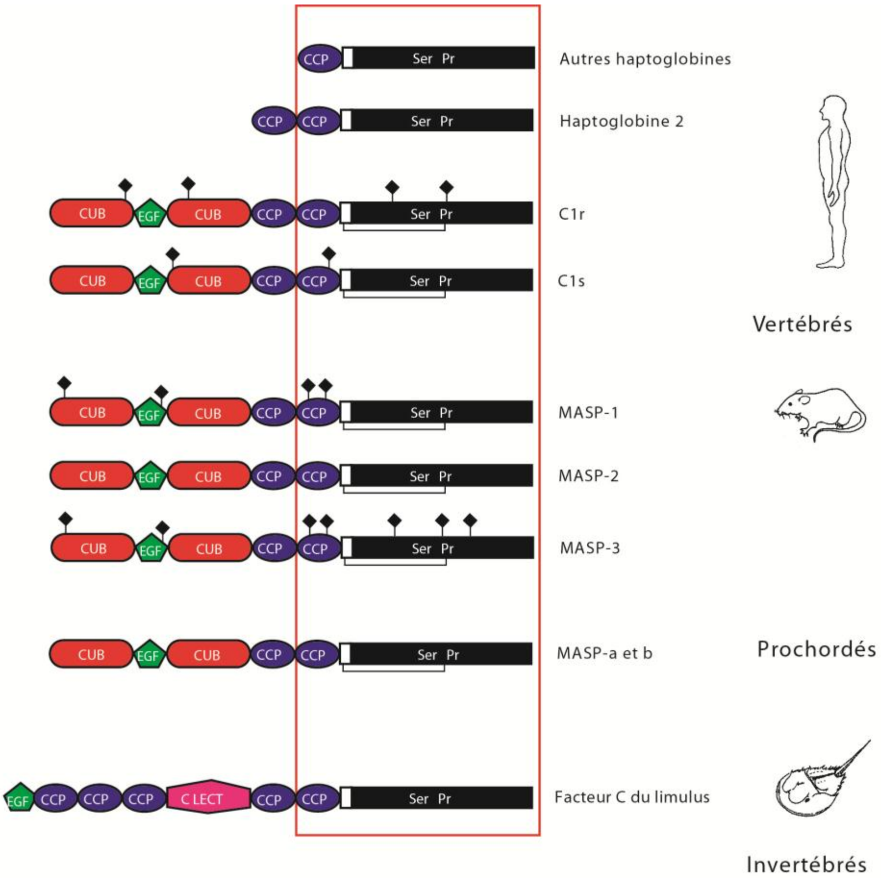
Cette structure a apporté un grand nombre de renseignements sur le fonctionnement de C1s :

(i) reconnaissance spécifique des substrats C2 et C4 : la grande spécificité de C1s serait due à une combinaison d'éléments de restriction localisés autour de l'entrée de la gorge du site actif, et d'éléments de reconnaissance spécifiques présents sur le module CCP2. Cette hypothèse est en accord avec l'étude fonctionnelle que nous avons menée en parallèle (Rossi et al., 1998).

Au niveau du domaine protéase à sérine, contrairement aux protéases digestives telles que la trypsine, l'accessibilité du site actif est restreinte par un ensemble de motifs structuraux. C'est une caractéristique qui est observée également dans les protéases de la coagulation. Un élément majeur de l'occlusion de l'entrée du site actif de C1s est la lysine 614 (192 dans la chymotrypsine) qui se positionne à l'entrée du site actif et restreint ainsi son accès aux 3 substrats de C1s, C2, C4 et C1-inhibiteur.

(ii) agencement structural CCP2/SerPr : l'orientation du module CCP2, du côté opposé à l'entrée du site actif, et perpendiculairement au domaine protéase à sérine est en accord avec le modèle que nous avions élaboré (Rossi et al., 1995). Sa forme globale s'apparente à celle d'une massue, le module CCP étant rendu solidaire du domaine protéase à sérine par l'interpénétration de segments peptidiques riches en proline et tyrosine à l'interface des deux entités protéiques (Figure 8B). On retrouve vraisemblablement le même type d'assemblage chez d'autres membres de la famille CCP-SerPr, pour laquelle nous avons mis en évidence l'existence d'une signature commune au niveau de la séquence primaire (Gaboriaud et al., 1998). Cette signature a été conservée au cours de l'évolution puisqu'on la retrouve chez les invertébrés, sur le facteur C de la limule, une molécule qui combine à la fois les propriétés d'un système de défense et de coagulation sanguine. (Figure 9)

(iii) **fonctionnement du complexe C1 :** Le prolongement du domaine SerPr par un bras rigide, le module CCP2, positionne le site actif de C1s à distance du reste de la molécule et individualise ainsi la région catalytique et la région d'interaction. Ainsi, du fait de l'existence d'une charnière mobile entre les deux modules CCP, on peut imaginer un mécanisme du déclenchement de la voie classique s'effectuant par pivotement de la région catalytique de C1s par rapport au reste de la molécule. (Figure 10)



CCP
Ser Pr
Autres haptoglobines
CCP
CCP
Ser Pr
Haptoglobine 2
CUB
EGF
CUB
CCP
CCP
Ser Pr
C1r
CUB
EGF
CUB
CCP
CCP
Ser Pr
C1s
Vertébrés
CUB
EGF
CUB
CCP
CCP
Ser Pr
MASP-1
CUB
EGF
CUB
CCP
CCP
Ser Pr
MASP-2
CUB
EGF
CUB
CCP
CCP
Ser Pr
MASP-3
CUB
EGF
CUB
CCP
CCP
Ser Pr
MASP-a et b
Prochordés
EGF
CCP
CCP
CCP
C LECT
CCP
CCP
Ser Pr
Facteur C du limulus
Invertébrés

Figure 9 : Structure modulaire des protéases de la famille CCP-SerPr.
CCP, module de type « complement Control Protein » CUB, module identifié initialement dans les protéines C1s et C1r, Uegf, Bone morphogenetic protein module. EGF, module de type « Epidermal Growtg Factor ». CLECT, domaine de type C- lectine. Ser Pr, domaine protéase à sérine.

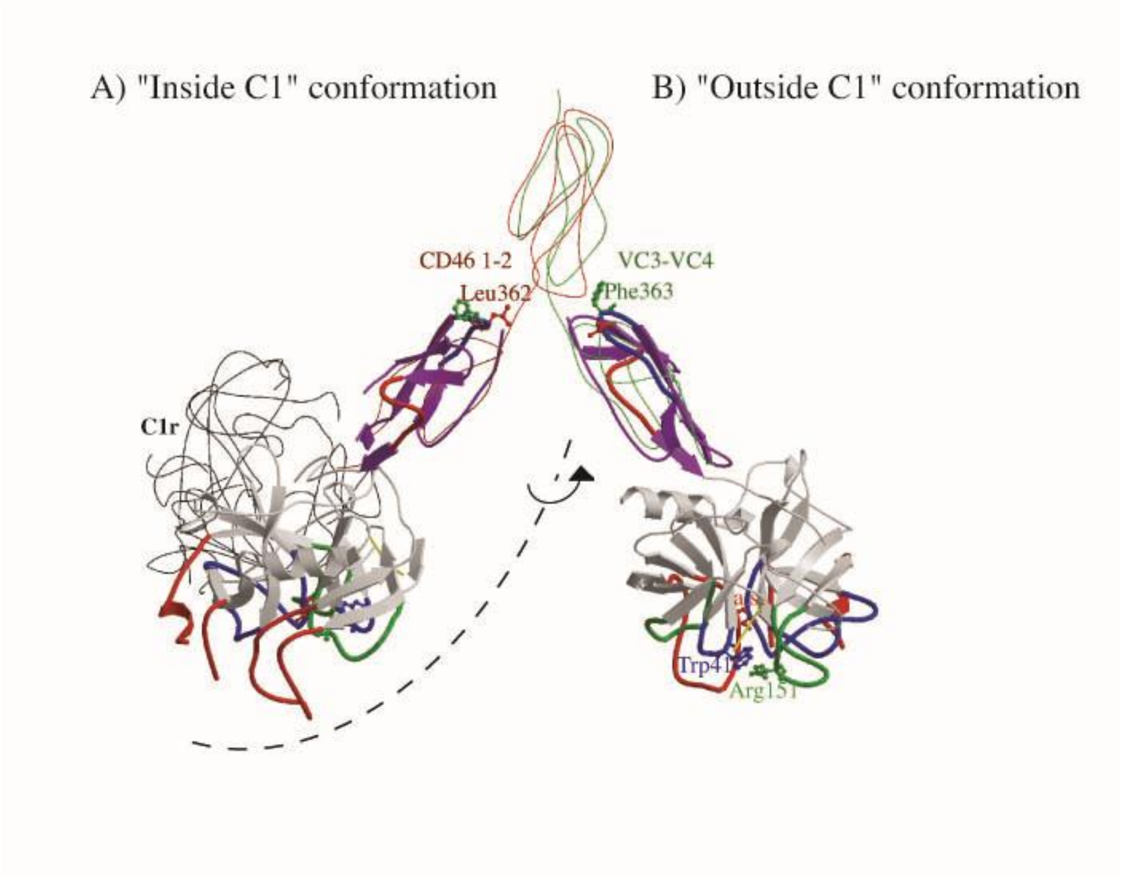


Figure 10 : Illustration du rôle potentiel du bras rigide CCP2 de C1s dans le pivotement du domaine catalytique SerPr pour l'activation de la cascade du complément.
L'activation de C1s par C1r (noir) dans le complexe C1 (conformation A), nécessite une orientation du domaine SP qui n'est pas compatible avec son abilité à cliver C2 et C4 : Les chaînes latérales de la Trp41 (bleu) , Arg151 (vert) et de la sérine de la triade catalytique (rouge) sont orientées face à C1r dans le complexe C1. Dans la conformation A, le module CCP2 de C1r est positionné en arrière du domaine SerPr. La conformation B montre après pivotement par l'intermédiaire de la région charnière entre le CCP1 et CCP2, une orientation préferentielle pour le clivage des substrats C2 et C4 en dehors du complexe C1. Le modèle a été construit sur la base de la conformation connue de paires de modules CCP, (de CD46 CCP1-2 en rouge et de VCP CCP3-4 en vert). Les chaînes hydrophobes exposées de la Leu362 (rouge) et Phe363 (vert) observées sur la structure sont proposées comme stabilisatrices de chacune des deux conformations.

3.1.5 - Mise en évidence de deux mécanismes distincts de reconnaissance des substrats C2 et C4 par C1s.

Publication : **P4** Rossi et al., *J.Biol.Chem.* (1998)

Communications : **C6, C10.**

Projet : L'étape suivante de mon projet de recherche a consisté à mettre à profit le modèle tridimensionnel obtenu par réticulation chimique pour étudier les relations structure-fonction de la région catalytique de C1s. Pour cela nous avons caractérisé la fonction des différents domaines de cette région.

Résultats : Nous avons montré que l'hydrolyse des substrats C2 et C4 par C1s entier ainsi que par le fragment catalytique CCP1-CCP2-SerPr et ses isoformes bi- et tri-antennaires est comparable. En revanche, nous avons pu mettre en évidence des mécanismes de reconnaissance très différents pour la protéolyse limitée des deux substrats C2 et C4 de C1s. Dans ce but, nous avons produit dans un système eucaryote (baculovirus/cellules d'insecte), des fragments recombinants tronqués de la région catalytique de C1s. Ces fragments comportent soit le deuxième module CCP2 relié au domaine protéase à sérine (CCP2-SerPr), soit uniquement le domaine protéase à sérine (SerPr).

Leur caractérisation fonctionnelle a montré que la coupure de C2 implique uniquement le domaine protéase à sérine, alors que celle de C4 fait intervenir des sites de reconnaissance auxiliaires localisés dans chacun des deux modules CCP (Rossi et al., 1998).

L’utilisation de la biologie moléculaire, et plus particulièrement d’un système d’expression dans les cellules d’insecte, a été pionnière dans le domaine des études structurales des molécules modulaires telles que C1s et C1r. C’est grâce à cette technique que nous avons réussi à franchir un pas important qui a conduit à de nombreuses études structure-fonction dans le laboratoire.

3.1.6 - Cartographie fonctionnelle de la région catalytique de C1s : Identification des déterminants moléculaires responsables de la spécificité de C1s pour ses deux substrats, C2 et C4.

Encadrement : Isabelle Bally - technicienne CEA.

Publication : P11 Bally et al., *J. Immunol.* (2005)

Communications : C14, C16, C17, C24

Projet : Dans le cadre de mon activité d'ATER à l’Université Joseph Fourier, j’ai réintégré, en octobre 1998, l’équipe du Laboratoire d’Enzymologie Moléculaire de l’Institut de Biologie Structurale. Mon projet de recherche a consisté à poursuivre le travail entrepris sur **la protéase C1s du complément**

avec pour objectif d'élaborer une **cartographie fonctionnelle** de sa région catalytique. Par comparaison avec d'autres protéases à sérine homologues de C1s (C1r, MASP-1, MASP-2) (Figure 5), nous avons conçu une série de mutants de C1s comportant, soit une modification ponctuelle, soit de petites délétions (jusqu'à 4 acides aminés) sur les modules CCP1, CCP2 ou le domaine protéase à sérine SP (Figure 11).

Résultats : Plus de vingt mutants ont été produits dans des cellules d'insecte et purifiés. Leur caractérisation fonctionnelle, notamment leur activité enzymatique vis-à-vis des substrats C2 et C4, a mis en évidence deux sites importants dans la spécificité de C1s vis-à-vis de C4.

(i) Un site situé **à l'interface des deux modules CCP**. Les mutations ponctuelles Q340K, P341I, V342K et D343N, réduisent l'efficacité de la protéolyse de C4 de 70 % et n'ont pas d'effet sur la protéolyse de C2. Ces résultats soulignent l'importance de la flexibilité de la région charnière entre les deux modules CCP1 et CCP2 dans le positionnement efficace du substrat C4 par rapport au site actif localisé sur le domaine protéase à sérine de C1s (Bally et al., 2005). Ils confortent ceux obtenus précédemment (Rossi et al., 1998) indiquant que C1s possède des sites auxiliaires de reconnaissance de C4 sur les deux modules CCP1-CCP2 tandis que seul le domaine protéase à sérine est nécessaire pour la coupure de C2.

(ii) le deuxième site identifié concerne la **lysine 614 du site actif du domaine SP** identifiée comme élément restrictif de l'entrée du site actif de C1s. Sa mutation en acide glutamique (présent à la fois dans C1r et dans la trypsine) réduit d'environ 50% l'activité enzymatique de C1s vis-à-vis de C4 et de C2 (résultats non publiés) (Figure 11).

Sur l'ensemble des régions analysées, nous avons eu la surprise de n'en identifier que deux et avions supposé que la stratégie utilisée n'était pas adaptée à ce type d'étude.

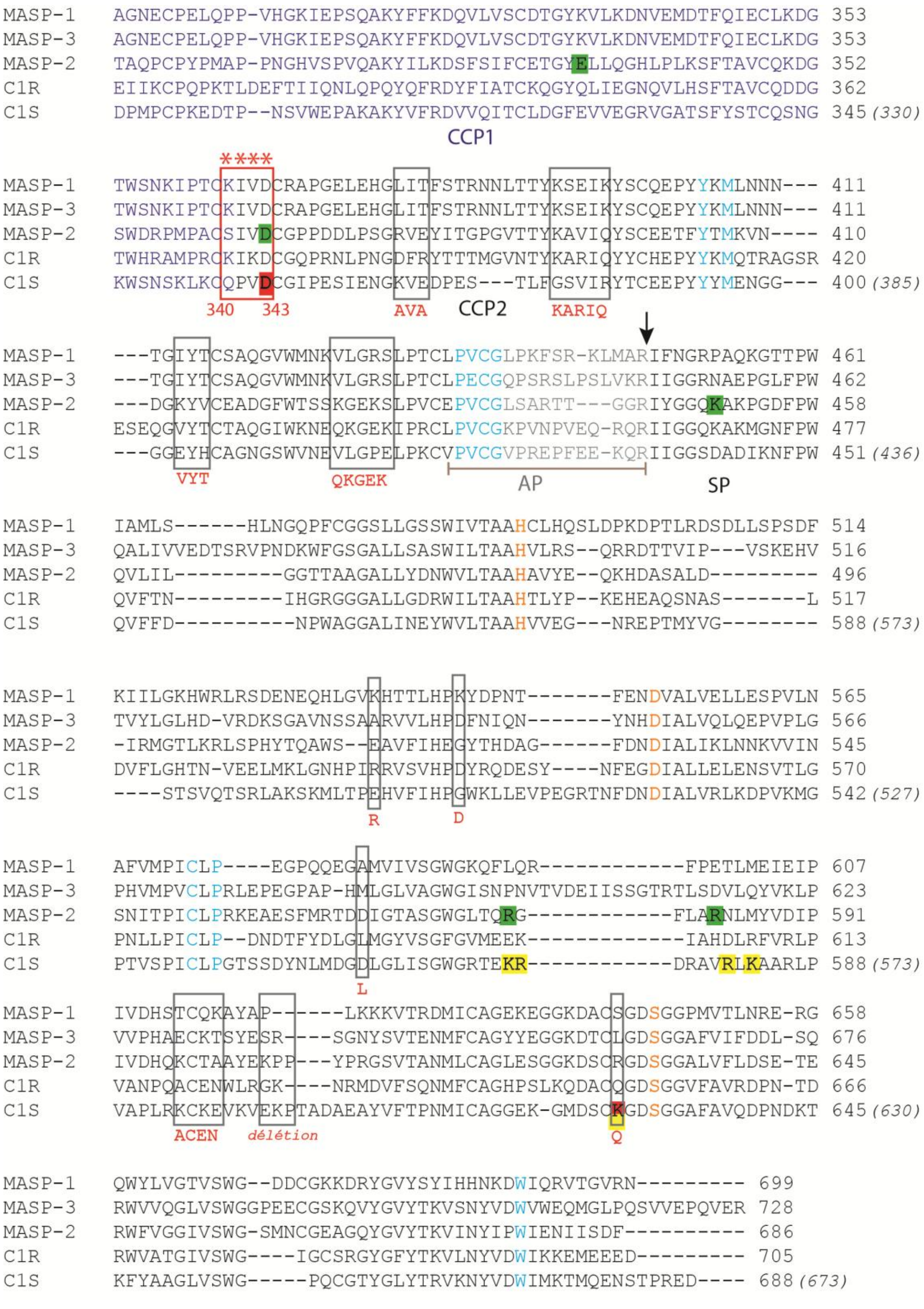
Depuis ce travail, plusieurs équipes ont cherché à identifier les résidus impliqués dans la reconnaissance de C4 par C1s et MASP-2 (dont la spécificité est analogue à celle de C1s). Dans l'état actuel des connaissances, un sous-site a été identifié sur le domaine SerPr de C1s par Duncan et al., en 2012, comportant 4 résidus basiques (surligné en jaune sur la Figure 11), que nous n'avions pas analysé en premier lieu.

En revanche, les deux autres sites de fixation de C4 identifiés à ce jour correspondent aux deux résidus que nous avons identifiés en 2005, en effet :

(i) Wijeyewickrema et al., en 2014 ont montré par une étude similaire utilisant la mutagenèse dirigée que la Lysine 614 du domaine SerPr est un élément très important dans la catalyse enzymatique de C4 par C1s.

(ii) Kidmose et al., en 2012, ont identifié à partir de la structure cristallographique du complexe MASP-2/C4, un acide aspartique du module CCP2 de MASP-2 (identique dans C1s) comme point d'ancrage de C4 sur l'enzyme. Ceci valide nos résultats montrant l'implication de l'acide aspartique 343 de C1s dans la fixation de C4.

Ces résultats montrent que, de façon étonnante, les sites d'interaction de MASP-2 avec C4 au niveau des modules CCP1 et 2 sont réduits à deux résidus, l'un sur le CCP1 et l'autre sur le CCP2 dans la région charnière (E et D, respectivement, entourés en vert, Figure 11). C'est probablement également le cas pour C1s, l'un de ces deux résidus, celui du CCP2, étant équivalent à celui que nous avons identifié. Notre stratégie n'était finalement pas si incertaine…



```
MASP-1   AGNECPELQPP-VHGKIEPSQAKYFFKDQVLVSCDTGYKVLKDNVEMDTFQIECLKDG 353
MASP-3   AGNECPELQPP-VHGKIEPSQAKYFFKDQVLVSCDTGYKVLKDNVEMDTFQIECLKDG 353
MASP-2   TAQPCPYPMAP-PNGHVSPVQAKYILKDSFSIFCETGYELLQGHLPLKSFTAVCQKDG 352
C1R      EIIKCPQPKTLDEFTIIQNLQPQYQFRDYFIATCKQGYQLIEGNQVLHSFTAVCQDDG 362
C1S      DPMPCPKEDTP--NSVWEPAKAKYVFRDVVQITCLDGFEVVEGRVGATSFYSTCQSNG 345 (330)

MASP-1   TWSNKIPTCKIVDCRAPGELEHGLITFSTRNNLTTYKSEIKYSCQEPYYKMLNNN--- 411
MASP-3   TWSNKIPTCKIVDCRAPGELEHGLITFSTRNNLTTYKSEIKYSCQEPYYKMLNNN--- 411
MASP-2   SWDRPMPACSIVDCGPPDDLPSGRVEYITGPGVTTYKAVIQYSCEETFYTMKVN---- 410
C1R      TWHRAMPRCKIKDCGQPRNLPNGDFRYTTTMGVNTYKARIQYYCHEPYYKMQTRAGSR 420
C1S      KWSNSKLKCQPVDCGIPESIENGKVEDPES---TLFGSVIRYTCEEPYYYMENGG--- 400 (385)

MASP-1   ---TGIYTCSAQGVWMNKVLGRSLPTCLPVCGLPKFSR-KLMARIFNGRPAQKGTTPW 461
MASP-3   ---TGIYTCSAQGVWMNKVLGRSLPTCLPECGQPSRSLPSLVKRIIGGRNAEPGLFPW 462
MASP-2   ---DGKYVCEADGFWTSSKGEKSLPVCEPVCGLSARTT---GGRIYGGQKAKPGDFPW 458
C1R      ESEQGVYTCTAQGIWKNEQKGEKIPRCLPVCGKPVNPVEQ-RQRIIGGQKAKMGNFPW 477
C1S      ---GGEYHCAGNGSWVNEVLGPELPKCVPVCGVPREPFEE-KQRIIGGSDADIKNFPW 451 (436)

MASP-1   IAMLS------HLNGQPFCGGSLLGSSWIVTAAHCLHQSLDPKDPTLRDSDLLSPSDF 514
MASP-3   QALIVVEDTSRVPNDKWFGSGALLSASWILTAAHVLRS--QRRDTTVIP---VSKEHV 516
MASP-2   QVLIL---------GGTTAAGALLYDNWVLTAAHAVYE--QKHDASALD--------- 496
C1R      QVFTN---------IHGRGGGALLGDRWILTAAHTLYP--KEHEAQSNAS-------L 517
C1S      QVFFD----------NPWAGGALINEYWVLTAAHVVEG---NREPTMYVG-------- 588 (573)

MASP-1   KIILGKHWRLRSDENEQHLGVKHTTLHPKYDPNT-------FENDVALVELLESPVLN 565
MASP-3   TVYLGLHD-VRDKSGAVNSSAARVVLHPDFNIQN-------YNHDIALVQLQEPVPLG 566
MASP-2   -IRMGTLKRLSPHYTQAWS--EAVFIHEGYTHDAG------FDNDIALIKLNNKVVIN 545
C1R      DVFLGHTN-VEELMKLGNHPIRRVSVHPDYRQDESY----NFEGDIALLELENSVTLG 570
C1S      ----STSVQTSRLAKSKMLTPEHVFIHPGWKLLEVPEGRTNFDNDIALVRLKDPVKMG 542 (527)

MASP-1   AFVMPICLP----EGPQQEGAMVIVSGWGKQFLQR------------FPETLMEIEIP 607
MASP-3   PHVMPVCLPRLEPEGPAP-HMLGLVAGWGISNPNVTVDEIISSGTRTLSDVLQYVKLP 623
MASP-2   SNITPICLPRKEAESFMRTDDIGTASGWGLTQRG------------FLARNLMYVDIP 591
C1R      PNLLPICLP--DNDTFYDLGLMGYVSGFGVMEEK-------------IAHDLRFVRLP 613
C1S      PTVSPICLPGTSSDYNLMDGDLGLISGWGRTEKR------------DRAVRLKAARLP 588 (573)

MASP-1   IVDHSTCQKAYAP------LKKKVTRDMICAGEKEGGKDACSGDSGGPMVTLNRE-RG 658
MASP-3   VVPHAECKTSYESR----SGNYSVTENMFCAGYYEGGKDTCLGDSGGAFVIFDDL-SQ 676
MASP-2   IVDHQKCTAAYEKPP---YPRGSVTANMLCAGLESGGKDSCRGDSGGALVFLDSE-TE 645
C1R      VANPQACENWLRGK----NRMDVFSQNMFCAGHPSLKQDACQGDSGGVFAVRDPN-TD 666
C1S      VAPLRKCKEVKVEKPTADAEAYVFTPNMICAGGEK-GMDSCKGDSGGAFAVQDPNDKT 645 (630)

MASP-1   QWYLVGTVSWG--DDCGKKDRYGVYSYIHHNKDWIQRVTGVRN--------- 699
MASP-3   RWVVQGLVSWGGPEECGSKQVYGVYTKVSNYVDWVWEQMGLPQSVVEPQVER 728
MASP-2   RWFVGGIVSWG-SMNCGEAGQYGVYTKVINYIPWIENIISDF---------- 686
C1R      RWVATGIVSWG----IGCSRGYGFYTKVLNYVDWIKKEMEED--------- 705
C1S      KFYAAGLVSWG-----PQCGTYGLYTRVKNYVDWIMKTMQENSTPRED---- 688 (673)
```

Figure 11 : Mutations réalisées sur C1s pour l'étude de la cartographie fonctionnelle.

Séquence de la région catalytique de protéases homologues des voies lectine et classique du complément comprenant les modules CCP1, CCP2, le peptide d'activation AP et le domaine SP. Numérotation des protéines avec peptide signal et sans peptide signal en parenthèse. La triade catalytique est indiquée en orange, les résidus identifiés comme signature de la famille CCP-SP sont indiqués en bleu cyan. Les régions mutées sont encadrées et les mutations effectuées sont désignées en rouge. Les résidus surlignés en rouge ont été identifiés par notre équipe (Bally et al., 2005), ceux en jaune par Duncan et al., 2012. Les résidus de MASP-2 surlignés en vert entrent en jeu dans la reconnaissance de C4 par la protéase et ont été identifiés à partir de la structure cristallographique du complexe CCP1-CCP2-SerPr MASP-2/C4 (Kidmose et al., 2012).

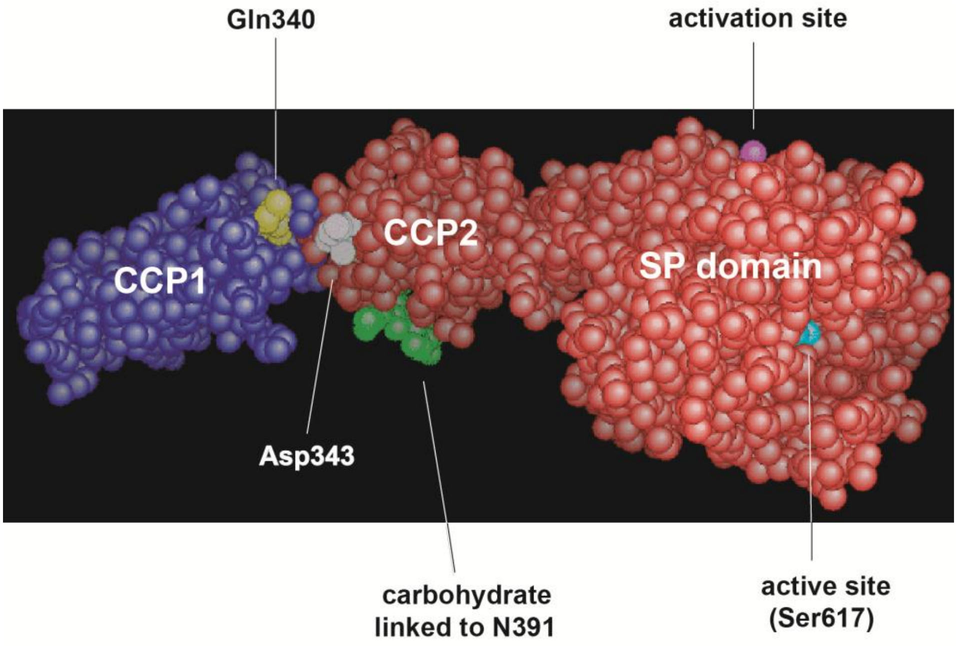


Figure 12 : Modèle tridimensionnel de la région catalytique de C1s.
Le module CCP1 est représenté en bleu, le module CCP2 et le domaine SP en rouge. Les chaînes latérales des résidus identifiés comme importants pour la coupure de C4 par C1s dans la région charnière entre le CCP1 et le CCP2 (Gln 340 et Asp 343) sont indiqués en jaune et gris, respectivement. Les positions des sites de coupure (Ser 617) et d'activation (Phe 417) sont représentées en rose et bleu respectivement. Le site de glycosylation est identifié en vert.

3.1.7 - Régulation de l'activité enzymatique de C1s par C1-inhibiteur – une interaction atypique avec l'héparine.

Encadrement : Marina Lamairia et Sarah Ancelet – BTS Anabiotech

Publication : P16 Rossi et al., *J. Immunol.* (2010)

Communication : C30

Projet : C1-inhibiteur est l'inhibiteur physiologique de C1s, C1r, MASP-1 et -2 et de certaines protéases à sérine de la coagulation telles que la kallicréine plasmatique, et les facteurs XI et XII. C'est un inhibiteur de la famille des SERPIN, SERine Protease INhibitor qui fonctionne comme un substrat « suicide » se fixant de manière irréversible au site actif des enzymes ciblées (Figure 22 et 23). C1- inhibiteur est l'élément le plus grand de la famille des serpines car il possède en plus de son domaine Serpine C-terminal, seul domaine responsable de son activité inhibitrice, une longue extension N-terminale (env. 100 résidus) dont la fonction n'est pas encore clairement définie. C'est également une protéine extrêmement glycosylée qui comprend entre 13 et 20 sites potentiels de O- et N- glycosylation dont seulement 3 N-glycosylations sur le domaine serpine. Son action inhibitrice peut être potentialisée par l'héparine ce qui est le cas pour C1s et le facteur XI mais pas pour ses autres cibles.

Avec pour objectif de définir comment cette potentialisation est réalisée au niveau moléculaire dans l'inhibition de C1s, et de définir si les sucres interviennent dans ce mécanisme, nous avons exprimé dans les cellules d'insectes le domaine serpine de C1-inhibiteur (C1-inhibiteur tronqué de ses 97 premiers résidus) en éliminant un à un les trois sucres N-liés en position 216, 231 et 330 (Figure 13).

Résultats : **(i)** Le fragment aglycosylé en position 330 n'a pu être caractérisé car son rendement d'expression s'est trouvé dramatiquement réduit par rapport au fragment serpine sauvage indiquant que ce sucre est probablement crucial pour le repliement et/ou la stabilité de C1-inhibiteur. Par contre, l'élimination des deux autres sucres en position 216 et 231 a permis de montrer que ces derniers n'ont aucune implication dans l'inhibition de C1s et dans l'inhibition de l'activation du complexe C1.

(ii) L'interaction de C1-inhibiteur et de C1s avec des fragments d'héparine de taille variable (2 à 50 unités saccharidiques) a été étudiée par résonance plasmonique de surface (SPR) et thermal shift assay (TSA) montrant que la taille minimale d'interaction avec C1 inhibiteur et C1s est respectivement de six et huit unités oligosaccharidiques. Ces valeurs sont en accord avec les résultats d'inhibition qui montrent 50% de la potentialisation maximale pour un fragment d'héparine de 10 unités saccharidiques. Dans leur ensemble ces résultats nous ont permis de proposer un modèle d'interaction révélant notamment comment l'héparine stimule son activité inhibitrice des protéases par un mécanisme inhabituel de type « sandwich ». Ces résultats permettent de mieux comprendre le fonctionnement de cet inhibiteur qui joue un rôle essentiel dans le contrôle de l'inflammation (Rossi et al., 2010a).

L'un des objectif de ce travail, que nous avons entrepris dès 2005, était également d'obtenir la structure tridimensionnelle du complexe C1s/C1-inhibiteur en produisant et purifiant les fragments Serpine de C1-inhibiteur et CCP1-CCP2-SerPr de C1s. Après de nombreux essais, nous n'avons pas été en mesure de purifier une fraction assez homogène du complexe pour pouvoir la cristalliser. Les serpines sont des molécules très instables en solution. Selon les conditions, elles adoptent plusieurs types de conformations inactives qui peuvent être irréversibles (Figure 23, Gettins, 2002). C'est un phénomène que nous avons observé tout particulièrement sur le domaine serpine de C1-inhibiteur délété de sa partie N-terminale. En 2007, Beinrohr et al., ont rencontré le même obstacle et ont publié, simultanément à nos essais, la structure d'une forme latente (une des formes inactives) de C1-inhibiteur qui n'apporte pas beaucoup d'informations sur la spécificité de C1-inhibiteur pour C1s.

Ils ont néanmoins émis l'hypothèse de l'existence d'un mécanisme sandwich pour la potentialisation de l'inhibition de C1s par C1-inhibiteur que nous avons confirmé par ces études fonctionnelles.

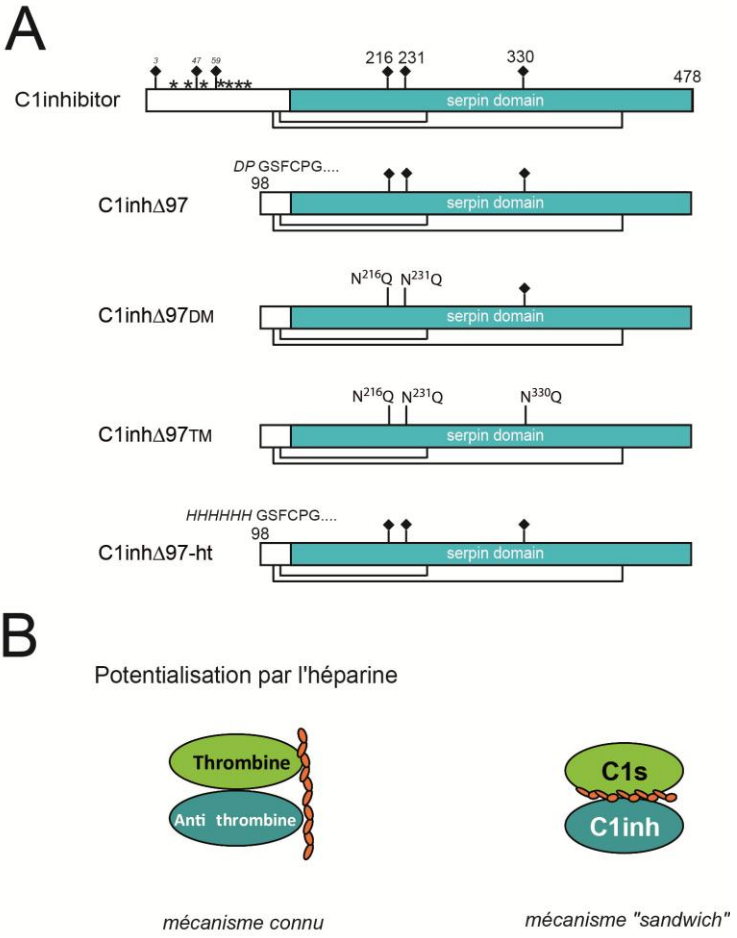

A
3
47
59
216 231
330
478
C1inhibitor
serpin domain
DP GSFCPG....
98
C1inhΔ97
N216Q N231Q
C1inhΔ97DM
N216Q N231Q N330Q
C1inhΔ97TM
HHHHHH GSFCPG....
98
C1inhΔ97-ht
B
Potentialisation par l'héparine
Thrombine
Anti thrombine
C1s
C1inh
mécanisme connu
mécanisme "sandwich"

Figure 13 : Fragments serpine de C1-inhibiteur produits dans les cellules d'insectes.

(A) *Le domaine serpine (bleu) de C1-inhibiteur a été produit dans le système baculovirus/cellules d'insectes.les (◆) et (*) correspondent aux N- et O-glycosylations, respectivement. Les N-glycosylations ont été supprimées par mutagenèse dirigée des acides aminés N en Q. Le fragment C1inhΔ97-ht a été produit avec un His-TAG en N-terminal.*

(B) *Représentation schématique du complexe serpine/protéase cible dans le cas connu de la thrombine avec l'anti-thrombine et dans le cas de l'étude Rossi et al. J. Immunol. 2010.*

3.2 - Les protéases MASP de la voie lectine - Un système ancestral majeur de défense innée contre l'infection.

Les nombreux travaux effectués sur les voies “ classique ” et “ alterne ” du complément définissent à l'heure actuelle assez précisément les mécanismes impliqués dans l'activation et les effets biologiques de ce système. Au début des années 2000, ce n'était pas le cas de la **“ voie lectine ”,** troisième voie d'activation du complément, dont la découverte est plus récente (Matsushita and Fujita, 1992, Ji et al., 1993, Thiel et al., 1997). La question fondamentale à laquelle plusieurs équipes ont tenté de répondre au cours des années est de savoir quel est le rôle précis de cette voie d'activation dans la réponse immune innée.

La « voie lectine » est déclenchée par un complexe protéasique (MBL-MASP ou ficoline-MASP) qui présente *a priori* certaines similarités avec C1, le complexe responsable de l'activation de la voie classique du complément (Figures 1 et 2). La fonction de reconnaissance des complexes MBL-MASP est en effet réalisée par la **M**annan-**B**inding **L**ectin, une protéine oligomérique de la famille des "**collectines**" composée de tiges collagènes prolongées par des domaines homotrimériques de la famille "C-type lectin", dont la structure d'ensemble rappelle celle de la sous-unité C1q de C1 (Hoppe and Reid, 1994, Turner, 1996) (Figure 3). De même, la fonction catalytique des complexes **MBL-MASP** est assurée par deux protéases à sérine modulaires MASP-1 et MASP-2 (**M**BL-**A**ssociated **S**erine **P**roteases) qui présentent le même type d'organisation structurale que C1r et C1s, les protéases du complexe C1 (Arlaud et al., 1987).

D'une façon générale, les collectines exercent une fonction de connexion entre des particules présentant à leur surface des oligosaccharides (telles que les microorganismes et les allergènes) et le système immunitaire de l'hôte, et remplissent de ce fait des fonctions importantes dans la **défense innée chez les mammifères** (Holmskov et al., 1994). Le très grand intérêt du complexe MBL-MASP réside dans la capacité de la **MBL** de reconnaître des oligosaccharides contenant des résidus mannose ou N-acétylglucosamine, présents à la surface d'un grand nombre de micro-organismes pathogènes. Cette propriété confère à la MBL les caractéristiques d'un "**anticorps universel**" ou un "**ante-antibody**" (Turner, 1996). Le complexe MBL-MASP présente donc un spectre de reconnaissance très large, qui lui confère un rôle majeur dans la protection contre l'infection bactérienne et virale. Différentes études réalisées chez des individus présentant un taux faible de MBL, ou déficients par suite de mutations génétiques, révèlent chez ces sujets une susceptibilité accrue aux infections récurrentes, et étayent donc cette hypothèse (Turner et al., 1991, Summerfield et al., 1995). La MBL a été identifiée chez de nombreuses espèces de vertébrés, et des clones d'ADNc correspondant à deux protéines de type MASP-1 ont été isolés d'un ascidien (*Halocynthia roretzi*), espèce intermédiaire entre les vertébrés et les invertébrés (Ji et al., 1997), ce qui suggère que la voie lectine pourrait représenter **le mécanisme d'activation du complément le plus ancien**, et donc un **système de défense ancestral** (Figure 2).

Alors qu'on disposait d'informations précises sur la structure tridimensionnelle et la fonction de la MBL (Sheriff et al., 1994)**,** les connaissances étaient beaucoup plus imprécises en ce qui concerne les deux protéases MASP-1 et MASP-2 du complexe MBL-MASP. Initialement, une seule protéase (MASP-1) associée à la MBL avait été identifiée (Matsushita and Fujita, 1992), et ses caractéristiques fonctionnelles laissaient supposer qu'elle combinait les propriétés de C1r et de C1s, respectivement responsables de l'activation de C1 et de son activité protéolytique. La découverte de MASP-2 et sa caractérisation fonctionnelle (Thiel et al., 1997) ont suggéré au contraire que ces deux protéases, au moins, sont impliquées dans les complexes MBL-MASP où elles pourraient être fonctionnellement équivalentes à C1r (MASP-1) et C1s (MASP-2).

Le fait que ces protéines soient présentes en quantités très faibles dans le sérum humain (quelques microgrammes par ml), et la difficulté à les séparer totalement l'une de l'autre rendait cependant difficile une caractérisation fonctionnelle définitive, c'est pourquoi nous avons entrepris de caractériser les deux **protéases à sérine MASP-1** et **MASP-2 recombinantes produites dans les cellules d'insectes**.

3.2.1 - Caractérisation structurale et fonctionnelle des protéases MASP-1 et MASP-2

Collaboration interne : Sandor Cseh, post-doctorant.

Publications : P9 Rossi V. et al., *J Biol Chem.* (2001), **P8** Thielens et al., *J. Immunol.* (2001),

P13 Kerr et al., *Mol. Immunol.* (2008).

Communications : C18, C22, C23

Projet : MASP-1 et MASP-2 entières et deux fragments tronqués CCP1-CCP2-SerPr et CCP2-SerPr de leur région catalytique comportant respectivement, deux ou un seul module CCP adjacents au domaine protéase à sérine, ont donc été produits dans les cellules d'insecte.

Résultats : Les résultats obtenus indiquent que :

(i) MASP-2 s'autoactive et **clive spécifiquement les mêmes substrats que C1s (C2 et C4)** avec des sites de reconnaissance homologues, sur les modules CCP et le domaine protéase à sérine pour la reconnaissance de C4, et uniquement sur le domaine protéase à sérine pour C2. D'autre part, cette étude indique que **MASP-2 est plus active que C1s**, respectivement de 3 et 23 fois, pour la protéolyse de C2 et C4. Ces résultats nous ont permis de clarifier les rôles respectifs de MASP-2 et

MASP-1 qui n’étaient pas alors clairement établis. Parmi les MASPs, c’est uniquement MASP-2 qui est responsable du déclenchement de la voie lectine d’activation du complément.

(ii) Pour ce qui concerne **MASP-1**, en revanche, à la suite de ce travail, la question de son rôle physiologique restait encore en suspens car son activité est très faible pour la protéolyse de C2 et qu’elle ne présente aucune activité protéolytique vis-à-vis de C4 (Rossi et al., 2001). Contrairement à C1s, MASP-2 est également capable de couper C3 mais cette activité est très faible et n’est donc probablement pas significative d’un point de vue physiologique. Ces résultats nous ont également permis d’observer que **la spécificité de MASP-2 est plus large que celle de C1s**. Au cours de l’évolution les protéases modulaires du complément ont donc évolué vers des enzymes à spécificité restreinte.

Des études visant à caractériser le type d’association et la stoechiométrie du complexe entre MASP-1, MASP-2 et la MBL ont également été menées en parallèle dans le laboratoire. Les résultats indiquent que contrairement à C1s et C1r, qui s’associent en tétramère dans le complexe C1, MASP-1 et MASP-2 interagissent sous forme d’homodimères avec la MBL, pour former des complexes distincts (Thielens et al., 2001).

Ces observations permettent d’établir un modèle des complexes MBL-MASP qui est très différent de celui de C1, contrairement à ce qui était envisagé jusqu’alors. Le complexe MBL-MASP-2 comprendrait ainsi une molécule de MBL associée à un dimère de MASP-2 qui aurait à la fois les propriétés d’auto-activation de C1r, et la capacité d’activer la cascade protéolytique du complément par coupure de C2 et C4 comme C1s, mais de façon beaucoup plus efficace. Cette efficacité accrue confère à MASP-2 un rôle fonctionnel important dans l’immunité innée, importance qui jusqu’alors pouvait être mise en doute du fait de la faible concentration de MASP-2 par rapport à C1s dans le sérum humain (0,5 mg/l, contre 50 mg/l pour C1s).

Notre étude a permis de définir le rôle de MASP-2 comme étant la protéase homologue de C1s, responsable du déclenchement de la voie lectine. En revanche le rôle de MASP-1 est resté énigmatique jusqu’à 2012. Heja et al., en 2012 ont démontré en développant des inhibiteurs spécifiques de MASP-1 ou MASP-2 que contrairement à ce qu’on avait pu imaginer au départ d’après des résultats d’expériences menées in vitro, MASP-1 est essentielle à l’activation de la voie lectine et rempli le rôle de C1r dans le complexe C1 en s’autoactivant pour activer ensuite MASP-2. Cette caractéristique a été validée à la même époque par une étude chez un patient ayant une mutation non sens dans le gène codant MASP-1/MASP-3 (Degn et al., 2012).

3.2.2 - Identification des déterminants structuraux responsables de l'efficacité accrue de MASP-2 vis-à-vis de ses substrats C2 et C4 – Caractérisation de molécules recombinantes chimères de C1s et MASP-2

Encadrement : Florence Teillet, DEA

Publication : P12 Rossi V. et al., *J.Biol.Chem. (*2005), **P13** Kerr et al*., Mol. Immunol.*(2008).

Communications : C25, C26, C28

Projet : MASP-2 ayant été identifiée comme la protéase du complexe activateur de la voie lectine responsable de la coupure de C4 et C2, avec une efficacité de coupure supérieure à celle de C1s, nous avons poursuivi l'étude dans le but d'identifier les régions de MASP-2 responsables de cette meilleure efficacité. Nous avons conçu par mutagenèse dirigée des hybrides de la région catalytique de C1s/MASP-2 dans lesquels les modules CCP ou le domaine protéase à sérine de C1s ont été échangés avec leur domaine correspondant de MASP-2 (Figure 14).

Résultats : Ces hybrides ont permis de montrer

(i) que **l'autoactivation de MASP-2 est engendrée par son domaine protéase à sérine** puisque la chimère C1s $_{\text{(MASP-2 SP)}}$ s'autoactive, tandis que C1s $_{\text{(MASP-2 CCP1/2)}}$ reste proenzyme ;

(ii) que **l'efficacité accrue de MASP-2** pour les substrats C2 et C4 est due à une meilleure affinité de son domaine SerPr pour C2 à laquelle s'ajoute une affinité accrue des modules CCP1 et CCP2 pour C4 ;

(iii) que les modules CCP de C1s n'ont pas d'implication dans la formation du complexe C1 et dans son activation, puisque la protéine chimère C1s $_{\text{(MASP-2 CCP1/2)}}$ conserve les mêmes propriétés que C1s à cet égard. (Rossi et al., 2005)

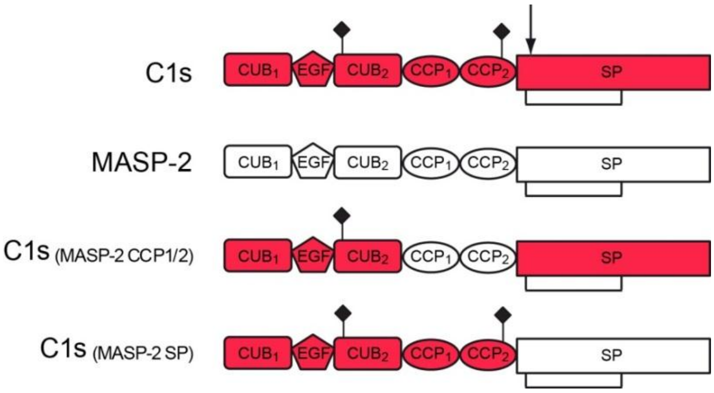


Figure 14 : Chimères de C1s et MASP-2 exprimées dans le système baculovirus/cellules d'insectes. Structure modulaire des 2 protéases C1s et MASP-2 ainsi que des chimères de leur région catalytique (CCP1-2-SerPr). Les protéines chimères ont été conçues en insérant par mutagenèse dirigée des sites de restriction dans le gène de C1s et MASP-2 et par échange de cassette d'ADN entre les deux gènes. Les modules sont définis legende figure 5. La flèche indique le site d'activation soit par coupure par C1r pour C1s et C1s (MASP-2 CCP1/2) soit par autoactivation pour MASP-2 et C1s(MASP-2 SerPr). L'unique pont disulfure représenté est celui reliant le peptide d'activation au domaine protéase à sérine (Rossi el al., J.Biol.Chem. 2005).

Parallèlement, j'ai été contactée par l'équipe du Pr. Robert Pike de l'Université de Monash en Australie pour collaborer à une étude visant à déterminer la spécificité de MASP-2 pour des substrats peptidiques exprimés par la technique du « Phage display ». Cette méthode nous a permis d'identifier les sous-sites S2 et S3 du site actif de MASP-2, qui sont cruciaux pour le positionnement du substrat et sa coupure protéolytique. Ainsi, les peptides présentant des résidus glycine en position P2 et leucine ou un autre acide aminé hydrophobe en position P3 sont des substrats préférentiels de MASP-2. Cette étude a également permis de confirmer ce que nous avions montré dans le travail publié précédemment (Rossi et al., 2001), à savoir que l'efficacité de MASP-2 pour l'hydrolyse des peptides présentant la séquence spécifique de C2, C4 ou C1 inhibiteur est 1000 fois supérieure à celle de C1s (Kerr et al., 2008).

3.2.3 - L'énigme MASP-3

Encadrement : Ludovic Lenclume, stagiaire M2

Publication : P18 Gaboriaud. et al., *Plos One* (2013)

Projet : Troisième membre de la famille MASP, MASP-3 est une protéase à sérine homologue de MASP-1 et MASP-2 dont le rôle reste une énigme depuis sa découverte en 2001 (Dahl et al., 2001). Cette protéase ne clive en effet aucune des molécules impliquées dans les cascades protéolytiques du complément, et l'hypothèse de son rôle potentiel dans un autre système, celui de la coagulation sanguine, a été émise.

Résultats : Dans l'objectif de caractériser sa spécificité, nous avons utilisé une série de substrats synthétiques spécifiques des protéines de la coagulation et pu mettre en évidence que MASP-3 présente une très faible activité de type thrombine. L'affinité de MASP-3 pour ces substrats étant très faible, au moins 10 fois inférieure à celle de la thrombine, on peut imaginer que cette enzyme nécessite un partenaire pour être totalement active.

La résolution récente par notre équipe (Gaboriaud et al., 2013) de la structure de MASP-3 dans sa forme active, en complexe avec un inhibiteur bactérien de protéases à sérine, l'écotine, valide ce constat en montrant que l'enzyme nécessite pour son activation, à la fois la coupure d'une liaison peptidique (cas classique chez les protéases à sérine) mais également un changement conformationnel drastique qui ne pourrait être induit que par association avec un activateur.

Il a été proposé que MASP-3 ait un rôle de régulateur de l'activation de la voie lectine par compétition avec les protéases MASP-2 et MASP-1, mais ceci n'a pas été démontré et l'énigme MASP-3 reste donc encore entière à ce jour.

Chapitre IV

Le complexe C1 – Modèle revisité

C1 est un assemblage macromoléculaire très complexe et les tentatives multiples dans le laboratoire pour obtenir une structure tridimenssionelle sont longtemps restées vaines. Jusqu'à l'étude présentée ici, le modèle conçu par Gérard Arlaud en 1984 était le seul dont nous disposions. A partir des années 2000, l'ensemble des données obtenues dans le laboratoire sur la structure et la fonction de fragments des protéases C1r et C1s, ont finalement conduit à cette étude dans laquelle l'association du tétramère de protéases avec C1q a été disséqué par mutagénèse dirigée. Les résultats obtenus ont abouti à un modèle qui correspond à ce jour à l'image la plus précise du complexe C1.

4.1 - Identification des sites d'assemblage du complexe C1 – Le modèle de C1 revisité.

Publication : P14 Bally et al., *J. Biol. Chem.* (2009)

Communication : C29

Projet : Le complexe C1 est formé par l'assemblage de la protéine de reconnaissance C1q et d'un tétramère calcium-dépendant de deux protéases à sérine homologues C1r et C1s, C1s-C1r-C1r-C1s (Figures 1 et 16). C1q est une molécule hétéro-hexamérique semblable à un bouquet de 6 tulipes dont la région globulaire correspond à la fleur et reconnaît les cibles à éliminer, les tiges étant des régions collagènes associées au tétramère $C1r_2C1s_2$. La complexité de cette association multimoléculaire et sa flexibilité en solution sont des obstacles importants pour l'obtention d'une structure tridimensionnelle. C'est pourquoi, nous avons opté pour une approche de dissection moléculaire de l'interaction du tétramère $C1r_2C1s_2$ avec C1q pour affiner le modèle dont nous disposions et qui datait de 1987 (Arlaud et al., 1987)

Résultats : La résolution par diffraction aux rayons X de la structure du fragment N-terminal CUB1-EGF de C1s avait conduit à un modèle de l'interaction du tétramère avec C1q selon lequel les tiges collagènes interagissaient avec l'interface C1r/C1s par des interactions ioniques impliquant des résidus acides apportés par le module EGF de C1r (Gregory et al., 2003). Afin d'identifier les résidus de C1s et C1r impliqués dans cette interaction avec C1q, une série de mutants ont été exprimés dans les cellules d'insectes puis testés pour l'interaction avec C1q par résonance plasmonique de surface (SPR). Sur la base des résultats obtenus, nous avons proposé **un modèle revisité** de C1 (Figure 16 C) dans lequel les domaines d'interaction CUB1-EGF-CUB2 de C1r et C1s sont entièrement enfouis dans le cône formé par les tiges collagènes de C1q et fournissent 6 sites d'interaction acides avec six résidus basiques portés chacun par une tige de C1q (Figure 15). Ces sites d'interaction, également impliqués dans la fixation du calcium par les domaines CUB correspondants, sont répartis en deux groupes, un de faible affinité sur le module CUB1 de C1s et deux de forte affinité sur les modules CUB1 et CUB2 de C1r. Les sites identifiés sont homologues de ceux identifiés dans CUB1-EGF-CUB2 de MASP-3 pour l'interaction avec la MBL ou les Ficolines (Teillet et al., 2008).

L'association des protéases au centre du cône défini par les tiges collagènes de C1q est une information qui **réfute le modèle accepté depuis les années 1980** dans lequel les protéases s'enroulaient autour du cône (Figure 16). Dans ce cas de figure l'activation du complément après fixation du complexe C1 sur une cible (bactérie par exemple), induit un changement conformationnel dans C1q qui se transmet des régions globulaires aux tiges collagènes sur lesquelles se trouvent les sites de liaison au tétramère. Ce changement induit l'autoactivation de C1r et l'activation successive de C1s par C1r. C1s peut alors présenter de façon efficace ses régions catalytiques aux substrats C2 et C4 en dehors du cône formé par le reste du complexe C1 (Figure 16D).

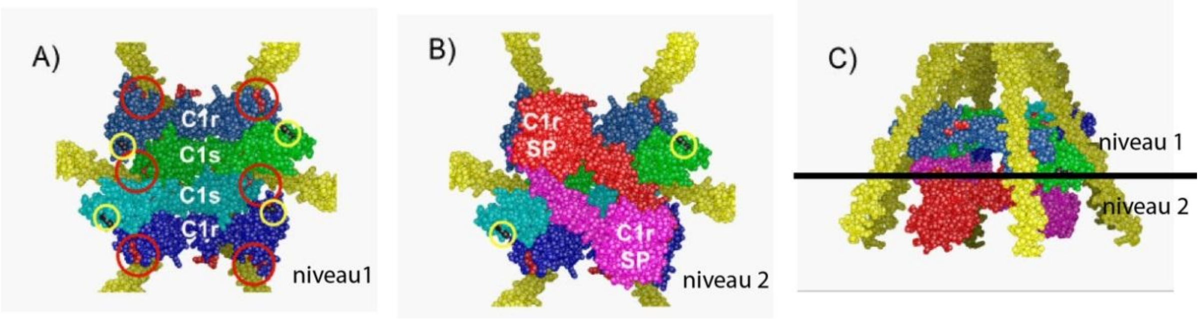


Figure 15 : Modèle 3-D de l'assemblage du complexe C1
D'après Bally et al.,J.Biol.Chem. 2009. (A), vue de dessous du niveau 1 de l'assemblage du tétramère C1s-C1r-C1r-C1s avec C1q. Les sites d'Interaction entre régions CUB1-EGF-CUB2 de C1r (bleu) et C1s (cyan et vert) avec les tiges collagènes de C1q sont représentés par des cercles rouges et les extrémités C-terminales de chaque région (CUB2) sont localisées par des cercles jaunes. (B), vue de dessous du niveau 2 de l'assemblage, positionnement des domaines catalytiques CCP1-CCP2-SP de C1r en rouge et rose. Les domaines similaires de C1s émergeant des extrémités identifiées par un cercle jaune ne sont pas montrés pour plus de clarté. SP, serine protease domain. (C), vue de côté de l'assemblage du complexe C1

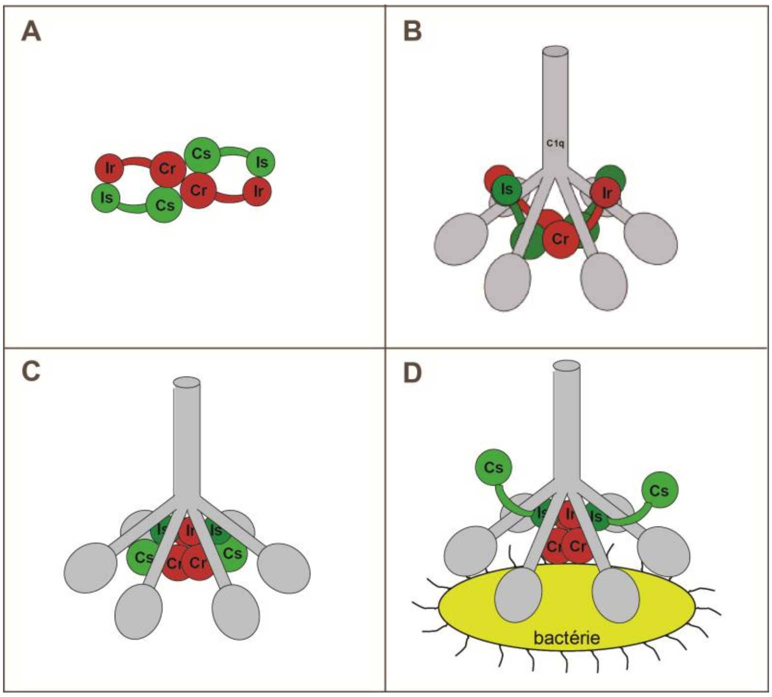


Figure 16 : Représentations schématiques du complexe C1.
(A) *Schéma du tétramère protéasique calcium dépendant C1s2-C1r2 sous sa forme proenzyme. Les régions catalytiques de C1r et C1s (Cr et Cs) interagissent entre elles au centre d'une structure en 8. Les régions d'interaction (Ir et Is) interagissent deux à deux aux deux extrémités opposées du complexe (Villiers et al., 1985, Weiss et al., 1986)* ***(B)*** *Modèle du complexe C1 tel que défini par (Colomb et al., 1984, Arlaud et al., 1987). Le cœur du complexe protéasique formé des domaines catalytiques de C1r et C1s est orienté au centre du cône formé par C1q Le tétramère s'associe aux bras collagènes de 2 tiges/6 par l'intermédiaire des régions d'interactions de C1r et C1s.****(C)*** *Modèle revisité version 2009 selon Bally et al. , 2009. L'ensemble du tétramère protéasique est enfoui dans le cône ménagé par les bras collagènes de C1q, en deux niveaux, un niveau 1 proche de l'amorce du cône contenant les régions d'interaction de C1r et C1s et les régions catalytiques interagissant entre elles en-dessous dans un niveau 2.* ***(D)*** *Représentation de C1 sous forme activée. Lorsque C1q se fixe sur un activateur, par exemple une bactérie, le complexe protéasique est activé et les régions catalytiques de C1s activé s'orientent en dehors du complexe C1 pour pouvoir procéder à la coupure de ses substrats.*

Chapitre V

Aparté sur C9 et AIL

Stage post-doctoral

Ce chapitre décrit un intermède de 18 mois à l'université de Kansas City (Missouri) dans le laboratoire d'Alfred Esser pour un stage post-doctoral qui a été écourté du fait de l'obtention d'un poste d'ATER à l'université Joseph Fourier. Ce stage a été l'occasion pour moi d'aborder de nouvelles techniques : la mutagénèse dirigée et l'expression dans les cellules de mammifères et les bactéries. J'ai également eu l'occasion d'aborder l'étude de C9, une molécule très intéressante, interagissant avec les membranes. Cette molécule est très difficile à manipuler car cytotoxique, responsable de la formation de pore membranaire au sein du complexe d'attaque membranaire, une étape finale de la lyse cellulaire induite par l'action du complément.

Le **laboratoire de Biologie Cellulaire et Biophysique de l'Université du Missouri Kansas-City** est spécialisé dans les aspects biophysiques des **interactions entre les protéines du complément et les membranes** (artificielles ou cellulaires). Ce laboratoire s'intéresse particulièrement à l'étape finale de l'activation du système complémentaire humain, c'est-à-dire au Complexe d'Attaque Membranaire **(MAC)** responsable de la **lyse cellulaire.** L'objectif du laboratoire est de comprendre les mécanismes moléculaires qui régissent l'assemblage du MAC conduisant à la lyse cellulaire.

La lyse cellulaire résultant de l'activation des voies classique, lectine ou alterne du complément est provoquée par l'insertion dans la membrane de la cellule cible d'un complexe d'attaque membranaire C5b-9 assemblé à partir des protéines solubles C5b à C9 (Figure 1). Le mécanisme d'insertion de ces protéines dans les membranes nécessite donc des changements conformationnels permettant leur passage d'un état soluble à un état de fixation membranaire.

La protéine modulaire C9 qui a fait l'objet de mon étude est homologue des composants C6, C7 et C8 du complément et de la **perforine**, une protéine à propriétés cytotoxiques présente dans le lymphocyte T. On ne connaît pas précisément par quel type d'interaction avec la membrane lipidique C9 conduit à la lyse des micro-organismes. Deux types de modèles ont été proposés : (i) un modèle dans lequel plusieurs molécules de C9 s'organiseraient en pores trans-membranaires et (ii) un modèle dans lequel l'interaction d'un nombre limité de molécules de C9 conduirait à la lyse cellulaire par une désorganisation de la membrane.

Afin d'examiner ces deux possibilités, mon sujet de recherche a comporté deux parties : (i) l'étude de l'implication des ponts disulfures de C9 dans la toxicité du MAC et (ii) la topographie de l'insertion membranaire de C9 à l'aide de mutants de glycosylation**.**

5.1 - Implication des ponts disulfures de C9 dans la toxicité du MAC.

Publication : **P15** Rossi et al., *Mol. Immunol.* (2010)

Communications : C11, C13.

Projet : C9 comporte une structure en mosaïque à 5 modules contenant 24 cystéines formant 12 ponts disulfures. Trois ponts disulfures C121-C160, C233-C234 et C359-C384, ne sont pas impliqués dans la stabilisation structurale des modules de C9 (Figure 17). Ces ponts disulfures jouent un rôle primordial dans la polymérisation de C9 et interviendraient dans des échanges thiols-ponts disulfures lors de l'agencement de C9 dans des canaux trans-membranaires. La localisation précise des cystéines impliquées dans ces échanges n'était pas établie au début de ce travail.

Résultats : Dans le but de déterminer le rôle de ces trois ponts disulfures dans la fonction cytotoxique de C9, j'ai produit des mutants des cystéines de la protéine dans des cellules Cos7 et, pour certains, dans des cellules d'insectes. Ces mutants comportent, soit une seule mutation (C121S, C160S, C233T, C234T, C359S et C384S), soit certaines combinaisons dans lesquelles sont éliminés : (i) un pont disulfure (C121S-C160S, C233T-C234T, C359S-C384S), (ii) deux ponts disulfures (C121S-C160S/C233T-C234T), ou (iii) les trois (C121S-C160S/C233T-C234T/C359S-C384S) (Figure 17). J'ai montré que le premier pont disulfure C121-C160 est essentiel au repliement et à la sécrétion de C9 tandis que l'élimination des deux autres ponts C233-C234 et C359-C384 n'affecte pas la cytotoxicité de C9. Il découle donc de ces résultats que s'il y a échange thiols-ponts disulfures lors de la polymérisation de C9, il ne peut s'agir que du doublet de cystéines C121S-C160S, qui est également essentiel au repliement de la protéine.

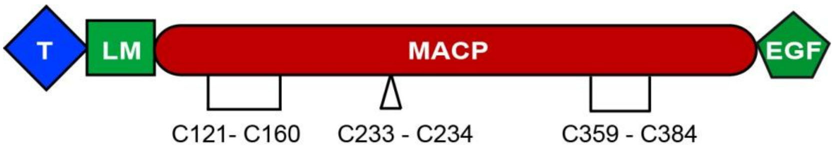


Figure 17: Représentation schématique de la structure modulaire de C9.
T, module Thrombospondine 1, LM, module de type "LDL receptor", MACP, protéine du MAC homologue de la perforine, EGF, module de type facteur de croissance épidermique. Parmi les 12 ponts disulfures de C9, sont représentées uniquement les demi-cystines C121-C160, C233-C234 et C359-C384 qui nous ont interessées.

5.2 - Topographie de l'insertion membranaire de C9 - utilisation de mutants de glycosylation.

Publication : **P15** Rossi et al., *Mol. Immunol.* (2010)

D'après un modèle élaboré par Peitsch et al., en 1990, l'activité cytotoxique de C9 serait due à sa polymérisation sous forme de pores trans-membranaires. En effet, C9 comporterait deux segments susceptibles de former deux hélices trans-membranaires (en position 292-308 et 313-333), le reste de la molécule en N- et C-terminal étant orienté vers l'extérieur de la cellule.

Afin d'étudier ce modèle d'insertion membranaire de C9 j'ai construit des mutants glycosylés de la protéine. Un premier mutant a été conçu dans lequel les deux sites naturels de glycosylation de C9 ont été supprimés et un nouveau site de glycosylation a été créé en position 311. D'après le modèle, cette mutation se situe dans le segment intracellulaire de la protéine (Figure 18). L'activité cytolytique de ce mutant est comparable à celle de C9 sérique, ce qui indique que la translocation membranaire de ce segment est peu vraisemblable, le segment glycosylé ne pouvant traverser la membrane cellulaire. Des expériences de cytométrie de flux nous ont permis de conforter cette hypothèse car elles ont montré que C9 membranaire peut être détecté par un anticorps spécifique de la région 308-316, ce qui serait impossible si le segment était intracellulaire. La région étudiée ne pouvant être orientée du côté cytoplasmique de la cellule, cette étude démontre que l'insertion membranaire de C9 ne se fait pas selon le modèle proposé.

D'après Taylor et al., 1994, C9 exercerait son activité cytotoxique sous forme polymérique, la polymérisation impliquant la région N-terminale de la protéine (à partir du résidu 23). Afin de vérifier cette hypothèse, un mutant comportant un site unique de glycosylation en position 26 a été créé. L'activité cytolytique de ce mutant est comparable à celle de C9 sérique ce qui indique que cette région n'est pas essentielle à la fonction de la protéine. Le travail réalisé sur C9 a permis dans son ensemble de refuter le modèle selon lequel C9 forme un pore transmembranaire et propose un modèle de déstabilisation de la bicouche lipidique par une interaction longitudinale de C9 du côté externe de la membrane (Rossi et al., 2010b).

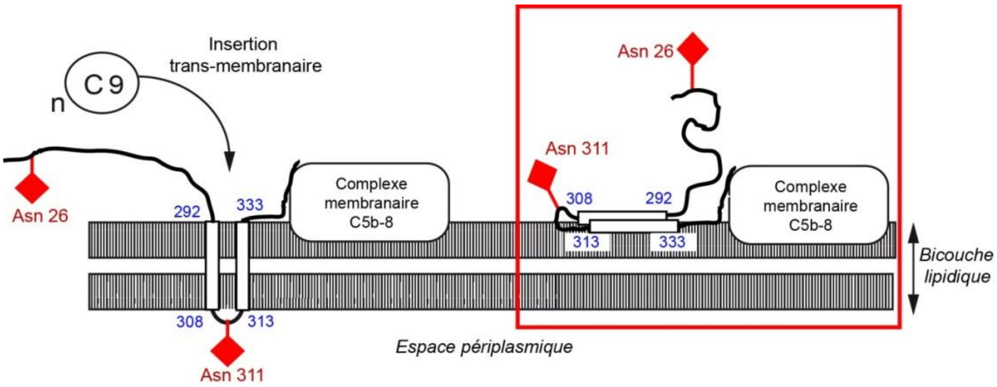


Figure 18 : Topographie de l'interaction de C9 avec la membrane par des mutants de glycosylation.
Modèle de la perméabilisation de la membrane bactérienne par C9 proposé par Peitsch et al., 1990 (côté gauche). C9 traverse la bicouche lipidique de la membrane bactérienne par l'intermédiaire de deux hélices alpha (292-308) et (313-333), les acides aminés 308 à 313 étant localisés sur la face interne du périplasme. Les deux sites de N-glycosylation insérés par mutagenèse dirigée en position 26 et 311 sont indiqués en rouge. Modèle revisité par nos soins(côté droit) : les deux sites de N-glycosylation ne modifient pas l'activité cytolytique de C9 ce qui (entre autres) valide le modèle encadré en rouge que nous avons publié dans Rossi et al., Mol. Immunol.2010. Dans ce modèle les deux hélices alpha interagissent avec la face externe de la membrane périplasmique conduisant à la lyse bactérienne par déstabilisation de la membrane.

5.3 - Etude de la protéine membranaire AIL - Construction d'un système membranaire artificiel

Communication : C12

Le second projet de recherche de mon stage post-doctoral a porté sur les relations structure-fonction de la protéine membranaire **AIL de *Yersinia enterocolitica***. Cette bactérie pathogène est à l'origine chez l'être humain d'infections gastro-intestinales qui peuvent conduire à des septicémies ou des arthrites. La protéine Ail remplit un rôle important dans la **pathogénicité de *Yersinia enterocolitica***, non seulement parce qu'elle a des propriétés d'adhérence et de pénétration membranaire, mais également parce qu'elle est responsable de la **résistance de la bactérie à l'élimination par le système du complément**. D'autres facteurs de virulence homologues ont également été identifiés chez diverses bactéries, telles que lom (chez *E. coli* après infection par le phage λ), pag c et rck (chez *Salmonella thyphimurium*) ou ompx (chez *Enterobacter cloacoae*). A ce jour, la structure de ces protéines de surface n'est pas connue. Cependant un modèle de Ail (17 kDa), basé sur des expériences de mutagenèse dirigée (Beer and Miller, 1992), décrit la protéine comme comportant 8 segments trans-membranaires d'environ 10 acides aminés et n'ayant aucune extension en N- ou C-terminal.

Afin d'étudier les relations structure-fonction de Ail, la protéine doit être isolée de la membrane bactérienne et réinsérée dans un système de vésicules membranaires artificielles. Ma contribution à ce sujet a consisté à produire une protéine comportant deux séquences « étiquettes » en N- et C-terminal (His-Tag et IRS-Tag respectivement) de manière à pouvoir la réorienter dans un système membranaire artificiel. La protéine a été produite avec succès grâce à un système d'expression inductible (PASK-75), mais elle est majoritairement périplasmique. Un clonage successif dans un autre vecteur sous le contrôle du promoteur naturel de Ail a été effectué mais la caractérisation de la protéine produite et la suite de l'étude n'ont pu être menées à terme.

Ce stage post-doctoral était programmé pour s'achever en **Avril 1999.** En raison de mon recrutement à un poste d'ATER en octobre 1998, la phase ultime d'exploitation des résultats a dû être écourtée.

Chapitre VI

Protéines de l'immunité innée à l'interface hôte-pathogène

Etude du récepteur CR1 et de la thrombospondine 1

Depuis 2010, le laboratoire d'enzymologie moléculaire a été réorganisé et est maintenant dirigé par Nicole Thielens. Son thème de recherche global porte sur l'étude des protéines de l'immunité innée à l'interface hôte-pathogènes. Dans ce cadre, ma thématique de recherche a évolué et s'est maintenant tournée vers l'étude des relations structure fonction qui entrent en jeu dans l'interaction des collagènes de défense (C1q, MBL, Ficolines) avec des récepteurs présents à la surface des cellules phagocytaires et impliqués dans la phagocytose.

Dans un premier temps, dans le cadre de la thèse de Mickael Jacquet de 2009 à 2012, nous nous sommes intéressés plus particulièrement à l'étude des interactions mettant en jeu deux récepteurs des collagènes de défense du complément, CR1 et la thrombospondine 1 (TSP1).

Les collagènes de défense du complément humain, C1q, MBL et ficolines, font partie de l'arsenal du système immunitaire de l'hôte impliqué dans la première ligne de défense anti-microbienne (Matsushita and Fujita, 2001, Bohlson et al., 2007) Ces molécules oligomériques solubles sont composées de tiges collagènes prolongées par des domaines globulaires qui, dans le cas de la MBL et des ficolines, sont des domaines de type lectine. Cette organisation structurale confère aux collagènes de défense un rôle bivalent. Leurs régions globulaires à une extrêmité de la protéine ciblent spécifiquement les signaux de dangers à la surface des pathogènes (PAMPS) et des cellules du SOI altéré (ACAMPS). Les tiges collagènes, à l'autre extrémité, interagissent avec un ensemble d'effecteurs immuns conduisant à l'activation du complément par les protéases à sérine associées (C1s, C1r, et MASPs) et/ou à l'association avec des protéines de surface ou des récepteurs sur les cellules phagocytaires comme la calréticuline (CRT), le récepteur du complément CR1/CD35, ou autres récepteurs tels que LRP-1/CD91 (Figure 19).

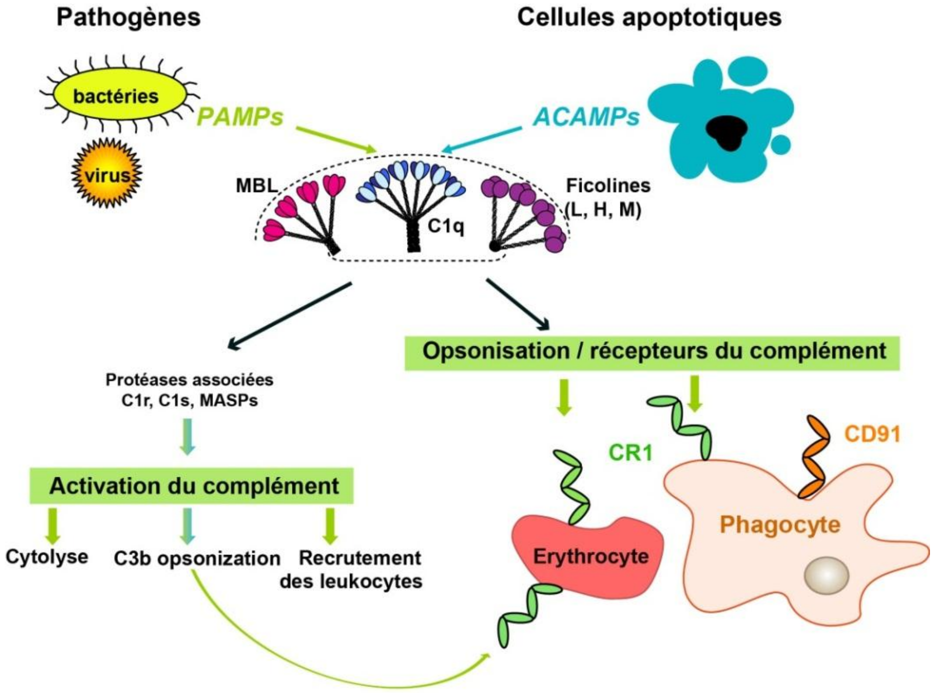

Pathogènes
Cellules apoptotiques
bactéries
virus
PAMPs
ACAMPs
MBL
C1q
Ficolines
(L, H, M)
Opsonisation / récepteurs du complément
Protéases associées
C1r, C1s, MASPs
Activation du complément
Cytolyse
C3b opsonization
Recrutement
des leukocytes
CR1
CD91
Erythrocyte
Phagocyte

Figure 19 : Les collagènes de défense C1q, MBL et ficolines, senseurs de l'immunité innée.
C1q, la MBL et les ficolines reconnaissent par l'intermédiaire de leurs domaines globulaires des motifs moléculaires exprimés à la surface des pathogènes et des cellules du soi altéré (PAPMs et ACAMPs). Cette fixation multivalente déclenche deux types de mécanismes effecteurs de l'immunité qui conduisent à l'élimination de l'agent ciblé : (i) les collagènes de défense activent le complément par l'intermédiaire des protéases à sérine associées, C1r, C1s et MASPs et (ii) ils servent de molécules de pontage entre les agents à éliminer et les cellules effectrices qui expriment à leur surface des récepteurs spécifiques, tel que le récepteur CR1 du complément ou CD91.

6.1 - Etude du récepteur CR1

Financement : *ANR PIRIBIO 2010-2013*

Collaboration interne : Gian Luca Cioci, post-doctorant.

Encadrement : Mickael Jacquet, thèse IRTELIS (CEA), soutenue le 24 septembre 2012

Publication : P17 Jacquet et al., *J. Immunol* (2013)

Communication C31

Projet : CR1 est une molécule polymorphique ayant un double rôle dans l'immunité innée : (i) un rôle de **régulateur** visant à contrôler les effets inflammatoires engendrés par l'activation du complément et (ii) un rôle de **récepteur** des particules adsorbées par les opsonines du complément. Une implication de CR1 dans diverses maladies auto-immunes et inflammatoires a été mis en évidence et suscite un intérêt croissant dans le diagnostic et la thérapie de ces maladies (Khera and Das, 2009).

L'interaction avec les particules opsonisées conduit à deux types d'effets biologiques selon le type cellulaire à la surface duquel CR1 est exprimé. Sur les érythrocytes, CR1 a un rôle de molécule de pontage et intervient dans la clairance des agents étrangers du flux sanguin vers le foie, tandis que CR1 exprimé à la surface des monocytes, macrophages ou neutrophiles induit la phagocytose des particules opsonisées (Figure 17).

Du fait de sa large dispersion dans le sang sur les érythrocytes, CR1 est la principale cible de *Plasmodium falciparum*, le parasite responsable de la malaria. La molécule pFEMP1 exprimée à la surface du parasite interagit spécifiquement avec CR1, conduisant à une agrégation des érythrocytes par formation de « rosettes » dans les vaisseaux sanguins. La fixation de pFEMP1 est différente selon le phénotype de CR1 et conduit à des formes de malaria plus ou moins sévères. Trois types de polymorphismes différents ont été décrits pour CR1 : un polymorphisme lié à la taille, un polymorphisme concernant la quantité exprimée à la surface des cellules, et un polymorphisme concernant deux mutations utilisées comme marqueurs de groupe sanguin, le groupe « Knops » (KN) (Moulds et al., 2001).

CR1 présente une homologie avec les protéases C1s et MASP-2 que j'ai étudiées jusqu'à présent car cette protéine est composée exclusivement, dans sa partie extracellulaire de 30 modules CCP. Les 28 modules les plus externes sont répartis en 4 régions LHR (Long Homologous Repeats) A à D sur lesquelles ont été localisés les sites d'interaction des opsonines C3b et C4b et le site d'interaction avec la molécule PfEMP1 du parasite *P. falciparum*. Si les opsonines C3b et C4b sont connues depuis longtemps comme ligands de CR1, ce n'est que récemment qu'il a été montré que

C1q et la MBL sont également capables d'interagir avec CR1 (Klickstein et al., 1997, Ghiran et al., 2000).

Le projet de thèse de Mickael Jacquet a été d'identifier les déterminants structuraux qui régissent l'interaction entre CR1 et les collagènes de défense. Nous avons produit dans les cellules d'insectes, un fragment comprenant les modules CCP22 à 29 (qui correspond à la partie LHR-D) ainsi que toute une série de fragments tronqués de cette région. Ces fragments ont été utilisés pour des tests d'interaction par résonance plasmonique de surface (SPR) couplés à des tests d'interaction en phase solide.

Résultats : Une partie du travail, publiée en 2013 (Jacquet et al., 2013) met en évidence un rôle de la **Ficoline L** comme opsonine capable d'interagir avec CR1 au niveau du LHRD comme c'est le cas pour C1q et la MBL. Ces résultats valident, grâce à des expériences « in vitro » réalisées avec des protéines purifiées, ce qui avait été montré pour C1q et la MBL à partir d'extraits cellulaires. D'autre part, l'utilisation de fragments tronqués de CR1 pointe le rôle déterminant **des modules CCP24 et 25** dans l'interaction de CR1 avec la MBL et la Ficoline L. Au niveau des collagènes de défense, le site de fixation de CR1 fait intervenir une lysine (K55 dans la MBL) impliquée dans l'interaction avec les protéases associées ou un site proche de cette lysine.

La caractérisation de l'interaction de CR1 avec C1q s'est avérée plus difficile car elle fait très probablement intervenir des sites à la fois au niveau de la région collagène et au niveau des régions globulaires de C1q. D'autre part, les tests d'interaction n'ont pu être réalisés que lorsque CR1 est fixé à une surface, ce qui est très probablement relié à son organisation physiologique sous forme de clusters à la surface de la membrane des cellules (Klickstein et al., 1988). Pour ces raisons d'orientation spécifique de CR1 nous avons également réalisé des essais sur un système cellulaire (cellules de mammifères) à la surface duquel nous avons produit le fragment CCP22 à 29 pour étudier ses propriétés d'interaction par cytométrie de flux et immunofluorescence. Bien que le fragment soit exprimé à la surface cellulaire sous forme de clusters, ce qui est décrit pour CR1, nous n'avons pas pu mettre en évidence, de fixation de C1q et de la MB dans nos conditions expérimentales.

Une deuxième partie du projet, qui devrait être finalisée sous peu, a consisté à identifier plus précisément les acides aminés impliqués dans l'interaction des modules CCP24 et CCP25 de CR1 avec les collagènes de défense. Nous avons émis l'hypothèse que cette interaction est électrostatique car elle dépend fortement de la concentration saline du milieu et qu'elle fait intervenir une lysine de la MBL impliquée également dans la fixation des protéases à sérines par le même type d'interaction. La comparaison de séquence de l'ensemble des modules CCP de CR1 fait ressortir 5 résidus acides répartis sur les modules CCP24 et CCP25, que l'on ne retrouve pas dans les autres modules CCP et qui pourraient, d'après le modèle Figure 20, être impliqués dans l'interaction avec une lysine de charge opposée sur les collagènes de défense.

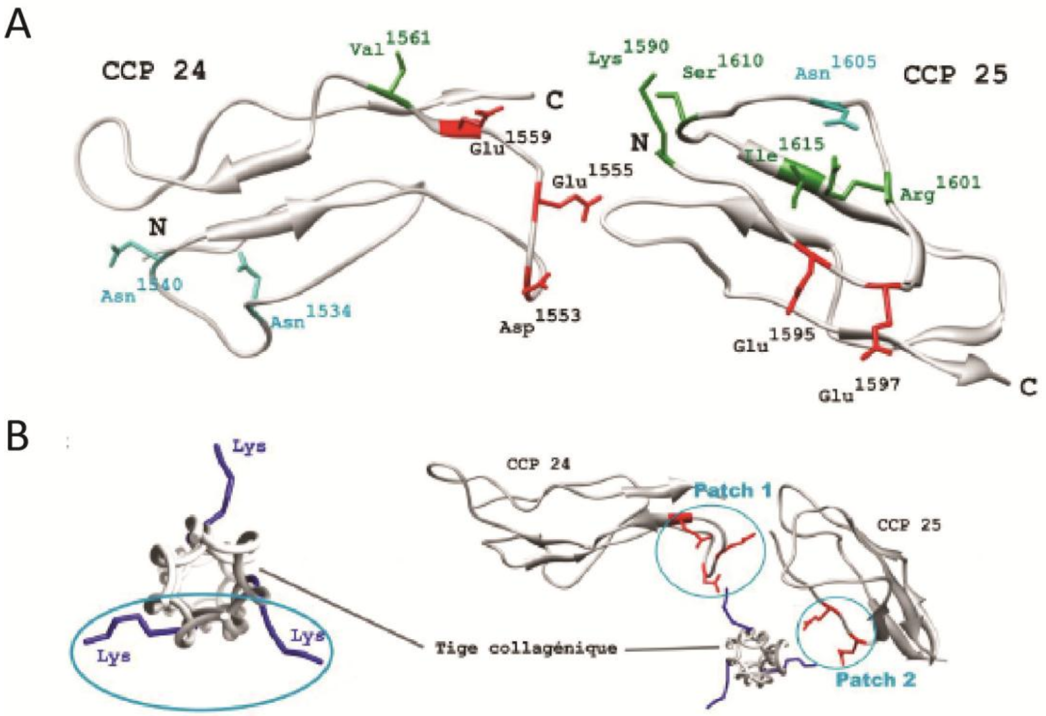

A
CCP 24
Val1561
C
Glu1559
Glu1555
N
Asn1540
Asn1534
Asp1553
Lys1590
Ser1610
Asn1605
CCP 25
N
Ile1615
Arg1601
Glu1595
Glu1597
C
B
Lys
Lys
Lys
Tige collagénique
CCP 24
Patch 1
CCP 25
Patch 2

Figure 20 : Représentation structurale des modules CCP24-25 de CR1.
***(A)** Modèle des modules CCP24 et 25 d'après (Soares et al., 2005) avec la localisation des résidus proposés pour l'interaction avec les collagènes de défense (en rouge). Les polymorphismes sont réprésentés en vert et les asparagines potentiellement N-glycosylées en cyan.* ***(B)** A gauche, représentation structurale de la triple hélice de collagène d'après (Gingras et al., 2011). A droite, modèle d'interaction des résidus acides des modules CCP24-25 de CR1 (en rouge) avec les lysines de la triple hélice de collagène représentées en bleu.*

L'utilisation d'un mutant de l'ensemble des résidus acides candidats en alanine n'a malheureusement pas validé le modèle. CR1 CCP24-25 muté interagit de la même manière avec C1q et la MBL que le fragment équivalent sauvage. D'autre part C1q, dont les lysines impliquées dans la fixation des protéases C1s et C1r sont mutées en alanine, garde une capacité d'interaction équivalente avec CR1 CCP24-25 muté ou sauvage.

Ces résultats laissent supposer que le site de fixation de CR1 à C1q, bien que proche du site d'interaction, est différent de celui des protéases C1r et C1s, contrairement à notre hypothèse initiale. D'après notre modèle, nous avons identifié d'autres résidus qui pourraient être responsables de cette interaction à la fois sur C1q et CR1 et qui devront être validés expérimentalement.

6.2 - La thrombospondine 1 : une plateforme d'ancrage des collagènes de défense

Encadrement : M2 BCI/UJF d'Isma Kermia et stage de 2 mois½- Welcome trust Université de Glasgow de David Rogers (2010)

Projet : La thrombospondine 1 (TSP1) est une molécule de la matrice extracellulaire impliquée dans de nombreuses fonctions cellulaires telles que la phagocytose et l'angiogenèse. Cette molécule interagit par sa région N-terminale avec la calréticuline et le récepteur membranaire CD91/LRP1 et engendre le désassemblage des adhérences focales. TSP1 et les collagènes de défense interagissant tout deux individuellement avec la calréticuline à la surface des cellules, nous nous sommes intéressés à l'interaction éventuelle de TSP1 seule avec les collagènes de défense.

Résultats : Les résultats obtenus indiquent que C1q et la MBL interagissent avec TSP1 et plus précisément, comme dans l'interaction TSP1/calréticuline, avec le domaine N-terminal de TSP1. Aucune information à ce sujet n'a été à ce jour rapportée dans la littérature scientifique et nous avons pour objectif de caractériser cette interaction tant au niveau structural qu'au niveau fonctionnel.

La première partie du projet a consisté à cloner et produire le fragment N-terminal de TSP1 dans *E.coli* (souche origami). Des essais préliminaires en SPR et tests en phase solide ont permis d'identifier les collagènes de défense comme de nouveaux ligands de cette molécule de la matrice extracellulaire.

Le projet a été suspendu et sera réévalué dans le contexte de la nouvelle thématique de groupe qui donne actuellement la priorité aux récepteurs membranaires impliqués dans l'efferocytose.

Ce projet soulève néanmoins plusieurs questions auxquelles nous souhaiterions répondre dans le futur : (i) Quel est **le rôle physiologique de l'interaction TSP1/collagènes de défense ?** On sait que la production de cette molécule à la surface des cellules, comme c'est le cas pour la calréticuline, est amplifiée en cours d'apotose. A-t-elle un rôle de signal « eat-me » comme c'est le cas pour la calréticuline ? Est-ce que cette molécule est un récepteur de molécules opsonisées par les collagènes de défense comme c'est le cas pour CR1 ? (ii) Nous nous attacherons également à **caractériser cette interaction avec les différents collagènes de défense,** comme nous l'avons fait pour CR1. Nous projetons d'utiliser les mêmes méthodes d'interaction en phase solide et de SPR pour déterminer les constantes d'association des deux partenaires de l'interaction. Parallèlement, dans le but de démontrer qu'il y a interaction des collagènes de défense à la surface des cellules normales ou apoptotiques nous projetons de faire de la colocalisation par microscopie de fluorescence, en collaboration avec Philippe Frachet qui développe ce type de méthodes dans notre groupe. TSP1 interagissant par son domaine N-terminal avec l'héparine, les intégrines $\alpha 4\beta 1$, et la calréticuline, l'identification de son site d'interaction avec les collagènes de défense pourra être effectuée par des expériences de compétition.

Chapitre VII

Projets de recherche

Dans ce chapitre je vais décrire mes deux projets de recherche en reproduisant les demandes de financement associées. Chaque partie sera précédée d'une petite introduction décrivant de façon plus détaillée le contexte du projet.

Le projet 1 est une collaboration ANR d'une durée de 3 ans sur les serpines bactériennes du microbiote intestinal. Le projet 2 concerne le rôle du récepteur phagocytaire LRP1 dans l'efferocytose médiée par les collagènes de défense et correspond plus précisément à ma contribution à la nouvelle thématique du groupe IRPAS qui a débuté en 2014.

7.1 - Projet 1 : Inhibiteurs bactériens de protéases à sérines : nouvelles fonctions et potentiel thérapeutique contre les maladies inflammatoires de l'intestin – ANR SerpinGUTARGET

Suite au travail réalisé sur C1-inhibiteur, nous avons été contactés par l'équipe d'Emmanuelle Maguin de l'INRA de Jouy en Josas à l'automne 2013 pour participer à une demande d'ANR sur un sujet qui cherche à déterminer le rôle des serpines produites par les bactéries commensales dans la protection du tractus digestif contre les maladies inflammatoire de l'intestin. Mon rôle sera de caractériser les propriétés inhibitrices de ces serpines.

7.1.1 - Introduction sur le fonctionnement d'une serpine

Les serpines sont des inhibiteurs protéiques très répandus que l'on retrouve dans les organismes procaryotes et eucaryotes (Irving et al., 2000, Rawlings et al., 2004). Ces inhibiteurs d'environ 300-350 résidus ont tous un repliement globulaire analogue formé de trois feuillets beta (A, B et C) et de 8-9 hélices alphas (ha-hI). L'achétype d'une serpine (Serpin A1) est représenté dans la Figure 21 (Law et al., 2006). La région responsable de l'interaction avec une protéase cible est la boucle RCL (Reactive Center Loop) qui est exposée à la surface de la structure de base. L'interaction avec la protéase conduit au clivage d'une liaison peptidique de cette boucle et le mécanisme d'hydrolyse reste incomplet, l'enzyme restant liée de façon irréversible à l'inhibiteur. Pour cela la serpine subit un changement conformationnel drastique par l'insertion d'une partie de la RCL clivée dans le feuillet beta A (en rouge sur la Figure 21).

Ce changement conformationnel est à l'origine de l'instabilité des serpines en solution qui rend leur étude in vitro très difficile. Elles peuvent adopter plusieurs types de conformations inactives et sont très sensibles aux mutations (Figure 22, Gettins, 2002). Ainsi, chez l'homme, plusieurs maladies appellées "serpinopathies" sont liées à une déficience fonctionnelle "conformationnelle" des serpines. L'amphysème par exemple est lié à un deficit fonctionnel de la serpine A1 (α1 antitrypsine), la thrombose à un déficit de la serpine C1 (antithrombine) et l'angiooedème à la serpine G1 (C1-inibiteur).

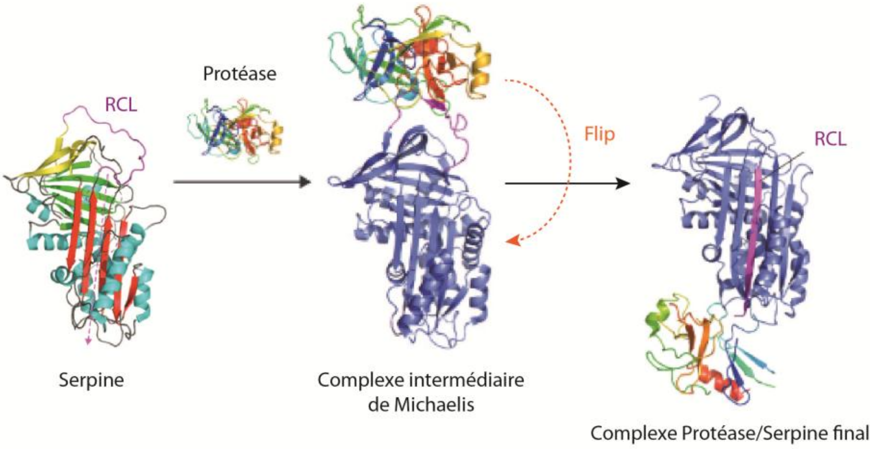

Protéase
RCL
Flip
RCL
Serpine
Complexe intermédiaire de Michaelis
Complexe Protéase/Serpine final

Figure 21 : Structure et mécanisme d'inhibition des serpines.
Adaptée de Law et al., 2006. Structure de la SERPINA1 (Protein Data Bank (PDB) code 1QLP) [Elliot et al., 1996]. Les feuillets A,B et C sont représentés en rouge, vert et jaune respectivement. Les Hélices (ha-hi)et la RCL sont représentées en bleu et magenta respectivement. Le trajet de l'insertion de la bouvcle RCL après coupure par la protéase est indiqué par une flèche pointillée en rouge (Flip). Le complexe intermédiaire de Michaelis correspond à la structure cristallographique de la SERPINA1 et de la trypsine inactive (PDB code 1OPH) (Dementiev et al., 2003) et le complexe final SERPINA1 /trypsine inactiv, à la structure (PDB code 1EZX) [Huntington et al., 2000].

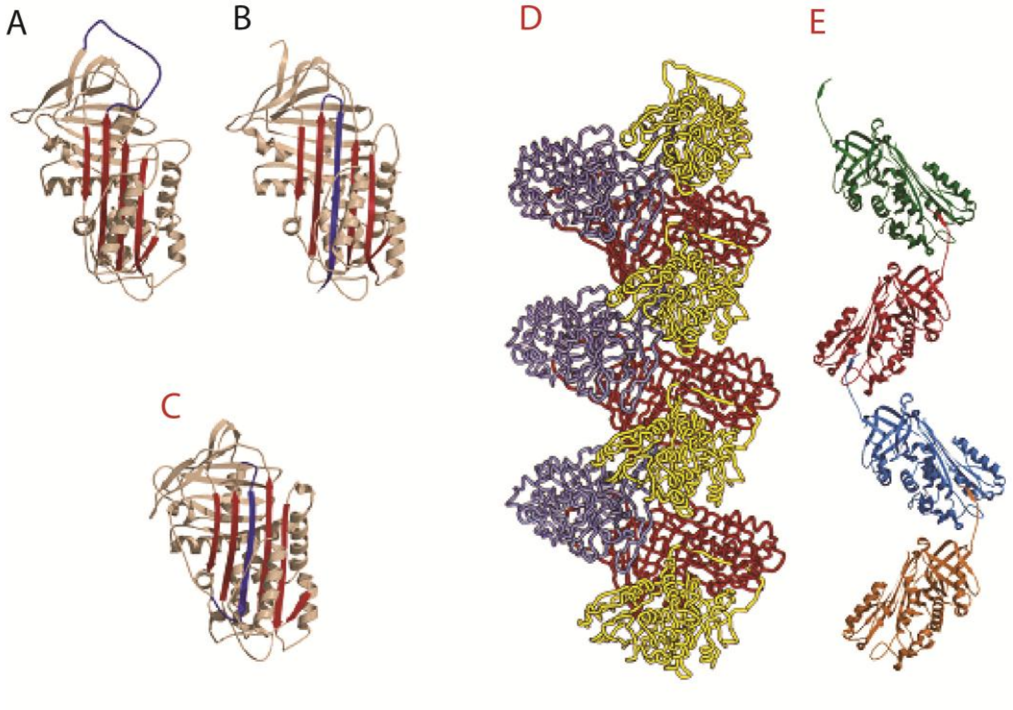


Figure 22 : Les différents états conformationnels des serpines.
Adapté de Gettins et al., 2002. ***(A)*** *conformation native, α1 antitrypsine (PDB code 1QLP) (Elliott et al., 1998),* ***(B)*** *conformation de l'α1 antitrypsine clivée (PDB code 7API) (Engh et al., 1989),* ***(C)*** *PAI-1 en conformation latente (PDB code 1C5G) (Turcker et al., 199), La boucle réactive RCL est représentée en bleu le feuillet beta en rouge.(D)et(E) sont deux types de structures polymériques.* ***(D)*** *structure d'un variant de l'a1antitrypsine impliquant l'insertion dela boucle RCL à l'état natif (non clivé) d'une molécule dans le feuillet beta A de la molécule suivante.* ***(E)*** *structure cristallographique d'un polymère d'a1 antitrypsine dans lequel le polymère est formé par l'insertion de la RCL clivée d'une molécule dans le feuillet beta A de la molécule adjacente. (Mast et al., 1992). Les conformations B, C, D et E sont inactives.*

7.1.2 - Demande de financement ANR SerpinGUTarget

Ce projet de recherche de 3 ans est financé depuis début 2015, mais notre intervention dans le projet est prévue à partir de 2016.

Contexte du projet : Les maladies inflammatoires chroniques de l'intestin (MICI), telles que la maladie de Crohn ou la colite ulcérative, sont des maladies invalidantes qui représentent un réel problème de santé publique dans nos sociétés « modernes ». Elles affectent plusieurs millions de personnes dans le monde et sont en augmentation constante depuis les années 50 du fait du mode de vie occidental et des améliorations de conditions d'hygiène et de régime alimentaire. Ces affections ont un impact important sur les patients car elles créent des perturbations dans leur capacité au travail et leur vie de famille. Les traitements qui existent actuellement sur le marché (mésalamine, Salazopyrine, Infliximab) ciblent principalement la réponse inflammatoire. Ces traitements sont onéreux (7000 euros /an/patient) ; ils ont des effets secondaires invalidants et sont inefficaces sur environ 40% des patients traités. Pour ces raisons, les thérapies actuelles nécessitent d'être améliorées et les cibles diversifiées. Nous avons été contacté par l'équipe d'Emmanuelle MAGUIN (EFI/MICALIS/ INRA Jouy en Josas) afin de participer, en tant que partenaire à une demande ANR, à un champ d'investigation destiné à développer une nouvelle approche thérapeutique contre les MICI, basée sur les inhibiteurs de protéases à sérine de type SERPIN (Serin Protease INhibitors). Ce projet comprend un consortium de 4 équipes dont deux de l'INRA de Jouy en Josas qui vont aborder les deux tiers du projet par des méthodes de métagénomique et génomique fonctionnelles. Les deux autres partenaires sont Grenoblois, et vont compléter ce projet par des études structurales (Josuan Marquez, EMBL) et fonctionnelles (IBS).

J'assure la reponsabilité scientifique du partenariat IBS.

Etat des connaissances : Le microbiote intestinal humain est constitué d'environ 100 000 milliards de bactéries. Ce nombre est presque 10 fois supérieur au nombre de cellules de notre corps, et le nombre de gènes bactériens est environ 150 fois plus important que celui de l'homme. Les rôles du microbiote dans la maturation du système immunitaire, le métabolisme de l'hôte et l'homéostasie intestinale sont désormais admis et ont amené à la nouvelle notion d'un « supra-organisme » humain dont la santé digestive résulte de l'intégration des caractéristiques humaines et des fonctions du microbiote bactérien (Ley et al., 2008) [1]. Les changements de microbiote (dysbioses) observés dans plusieurs pathologies (maladies inflammatoires de l'intestin, diabètes, obésité, etc) (Le Chatelier et al., 2013, Santos Rocha et al., 2012) [2,3] et les effets de greffes de microbiote digestif plaident en faveur d'un rôle majeur du microbiote sur la physiologie, le bien-être et la santé de l'hôte. L'analyse comparative du microbiote chez des individus sains ou souffrant de maladies inflammatoires chroniques de l'intestin a permis d'identifier une souche anti-inflammatoire *Faecalibacterium prausnitzii* comme étant sous-représentée dans le microbiote des individus souffrants de MICI. Des études complémentaires ont mis en évidence l'expression accrue de protéases à sérine humaines dans le tractus intestinal des sujets souffrant de MICI. Une tentative pionnière consistant à traiter avec un inhibiteur de protéases à sérine humain (Elafine) des souris dont l'inflammation du tractus intestinal a été induite chimiquement a montré un effet protecteur de ce type d'inhibiteur (Motta et al., 2012)[4]. Ce champ d'investigation constitue une avancée prometteuse dans le développement de nouveaux agents thérapeutiques.

Résumé des objectifs du projet ANR et implication des différents partenaires :

(i) Sélection et caractérisation *in vitro* des serpines bactériennes du microbiote humain.

Analyses biostatistiques et bioinformatiques afin de produire une liste de serpines bactériennes associées aux MICI et dont la quantité est très différente chez les individus sains par rapport aux individus souffrant de MICI (partenaire 1 et 2 / MICALIS). Choix de 25 serpines, clonage, expression et purification puis test sur des protéases à sérine impliquées dans les MICI (élastase de neutrophile, cathepsine G, Tryptase…) (partenaire 1/ MICALIS). **Caractérisation enzymatique des 6 serpines les plus efficaces pour déterminer leur efficacité inhibitrice par étude cinétique des complexes serpine/protéases (partenaire 4/ IBS).**

(ii) Validation *in vivo* de l'effet protecteur des serpines bactériennes contre l'inflammation de l'intestin. Conception des souches bactériennes exprimant les serpines d'intêret et validation de leur effet anti-inflammatoire dans le modèle de souris avec une MICI induite chimiquement. (équipe 1/MICALIS)

(iii) Etude des relations structure fonction des complexes serpine/protéases. Production et purification des 5 serpines les plus intéressantes (partenaire 1) pour des études structurales de cristallographie aux rayons X (partenaire 3) et optimisation. Production de mutants avec des efficacités améliorées.

Moyens demandés et Personnes impliquées :

	Responsable du projet	**Personnes impliquées**	**Projet et somme obtenue**
Partenaire 1 **Coordinateur**	**E. MAGUIN** **INRA-MICALIS** **Jouy en Josas**	**Total : 13** 4 chercheurs 4 ingénieurs 1 technicien 2 thésards 1 post doc 1 M2	*Sélection et caractérisation in vitro des serpines* ***311 608 €***

Partenaire 2	**N.PONS** **INRA-MGP** **Jouy en Josas**	**Total : 4** 1 IR 3 M2	*Sélection et caractérisation in vivo* ***10 058 €***
Partenaire 3	**J. MARQUEZ** **EMBL- Grenoble**	**Total : 4** 1 chercheur 2 IEs 1 post doc	*Détermination de la structure 3-D par cristallographie aux rayons X* ***63 739 €***
Partenaire 4	**V. ROSSI** **IBS- Grenoble**	**Total : 3** 3 chercheurs **N.Thielens (DR INSERM)** **C.Gaboriaud (DR CEA)** **V.Rossi (MCU, UJF)**	*Caractérisation par des méthodes biophysiques et biochimiques de l'inhibition des protéases à sérines par les serpines sélectionnées* ***44 999 €***

Implication du partenaire 4 dans le projet :

Notre contribution au projet ANR SerpinGUTARGET a pour objectif de **caractériser par des méthodes biophysiques et biochimiques les serpines bactériennes** qui auront été identifiées, sélectionnées et purifiées par le partenaire 1. La détermination des caractéristiques cinétiques du processus d'inhibition des protéases à sérine humaines par les serpines bactériennes sera abordée sous trois aspects :

(i) **Caractérisation de l'interaction protéase à serine/serpine :** la constante d'association entre les deux protéines sera déterminée en utilisant des substrats estérasiques spécifiques des protéases étudiées. La cinétique de clivage des substrats sera caractérisée par spectrofluorimètrie. Si nécessaire, l'association des serpines avec les protéases pourra être analysée par résonance plasmonique de surface (plateforme de SPR, IBS, gérée par N. Thielens de notre laboratoire) en utilisant des protéases à sérine inactivées par mutation de la sérine du site actif en alanine. Cette modification permet de garder une dynamique d'interaction de la protéase avec la serpine car elle empêche la fixation irréversible des deux molécules entre elles (mécanisme d'inhibition de type « substrat suicide » classique des serpines).

(ii) **Efficacité de l'inhibition :** Détermination de la stoechiométrie de la réaction et quantification des complexes covalents serpine/protéases à sérine par analyse sur gel SDS-PAGE, chromatographie d'exclusion diffusion et spectrométrie de masse MALDI (plateforme de spectrométrie de l'IBS).

(iii) **Stabilité des complexes serpine/protéase, et potentialisation par l'héparine.**

La stabilité des complexes sera caractérisée par Thermal Shift Assay et l'inactivation par polymérisation (phénomène courant chez les serpines) sera caractérisée par des techniques d'analyses en conditions natives (gels d'acrylamide, chromatographie et éventuellement centrifugation analytique (plateforme de l'IBS)). La potentialisation par l'héparine sera analysée par cinétique comparative (avec ou sans héparine) par spectofluorimétrie.

Bibliographie :

(1) Ley et al., (2008) Evolution of mammmals and their gut microbes. *Science*, 320: 1647-51.

(2) Le Chatelier et al., (2013) Richness of human gut microbiome correlates with metabolic markers. *Nature* 500: 541-6.

(3) Santos Rocha et al., (2012) Anti-inflammatory properties of dairy lactobacilli. Inflamm. Bowel Dis., 14: 657-66.

(4) Motta et al., (2012) Food-grade bacteria expressing elafin protect against inflammationand restore colon homeostasis. *Sci Transl.Med*. 4: 158ra144.

La demande comprend le financement pour un étudiant de Master2 recherche qui devrait prendre effet en janvier 2016.

7.2 - Projet 2 : Rôle et mécanisme d'action du récepteur CD91/LRP-1 dans l'efferocytose

7.2.1 - Description du projet de groupe : Les collagènes de défense dans l'efferocytose.

L'élimination des cellules mortes de l'organisme, un phénomène nommé efferocytose, est cruciale pour les processus physiologiques essentiels tels que l'homéostasie tissulaire, le contrôle de l'inflammation et la tolérance immune. D'après des études menées chez l'homme et dans des modèles animaux, il devient de plus en plus évident qu'un dysfonctionnement de l'efferocytose contribue à de nombreuses maladies, dont plus particulièrement les maladies auto-immunes. L'élimination rapide et efficace des cellules apoptotiques par des phagocytes empêche la libération de composants intracellulaires potentiellement dangereux pour l'organisme. De plus, l'"engulfment" des cellules apoptotiques par les cellules présentatrices de l'antigène induit un programme anti-inflammatoire et participe à la présentation d'antigènes du soi aux cellules T, qui est nécessaire au maintien de la tolérance immune. Inversement, une élimination tardive des cellules apoptotiques peut conduire à la génération de cellules nécrotiques et à la sécrétion de molécules de danger immunostimulantes. Ces molécules peuvent déclencher une réponse auto-immune telle que la production d'auto-anticorps dirigés contre des antigènes des cellules apoptotiques, incluant l'ADN, les protéines nucléaires et des composants du cytoplasme. L'accumulation de complexes immuns qui découlent de ce processus produit une inflammation chronique qui est la principale cause de dommages tissulaires dans de nombreuses maladies auto-immunes.

Ces maladies chroniques représentent un problème de santé publique important car elles sont invalidantes et conduisent parfois à la mort du patient, principalement du fait d'un diagnostic tardif et de l'absence de traitement efficace. De plus, étant donné que les maladies auto-immunes se développent plus particulièrement chez les patients de sexe féminin en âge de procréer, elles ont un impact considérable sur la famille et la société dans son ensemble. Pour cette raison, il est essentiel d'identifier de nouvelles cibles pour le développement d'outils diagnostiques ou curatifs des maladies inflammatoires et auto-immunes.

Bien que de nombreuses molécules associées à l'efferocytose aient été décrites à ce jour (signaux « eat me » des cellules apoptotiques, molécules pontantes solubles et molécules impliquées dans la signalisation et l'endocytose par les phagocytes), plusieurs études mettent en avant le rôle pivot des opsonines solubles telles que la protéine C1q du complément. C1q a été initialement identifié comme étant la protéine de reconnaissance capable de déclencher l'activation de la voie classique du complément, mais récemment son rôle dans l'élimination immunologiquement « silencieuse » des cellules du soi a été démontré. Une déficience homozygote en C1q est associée dans 90% des cas à des maladies auto-immunes comme le SLE (Lupus Erythémateux Systémique), révélant le rôle protecteur crucial de cette protéine (Walport et al., 1998). En accord avec ces observations cliniques, on observe chez les souris C1q-/- le développement d'une maladie de type lupus et une déficience dans l'élimination des cellules apoptotiques (Botto et al., 1998). Bien que C1q puisse induire l'élimination des cellules apoptotiques par la production de l'opsonine C3b par activation du complément, il est maintenant acquis que C1q est, comme C3b, une molécule de pontage entre les cellules apoptotiques et les phagocytes. Des études récentes ont également montré que C1q module les cellules présentatrices des antigènes « professionnelles » en induisant une polarisation fonctionnelle et l'expression de cytokines. Fait intéressant, bien que la plupart des protéines du complément soient produites dans le foie, C1q est synthétisé localement par différentes cellules dont les cellules myéloïdes (monocytes, macrophages et cellules dendritiques immatures). En accord avec ces observations, un traitement efficace de l'immunodéficience de C1q par la transplantation de cellules souches hématopoïétiques a été identifié chez les souris C1q -/- (Cortes-Hernandez et al., 2004) et récemment chez l'homme (Arkwright et al., 2014). Dans le SLE, il existe également des patients qui, sans posséder de gène de C1q non fonctionnel, développent la maladie du fait de la neutralisation fonctionnelle de C1q par des autoanticorps.

Le rôle de C1q dans le SLE peut apparaître paradoxal puisqu'une activation excessive du complément par des cibles endogènes telles que les cellules apoptotiques est un évènement clé dans les pathologies autoimmunes et inflammatoires. L'élimination complément-dépendante des cellules du soi nécessite donc un contrôle très équilibré de la réponse inflammatoire.

Objectif du projet de groupe : La place importante de C1q dans l'élimination des cellules apoptotiques a émergé ces 15 dernières années. Pour remplir cette fonction, C1q agit comme une molécule pontante entre les cellules apoptotiques et les cellules phagocytaires. C1q reconnait plusieurs types de cellules du soi modifié, à différent stades de mort cellulaire (apoptose précoce et tardive, cellules nécrotiques) et se fixe spécifiquement sur des molécules « eat me » soit directement, soit par l'intermédiaire d'autres molécules de reconnaissance telles que les pentraxines, les IgM et les auto-anticorps (Jeannin et al., 2008, Paidassi et al., 2009, Hochreiter-Hufford and Ravichandran, 2013). L'équipe de Philippe Frachet dans notre laboratoire a étudié depuis une dizaine d'année le rôle de C1q dans la reconnaissance et l'élimination des cellules apoptotiques. Dans ce cadre, des ligands de C1q ont été identifiés à la surface des cellules apoptotiques tels que la phosphatidylsérine (PS), l'ADN, la calréticuline de surface (ecto CRT) et la GAPDH (Païdassi et al., 2008, 2008b, 2011, Terrasse et al., 2012). Le site de fixation de la PS et de l’ADN sur C1q fait intervenir les têtes globulaires sur un site commun, identifié par cristallographie aux rayons X (Garlatti et al., 2010). Les deux autres ligands CRT et GAPDH, en revanche, font intervenir à la fois les régions globulaires et les queues collagènes, caractéristique que nous avons également observée pour d’autres ligands tels que le récepteur CR1.

Notre nouveau projet de groupe a pour objectif de décrypter les mécanismes moléculaires et cellulaires par lesquels C1q interagit **avec les récepteurs phagocytaires** pour induire une efferocytose efficace et contrôlée. Ces informations devraient nous permettre de comprendre comment l’équilibre délicat entre inflammation et tolérance immune est maintenu dans les individus sains comparativement aux individus atteints de maladies auto-immunes.

Une demande ANR (C1qEffero) portant sur le projet de groupe a été déposée en 2015, en collaboration avec l’équipe ‘Immunité Innée et Immunothérapie’ (Yves Delneste) du Centre de Recherche en Cancérologie Nantes-Angers, spécialiste de la biologie des cellules myéloïdes humaines. Elle est focalisée sur les voies d’élimination des cellules apoptotiques impliquant C1q et des récepteurs nouvellement décrits dans la littérature tels que **SREC-I** (scavenger receptor of endothelial cells -I), **RAGE** (Receptor for Advanced Glycation End-products), **CR3** (complement receptor 3), **DC-SIGN** (Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin) et **LAIR-1** (leukocyte-associated Ig-like Receptor 1). La plupart de ces récepteurs sont multifonctionnels et leur interaction avec C1q pourrait impliquer des complexes ternaires avec des molécules interagissant avec C1q décrites antérieurement comme **LRP1/CD91** (Low-density lipoprotein receptor-Related Protein 1), **CRT** (calreticulin) et **gC1qR** (globular head of C1q Receptor) (Figure 23).

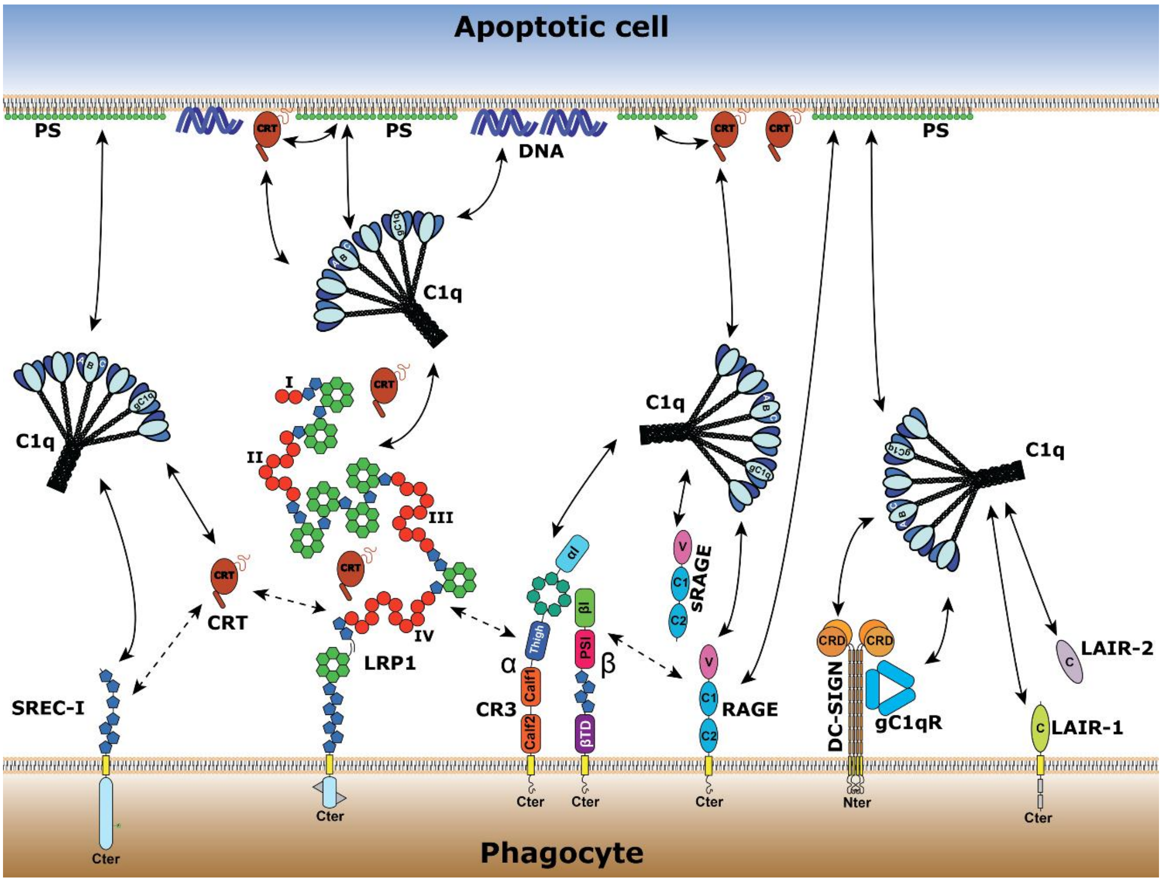

Apoptotic cell
PS
CRT
PS
DNA
CRT
CRT
PS
C1q
C1q
I
CRT
II
III
C1q
C1q
gC1q
CRT
CRT
CRT
IV
LRP1
SREC-I
gI
Thigh
Calf1
Calf2
α
CR3
βI
PSI
βTD
β
V
C1
C2
sRAGE
V
C1
C2
RAGE
CRD
CRD
DC-SIGN
gC1qR
LAIR-2
C
LAIR-1
Cter
Cter
Cter
Cter
Cter
Nter
Cter
Phagocyte

Figure 23 : Vue simplifiée du réseau d'interaction de l'efferocytose médiée par C1q (ANR C1qEffero 2015).
Les flèches pleines et en pointillés indiquent respectivement les interactions décrites dans la literature et celles qu'il faudra confirmer. LRP1 peut également être retrouvé à la surface des cellules dendritiques mais celà n'est pas montré pour plus de clarté. Les modules EGF (SREC-I, LRP1, CR3) sont représentés sous forme de pentagones bleus, les domaines β-propellers (LRP1, CR3) à 6 ou 8 hélices sont représentés par des hexa- ou heptagones en vert. Les modules LDL receptor type A (LA) repeats également appelés CR (LRP1) sont représentés par des cercles rouges et forment 4 clusters numérotés de I à IV. V, C1, C2 (RAGE) et C (LAIR) sont des domaines de type immunoglobulin (Ig)-like. Le domaine de reconnaissance des carbohydrates (CRD) de DC-SIGN est indiqué en orange. Les phosphatidylsérines exposées à la surface des cellules apoptotiques (AC) sont schématisées par des cercles verts.

Situation de mon projet dans la thématique du groupe :

Au niveau des cellules apoptotiques l'implication de la CRT en tant que signal "eat me" reconnu par C1q a été mise en évidence relativement récemment. En revanche, à la surface des phagocytes, elle a été décrite il y a plus de 20 ans comme un récepteur des régions collagènes de C1q, appelé cC1qR, proposé ensuite comme corécepteur du récepteur endocytaire LRP1/CD91 et impliqué dans la capture des cellules apoptotiques opsonisées (Malhotra et al., 1993). Il a été rapporté également que la CRT serait attachée à la surface des neutrophiles apoptotiques et des cellules cancéreuses par l'intermédiaire de LRP1 (Donnelly et al., 2006, Garg et al., 2012). L'interaction de C1q avec un complexe LRP1/CRT pourrait jouer un rôle bivalent à la fois dans la reconnaissance (côté cellules apoptotiques) et dans la clairance des cellules apoptotiques (côté cellules phagocytaires). Ces observations montrent que le rôle de LRP1 et de la CRT dans l'efferocytose C1q-dépendante est beaucoup plus complexe que ce qui avait pu être imaginé dans un premier temps et nécessite d'être clarifié.

Par ailleurs, deux études menées en partie au laboratoire ont montré une interaction directe de C1q et de la MBL avec LRP1 indépendamment de la CRT. Dans le cas de C1q, cette interaction fait intervenir à la fois les régions collagènes et les régions globulaires. L'interaction avec les régions collagènes est calcium-dépendante, inhibée par une concentration saline élevée (0.65 M), et suit un modèle d'interaction 1 :1, tandis que celle avec les régions globulaires n'est pas dépendante de la force ionique et suit un modèle d'interaction en deux étapes (Duus et al., 2010a). C'est probablement, comme nous l'avions observé pour le récepteur CR1, sur le site d'interaction avec les protéases à sérine du complexe C1 ou un site proche que se ferait l'interaction des régions collagènes de C1q avec LRP1.

Dans le cas de la MBL, cette hypothèse est également envisageable du fait de l'inhibition de l'interaction MBL/LRP1 par la protéase MASP-3 et par la mutation d'une lysine de la MBL impliquée dans l'association avec les protéases MASP (Duus et al., 2010b).

Au niveau de LRP1, des compétiteurs tels que RAP (Receptor Associated Protein), une protéine chaperone associée à LRP1 dans le réticulum endoplasmique, et la calréticuline (réduisent l'interaction avec C1q et la MBL, ce qui laisse supposer que cette dernière implique les clusters II et IV sur lesquels se fixe RAP (Neels et al., 1999).

Mon projet de recherche vise donc à décrypter au niveau moléculaire et cellulaire le rôle de LRP1 dans la phagocytose des cellules apoptotiques et l'implication des collagènes de défense (C1q, MBL) et de la CRT dans ce processus. Le travail commencé en 2014 vise à développer les outils moléculaires et cellulaires qui seront utilisés pour les différents aspects du projet présentés dans la demande ANR C1qEffero. Cette partie du travail est actuellement financée par l'université et la demande de financement est reproduite ci-après.

7.2.2 - Demande de financement AGIR (Université Joseph Fourier)

Comment le récepteur LRP1/CD91 est-il impliqué dans la tolérance immune ?

Contexte du projet : Le groupe IRPAS de l'Institut de Biologie Structurale (UMR 5075 UJF-CNRS-CEA), dans lequel j'exerce mon activité d'enseignant-chercheur en biologie a été créé il y a 3 ans. Sa thématique globale porte sur la réponse immunitaire aux pathogènes et aux cellules du soi altéré. Les trois équipes qui constituent le groupe élaborent actuellement un projet commun visant à étudier **le rôle des collagènes de défense dans le maintien de la tolérance immune**. Cette dynamique au sein du groupe est l'opportunité pour moi de développer un projet centré sur un récepteur d'endocytose multifonctionnel de la famille des Low Density Lipoprotein (LDL)-récepteurs, le LDL receptor-related protein 1 (LRP1) également appelé CD91. Ce récepteur joue un rôle majeur dans les processus de phagocytose des éléments du soi modifié (cellules apoptotiques et nécrotiques, débris cellulaires, LDL) et du « non soi » (virus, bactéries). **L'objectif de ce projet est d'étudier l'implication du récepteur LRP1 dans la phagocytose des cellules apoptotiques via les collagènes de défense par une approche intégrée allant de la molécule à la cellule.**

Etat des connaissances : Dans l'organisme, l'élimination efficace des éléments du soi modifié telles que les cellules apoptotiques est cruciale pour la tolérance immune. Les collagènes de défense C1q et MBL (Mannan-Binding Lectin) occupent une place prédominante dans cette tolérance en tant qu'effecteurs solubles de l'immunité innée et ont un rôle central dans la reconnaissance des éléments à éliminer et la phagocytose qui en découle (Galvan et al., 2012)[1]. Un déficit en C1q ou en MBL conduit à une susceptibilité accrue aux maladies autoimmunes telles que le lupus erythémateux (SLE), une conséquence de la production par l'organisme d'autoanticorps contre des antigènes pouvant être issus de cellules endommagées non éliminées suffisamment efficacement (Rahman and Isenberg, 2008)[2]. La phagocytose est induite par l'interaction des collagènes de défense avec des récepteurs différents selon le contexte physiologique. Le projet porte sur l'un de ces récepteurs, LRP1/CD91.

Le récepteur phagocytaire LRP1/CD91 : LRP1 est un membre de la famille des 'LDL récepteurs' également appelés récepteurs « scavenger » du fait de leur aptitude à phagocyter un grand nombre de ligands. Il est l'un des éléments le plus complexe de la famille. D'un poids moléculaire de 600 kDa, il comprend une extension extracellulaire multimodulaire divisée en 4 clusters (modules de type « complément ») impliqués dans la fixation de nombreux ligands, dont plus d'une trentaine induisent la phagocytose (Fig 1A).

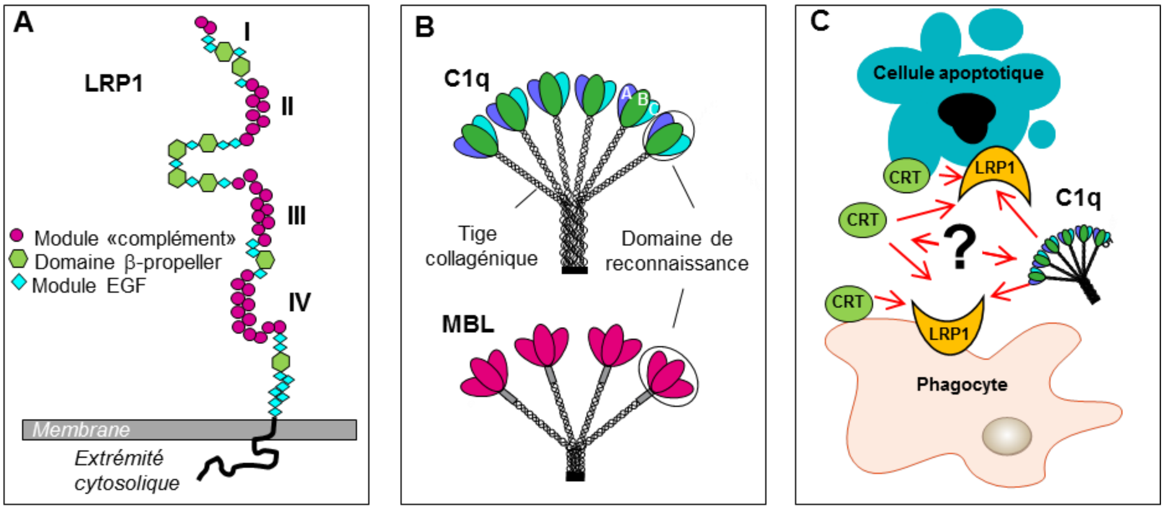


Figure 1. (A) Représentation schématique de la structure modulaire de LRP1/CD91, les 4 clusters impliqués dans la fixation des ligands sont indiqués en chiffres romains. (B) Collagènes de défense, C1q et MBL. (C) Questions posées dans le projet : Quels sont les partenaires de LRP1 impliqués dans la phagocytose des cellules apoptotiques ? Relations structure/fonction des interactions ? CRT, calréticuline.

Les collagènes de défense comme effecteurs de l'élimination des corps apoptotiques - L'efferocytose : Les collagènes de défense C1q et MBL sont des protéines oligomériques solubles formées de tiges collagéniques prolongées par des domaines globulaires agissant comme senseurs et effecteurs de l'immunité innée (Fig. 1B). Elles interviennent comme des opsonines « pontantes » fixant d'un côté les éléments nocifs pour l'organisme et de l'autre des récepteurs à la surface des cellules phagocytaires (Fig 1C). On ne connait pas précisément comment les collagènes de défense sont impliqués dans la phagocytose via le récepteur LRP1. Quelques études ont décrit que l'élimination des cellules apoptotiques par l'intermédiaire de C1q et de la MBL nécessite sur les phagocytes la formation d'un complexe ternaire comprenant LRP1 et la calréticuline (CRT), une molécule multifonctionnelle initialement connue pour son rôle de chaperone et récemment caractérisée comme signal « eat me » sur les cellules en apoptose (Gardai et al., 2005)[3]. D'autres résultats montrent que les collagènes de défense interagissent directement avec LRP1 (Duus et al., 2010a)[4],(Duus et al., 2010b)[5]. D'autre part, il a été montré récemment que des cellules apoptotiques expriment également à leur surface LRP1 en association avec la CRT (Garg et al., 2012)[6]. **L'objet de ce projet sera de décrypter au niveau moléculaire le réseau d'interactions dans lequel LRP1 est impliqué et conduit à la phagocytose des cellules apoptotiques.**

__Méthodologies :__ **(i) Dissection moléculaire de LRP1.** En nous basant sur mes compétences acquises par l'étude des molécules modulaires du complément humain associées à C1q et à la MBL (protéases C1r, C1s, MASPs, récepteur CR1/CD35 (Jacquet et al., 2013)[7], la stratégie de dissection va consister à produire par voie recombinante dans un système cellulaire eucaryote (cellules de mammifères) **des**

formes tronquées de LRP1 (comprenant un ou plusieurs clusters de LRP1) soit sous forme soluble soit sous forme membranaire. En parallèle, disposant au laboratoire de la technologie de production sous forme recombinante de la CRT, la MBL et C1q, nous criblerons les sites d'interaction de ces ligands solubles avec des fragments et mutants. **(ii) Les propriétés d'interaction** des protéines produites seront testées par résonance plasmonique de surface (SPR) et ELISA ou par cytométrie en flux et microscopie confocale (interaction avec les cellules). D'autres ligands connus de LRP1 (TSP1, RAP, complexe C1s/C1inhibiteur) ou des collagènes de défense (C1s, C1r, MASPs) disponibles au laboratoire pourront être utilisés comme compétiteurs **(iii) La production de lignées cellulaires exprimant les fragments tronqués de LRP1** devrait permettre de caractériser les éléments moléculaires qui modulent les fonctions du récepteur à la surface cellulaire au cours de l'efferocytose.

Résultats attendus : Ce projet permettra le développement d'outils moléculaires et de stratégies visant à l'identification des régions et résidus protéiques responsables de l'interaction entre LRP1 et ses partenaires solubles conduisant à l'efferocytose. Le déchiffrage des mécanismes moléculaires et cellulaires par lesquels C1q et la MBL participent à l'élimination des cellules apoptotiques est essentiel pour mieux comprendre l'équilibre délicat entre tolérance immune et inflammation/autoimmunité dans des circonstances tant physiologiques que pathologiques.

Moyens demandés : 40 k€-pack 1: culture cellulaire et biologie moléculaire : 15 k**€**, Purification des protéines recombinantes : 5 k**€**, financement d'un stage de M2 et d'un BTS en alternance : 10 k**€**, Réactifs pour tests de phagocytose 7 k**€** et pour des missions (conférences) 3 k**€**.

Personnes du groupe impliquées :

Véronique Rossi	**MCU, UJF**	100 % du temps	*Coordination du projet, production de mutants et expériences d'interactions (ELISA, SPR, FACS)*
Catherine Wicker-Planquart	**CR, CNRS**	50 %	*Production de protéines recombinantes*
Isabelle Bally	**Ingénieur, CEA**	20 %	*Production de C1q recombinant*
Philippe Frachet	**MCU, UJF**	15 %	*Expertise sur la phagocytose des cellules apoptotiques*
Nicole Thielens	**DR, INSERM**	15 %	*Expertise sur la mesure d'interactions par SPR, sur MBL et C1q recombinants*

Situation du projet dans le contexte local, national et international :

Contexte local : Ce projet permettra d'initier un nouvel axe de recherche, en lien étroit avec le projet de groupe émergeant sur C1q et tolérance immune, et de **développer des outils moléculaires en support du projet de groupe.** Les connaissances acquises sur LRP1 compléteront utilement celles obtenues sur différents récepteurs impliqués dans l'efferocytose par les autres membres du groupe. Le projet proposé bénéficiera des compétences pluridisciplinaires du groupe, pour la production des collagènes de défense recombinants, l'analyse d'interactions moléculaires par SPR (N. Thielens), les tests de phagocytose de cellules apoptotiques (P. Frachet), et éventuellement des approches structurales, ce qui garantit sa faisabilité.

Contexte national **:** Ce projet pourra bénéficier de la collaboration initiée récemment au sein du groupe avec l'équipe 'Immunité Innée et Immunothérapie' d'Yves Delneste du Centre de Recherche en Cancérologie Nantes-Angers, spécialiste de la biologie des cellules myéloïdes humaines de l'immunité. Cette collaboration permettra d'étudier les conséquences de l'interaction C1q/MBL-LRP1 en termes d'orientation du phénotype des cellules phagocytaires et de la sécrétion de cytokines pro- ou anti-inflammatoires.

Contexte international **:** LRP1 est une molécule peu étudiée sous l'angle de son interaction avec les collagènes de défense. L'aspect pluridisciplinaire à forte coloration structurale est très peu abordé au niveau international étant donné la complexité du récepteur. Notre savoir-faire est un atout majeur dans l'avancée des connaissances dans ce domaine. En ce qui concerne C1q, c'est également une molécule complexe assemblée à partir de 3 chaînes polypeptidiques codées par 3 gènes différents. Notre groupe a été le premier et est actuellement le seul à notre connaissance à le produire sous forme entière recombinante fonctionnelle(Bally et al., 2013a)[8]. Ce savoir-faire nous confère un avantage dans l'étude des multiples fonctions effectrices et de reconnaissance de cette protéine, en nous permettant notamment de découpler ces propriétés par la production de mutants.

Bibliographie :

(1) Galvan et al. (2012) C1q and phagocytosis: the perfect complement to a good meal. *J Leukoc. Biol*. 92: 489-497.

(2) Rahman et al. (2008) Systemic lupus erythematosus*. N. Engl. J. Med.* 358: 929-39.

(3) Gardai et al. (2005) Cell-surface calreticulin initiates clearance of viable or apoptotic cells through trans-activation of LRP on the phagocyte. *Cell* 123: 321-334

(4) Duss et al., (2010) Direct interaction between CD91 and C1q *FEBS J.* 277(17):3526-37

(5) Duss et al. (2010) CD91 interacts with mannan-binding lectin (MBL) through the MBL-associated serine protease binding site. *FEBS J.* 277(23):4956-64

(6) Garg et al. (2012) A novel pathway combining calreticulin exposure and ATP secretion in immunogenic cancer cell death. EMBO J. 31(5):1062-79

(7) Jacquet …and Rossi (2013) Deciphering complement receptor type 1 interactions with recognition proteins of the lectin complement pathway. *J Immunol.* 190(7):3721-31

(8) Bally et al. (2013) Expression of recombinant human complement C1q allows identification of the C1r/C1s-binding sites. PNAS 110(21):8650-5

7.1.3 - Description détaillée de la structure multimodulaire de LRP1 et de son mode d'interaction calcium-dépendant.

LRP1 est un membre de la famille des récepteurs endocytaires de type LDL récepteurs. Ce récepteur est composé de deux sous-unités associées de manière non covalente : la sous-unité α (515 kDa) correspondant à la partie extracellulaire N-terminale, et la sous-unité β (85 kDa) comprenant un domaine transmembranaire et une région cytoplasmique impliquée dans l'interaction avec des molécules de signalisation (Figure 24). LRP1 peut fixer et internaliser une variété très vaste de ligands (plus d'une cinquantaine identifiés à ce jour) impliqués dans différentes fonctions biologiques (Lillis et al., 2008). LRP1 participe par exemple au métabolisme lipidique (ApoE, lipoproteins…) et au maintien de l'homéostasie sanguine par l'élimination des complexes protéases/inhibiteurs (α2M, Thrombin/antithrombin, C1s/C1-Inh...). En plus des fonctions associées à la phagocytose des éléments du soi modifié qui nous intéresse plus particulièrement, LRP1 joue un rôle protecteur dans les neurones et la migration des cellules par désassemblage des adhésions focales (CRT, thrombospondine 1, fibronectine). Tous les ligands extracellulaires de LRP1 identifiés à ce jour sont structuralement différents mais interagissent pour la plupart avec les mêmes régions du récepteur.

Structure modulaire de LRP1 :

LRP1 contient 31 modules CR (cystein rich complement type repeat) également nommés LA (LDL receptor class A repeats), que l'on retrouve chez tous les membres de la famille des LDL récepteurs. Ces modules CR forment 4 clusters (I, II, III et IV) de taille différente, contenant respectivement 2, 8, 10 et 11 répétitions espacées par des modules EGF et des structures en hélice appelées beta-propellers (Dieckmann et al., 2010, Figure 24). Les clusters de CR ont un rôle d'interaction directe avec les ligands de LRP1, les autres modules ayant probablement un rôle dans l'espacement et l'orientation des sites d'interaction. Les modules CR contiennent environ 40 acides aminés et sont stabilisés par 3 ponts disulfures. Leur repliement consiste en un motif de type « β-hairpin » N-terminal et un site de fixation du calcium très conservé dans la partie C-terminale du module.

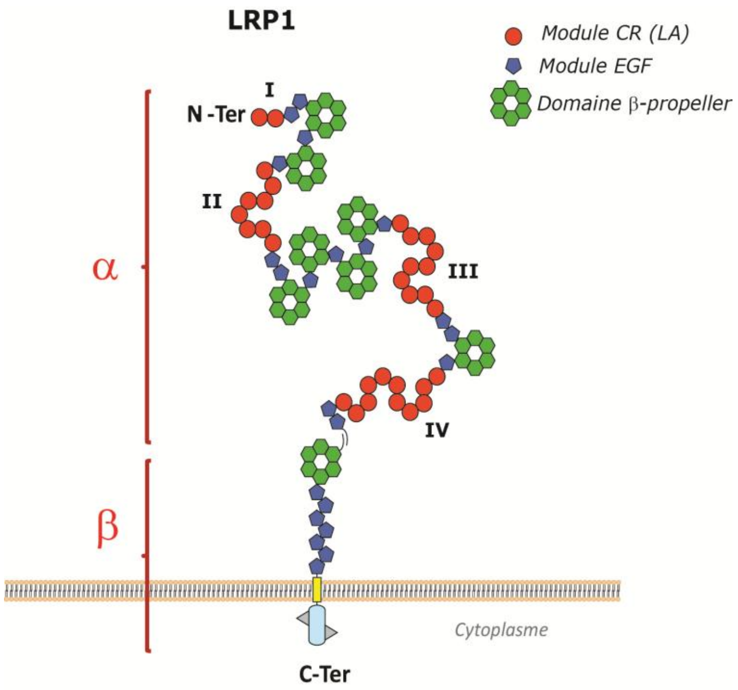


Figure 24 : Représentation schématique de la structure modulaire de LRP1/CD91

Fixation des ligands de LRP1 – le cas de RAP :

Etant donné le nombre important de ligands susceptibles de se fixer sur LRP1, la protéine, dès sa synthèse dans le réticulum endoplasmique, est protégée par la chaperone RAP (Receptor Associated Protein), une protéine escorte qui se fixe de façon pH- et calcium-dépendante aux clusters II et IV préférentiellement, ce qui est le cas de la majorité des autres ligands de LRP1 (Bu, 2001, Neels et al., 1999). Andersen et Moestrup décrivent dans une revue récente la façon dont la fonction d'interaction des récepteurs endocytaires avec leurs ligands est régulée par deux caractéristiques communes, l'implication d'ions calcium et l'interaction pH-dépendante (Andersen and Moestrup, 2014). Dans le cas de LRP1, cette dépendance au calcium peut-être expliquée par comparaison avec la structure du complexe CR3-4 du LDL récepteur en interaction avec RAP résolue par cristallographie aux rayons X (Fisher et al., 2006). Dans chaque CR qui fixe RAP, un ion calcium participe à la coordination de deux résidus aspartates engagés dans une interaction éléctrostatique avec une lysine de RAP (Figure 25). Ce motif d'interaction est également retrouvé dans deux autres complexes de récepteurs endocytaires avec leur ligand , LRP8 et le VLDL respectivement avec le rhinovirus humain Vp1 et la reeline (Yasui et al., 2010, Verdaguer et al., 2004). (Figure 25)

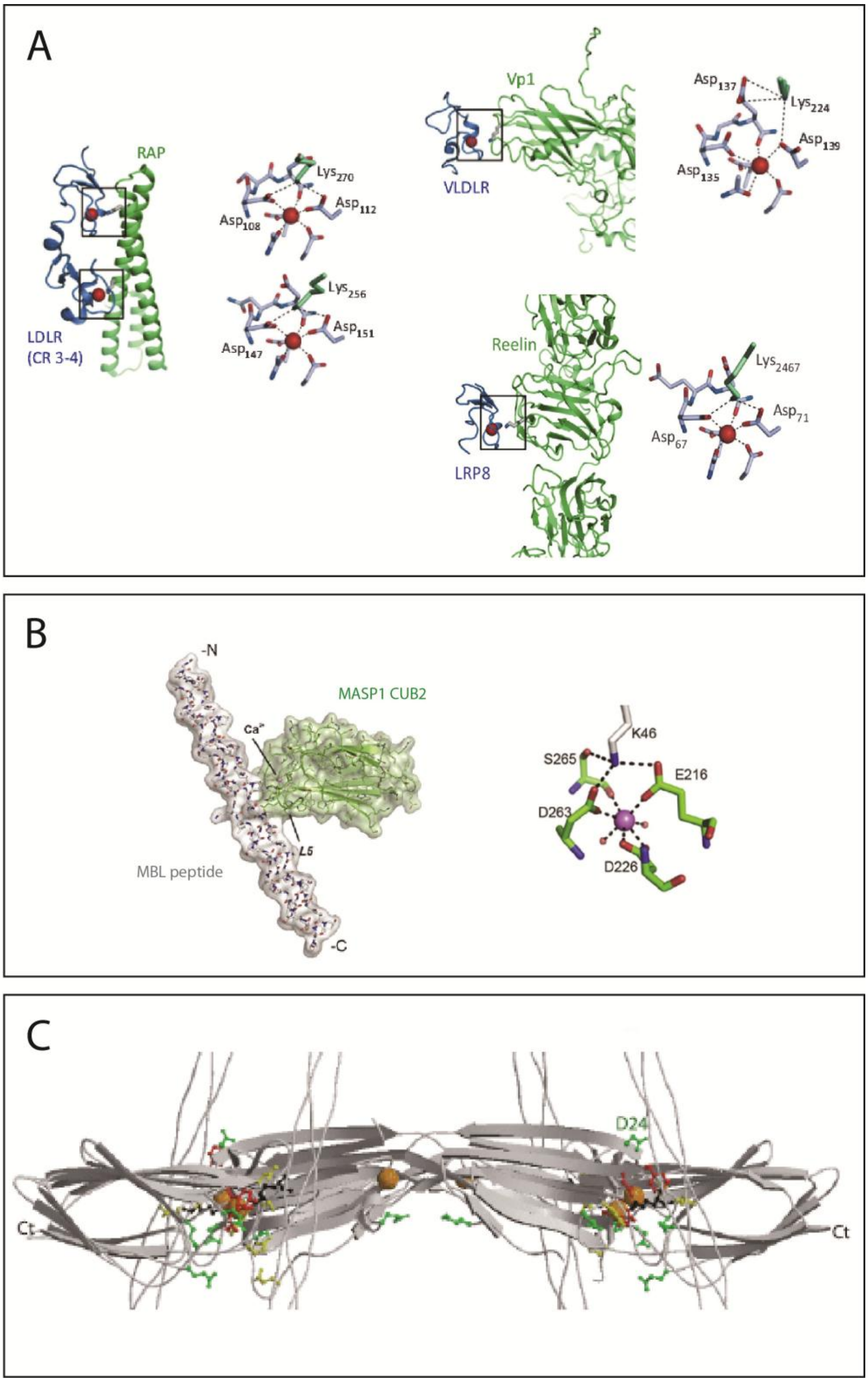

A
RAP
LDLR
(CR 3-4)
Lys270
Asp108
Asp112
Lys256
Asp147
Asp151
Vp1
VLDLR
Asp137
Lys224
Asp139
Asp135
Reelin
LRP8
Lys2467
Asp71
Asp67
B
-N
MASP1 CUB2
Ca2+
L5
MBL peptide
-C
K46
S265
E216
D263
D226
C
D24
Ct
Ct

Figure 25 : Modèle commun d'interaction calcium-dépendante entre les récepteurs à CR et leur ligand et entre les modules CUB de MASP et la MBL.

*A, D'après la revue de Andersen et Moestrup, 2014, Bases structurales de la reconnaissance des ligands calcium-dépendante par des récepteurs comportant des modules CR. **a**, Low-densityLipoprotein (LDL) receptor-receptor associated protein (RAP. **b**, very low, density lipoprotein receptor (VLDLR)- human rhinovirus capsid protein Vp. **c**, LDLR-related protein 8 (LRP8)-reelin. Les ions calcium sont représentés par des sphères rouges. **B,** structure du complexe MBL/CUB2 de MASP1 de rat. **C,** modèle de l'interaction d'un dimère CUB1-EGF de MASP1 avec la MBL. Les résidus acides sont représentés en vert, les ponts disulfures en jaune, l'ion calcium en orange et les lysines de la MBL participant à l'interaction en noir.*

Les complexes MBL/MASP et C1q/C1r2C1s2 : un autre cas de fixation dépendante du calcium :

La fixation calcium-dépendante décrite par Andersen et collaborateurs pour Les LDL-récepteurs est comparable à celle de plusieurs autres complexes calcium-dépendants impliquant des modules qui, bien que structuralement différents, fixent le calcium et définissent par les acides aminés acides qui coordonnent cet ion un patch de charges négatives qui interagit avec un résidu basique (lysine ou arginine) de leur ligand. C'est probablement le cas des interactions entre les modules CUB et leurs ligands qui présentent des homologies structurales de fixation du calcium similaires chez tous les modules de la famille (Gaboriaud et al., 2011). La résolution structurale d'un complexe entre un peptide collagénique mimant le site d'interaction des MASPs de la MBL de rat et le module CUB2 de MASP1 valide cette hypothèse. L'interaction de la MBL humaine avec la région CUB1-EGF-CUB2 de MASP3 (qui est identique dans MASP1) dont la structure a été résolue par notre équipe s'effectue probablement de la même manière et a mené à un modèle d'interaction (Teillet et al., 2008) (Figure 25) qui a été utilisé comme base de réflexion pour l'élaboration du modèle homologue de C1 décrit dans le chapitre IV.

L'ensemble de ces données indique que l'interaction calcium-dépendante des collagènes de défense avec leur ligand suit un modèle similaire à celui des LDL récepteurs. On pourrait donc supposer que l'interaction directe des collagènes de défense avec LRP1 est réalisée de façon homologue.

7.1.4 - Résultats préliminaires

Communication : C32

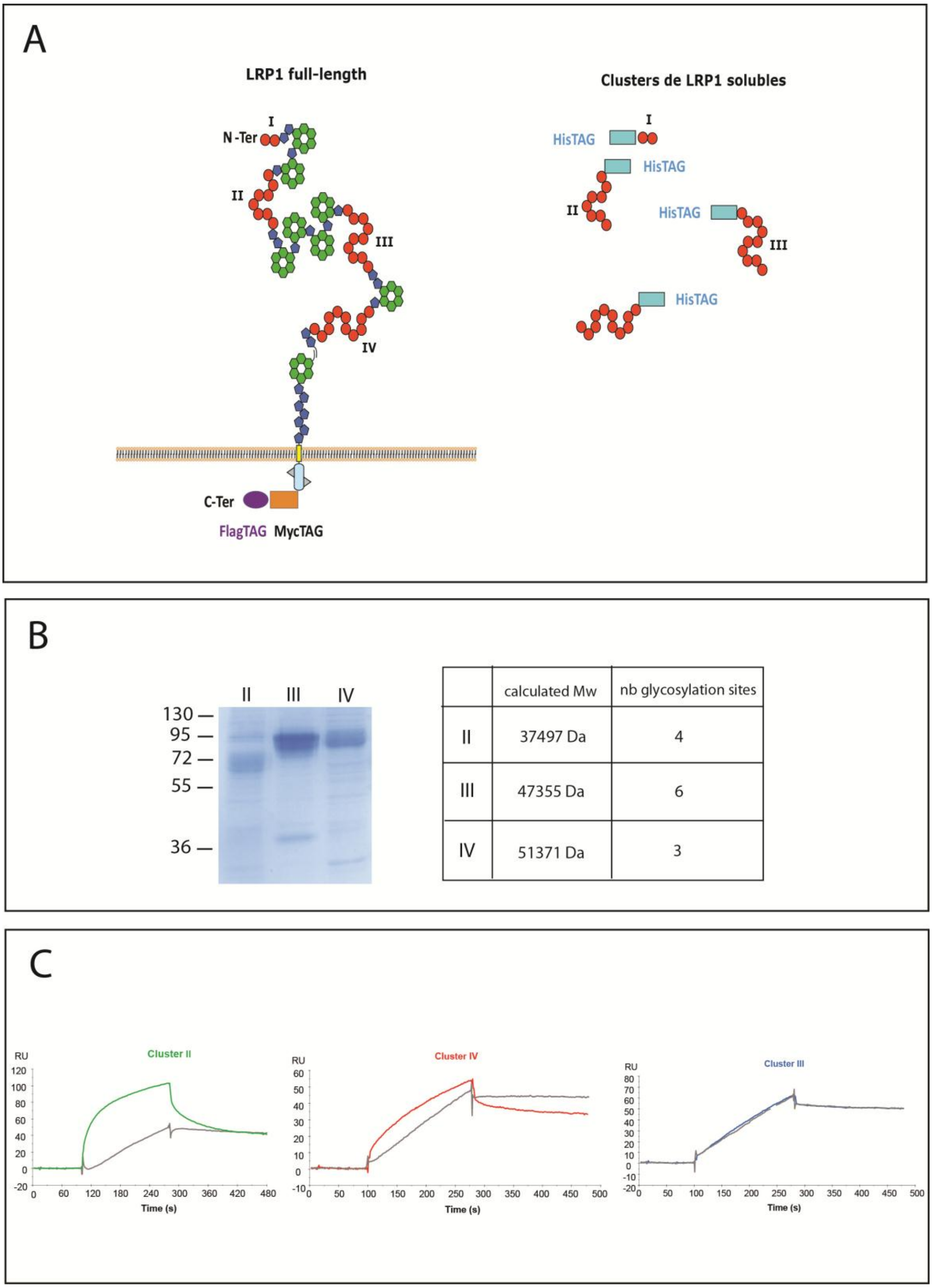
Suite à la demande de financement AGIR, nous avons obtenu 20 k**€** pour deux ans (2014-2015) et je vais décrire ici l'avancée du projet.

Production et purification des clusters de LRP1 sous forme soluble :

Les clusters I, II, III et IV ont été conçus par amplification PCR des fragments correspondant (Figure 26 A) à partir du clone codant LRP1 full-length (Origène) et insertion dans le vecteur pcDNA3.1. La production de ces clusters dans les cellules HEK 293 F a été effectuée par expression à partir de transfectants stables. L'analyse des surnageants de culture, par western blot et immunorévélation avec un anticorps anti His-TAG, a montré une expression des clusters II, III et IV tandis qu'aucun fragment correspondant au cluster I n'a pu être détecté. La purification des clusters II, III et IV à partir de 250 ml de surnageant de culture sur une colonne d'affinité His-Select a permis d'obtenir quelques centaines de microgrammes des fractions purifiées présentées sur la Figure 24B.

Les clusters migrent en condition réduite sous la forme d'une bande majoritaire qui est nettement supérieure à la taille attendue d'après la séquence peptidique, ce qui est probablement dû à la présence de plusieurs ponts disulfure et sites de N-glycosylation. La purification n'est cependant pas assez efficace et nous projetons de l'optimiser avec une colonne d'affinité Sépharose-RAP. Dans ce but nous avons produit RAP à partir d'une construction dans pET22B+ comportant une étiquette Histidine du côté N-terminal.

Bien que les clusters ne soient pas entièrement purs, ils ont été testés pour leur interaction avec C1q en SPR. La figure 26 C présente les résultats préliminaires de l'interaction entre les clusters solubles et C1q recombinant sauvage (wt) fixé sur une surface CM5. Les 3 clusters interagissent avec C1q. Les interactions des clusters II et IV sont inhibées par un excès molaire de RAP d'un facteur 2 et conduisent à une courbe dont la forme s'apparente à celle de l'interaction du cluster III avec C1q qui n'est pas inhibée par RAP. Il a été reporté que RAP se fixe préférentiellement sur les clusters II et IV et dans une moindre mesure sur le cluster III (Neels et al., 1999). Par ailleurs, Bu et Rennke, 1996, ont montré que l'expression des 3 clusters solubles de LRP1 est améliorée lorsque RAP est cotransfecté dans des cellules de glioblastome humain U87F. Nos observations pourraient être en accord avec ces données si C1q se fixait sur les trois clusters par une interaction impliquant sur certains sites les régions globulaires et sur d'autres les régions collagènes. Cette hypothèse est à confirmer avec des clusters de plus grande pureté et en utilisant C1q ainsi que ses régions globulaires et de type-collagène.


A
LRP1 full-length
I
N-Ter
II
III
IV
C-Ter
FlagTAG
MycTAG
Clusters de LRP1 solubles
HisTAG
I
II
III
B
II
III
IV
130 —
95 —
72 —
55 —
36 —
C
Cluster II
Cluster IV
Cluster III
RU
Time (s)
120
100
80
60
40
20
0
-20
0
60
120
180
240
300
360
420
480
60
50
40
30
20
10
0
-10
50
100
150
200
250
300
350
400
450
500
80
70
-20

	calculated Mw	nb glycosylation sites
II	37497 Da	4
III	47355 Da	6
IV	51371 Da	3

Figure 26 : Production des clusters solubles de LRP1 et caractérisation de l'interaction avec C1q

A, *représentation schématique de LRP1 (construction Origène, comportant une étiquette Flaget Mycen C-terminal) et des clusters solubles produits dans les cellules HEK 293F.* ***B,*** *Analyse par SDS-PAGE (12,5%) des clusters II, III et IV. (9 µg en conditions réductrices, coloration au bleu de coomassie ;* ***C,*** *Analyse de l'interaction des clusters avec C1q recombinant par SPR. Les différents clusters (600 nM) sont injectés sur une sensor shipCM5 sur laquelle a été fixé C1q recombinant (18428 RU) dans du tampon TEA 50 mM, NaCl 145 mM, $CaCl_2$ 2mM, P20 0.005%. Les courbes grises correspondent à l'injection d'un mélange contenant les clusters et un excès molaire de RAP d'un facteur 2 (1.2 µM).*

Développement d'outils cellulaires :

Pour associer les informations obtenues *in vitro* à une fonction de LRP1, nous développons également des outils cellulaires. La plupart des cellules produisent LRP1, ce qui constitue un obstacle pour l'expression de LRP1 modifié par voie recombinante. La plupart des études qui ont été menées sur les relations structure-fonction de LRP1 au niveau cellulaire en utilisant des fragments de LRP1 ont nécessité d'inhiber la production de LRP1 endogène ou de transfecter une souche LRP1$^-$. Dans un premier temps, pour pouvoir produire les clusters recombinants sur la membrane cytoplasmique, nous avons utilisé une souche CHO LRP1$^-$ (fournie par le Dr. Guojun BU). Ces clusters membranaires sont l'équivalent des clusters solubles produits par ailleurs, mais associés à la membrane par la chaîne beta et sont appelés minirécepteurs. Des essais de transfection stables dans les cellules CHO LRP1$^-$ ont été expérimentés sur LRP1 full-length et une construction du mini récepteur IV construite d'après (Li et al., 2002) par le Dr. Stéphane Dedieu (Figure 27 A).

L'obtention de transfectants stables s'est avérée plus compliquée pour LRP1 full-length pour lequel le taux de transfection ne dépasse pas 60% comparativement à celles transfectés par le minirécepteur IV pour lesquelles il dépasse 80%. L'efficacité de transfection est quantifiée par cytométrie en flux et visualisée par immunofluorescence comme (Figure 27C). Sur les images d'immunofluorescence on peut observer que LRP1 full-length et le minirécepteur IV sont produits et adressés efficacement en surface de la membrane cytoplasmique. Le même type d'image est obtenu sur les macrophages (Mono Mac6) dans lesquels l'expression de LRP1 est induite par les LPS (Duus et al., 2010a).

Un essai préliminaire de fixation de C1q sur les cellules transfectées a été réalisé et analysé par cytométrie en flux (Figure 27D). Les courbes du panneau de gauche pourraient indiquer une tendance vers la fixation de C1q sur LRP1, le minirécepteur IV fixant moins de C1q que le LRP-full length ce qui pourrait être interprété par le fait que LRP1 possède plus de sites de fixations de C1q que le minirécepteur seul. Cependant, si cette manipulation est reproductible, la quantité de C1q fixé reste très faible comparativement à la fixation basale de C1q sur les cellules CHO n'exprimant pas LRP1 (panneau de gauche) et sera peut être difficile à confirmer de cette manière. Les minrécepteurs I, II et III (Figure 27B) sont en cours de clonage à partir de la construction du minirécepteur IV dans laquelle la portion contenant le module EGF suivi des 10 modules CR a été délétée. L'insertion par mutagenèse dirigée des parties codantes des récepteurs I, II, III devrait nous permettre d'obtenir des minirécepteurs comparables au minirécepteur IV.

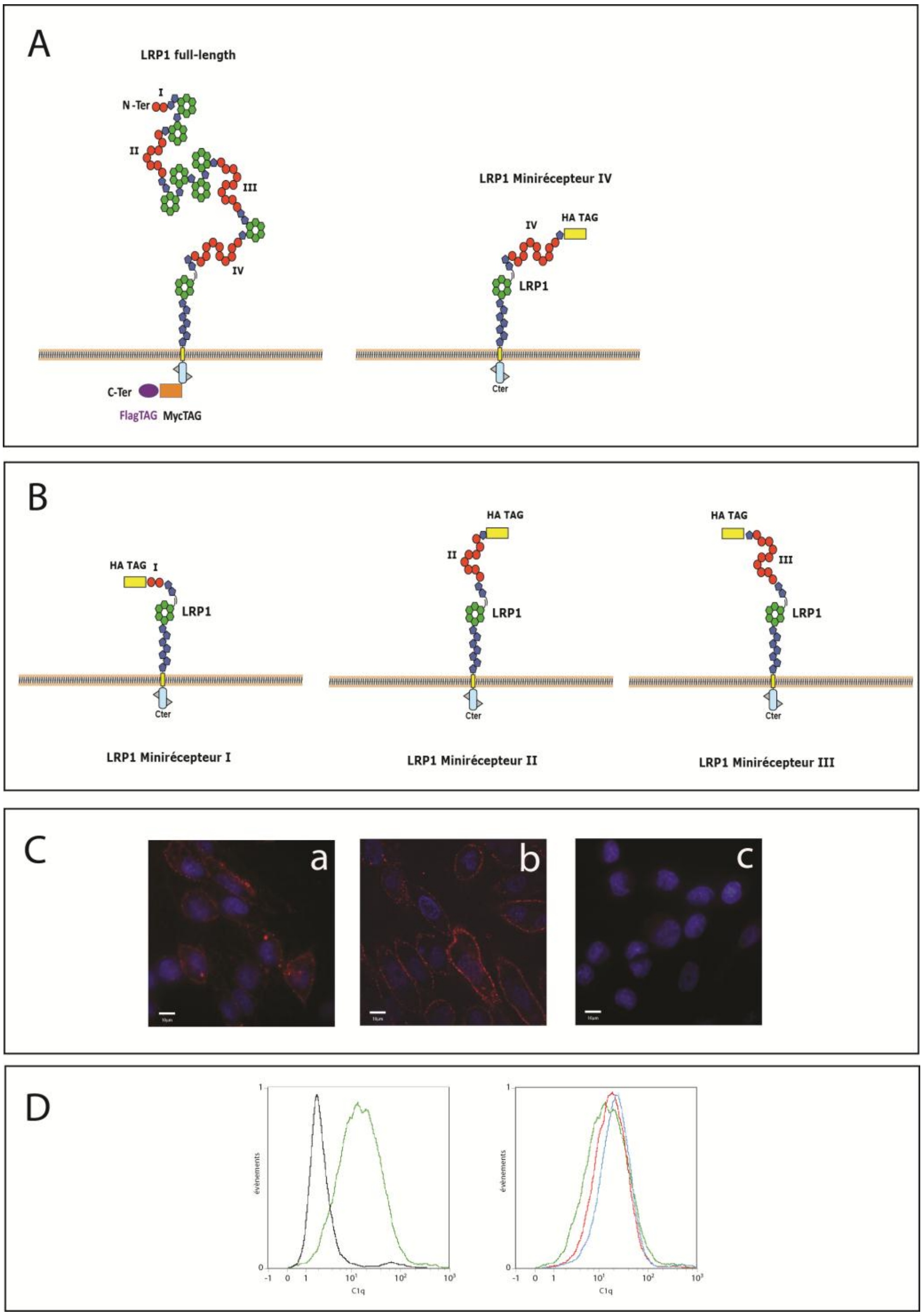

A
LRP1 full-length
N-Ter
I
II
III
IV
C-Ter
FlagTAG
MycTAG
LRP1 Minirécepteur IV
HA TAG
IV
LRP1
Cter
B
HA TAG
I
LRP1
Cter
LRP1 Minirécepteur I
HA TAG
II
LRP1
Cter
LRP1 Minirécepteur II
HA TAG
III
LRP1
Cter
LRP1 Minirécepteur III
C
a
b
c
D
évènements
C1q
1
0
-1
0
1
10^1
10^2
10^3

Figure 27 : Minirécepteurs de LRP1 - Outils cellulaires
***A**, LRP1 full-length et minirécepteur IV utilisés pour les premiers essais d'expression dans les cellules CHO LRP1null.*
***B**, Minirécepteurs en cours de clonage. **C**, analyse de l'expression de LRP1 full-length et minirécepteur IV à la surface des cellules par immunofluorescence. Les cellules CHO-LRP1$^-$ (C-c), CHO-LRP1 full-length (C-a) et CHO-LRP1 minirécepteur IV (C-b) ont été fixées en PFA 4%. LRP1 full-length est immunomarqué en rouge avec un mAb anti-LRP1 (clone A2MR-alpha2, BD pharmingen) et CHO-LRP1 minirécepteur IV est immunomarqué en rouge avec un mAb anti HA (clone HA.11, Covance). Les cellules témoins CHO-LRP1$^-$ sont immunomarquées en rouge avec les 2 anticorps (anti-LRP1 et anti-HA). Images de la plateforme d'imagerie cellulaire M4D, IBS **D**, analyse par cytométrie en flux de l'interaction de C1q avec les cellules CHO exprimant LRP1 full-length ou le minirécepteur IV. Panneau de gauche, fixation de base de C1q sur les cellules CHO LRP1$^-$ (en vert) comparativement à un témoin négatif (incubation des cellules CHO LRP1 null uniquement avec l'anticorps anti-C1q et l'anticorps secondaire couplé au FITC) (en noir). Panneau de droite, fixation comparée de C1q sur les différentes lignées : fixation de base sur les cellules CHO LRP1$^-$ (vert), sur CHO LRP1 full-length (bleu) et les CHO minirécepteur IV (rouge). Acquisition au cytomètre de la plateforme M4D, IBS.*

7.1.4 - Perspectives

On peut considérer qu'à l'issue de notre contrat de deux ans, fin 2015, la plupart des outils nécessaires à la première phase du projet auront été rassemblés. Les objectifs à atteindre restent cependant multiples.

A court terme, il faudra purifier efficacement les clusters solubles de LRP1. Pour cet objectif nous projetons d'utiliser une colonne d'affinité RAP qui sera sans doute utile pour les clusters II et IV mais qui pourrait s'avérer inutile pour le cluster III qui, d'après la littérature, aurait moins d'affinité pour RAP. La quantité de cluster III produite par les cellules HEK293 F est plus importante que pour les clusters II et IV. Une optimisation par cotransfection avec un vecteur exprimant RAP pourrait améliorer l'expression et dans ce but, nous avons cloné RAP à partir du vecteur pET22B dont nous disposions dans pcDNA3.1 (Néo). Ceci a nécessité l'insertion préalable du peptide signal de RAP par mutagenèse dirigée. Les clusters solubles seront ensuite testés pour leur interaction avec RAP, la CRT, les collagènes de défense (C1q, MBL et ficolines) et pour la formation possible de complexes ternaires CRT/LRP1/C1q ou autres. Etant donné que nous disposons maintenant de C1q recombinant dont l'expression a été réalisée dans le laboratoire (Bally et al., 2013b), il sera possible d'utiliser différents mutants pour cibler spécifiquement les sites d'interaction de C1q avec LRP1. On peut s'attendre à ce qu'une lysine des régions collagènes de C1q, proche ou impliquée directement dans la fixation du tétramère C1r2s2 ne puisse plus interagir avec LRP1 comme rapporté par (Duus et al., 2010a) par des expériences de compétition. C'est probablement le cas si le mode d'interaction calcium – dépendant est homologue à celui décrit par Gaboriaud et al., 2011 et Andersen and Moestrup, 2014. Le mutant des lysines qui sont responsables de la fixation de C1q au tétramère C1r2s2 est actuellement disponible au laboratoire. Le même type d'étude pourra être effectué avec la MBL et la ficoline-2 qui sont également produites au laboratoire. Au niveau cellulaire, si nous ne parvenons pas à reproduire l'expérience de cytométrie en flux (Figure 27D) il faudrait reconsidérer les expériences en fonction des résultats que nous aurons obtenus *in vitro*. Si l'interaction de C1q avec LRP1 au niveau des cellules phagocytaires a pu être montrée (Duus et al., 2010a), rien à l'heure actuelle ne confirme une interaction directe sur des cellules non phagocytaires. L'interaction de C1q avec LRP1 au niveau des cellules apoptotiques pourra être étudiée en induisant l'apoptose sur les cellules CHO transfectées. Si c'est le cas, on pourrait imaginer qu'une autre protéine, induite lors de l'apoptose, permet cette interaction.

A plus long terme, nous nous consacrerons à l'identification des régions des clusters impliquées dans l'interaction avec C1q et les autres collagènes de défense du complément, ainsi qu'avec la CRT. Les sites d’interaction pourront être identifiés ,après identification du nombre minimal de modules CR de LRP1 impliqués dans les interactions, par l’étude structurale de complexes de ces fragments avec leurs ligands entiers ou tronqués. Compte-tenu des difficultés liées à l’obtention de la structure cristallographique aux rayons X de complexes protéiques, des études complémentaires utilisant d’autres techniques telles que le SAXS ou la RMN pourront être utilisées. Si les fragments peuvent être produits en quantité suffisante, la comparaison par RMN de molécules isolées ou en complexe pourrait également nous permettre de localiser les acides aminés impliqués dans l’interaction de LRP1 avec ses ligands.

Au niveau cellulaire, l’implication de LRP1 et des collagènes de défense dans l’efferocytose devra être démontrée. Deux chercheurs de notre groupe, Jean-Baptiste Reiser et Jean-Philippe Kleman s’intéressent depuis la mise en place de la thématique de groupe aux voies de signalisation activées lors de l’efferocytose via LRP1. Des résultats sur *C. elegans* ont montré que deux voies régulent le processus d’efferocytose. L’une de ces voies implique un récepteur transmembranaire CED-1 qui reconnait un ligand non identifié sur les cellules apoptotiques et recrute la protéine CED-6 qui se fixe sur un motif NPxY sur la partie cytoplasmique de CED-1 (Kinchen et al., 2005). Cette interaction active CED-10, une GTPAse de type Rac qui déclenche la réorganisation de l’actine nécessaire à la phagocytose. LRP-1 serait chez les mammifères un homologue de CED-1. Cette hypothèse s’appuie sur plusieurs observations : (i) LRP1 est un récepteur transmembranaire contenant un motif NPxY intracellulaire qui interagit avec une forme phosphorylée d’un homologue de CED-6 chez les mammifères, la protéine GULP (Ranganathan et al., 2004, Su et al., 2002). (ii) L’utilisation de formes chimères du récepteur LRP1 a par ailleurs montré que son extension cytoplasmique contient les éléments nécessaires à la phagocytose et renforce cette hypothèse(Patel et al., 2003).

Les modèles cellulaires développés seront utiles à cet aspect du projet.

A plus long terme, des expériences pourront être réalisées en collaboration avec l’équipe d’Yves Delneste pour évaluer le rôle de LRP1 dans l’efferocytose dans un contexte plus « physiologique » (phagocytes dérivés de cellules primaires humaines par exemple).

Yves Delneste est partenaire de notre projet ANR C1qEffero 2015, il dirige l’équipe d’immunité et immunothérapie du centre de recherche sur le cancer du centre Nantes-Angers (CRCNA – unité 892 - Unité CNRS 6299) en connection avec l’hôpital universitaire d’Angers.

Références bibliographiques

Andersen, C.B.F., and Moestrup, S.K. (2014). How calcium makes endocytic receptors attractive. Trends Biochem. Sci. *39*, 82–90.

Arkwright, P.D., Riley, P., Hughes, S.M., Alachkar, H., and Wynn, R.F. (2014). Successful cure of C1q deficiency in human subjects treated with hematopoietic stem cell transplantation. J. Allergy Clin. Immunol. *133*, 265–267.

Arlaud, G.J., Colomb, M.G., and Gagnon, J. (1987). A functional model of the human C1 complex Emergence of a functional model. Immunol. Today *8*, 106–111.

Bally, I., Rossi, V., Thielens, N.M., Gaboriaud, C., and Arlaud, G.J. (2005). Functional Role of the Linker between the Complement Control Protein Modules of Complement Protease C1s. J. Immunol. *175*, 4536–4542.

Bally, I., Rossi, V., Lunardi, T., Thielens, N.M., Gaboriaud, C., and Arlaud, G.J. (2009). Identification of the C1q-binding Sites of Human C1r and C1s: A REFINED THREE-DIMENSIONAL MODEL OF THE C1 COMPLEX OF COMPLEMENT. J. Biol. Chem. *284*, 19340–19348.

Bally, I., Ancelet, S., Moriscot, C., Gonnet, F., Mantovani, A., Daniel, R., Schoehn, G., Arlaud, G.J., and Thielens, N.M. (2013a). Expression of recombinant human complement C1q allows identification of the C1r/C1s-binding sites. Proc. Natl. Acad. Sci. *110*, 8650–8655.

Bally, I., Ancelet, S., Moriscot, C., Gonnet, F., Mantovani, A., Daniel, R., Schoehn, G., Arlaud, G.J., and Thielens, N.M. (2013b). Expression of recombinant human complement C1q allows identification of the C1r/C1s-binding sites. Proc. Natl. Acad. Sci. *110*, 8650–8655.

Beer, K.B., and Miller, V.L. (1992). Amino acid substitutions in naturally occurring variants of ail result in altered invasion activity. J. Bacteriol. *174*, 1360–1369.

Beinrohr, L., Harmat, V., Dobo, J., Lorincz, Z., Gal, P., and Zavodszky, P. (2007). C1 Inhibitor Serpin Domain Structure Reveals the Likely Mechanism of Heparin Potentiation and Conformational Disease. J. Biol. Chem. *282*, 21100–21109.

Bohlson, S.S., Fraser, D.A., and Tenner, A.J. (2007). Complement proteins C1q and MBL are pattern recognition molecules that signal immediate and long-term protective immune functions. Mol. Immunol. *44*, 33–43.

Bork, P., Downing, A.K., Kieffer, B., and Campbell, I.D. (1996). Structure and distribution of modules in extracellular proteins. Q. Rev. Biophys. *29*, 119–167.

Botto, M., Dell'Agnola, C., Bygrave, A.E., Thompson, E.M., Cook, H.T., Petry, F., Loos, M., Pandolfi, P.P., and Walport, M.J. (1998). Homozygous C1q deficiency causes glomerulonephritis associated with multiple apoptotic bodies. Nat. Genet. *19*, 56–59.

Bu, G. (2001). The roles of receptor-associated protein (RAP) as a molecular chaperone for members of the LDL receptor family. Int. Rev. Cytol. *209*, 79–116.

Bu, G., and Rennke, S. (1996). Receptor-associated protein is a folding chaperone for low density lipoprotein receptor-related protein. J. Biol. Chem. *271*, 22218–22224.

Le Chatelier, E., Nielsen, T., Qin, J., Prifti, E., Hildebrand, F., Falony, G., Almeida, M., Arumugam, M., Batto, J.-M., Kennedy, S., et al. (2013). Richness of human gut microbiome correlates with metabolic markers. Nature *500*, 541–546.

Colomb, M.G., Arlaud, G.J., and Villiers, C.L. (1984). Structure and activation of C1: current concepts. Complement Basel Switz. *1*, 69–80.

Cortes-Hernandez, J., Fossati-Jimack, L., Petry, F., Loos, M., Izui, S., Walport, M.J., Cook, H.T., and Botto, M. (2004). Restoration of C1q levels by bone marrow transplantation attenuates autoimmune disease associated with C1q deficiency in mice. Eur. J. Immunol. *34*, 3713–3722.

Dahl, M.R., Thiel, S., Matsushita, M., Fujita, T., Willis, A.C., Christensen, T., Vorup-Jensen, T., and Jensenius, J.C. (2001). MASP-3 and its association with distinct complexes of the mannan-binding lectin complement activation pathway. Immunity *15*, 127–135.

Degn, S.E., Jensen, L., Hansen, A.G., Duman, D., Tekin, M., Jensenius, J.C., and Thiel, S. (2012). Mannan-Binding Lectin-Associated Serine Protease (MASP)-1 Is Crucial for Lectin Pathway Activation in Human Serum, whereas neither MASP-1 nor MASP-3 Is Required for Alternative Pathway Function. J. Immunol. *189*, 3957–3969.

Degn, S.E., Kjaer, T.R., Kidmose, R.T., Jensen, L., Hansen, A.G., Tekin, M., Jensenius, J.C., Andersen, G.R., and Thiel, S. (2014). Complement activation by ligand-driven juxtaposition of discrete pattern recognition complexes. Proc. Natl. Acad. Sci. *111*, 13445–13450.

Dementiev, A., Simonovic, M., Volz, K., and Gettins, P.G.W. (2003). Canonical Inhibitor-like Interactions Explain Reactivity of 1-Proteinase Inhibitor Pittsburgh and Antithrombin with Proteinases. J. Biol. Chem. *278*, 37881–37887.

Dieckmann, M., Dietrich, M.F., and Herz, J. (2010). Lipoprotein receptors – an evolutionarily ancient multifunctional receptor family. Biol. Chem. *391*.

Donnelly, S., Roake, W., Brown, S., Young, P., Naik, H., Wordsworth, P., Isenberg, D.A., Reid, K.B.M., and Eggleton, P. (2006). Impaired recognition of apoptotic neutrophils by the C1q/calreticulin and CD91 pathway in systemic lupus erythematosus. Arthritis Rheum. *54*, 1543–1556.

Duncan, R.C., Mohlin, F., Taleski, D., Coetzer, T.H., Huntington, J.A., Payne, R.J., Blom, A.M., Pike, R.N., and Wijeyewickrema, L.C. (2012). Identification of a Catalytic Exosite for Complement Component C4 on the Serine Protease Domain of C1s. J. Immunol. *189*, 2365–2373.

Duus, K., Hansen, E.W., Tacnet, P., Frachet, P., Arlaud, G.J., Thielens, N.M., and Houen, G. (2010a). Direct interaction between CD91 and C1q: Direct interaction between CD91 and C1q. FEBS J. *277*, 3526–3537.

Duus, K., Thielens, N.M., Lacroix, M., Tacnet, P., Frachet, P., Holmskov, U., and Houen, G. (2010b). CD91 interacts with mannan-binding lectin (MBL) through the MBL-associated serine protease-binding site: CD91 interacts with mannan-binding lectin. FEBS J. *277*, 4956–4964.

Elliott, P.R., Abrahams, J.-P., and Lomas, D.A. (1998). Wild-type α 1-antitrypsin is in the canonical inhibitory conformation. J. Mol. Biol. *275*, 419–425.

Engh, R., Löbermann, H., Schneider, M., Wiegand, G., Huber, R., and Laurell, C.-B. (1989). The S variant of human α1-antitrypsin, structure and implications for function and metabolism. Protein Eng. *2*, 407–415.

Fisher, C., Beglova, N., and Blacklow, S.C. (2006). Structure of an LDLR-RAP Complex Reveals a General Mode for Ligand Recognition by Lipoprotein Receptors. Mol. Cell *22*, 277–283.

Forneris, F., Wu, J., and Gros, P. (2012). The modular serine proteases of the complement cascade. Curr. Opin. Struct. Biol. *22*, 333–341.

Gaboriaud, C., Rossi, V., Fontecilla-Camps, J.C., and Arlaud, G.J. (1998). Evolutionary conserved rigid module-domain interactions can be detected at the sequence level: the examples of complement and blood coagulation proteases. J. Mol. Biol. *282*, 459–470.

Gaboriaud, C., Rossi, V., Bally, I., Arlaud, G.J., and Fontecilla-Camps, J.C. (2000). Crystal structure of the catalytic domain of human complement C1s: a serine protease with a handle. EMBO J. *19*, 1755–1765.

Gaboriaud, C., Gregory‐Pauron, L., Teillet, F., Thielens, N.M., Bally, I., and Arlaud, G.J. (2011). Structure and properties of the Ca $^{2+}$ -binding CUB domain, a widespread ligand-recognition unit involved in major biological functions. Biochem. J. *439*, 185–193.

Gaboriaud, C., Gupta, R.K., Martin, L., Lacroix, M., Serre, L., Teillet, F., Arlaud, G.J., Rossi, V., and Thielens, N.M. (2013). The Serine Protease Domain of MASP-3: Enzymatic Properties and Crystal Structure in Complex with Ecotin. PLoS ONE *8*, e67962.

Galvan, M.D., Greenlee-Wacker, M.C., and Bohlson, S.S. (2012). C1q and phagocytosis: the perfect complement to a good meal. J. Leukoc. Biol. *92*, 489–497.

Gans, P., Rossi, V., Gaboriaud, C., Bally, I., Hernandez, J.-F., Blackledge, M.J., and Arlaud, G.J. (1998). NMR structures of the C-terminal end of human complement serine protease C1s. Cell. Mol. Life Sci. CMLS *54*, 171–178.

Gardai, S.J., McPhillips, K.A., Frasch, S.C., Janssen, W.J., Starefeldt, A., Murphy-Ullrich, J.E., Bratton, D.L., Oldenborg, P.-A., Michalak, M., and Henson, P.M. (2005). Cell-Surface Calreticulin Initiates Clearance of Viable or Apoptotic Cells through trans-Activation of LRP on the Phagocyte. Cell *123*, 321–334.

Garg, A.D., Krysko, D.V., Verfaillie, T., Kaczmarek, A., Ferreira, G.B., Marysael, T., Rubio, N., Firczuk, M., Mathieu, C., Roebroek, A.J., et al. (2012). A novel pathway combining calreticulin exposure and ATP secretion in immunogenic cancer cell death. EMBO J. *31*, 1062–1079.

Garlatti, V., Chouquet, A., Lunardi, T., Vives, R., Paidassi, H., Lortat-Jacob, H., Thielens, N.M., Arlaud, G.J., and Gaboriaud, C. (2010). Cutting Edge: C1q Binds Deoxyribose and Heparan Sulfate through Neighboring Sites of Its Recognition Domain. J. Immunol. *185*, 808–812.

Gettins, P.G. (2002). Serpin structure, mechanism, and function. Chem. Rev. *102*, 4751–4804.

Ghiran, I., Barbashov, S.F., Klickstein, L.B., Tas, S.W., Jensenius, J.C., and Nicholson-Weller, A. (2000). Complement receptor 1/CD35 is a receptor for mannan-binding lectin. J. Exp. Med. *192*, 1797–1808.

Gingras, A.R., Girija, U.V., Keeble, A.H., Panchal, R., Mitchell, D.A., Moody, P.C.E., and Wallis, R. (2011). Structural Basis of Mannan-Binding Lectin Recognition by Its Associated Serine Protease MASP-1: Implications for Complement Activation. Structure *19*, 1635–1643.

Gregory, L.A., Thielens, N.M., Arlaud, G.J., Fontecilla-Camps, J.C., and Gaboriaud, C. (2003). X-ray Structure of the Ca2+-binding Interaction Domain of C1s: INSIGHTS INTO THE ASSEMBLY OF THE C1 COMPLEX OF COMPLEMENT. J. Biol. Chem. *278*, 32157–32164.

Hedstrom, L. (2002). Serine Protease Mechanism and Specificity. Chem. Rev. *102*, 4501–4524.

Heja, D., Kocsis, A., Dobo, J., Szilagyi, K., Szasz, R., Zavodszky, P., Pal, G., and Gal, P. (2012). Revised mechanism of complement lectin-pathway activation revealing the role of serine protease MASP-1 as the exclusive activator of MASP-2. Proc. Natl. Acad. Sci. *109*, 10498–10503.

Hochreiter-Hufford, A., and Ravichandran, K.S. (2013). Clearing the Dead: Apoptotic Cell Sensing, Recognition, Engulfment, and Digestion. Cold Spring Harb. Perspect. Biol. *5*, a008748–a008748.

Holmskov, U., Malhotra, R., Sim, R.B., and Jensenius, J.C. (1994). Collectins: collagenous C-type lectins of the innate immune defense system. Immunol. Today *15*, 67–74.

Hoppe, H.J., and Reid, K.B. (1994). Collectins--soluble proteins containing collagenous regions and lectin domains--and their roles in innate immunity. Protein Sci. Publ. Protein Soc. *3*, 1143–1158.

Irving, J.A., Pike, R.N., Lesk, A.M., and Whisstock, J.C. (2000). Phylogeny of the serpin superfamily: implications of patterns of amino acid conservation for structure and function. Genome Res. *10*, 1845–1864.

Jacquet, M., Lacroix, M., Ancelet, S., Gout, E., Gaboriaud, C., Thielens, N.M., and Rossi, V. (2013). Deciphering Complement Receptor Type 1 Interactions with Recognition Proteins of the Lectin Complement Pathway. J. Immunol. *190*, 3721–3731.

Jeannin, P., Jaillon, S., and Delneste, Y. (2008). Pattern recognition receptors in the immune response against dying cells. Curr. Opin. Immunol. *20*, 530–537.

Ji, X., Azumi, K., Sasaki, M., and Nonaka, M. (1997). Ancient origin of the complement lectin pathway revealed by molecular cloning of mannan binding protein-associated serine protease from a urochordate, the Japanese ascidian, Halocynthia roretzi. Proc. Natl. Acad. Sci. *94*, 6340–6345.

Ji, Y.-H., Fujita, T., Hatsuse, H., Takahashi, A., Matsushita, M., and Kawakami, M. (1993). Activation of the C4 and C2 components of complement by a proteinase in serum bactericidal factor, Ra reactive factor. J. Immunol. *150*, 571–578.

Kerr, F.K., Thomas, A.R., Wijeyewickrema, L.C., Whisstock, J.C., Boyd, S.E., Kaiserman, D., Matthews, A.Y., Bird, P.I., Thielens, N.M., Rossi, V., et al. (2008). Elucidation of the substrate specificity of the MASP-2 protease of the lectin complement pathway and identification of the enzyme as a major physiological target of the serpin, C1-inhibitor. Mol. Immunol. *45*, 670–677.

Khera, R., and Das, N. (2009). Complement Receptor 1: disease associations and therapeutic implications. Mol. Immunol. *46*, 761–772.

Kidmose, R.T., Laursen, N.S., Dobo, J., Kjaer, T.R., Sirotkina, S., Yatime, L., Sottrup-Jensen, L., Thiel, S., Gal, P., and Andersen, G.R. (2012). Structural basis for activation of the complement system by component C4 cleavage. Proc. Natl. Acad. Sci. *109*, 15425–15430.

Kinchen, J.M., Cabello, J., Klingele, D., Wong, K., Feichtinger, R., Schnabel, H., Schnabel, R., and Hengartner, M.O. (2005). Two pathways converge at CED-10 to mediate actin rearrangement and corpse removal in C. elegans. Nature *434*, 93–99.

Klickstein, L.B., Bartow, T.J., Miletic, V., Rabson, L.D., Smith, J.A., and Fearon, D.T. (1988). Identification of distinct C3b and C4b recognition sites in the human C3b/C4b receptor (CR1, CD35) by deletion mutagenesis. J. Exp. Med. *168*, 1699–1717.

Klickstein, L.B., Barbashov, S.F., Liu, T., Jack, R.M., and Nicholson-Weller, A. (1997). Complement receptor type 1 (CR1, CD35) is a receptor for C1q. Immunity *7*, 345–355.

Lacroix, M., Rossi, V., Gaboriaud, C., Chevallier, S., Jaquinod, M., Thielens, N.M., Gagnon, J., and Arlaud, G.J. (1997). Structure and assembly of the catalytic region of human complement protease C1r: a three-dimensional model based on chemical cross-linking and homology modeling. Biochemistry (Mosc.) *36*, 6270–6282.

Law, R.H., Zhang, Q., McGowan, S., Buckle, A.M., Silverman, G.A., Wong, W., Rosado, C.J., Langendorf, C.G., Pike, R.N., Bird, P.I., et al. (2006). An overview of the serpin superfamily. Genome Biol *7*, 216.

Ley, R.E., Hamady, M., Lozupone, C., Turnbaugh, P.J., Ramey, R.R., Bircher, J.S., Schlegel, M.L., Tucker, T.A., Schrenzel, M.D., Knight, R., et al. (2008). Evolution of Mammals and Their Gut Microbes. Science *320*, 1647–1651.

Li, Y., Lu, W., Schwartz, A.L., and Bu, G. (2002). Receptor-Associated Protein Facilitates Proper Folding and Maturation of the Low-Density Lipoprotein Receptor and Its Class 2 Mutants †. Biochemistry (Mosc.) *41*, 4921–4928.

Lillis, A.P., Van Duyn, L.B., Murphy-Ullrich, J.E., and Strickland, D.K. (2008). LDL Receptor-Related Protein 1: Unique Tissue-Specific Functions Revealed by Selective Gene Knockout Studies. Physiol. Rev. *88*, 887–918.

Malhotra, R., Willis, A.C., Jensenius, J.C., Jackson, J., and Sim, R.B. (1993). Structure and homology of human C1q receptor (collectin receptor). Immunology *78*, 341.

Matsushita, M., and Fujita, T. (1992). Activation of the classical complement pathway by mannose-binding protein in association with a novel C1s-like serine protease. J. Exp. Med. *176*, 1497–1502.

Matsushita, M., and Fujita, T. (2001). Ficolins and the lectin complement pathway. Immunol. Rev. *180*, 78–85.

Megyeri, M., Harmat, V., Major, B., Vegh, A., Balczer, J., Heja, D., Szilagyi, K., Datz, D., Pal, G., Zavodszky, P., et al. (2013). Quantitative Characterization of the Activation Steps of Mannan-binding Lectin (MBL)-associated Serine Proteases (MASPs) Points to the Central Role of MASP-1 in the Initiation of the Complement Lectin Pathway. J. Biol. Chem. *288*, 8922–8934.

Motta, J.-P., Bermúdez-Humarán, L.G., Deraison, C., Martin, L., Rolland, C., Rousset, P., Boue, J., Dietrich, G., Chapman, K., Kharrat, P., et al. (2012). Food-grade bacteria expressing elafin protect against inflammation and restore colon homeostasis. Sci. Transl. Med. *4*, 158ra144.

Moulds, J.M., Zimmerman, P.A., Doumbo, O.K., Kassambara, L., Sagara, I., Diallo, D.A., Atkinson, J.P., Krych-Goldberg, M., Hauhart, R.E., Hourcade, D.E., et al. (2001). Molecular identification of Knops blood group polymorphisms found in long homologous region D of complement receptor 1. Blood *97*, 2879–2885.

Neels, J.G., van den Berg, B.M., Lookene, A., Olivecrona, G., Pannekoek, H., and van Zonneveld, A.-J. (1999). The second and fourth cluster of class A cysteine-rich repeats of the low density lipoprotein receptor-related protein share ligand-binding properties. J. Biol. Chem. *274*, 31305–31311.

Paidassi, H., Tacnet-Delorme, P., Garlatti, V., Darnault, C., Ghebrehiwet, B., Gaboriaud, C., Arlaud, G.J., and Frachet, P. (2008). C1q Binds Phosphatidylserine and Likely Acts as a Multiligand-Bridging Molecule in Apoptotic Cell Recognition. J. Immunol. *180*, 2329–2338.

Païdassi, H., Tacnet-Delorme, P., Lunardi, T., Arlaud, G.J., Thielens, N.M., and Frachet, P. (2008). The lectin-like activity of human C1q and its implication in DNA and apoptotic cell recognition. FEBS Lett. *582*, 3111–3116.

Paidassi, H., Tacnet-Delorme, P., Arlaud, G.J., and Frachet, P. (2009). How phagocytes track down and respond to apoptotic cells. Crit. Rev. Immunol. *29*, 111–130.

Païdassi, H., Tacnet-Delorme, P., Verneret, M., Gaboriaud, C., Houen, G., Duus, K., Ling, W.L., Arlaud, G.J., and Frachet, P. (2011). Investigations on the C1q–Calreticulin–Phosphatidylserine Interactions Yield New Insights into Apoptotic Cell Recognition. J. Mol. Biol. *408*, 277–290.

Patel, M., Morrow, J., Maxfield, F.R., Strickland, D.K., Greenberg, S., and Tabas, I. (2003). The Cytoplasmic Domain of the Low Density Lipoprotein (LDL) Receptor-related Protein, but Not That of the LDL Receptor, Triggers Phagocytosis. J. Biol. Chem. *278*, 44799–44807.

Peitsch, M.C., Amiguet, P., Guy, R., Brunner, J., Maizel, J.V., and Tschopp, J. (1990). Localization and molecular modelling of the membrane-inserted domain of the ninth component of human complement and perforin. Mol. Immunol. *27*, 589–602.

Perona, J.J., and Craik, C.S. (1997). Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. J. Biol. Chem. *272*, 29987–29990.

Petillot, Y., Thibault, P., Thielens, N.M., Rossi, V., Lacroix, M., Coddeville, B., Spik, G., Schumaker, V.N., Gagnon, J., and Arlaud, G.J. (1995). Analysis of the N-linked oligosaccharides of human C1s using electrospray ionisation mass spectrometry. FEBS Lett. *358*, 323–328.

Rahman, A., and Isenberg, D.A. (2008). Systemic lupus erythematosus. N. Engl. J. Med. *358*, 929–939.

Ranganathan, S., Liu, C.-X., Migliorini, M.M., von Arnim, C.A.F., Peltan, I.D., Mikhailenko, I., Hyman, B.T., and Strickland, D.K. (2004). Serine and Threonine Phosphorylation of the Low Density Lipoprotein Receptor-related Protein by Protein Kinase Cα Regulates Endocytosis and Association with Adaptor Molecules. J. Biol. Chem. *279*, 40536–40544.

Rawlings, N., Tolle, D., and Barrett, A. (2004). Evolutionary families of peptidase inhibitors. Biochem J *378*, 705–716.

Ricklin, D., Hajishengallis, G., Yang, K., and Lambris, J.D. (2010). Complement: a key system for immune surveillance and homeostasis. Nat. Immunol. *11*, 785–797.

Rossi, V., Gaboriaud, C., Lacroix, M., Ulrich, J., Fontecilla-Camps, J.C., Gagnon, J., and Arlaud, G.J. (1995). Structure of the catalytic region of human complement protease C1s: study by chemical cross-linking and three-dimensional homology modeling. Biochemistry (Mosc.) *34*, 7311–7321.

Rossi, V., Bally, I., Thielens, N.M., Esser, A.F., and Arlaud, G.J. (1998). Baculovirus-mediated Expression of Truncated Modular Fragments from the Catalytic Region of Human Complement Serine Protease C1s EVIDENCE FOR THE INVOLVEMENT OF BOTH COMPLEMENT CONTROL PROTEIN MODULES IN THE RECOGNITION OF THE C4 PROTEIN SUBSTRATE. J. Biol. Chem. *273*, 1232–1239.

Rossi, V., Cseh, S., Bally, I., Thielens, N.M., Jensenius, J.C., and Arlaud, G.J. (2001). Substrate Specificities of Recombinant Mannan-binding Lectin-associated Serine Proteases-1 and -2. J. Biol. Chem. *276*, 40880–40887.

Rossi, V., Teillet, F., Thielens, N.M., Bally, I., and Arlaud, G.J. (2005). Functional Characterization of Complement Proteases C1s/Mannan-binding Lectin-associated Serine Protease-2 (MASP-2) Chimeras Reveals the Higher C4 Recognition Efficacy of the MASP-2 Complement Control Protein Modules. J. Biol. Chem. *280*, 41811–41818.

Rossi, V., Bally, I., Ancelet, S., Xu, Y., Fremeaux-Bacchi, V., Vives, R.R., Sadir, R., Thielens, N., and Arlaud, G.J. (2010a). Functional Characterization of the Recombinant Human C1 Inhibitor Serpin Domain: Insights into Heparin Binding. J. Immunol. *184*, 4982–4989.

Rossi, V., Wang, Y., and Esser, A.F. (2010b). Topology of the membrane-bound form of complement protein C9 probed by glycosylation mapping, anti-peptide antibody binding, and disulfide modification. Mol. Immunol. *47*, 1553–1560.

Santos Rocha, C., Lakhdari, O., Blottière, H.M., Blugeon, S., Sokol, H., Bermúdez-Humarán, L.G., Azevedo, V., Miyoshi, A., Doré, J., Langella, P., et al. (2012). Anti-inflammatory properties of dairy lactobacilli. Inflamm. Bowel Dis. *18*, 657–666.

Sheriff, S., Chang, C.Y., and Ezekowitz, R.A. (1994). Human mannose-binding protein carbohydrate recognition domain trimerizes through a triple alpha-helical coiled-coil. Nat. Struct. Biol. *1*, 789–794.

Soares, D.C., Gerloff, D.L., Syme, N.R., Coulson, A.F.W., Parkinson, J., and Barlow, P.N. (2005). Large-scale modelling as a route to multiple surface comparisons of the CCP module family. Protein Eng. Des. Sel. *18*, 379–388.

Su, H.P., Nakada-Tsukui, K., Tosello-Trampont, A.-C., Li, Y., Bu, G., Henson, P.M., and Ravichandran, K.S. (2002). Interaction of CED-6/GULP, an Adapter Protein Involved in Engulfment of Apoptotic Cells with CED-1 and CD91/Low Density Lipoprotein Receptor-related Protein (LRP). J. Biol. Chem. *277*, 11772–11779.

Summerfield, J.A., Ryder, S., Sumiya, M., Thursz, M., Gorchein, A., Monteil, M.A., and Turner, M.W. (1995). Mannose binding protein gene mutations associated with unusual and severe infections in adults. Lancet *345*, 886–889.

Sunyer, J.O., Zarkadis, I.K., and Lambris, J.D. (1998). Complement diversity: a mechanism for generating immune diversity? Immunol. Today *19*, 519–523.

Taylor, K.M., Morgan, B.P., and Campbell, A.K. (1994). Altered glycosylation and selected mutation in recombinant human complement component C9: effects on haemolytic activity. Immunology *83*, 501.

Teillet, F., Gaboriaud, C., Lacroix, M., Martin, L., Arlaud, G.J., and Thielens, N.M. (2008). Crystal Structure of the CUB1-EGF-CUB2 Domain of Human MASP-1/3 and Identification of Its Interaction Sites with Mannan-binding Lectin and Ficolins. J. Biol. Chem. *283*, 25715–25724.

Terrasse, R., Tacnet-Delorme, P., Moriscot, C., Perard, J., Schoehn, G., Vernet, T., Thielens, N.M., Di Guilmi, A.M., and Frachet, P. (2012). Human and Pneumococcal Cell Surface Glyceraldehyde-3-phosphate Dehydrogenase (GAPDH) Proteins Are Both Ligands of Human C1q Protein. J. Biol. Chem. *287*, 42620–42633.

Thiel, S., Vorup-Jensen, T., Stover, C.M., Schwaeble, W., Laursen, S.B., Poulsen, K., Willis, A.C., Eggleton, P., Hansen, S., Holmskov, U., et al. (1997). A second serine protease associated with mannan-binding lectin that activates complement. Nature *386*, 506–510.

Thielens, N.M., Cseh, S., Thiel, S., Vorup-Jensen, T., Rossi, V., Jensenius, J.C., and Arlaud, G.J. (2001). Interaction Properties of Human Mannan-Binding Lectin (MBL)-Associated Serine Proteases-1 and -2, MBL-Associated Protein 19, and MBL. J. Immunol. *166*, 5068–5077.

Turner, M.W. (1996). Mannose-binding lectin: the pluripotent molecule of the innate immune system. Immunol. Today *17*, 532–540.

Turner, M.W., Super, M., Singh, S., and Levinsky, R.J. (1991). Molecular basis of a common opsonic defect. Clin. Exp. Allergy J. Br. Soc. Allergy Clin. Immunol. *21 Suppl 1*, 182–188.

Verdaguer, N., Fita, I., Reithmayer, M., Moser, R., and Blaas, D. (2004). X-ray structure of a minor group human rhinovirus bound to a fragment of its cellular receptor protein. Nat. Struct. Mol. Biol. *11*, 429–434.

Villiers, C.L., Arlaud, G.J., and Colomb, M.G. (1985). Domain structure and associated functions of subcomponents C1r and C1s of the first component of human complement. Proc. Natl. Acad. Sci. *82*, 4477–4481.

Walport, M.J. (2001a). Complement. First of two parts. N. Engl. J. Med. *344*, 1058–1066.

Walport, M.J. (2001b). Complement. Second of two parts. N. Engl. J. Med. *344*, 1140–1144.

Walport, M.J., Davies, K.A., and Botto, M. (1998). C1q and systemic lupus erythematosus. Immunobiology *199*, 265–285.

Weiss, V., Fauser, C., and Engel, J. (1986). Functional model of subcomponent C1 of human complement. J. Mol. Biol. *189*, 573–581.

Wijeyewickrema, L.C., Duncan, R.C., and Pike, R.N. (2014). The Role of the Lys628 (192) Residue of the Complement Protease, C1s, in Interacting with Peptide and Protein Substrates. Front. Immunol. *5*.

Yasui, N., Nogi, T., and Takagi, J. (2010). Structural Basis for Specific Recognition of Reelin by Its Receptors. Structure *18*, 320–331.

Annexes

Publications des travaux de recherche

Publications du Chapitre III

***P1** - Rossi V. et al., **(1995)** Structure of the catalytic region of human complement protease C1s: study by chemical cross-linking and three-dimensional homology modeling*

***P2** - Pétillot Y. et al., **(1995)** Analysis of the N-Linked oligosaccharides of human C1s using electrospray ionisation mass spectrometry.*

***P3** - Lacroix M.et al., **(1997)** Structure and assembly of the catalytic region of human complement protease C1r: a three-dimensional model based on chemical cross-linking and homology modeling.*

***P4** - Rossi V. et al., **(1998)** Baculovirus-mediated expression of truncated modular fragments from the catalytic region of human complement serine protease C1s. Evidence for the involvement of both complement control protein modules in the recognition of the C4 protein substrate.*

***P5** - Gans P. et al., **(1998)** NMR structures of the C-terminal end of human complement serine protease C1s.*

***P6** - Gaboriaud C. et al., **(1998)** Evolutionary conserved rigid module - domain interactions can be detected at the sequence level: the examples of complement and coagulation proteases.*

***P7** - Gaboriaud C. et al., **(2000)** Crystal structure of the catalytic domain of complement C1s : a serine protease with a handle.*

***P8** - Thielens N.M. et al., **(2001)** Interaction properties of human mannan-binding lectin (MBL)-associated serine proteases-1 and 2, MBL-associated protein 19, and MBL.*

***P9** - Rossi V. et al., **(2001)**. Substrate specificities of recombinant mannan-binding lectin-associated serine proteases-1 and -2.*

***P11** - Bally I. et al., **(2005)** Functional Role of the Linker between the Complement Control Protein Modules of Complement Protease C1s.*

***P12** - Rossi V. et al., **(2005)** Functional characterization of complement proteases C1s/MASP-2 chimeras reveals the higher C4 recognition efficacy of the MASP-2 complement control protein modules*

***P13** - Kerr F.K. et al., **(2008)** Elucidation of the substrate specificity of the MASP-2 protease of the lectin complement pathway and identification of the enzyme as a major physiological target of the serpin, C1-inhibitor.*

***P16** - Rossi V. et al., **(2010)** Functional Characterization of the Recombinant Human C1 inhibitor Serpin Domain: Insights into Heparin Binding.*

***P18** – Gaboriaud C. et al., **(2013)**, Enzymatic properties of the serine protease domain of MASP-3 and its crystal structure in complex with ecotin.*

En noir, publications intégrales

Biochemistry **1995,** *34,* 7311−7321

P1

Structure of the Catalytic Region of Human Complement Protease C$\overline{1}$s: Study by Chemical Cross-Linking and Three-Dimensional Homology Modeling[†]

Véronique Rossi,[‡] Christine Gaboriaud,[§] Monique Lacroix,[‡] Jacques Ulrich,[‖] Juan Carlos Fontecilla-Camps,[§]
Jean Gagnon,[‡] and Gérard J. Arlaud*,[‡]

Laboratoire d'Enzymologie Moléculaire, Laboratoire de Cristallogénèse et Cristallographie des Protéines, and
Laboratoire de Spectrométrie de Masse des Protéines, Institut de Biologie Structurale Jean-Pierre Ebel (CEA-CNRS),
41, avenue des Martyrs, 38027 Grenoble Cedex 1, France

Received December 1, 1994; Revised Manuscript Received January 26, 1995[⊗]

ABSTRACT: C$\overline{1}$s is a multidomain serine protease that is responsible for the enzymic activity of C$\overline{1}$, the first component of the classical pathway of complement. Its catalytic region (γ-B) comprises two contiguous complement control protein (CCP) modules, IV and V (about 60 residues each), a 15-residue intermediary segment, and the B chain (251 residues), which is the serine protease domain. With a view to identify domain−domain interactions within this region, the γ-B fragment of C$\overline{1}$s, obtained by limited proteolysis with plasmin, was chemically cross-linked with the water-soluble carbodiimide 1-ethyl-3-[3-(dimethylamino)propyl]carbodiimide; then cross-linked peptides were isolated after CNBr cleavage and thermolytic digestion. N-Terminal sequence and mass spectrometry analyses allowed us to identify two cross-links between Lys 405 of module V and Glu 672 of the B chain and between Glu 418 of the intermediary segment and Lys 608 of the B chain. Three-dimensional modeling of the CCP modules IV and V and of the catalytic B chain was also carried out on the basis of their respective homology with the 16th and 5th CCP modules of complement factor H and type I serine proteases. The information provided by both the chemical cross-linking studies and the homology modeling enabled us to construct a three-dimensional model for the assembly of the C-terminal part of the γ-B region, comprising module V, the intermediary segment, and the B chain. This model shows that module V interacts with the serine protease B chain on the side opposite to both the activation site and the catalytic site. Functional implications of this interaction are discussed in terms of the possible role of module V in the specific recognition and positioning of C4, one of the two substrates of C$\overline{1}$s.

The catalytic subunit of human C1, the first component of the classical pathway of complement, is a Ca^{2+}-dependent tetrameric association (C1s−C1r−C1r−C1s) of two homologous multidomain serine proteases, C1r and C1s, which undergo sequential activation when C1 binds, through its third subcomponent C1q, to immune complexes or various nonimmune activators. Autoactivation of C1r yields an active enzyme, C$\overline{1}$r, which in turn converts proenzyme C1s into C$\overline{1}$s. The latter is a highly specific protease with trypsin-like specificity that is responsible for limited proteolysis of C4 and C2, the next components of the complement cascade [reviewed by Cooper (1985), Schumaker et al. (1987), and Arlaud et al. (1987)].

Human proenzyme C1s, a 673-residue polypeptide (Tosi et al., 1987; Mackinnon et al., 1987), is cleaved upon activation between Arg 422 and Ile 423 to yield two disulfide-linked chains, A (N-terminal) and B (C-terminal) (Spycher et al., 1986). C$\overline{1}$s is thought to be organized in two functional regions: (i) an interaction region (α) corresponding to the N-terminal half of the A chain, responsible for Ca^{2+}-dependent C1r−C1s interactions involved in assembly of the C1s−C1r−C1r−C1s tetramer; (ii) a catalytic region (γ-B) responsible for the enzymic activity of C$\overline{1}$, comprising the C-terminal part of the A chain (γ) and the B chain, connected to each other through a single disulfide bridge. The B chain (251 residues), which contains the active site, is homologous to the catalytic domains of type I serine proteases. The γ region comprises two "complement control protein" (CCP)[1] modules, IV and V (about 60 residues each), homologous to those mostly found, usually in multiple copies, in complement regulatory proteins (Reid et al., 1986). Module V bears, at position 391, a complex-type N-linked oligosaccharide (Y. Petillot, P. Thibault, N. M. Thielens, V. Rossi, M. Lacroix, B. Coddeville, G. Spik, V. N. Schumaker, J. Gagnon, and G. J. Arlaud, unpublished experiments). It is followed by a 15-residue intermediary segment which in proenzyme C1s connects module V to the B chain moiety and is C-terminal to the γ region in C$\overline{1}$s.

The γ-B fragment, which is obtained by limited proteolysis of C$\overline{1}$s with plasmin, is visualized by electron microscopy after rotary shadowing as a compact structure (Villiers et

---

[†] This is Publication No. 254 of the Institut de Biologie Structurale Jean-Pierre Ebel. This work was supported by the Commissariat à l'Energie Atomique and the Centre National de la Recherche Scientifique.

[*] To whom correspondence should be addressed.

[‡] Laboratoire d'Enzymologie Moléculaire.

[§] Laboratoire de Cristallogénèse et Cristallographie des Protéines.

[‖] Laboratoire de Spectrométrie de Masse des Protéines.

[⊗] Abstract published in *Advance ACS Abstracts,* May 1, 1995.

[1] Abbreviations: CCP, complement control protein module; EDC, 1-ethyl-3-[3-(dimethylamino)propyl]carbodiimide; EDTA, ethylenediaminetetraacetic acid; FAB, fast atom bombardment; HPLC, high-pressure liquid chromatography; PTH, phenylthiohydantoin; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis. The nomenclature of complement components is that recommended by the World Health Organization; activated components are indicated by an overhead bar, e.g., C$\overline{1}$s.

0006-2960/95/0434-7311$09.00/0 © 1995 American Chemical Society

P1

al., 1985). However, neutron scattering studies indicate that this fragment has a relatively large radius of gyration, not compatible with a single globular domain (Zaccaï et al., 1990). Moreover, studies by differential scanning calorimetry (Medved et al., 1989) give evidence for the presence of three independently folded domains, corresponding to the CCP modules IV and V and the B chain. From a functional point of view, studies indicate that the γ region may be involved in the binding of C4 (Matsumoto & Nagaki, 1986; Matsumoto et al., 1989). However, apart from homology modeling of the catalytic B chain (Carter et al., 1984), no information is available yet on the three-dimensional structure of the individual domains of the $C\overline{1}s$ γ-B region and on the assembly of this region.

With the view to identify domain−domain interactions within the catalytic region of $C\overline{1}s$, the γ-B fragment was submitted to chemical cross-linking with EDC, a zero-length cross-linking agent able to convert ionic bonds between carboxyl groups and ϵ-amino groups into covalent pseudopeptide bonds (Means & Feeney, 1971). At the same time, three-dimensional modeling of the individual domains from γ-B was also performed on the basis of their homology with proteins of known structures, namely, type I serine proteases and CCP modules. In a second step, the complementary information provided by both the chemical cross-linking and the homology modeling was utilized to construct a three-dimensional model of the C-terminal part of the γ-B region, including the CCP module V, the 15-residue intermediary segment, and the B chain.

EXPERIMENTAL PROCEDURES

Materials. Human plasmin was obtained from Kabi Vitrum, Stockholm, Sweden. Thermolysin from *Bacillus thermoproteolyticus* was from Boehringer Mannheim France S.A. EDC and diisopropyl fluorophosphate were purchased from Sigma.

Isolation of Fragment $C\overline{1}s$ γ-B. $C\overline{1}s$ was purified from human plasma as described previously (Arlaud et al., 1979) and subjected to limited proteolysis with plasmin as described by Thielens et al. (1990). Purification of the $C\overline{1}s$ γ-B fragment was realized by high-pressure hydrophobic interaction chromatography on a TSK-Phenyl 5PW column (7.5 mm × 75 mm) (LKB) as described previously (Thielens et al., 1990), except that the plasmin digest was dialyzed against 1.5 M $(NH_4)_2SO_4$/0.1 M Na_2HPO_4 (pH 7.4) prior to loading onto the column. Alternatively, purification was achieved by ion-exchange chromatography on a Mono Q HR 5/5 column (Pharmacia). In that case, the plasmin digest was dialyzed against 20 mM Na_2HPO_4 (pH 7.4) and loaded onto the column equilibrated in the same buffer, and elution was carried out by a linear NaCl gradient from 0 to 0.25 M in 20 min. The flow rate was 0.5 mL/min, and fragment γ-B was detected from its absorbance at 280 nm. The concentrations of purified $C\overline{1}s$ and its γ-B fragment were determined from A_{280} measurements by using values of E (1%, 1 cm) = 14.5 and 18.3, respectively (Thielens et al., 1990; Zaccai et al., 1990). The average molecular weights of $C\overline{1}s$ (79 800) and $C\overline{1}s$ γ-B (47 520) were determined experimentally by electrospray ionization mass spectrometry (Y. Petillot, P. Thibault, N. M. Thielens, V. Rossi, M. Lacroix, B. Coddeville, G. Spik, V. N. Schumaker, J. Gagnon, and G. J. Arlaud, unpublished experiments).

SDS−PAGE Analysis. Proteins were analyzed on 12.5% polyacrylamide gels as described by Laemmli (1970). Phosphorylase *b* (M_r 94 000), bovine serum albumin (M_r 67 000), ovalbumin (M_r 43 000), carbonic anhydrase (M_r 30 000), soybean trypsin inhibitor (M_r 20 100), and α-lactalbumin (M_r 14 400), all reduced and alkylated, were used as molecular weight markers. Protein staining was performed with Coomassie Blue, and gels were scanned using a Shimadzu Model CS 9000 gel scanner.

Chemical Cross-Linking. The isolated γ-B fragment (160 nmol) was dialyzed against 150 mM NaCl/20 mM 2-(*N*-morpholino)ethanesulfonic acid (pH 6.0), diluted to 0.2 mg/mL, and then incubated for 2 h at 30 °C in the presence of 10 mM EDC. The reaction mixture was dialyzed exhaustively against the same buffer without EDC and then against 0.5% acetic acid and freeze-dried. The cross-linked products were dissolved in 8 mL of 6 M guanidine hydrochloride/2 mM EDTA/0.4 M Tris-HCl (pH 8.5). Reduction of the disulfide bridges was achieved by incubation in the presence of 0.5% (v/v) 2-mercaptoethanol for 2 h at 25 °C, and alkylation was performed by a further incubation for 2 h at 25 °C in the presence of 4% (v/v) 4-vinylpyridine. The reduced and S-pyridylethylated material was dialyzed against 0.5% acetic acid and then freeze-dried. After dissolution in 6 M urea/0.2 M formic acid (2.6 mL), the cross-linked products were fractionated by high-pressure gel permeation on a TSK G-3000 SW column (7.5 mm × 600 mm) (LKB) equilibrated in the same medium. The flow rate was 1 mL/min, and peaks were detected by absorbance at 280 nm. Fractions were dialyzed against 0.5% acetic acid and freeze-dried.

CNBr Cleavage of the Cross-Linked γ-B Fragment. The reduced and S-pyridylethylated, cross-linked γ-B fragment (48 nmol) was dissolved in 70% formic acid (1.2 mL) containing CNBr (65 mmol) and kept in the dark for 24 h at 4 °C. The same treatment was applied to a reduced and S-pyridylethylated, non-cross-linked γ-B sample (31 nmol). After dilution 1:15 with water, the CNBr-cleavage peptides from both γ-B samples were freeze-dried, redissolved in 0.1% trifluoroacetic acid, and then subjected to high-pressure gel permeation on TSK G-3000 SW and TSK G-2000 SW columns (7.5 mm × 600 mm) equilibrated in 0.1% trifluoroacetic acid and used together in order of decreasing pore size, as described by Carter et al. (1983). The flow rate was 1 mL/min, and peptides were detected by absorbance at 215 nm. Fractions 1 and 3 from control γ-B and fraction X1 from cross-linked γ-B were further fractionated by reverse-phase HPLC on a C4 Deltapak column (3.9 mm × 150 mm) (Waters Associates). The column was equilibrated with a mixture of solutions A (0.1% trifluoroacetic acid) and B (acetonitrile/methanol/propan-2-ol, 1:1:1 v/v/v) in the ratio 95:5 (v/v) and then eluted with a linear gradient to give a final ratio of 20:80 (v/v) in 60 min. The flow rate was 0.8 mL/min, and peptides were detected from absorbance at 215 nm.

Thermolytic Cleavage of the Cross-Linked Peptides. Cross-linked CNBr-cleavage peptides contained in fraction X1A (2−3 nmol) were dissolved in 0.5 mL of 2 mM $CaCl_2$/0.1 M NH_4HCO_3 (pH 8.0) and digested by 5% (w/w) thermolysin for 4 h at 60 °C. After the reaction, the pH of the mixture was adjusted to 2.0 with trifluoroacetic acid, and the thermolytic peptides were separated by reverse-phase HPLC on a C4 column as described above.

P1

N-Terminal Sequence Analysis. N-Terminal sequence analyses were performed using an Applied Biosystems Model 477A protein sequencer, and amino acid PTH derivatives were identified and quantitated on-line with a Model 120A HPLC system, as recommended by the manufacturer. Proteins separated by SDS−PAGE were electrophoretically transferred to polyvinylidene difluoride membranes (Matsudaira, 1987), according to the protocol defined for ProBlott membranes (Applied Biosystems).

Mass Spectrometry Analyses. FAB mass spectrometry analyses were carried out on a VG analytical ZAB-SE double-focusing mass spectrometer, as described previously (Thielens et al., 1990). Peptides were dissolved in 5% acetic acid, and thioglycerol was used as the matrix.

Electrospray ionization mass spectra were obtained on an API III triple−quadrupole mass spectrometer (PE/Sciex, Thornhill, Ontario, Canada) equipped with a nebulizer-assisted electrospray (ionspray) source. The mass spectrometer was scanned from m/z 400 to 1800, with steps of 0.5 or 1 m/z unit. The dwell time was 2 ms, and the resolution was 1 mass unit. For each scan, an average of 5−10 scans were accumulated. Samples submitted to flow injection analysis were dissolved in methanol/water/acetic acid (25:74:1) (v/v/v) and introduced by means of a syringe pump into a Valco C6W injector. For liquid chromatography/mass spectrometry analyses, samples were loaded onto an home-made capillary column (Nucleosil C8, 0.25 mm × 150 mm) and eluted at a flow rate of 5 μL/min with a linear gradient of 5−90% acetonitrile in 0.1% trifluoroacetic acid over 60 min. The effluent was introduced directly into the electrospray source via a 50 μm × 70 cm fused silica capillary.

Peptide Synthesis. Peptide Q665−D673 was synthesized chemically by the stepwise solid-phase method (Barany & Merrifield, 1980) in an Applied Biosystems 430A automated synthesizer, using the *tert*-butyloxycarbonyl strategy. Deprotection of the side chains and cleavage of the peptide from the resin were performed by HF treatment (Tam et al., 1983). Purification was achieved by preparative reverse-phase HPLC on a 300 Å Deltapak C18 column (1.9 cm × 30 cm) (Waters Associates), and the mass of the purified peptide was checked by FAB mass spectrometry.

Computer-Assisted Three-Dimensional Homology Modeling. The program O (Jones et al., 1991) was used on an ESV/20 graphics station (Evans & Sutherland) to construct homology-based three-dimensional models of the protein modules of the γ-B fragment and to monitor their assembly interactively on the basis of the information provided by chemical cross-linking.

The protocol used for modeling the individual modules was similar to that described in detail by Greer (1990) in the case of mammalian serine proteases. A set of homologous three-dimensional structures is used as a reference template to follow the local degree of structural conservation or divergence (mainly found at surface loops and turns) within the family. The superposition of these reference structures is performed on a set of highly conserved residues (such as the catalytic triad in the case of serine proteases) and then refined iteratively on all equivalent Cα positions (distant by less than 3.5 Å in a first step and by less than 1.5 Å in a second step). The sequence of the protein to be modeled is aligned onto the sequences of the reference set, avoiding, as much as possible, insertions or deletions in the conserved structural core of the reference set of structures. The reference structure exhibiting the fewest insertions and deletions relative to the model to be constructed is then used as the primary template. The nonidentical residues of this template within the conserved regions are computationally mutated to the corresponding residues of the model. In areas where the three-dimensional structures of the reference set diverge, a single reference structure is used as a template if the number of residues is equivalent, selecting when possible the one with higher sequence identity. The remaining areas, where the lengths of the model loops differ from those in the reference structures, are computationally constructed using a set of loops of the same length from a database of high-resolution X-ray structures (Jones & Thirup, 1986), chosen according to the similarity to the reference set of structures. The Cα three-dimensional models obtained by this method were not energy-minimized.

The serine proteases used as reference structures for modeling the B chain were the following: bovine chymotrypsin (Birktoft & Blow, 1972); bovine trypsin (Bartunik et al., 1989); porcine elastase (Sawyer et al., 1978); porcine kallikrein (Bode et al., 1983); rat mast cell protease (Remington et al., 1988); *Streptomyces griseus* trypsin (Read & James, 1988); rat tonin (Fujinaga & James, 1987); human leucocyte elastase (Navia et al., 1989); bovine chymotrypsinogen (Wang et al., 1985). The structures chosen were all determined at a resolution equal to or better than 1.8 Å, and coordinates were obtained from the Protein Data Bank (Abola et al., 1987; Bernstein et al., 1977). The CCP modules IV and V were modeled from the coordinates of the homologous 5th, 15th, and 16th modules from human factor H (Norman et al., 1991; Barlow et al., 1992, 1993).

RESULTS

Isolation of Fragment γ-B and Chemical Cross-Linking with EDC. Human C$\overline{1}$s was submitted to limited proteolysis with plasmin, and purification of the γ-B fragment from the plasmin digest was carried out by either high-pressure hydrophobic interaction chromatography or ion-exchange chromatography, as described under Experimental Procedures. In both cases, fragment γ-B was eluted as a single homogeneous peak, and the purification yield was about 25%. Analysis of the purified γ-B fragment by SDS−PAGE under nonreducing conditions showed that it migrated as a single band of M_r approximately 55 000. Reduction and alkylation yielded the B chain (M_r 31 000) and the γ fragment, which was usually resolved into four bands with M_r values ranging from 24 000 to 26 000. N-Terminal sequence analysis of the material contained in these bands after SDS−PAGE analysis and electrotransfer onto polyvinylidene difluoride showed a single major sequence, Leu-Arg-Tyr-His-Gly..., indicating that all observed γ species arose from cleavage of the Lys−Leu bond at position 269 of the C$\overline{1}$s A chain, in agreement with previous findings (Thielens et al., 1990).

Preliminary cross-linking studies of the γ-B fragment with EDC were performed at pH 6.5 and at low protein concentration (0.2 mg/mL), in order to minimize intermolecular cross-links. Under these conditions, as judged from SDS−PAGE analysis under reducing conditions, incubation of γ-B for 2 h at 30 °C in the presence of increasing concentrations (1−20 mM) of EDC led to the progressive disappearance of both the B chain and the γ fragment. Concomitantly, the

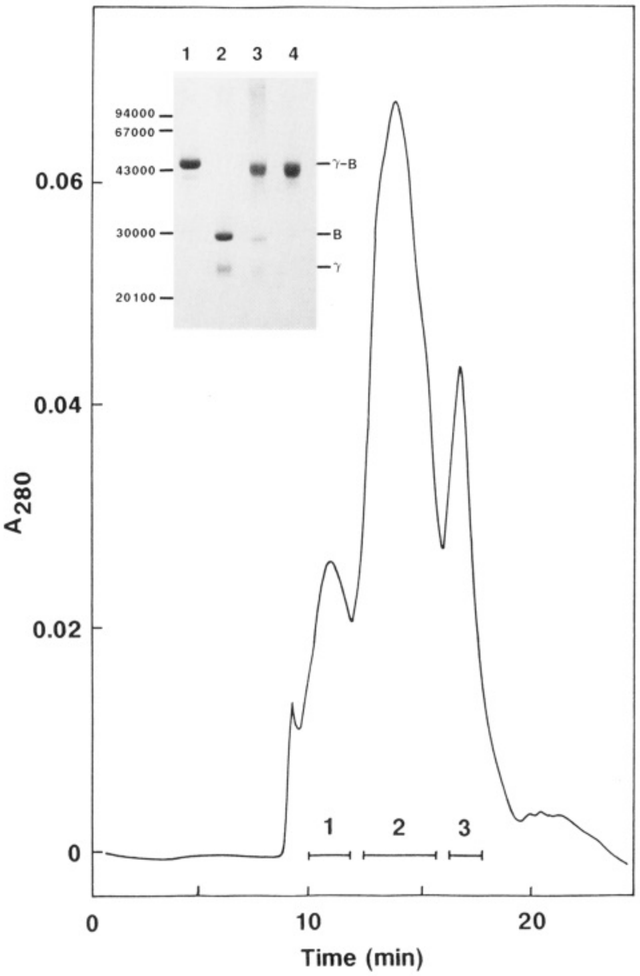


FIGURE 1: Isolation of the cross-linked γ-B species by high-pressure gel permeation. Fragment γ-B was cross-linked with EDC, reduced with 2-mercaptoethanol, and alkylated with 4-vinylpyridine as described under Experimental Procedures. The reduced and S-pyridylethylated, cross-linked material was fractionated by high-pressure gel permeation on a TSK G-3000 SW column equilibrated in 6 M urea/0.2 M formic acid. Proteins were detected by absorbance at 280 nm, and pools were made as indicated by bars. The insert shows SDS–PAGE analysis of (lane 1) the native γ-B fragment, unreduced; (lane 2) the native γ-B fragment, reduced; (lane 3) the reduced and S-pyridylethylated, cross-linked material loaded onto the column; and (lane 4) the cross-linked γ-B species isolated from fraction 2.

reaction yielded a first cross-linked γ-B species of apparent M_r 54 000, which in turn was progressively converted into a second species of apparent M_r 50 000, suggesting the occurrence of a two-step process. Several species of higher molecular weight (80 000–150 000), likely resulting from intermolecular cross-linking, also appeared at EDC concentrations above 10 mM. Decreasing the pH from 6.5 to 6.0 significantly increased the yield of the M_r 50 000 cross-linked species which, under the reaction conditions used for preparative experiments (see Experimental Procedures), reached 50–60% of the total species, as estimated from gel scanning.

One hundred sixty nanomoles of the γ-B fragment was submitted to chemical cross-linking with EDC, then reduced with 2-mercaptoethanol, and alkylated with 4-vinylpyridine, as described under Experimental Procedures. The reduced and S-pyridylethylated, cross-linked material was then fractionated by high-pressure gel permeation under denaturing conditions, as illustrated in Figure 1. This allowed us to separate the cross-linked γ-B species, contained in pool 2, from the above-described species of higher molecular weights, eluted in pool 1, and from the residual B chain and γ fragment, both found in pool 3. Although the cross-linked γ-B species obtained at this stage contained traces of the B chain, as shown on Figure 1, this material was used for the studies described below without further purification.

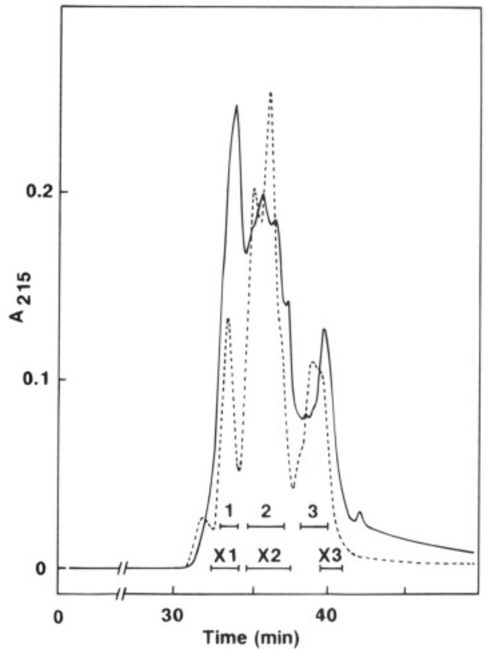


FIGURE 2: Initial fractionation by high-pressure gel permeation of the CNBr-cleavage peptides from cross-linked γ-B and from control γ-B. The CNBr-cleavage peptides from cross-linked γ-B (continuous line) and from control γ-B (dashed line) were separated on TSK G-3000 SW and TSK G-2000 SW columns as described under Experimental Procedures. Peptides were detected by absorbance at 215 nm, and pools were made as indicated by bars. Fractions 1, 2, and 3 refer to control γ-B, and fractions X1, X2, and X3 refer to cross-linked γ-B.

Isolation of Cross-Linked Peptides. In order to isolate cross-linked peptides containing cross-linking sites between the γ fragment and the B chain, the reduced and S-pyridylethylated, cross-linked γ-B species was submitted to CNBr cleavage as described under Experimental Procedures. The same treatment was also applied to a reduced and S-pyridylethylated but non-cross-linked control γ-B sample, and the CNBr-cleavage peptides from both γ-B samples were initially fractionated by high-pressure gel permeation, as illustrated in Figure 2. From the number of methionine residues present in the γ fragment and the B chain, CNBr cleavage was expected to yield 12 peptides (Figure 3).

As shown in Figure 2, the peptides from the control γ-B sample were separated into three major fractions, the contents of which were determined by N-terminal sequence analysis and electrospray ionization mass spectrometry. Fraction 1 only contained peptide E379–R422, the C-terminal peptide from fragment γ. Three peptides from the B chain (Y471–M487, L488–M526, D546–M601) and peptide P278–M378 from γ were found in fraction 2. The last fraction, 3, contained five peptides from the B chain (G527–M545, I602–M610, D611–M661, K622–M664, Q665–D673) and peptide L270–M277 from γ. For unknown reasons, peptide I423–M470, the N-terminal peptide from the B chain, could not be recovered from the column.

P1

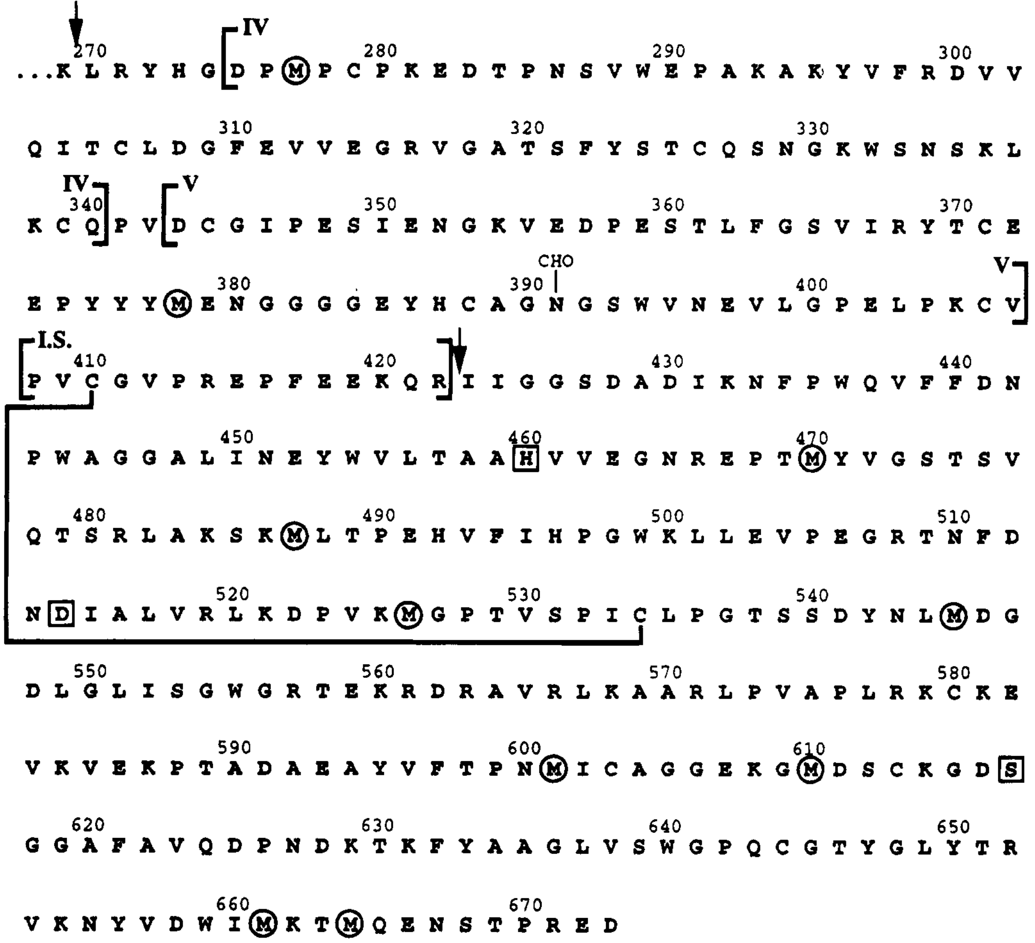

IV
270 280 290 300
...K L R Y H G D P M P C P K E D T P N S V W E P A K A K Y V F R D V V
310 320 330
Q I T C L D G F E V V E G R V G A T S F Y S T C Q S N G K W S N S K L
IV V
340 350 360 370
K C Q P V D C G I P E S I E N G K V E D P E S T L F G S V I R Y T C E
CHO V
380 390 400
E P Y Y Y M E N G G G G E Y H C A G N G S W V N E V L G P E L P K C V
I.S.
410 420 430 440
P V C G V P R E P F E E K Q R I I G G S D A D I K N F P W Q V F F D N
450 460 470
P W A G G A L I N E Y W V L T A A H V V E G N R E P T M Y V G S T S V
480 490 500 510
Q T S R L A K S K M L T P E H V F I H P G W K L L E V P E G R T N F D
520 530 540
N D I A L V R L K D P V K M G P T V S P I C L P G T S S D Y N L M D G
550 560 570 580
D L G L I S G W G R T E K R D R A V R L K A A R L P V A P L R K C K E
590 600 610
V K V E K P T A D A E A Y V F T P N M I C A G G E K G M D S C K G D S
620 630 640 650
G G A F A V Q D P N D K T K F Y A A G L V S W G P Q C G T Y G L Y T R
660 670
V K N Y V D W I M K T M Q E N S T P R E D

FIGURE 3: Amino acid sequence of the C$\overline{1}$s γ-B fragment. The amino acid numbering is that of intact C1s (Mackinnon et al., 1987; Tosi et al., 1987). The site of plasmin cleavage generating the γ-B fragment and the activation site of C1s defining fragment γ and the B chain are indicated by arrows. Methionine residues are circled, and active site residues are squared. The boundaries of the CCP modules IV and V and of the intermediary segment (I.S.) are indicated as defined in this study. The only disulfide bridge shown is that connecting, through half-cystine residues 410 and 534, the γ fragment to the B chain (Hess et al., 1991). CHO indicates the N-linked oligosaccharide.

Separation of the CNBr-cleavage peptides from the cross-linked γ-B sample by the same method also yielded three major fractions, but the elution profile showed marked differences (Figure 2). In particular, the first fraction, X1, was significantly increased, whereas noticeable differences in the shape and/or elution position of the following fractions, X2 and X3, were also observed. N-Terminal sequence analysis of fraction X1 indicated that, in addition to peptide E379–R422 also found in the corresponding fraction 1, it contained equivalent amounts of peptides I602–M610 and Q665–D673, as well as lower amounts of peptides P278–M378 and L488–M526. The presence of the latter two peptides is likely a consequence of contamination by the neighboring fraction X2, which showed essentially the same peptide contents as fraction 2. In contrast, fraction X3 was found to lack peptides I602–M610 and Q665–D673, identified in the corresponding fraction 3.

The above observations strongly suggested that fraction X1 contained cross-linked material, involving at least peptides I602–M610 and Q665–D673 from the B chain. Therefore, this was further analyzed by reverse-phase HPLC, yielding two major fractions, as shown in Figure 4. N-Terminal sequence analysis of fraction X1A yielded three sequences corresponding to peptides E379–R422, I602–M610, and Q665–D673. Although this fraction clearly showed two components, separate analyses indicated that both contained equivalent proportions of the three peptides. Comparative chromatographic analysis by the same method of the corresponding peptides obtained from CNBr cleavage of the non-cross-linked γ-B sample indicated that peptide E379–R422 normally eluted later than fraction X1A (see Figure 4), whereas peptides I602–M610 and Q665–D673 both eluted in the injection peak. It appeared likely, therefore, that fraction X1A contained cross-linked material involving these three peptides.

N-Terminal sequence analysis of the heterogeneous fraction X1B yielded two major sequences, corresponding to peptides P278–M378 and L488–M526. As shown in Figure 4, the corresponding peptides obtained from the untreated γ-B sample eluted on both sides of peak X1B. Analysis of fraction X1B by electrospray ionization mass spectrometry showed no significant m/z peak attributable to peptide L488–M526 but revealed a coherent series of peaks with mass values differing from the expected mass of peptide P278–M378 by increments of 28 and/or 156, corresponding respectively to the addition of a formyl group and to the fixation of an EDC molecule. It appeared likely, therefore, that fraction X1B contained various formylated and EDC-modified forms of peptide P278–M378, coexisting with, rather than cross-linked to, peptide L488–M526, the latter being probably also chemically modified. Although the presence of minor cross-linked species in fraction X1B could not be entirely ruled out, further investigations were conducted solely on the material contained in fraction X1A.

P1

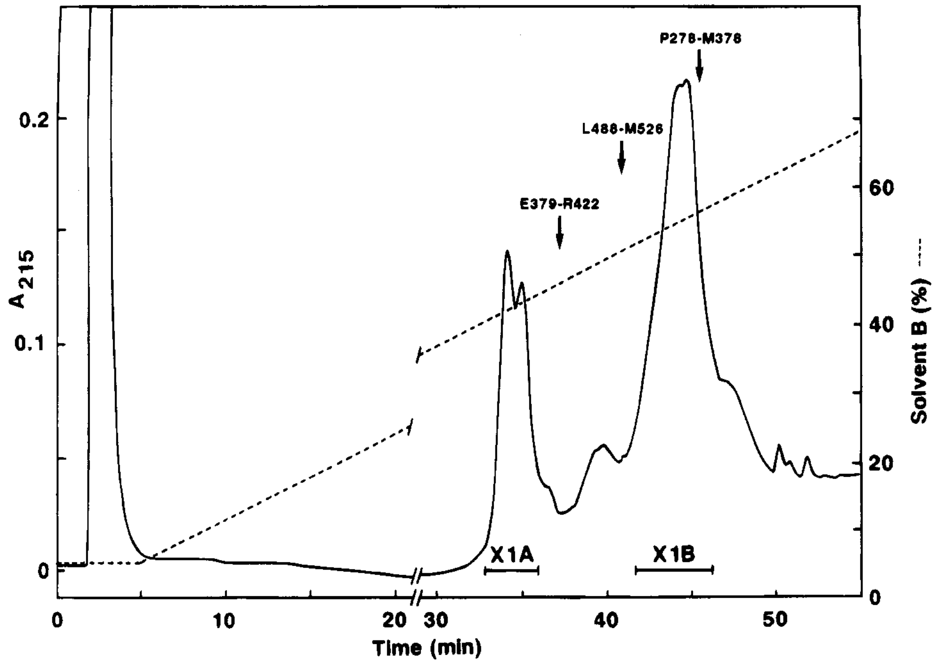


FIGURE 4: Analysis by reverse-phase HPLC of the peptides contained in fraction X1. The peptides contained in fraction X1 (see Figure 2) were analyzed by reverse-phase HPLC on a C4 column, as described under Experimental Procedures. Peptides were detected by absorbance at 215 nm, and pools were made as indicated by bars. The arrows indicate the elution positions of peptides E379−R422, L488−M526, and P278−M378 obtained from CNBr cleavage of the non-cross-linked γ-B sample.

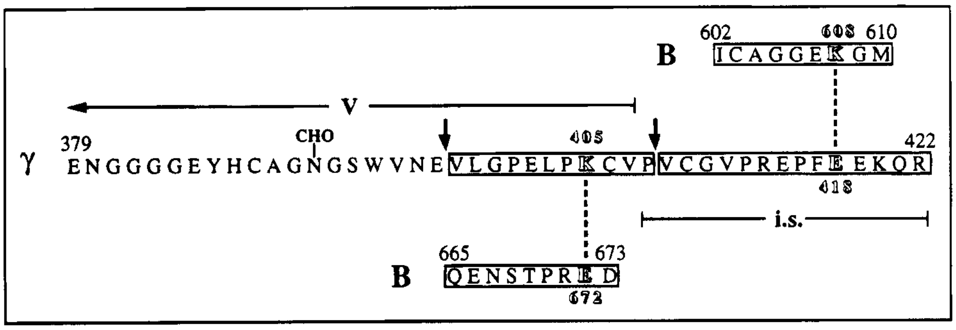


FIGURE 5: Proposed structure of the cross-linked peptides isolated from fraction X1A. Peptides I602−M610 and Q665−D673 both originate from the B chain, and peptide E379−R422 is from the γ fragment. The portions of peptide E379−R422 corresponding to module V and the intermediary segment are indicated. The major sites of cleavage by thermolysin are shown by arrows. The cross-linked peptides isolated after thermolytic cleavage are boxed, and the proposed cross-linking sites are represented by dotted lines. CHO indicates the N-linked oligosaccharide.

Identification of the Cross-Linking Sites. The presence of three peptides in fraction X1A suggested the occurrence of at least two distinct cross-linking sites. With the view to generate short peptides containing only one of these sites, the peptide material contained in fraction X1A was submitted to thermolytic cleavage, as described under Experimental Procedures. The digest was then fractionated by reverse-phase HPLC (data not shown), yielding six major fractions, the peptide contents of which were determined by both N-terminal sequence and mass spectrometry analyses. The cleavage reaction was not complete, as shown by the presence of residual uncleaved material, eluting in the latest fractions, 5 and 6. Both fractions 1 and 2 contained non-cross-linked peptides: the major component of the heterogeneous fraction 1 was peptide E379−E397, originating from the N-terminal end of peptide E379−R422, and fraction 2 contained peptide V409−R422, derived from the C-terminal end of this same peptide (Figure 5).

In contrast, N-terminal sequence analysis of fraction 3 yielded two sequences, corresponding to peptides I602−M610 and V409−R422. Quantitative analysis of the sequence data (Figure 6A) revealed that the PTH derivative of lysine, expected at cycle 7 of the sequence of peptide I602−M610, was not detected, clearly indicating that this residue was involved in the cross-link. Consistent with this conclusion, the only other lysine residue present in the two sequences, occurring at position 420 of peptide V409−R422, was recovered with the expected yield. Of the three acidic amino acids present in peptide V409−R422 (Glu 415, Glu 418, Glu 419), Glu 418 was recovered with the lowest yield (Figure 6A). The decrease was slight but significant, as it was observed consistently upon repeated sequence analyses performed on peptide material obtained from two separate experiments. Sequence analysis of the free form of peptide V409−R422, isolated from fraction 2 (Figure 6B), allowed direct comparison of the yields of Glu 418 in both cases, indicating that the relative decrease was indeed 40−50%. Analysis of the material contained in fraction 3 by both FAB and electrospray ionization mass spectrometry gave evidence for the presence of a species of molecular mass 2701.5 ± 1.0 (about one-third of the total peptide material), corresponding to peptides I602−M610 and V409−R422 cross-

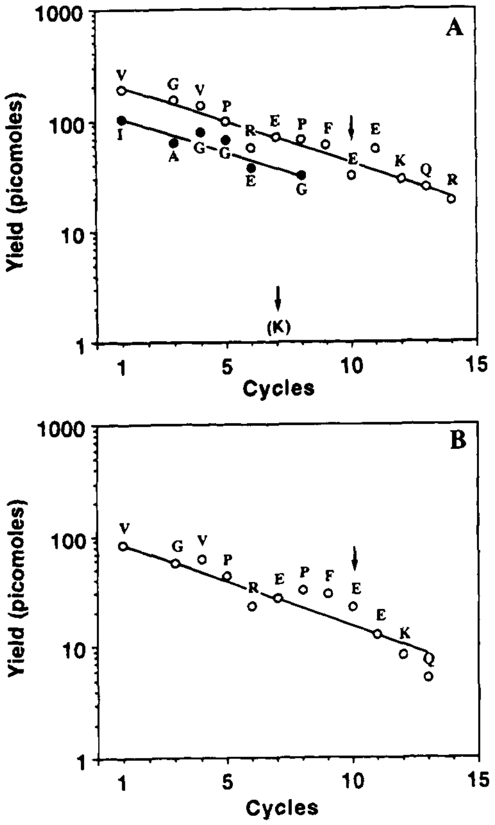


FIGURE 6: Quantitative analysis of the Edman degradation of the cross-linked peptides V409−R422/I602−M610. (A) Sequence analysis of the cross-linked material isolated from the thermolytic cleavage fraction 3, showing the sequences of both peptides V409−R422 (○) and I602−M610 (●). (B) Comparative sequence analysis of the free form of peptide V409−R422, isolated from the thermolytic cleavage fraction 2. The PTH derivatives of *S*-(pyridylethyl)cysteine, expected at cycle 2 of both peptides V409−R422 and I602−M610, and of homoserine, expected at cycle 9 of peptide I602−M610, were effectively identified but not quantified. Arrows indicate residues Lys 608 and Glu 418 involved in the cross-link.

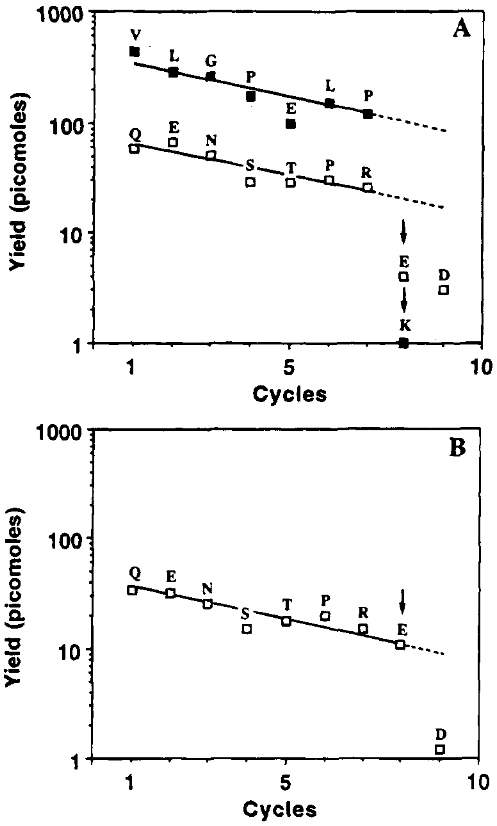


FIGURE 7: Quantitative analysis of the Edman degradation of the cross-linked peptides V398−P408/Q665−D673. (A) Sequence analysis of the cross-linked material isolated from fraction 4 of Figures 5, showing the sequences of both peptides V398−P408 (■) and Q665−D673 (□). (B) Comparative sequence analysis of the synthetic peptide Q665−D673. The PTH derivative of *S*-(pyridylethyl)cysteine, expected at cycle 9 of peptide V398−P408, was effectively identified but not quantified. Arrows indicate residues Lys 405 and Glu 672 involved in the cross-link.

linked through a pseudopeptide bond (expected mass of the species containing an homoserine at position 610 = 2701.2). Two other species corresponding to the free form of peptide V409−R422, either unmodified (observed mass = 1778.5 ± 0.5; expected mass = 1779.1) or EDC-modified (observed mass = 1935.1 ± 0.2; expected mass = 1935.1), each representing about one-third of the total peptide material, were also observed. Consistent with this result, the sequence analysis shown in Figure 6A also indicated the presence, beside the cross-linked species, of the free form of peptide V409−R422, estimated at about 50% of the peptide material. It became clear that one-half to two-thirds of the material submitted to sequence analysis consisted of the free form of peptide V409−R422, providing an explanation for the relative abundance of Glu 418. It was concluded, therefore, that peptides I602−M610 and V409−R422 were cross-linked between the ϵ-amino group of Lys 608 and the γ-carboxyl group of Glu 418, as illustrated in Figure 5.

Again, Edman degradation of the following thermolytic cleavage fraction 4 yielded two sequences, corresponding to peptides V398−P408 and Q665−D673 (see Figure 5). Quantitative analysis of the sequence data (Figure 7A) showed that the overall yield of peptide V398−P408 was much higher than that of peptide Q665−D673, suggesting partial cyclization of the N-terminal glutamine 665 to pyroglutamic acid. Only trace amounts of the PTH derivative of lysine were detected at cycle 8 of peptide V398−P408, indicating that Lys 405, the only lysine residue present in the two sequences, was involved in the cross-link. In the case of peptide Q665−D673, two acidic amino acids, Glu 672 and the C-terminal Asp 673, were recovered with low yields. With a view to discriminate between these residues, sequence analysis was performed on an equivalent amount of the corresponding synthetic peptide (Figure 7B). Comparison of the two analyses indicated that the yield of Glu 672 was decreased by about 85% in the cross-linked sample, whereas that of Asp 673 was equivalent in both cases. It was concluded, therefore, that peptides V398−P408 and Q665−D673 were linked to each other through a pseudopeptide bond involving Lys 405 and Glu 672. Further analysis of fraction 4 by electrospray ionization mass spectrometry clearly showed that peptides V398−P408 and Q665−D673 were cross-linked. Indeed, as summarized in Table 1, several species were detected, corresponding to the unmodified cross-linked species and to modified forms arising from either modification with EDC or cyclization of glutamine 665 to pyroglutamic acid or from both modifications. Significant amounts of the unmodified free form of peptide V398−P408 (observed mass = 1256.8 ± 0.1; expected mass = 1256.6) were also observed, indicating that, as found above for the other site, cross-linking was incomplete.

Homology Modeling of the Protein Modules of the C$\overline{1}$s γ-B Region. Bovine chymotrypsin was chosen as the primary template for modeling the catalytic B chain of C$\overline{1}$s.

P1

Table 1: Identification by Electrospray Ionization Mass Spectrometry of the Cross-Linked Peptides V398−P408/Q665−D673

species	exptl mass	calcd mass
cross-linked peptides V398−P408/Q665−D673	2313.2 ± 0.2	2313.6
cross-linked peptides V398−P408/Q665−D673 + 1 EDC	2469.6 ± 0.5	2469.6
cross-linked peptides V398−P408/Q665−D673 with pyroglutamic acid	2296.4 ± 0.5	2296.6
cross-linked peptides V398−P408/Q665−D673 with pyroglutamic acid + 1 EDC	2452.5 ± 0.8	2452.6

The sequence alignment used as a guideline was similar to that defined by Greer (1990). The two proteases show 33% sequence identity, and as much as 74% of their residues were considered to be structurally conserved, i.e., superimposed within the limit of 1.5 Å. Compared with chymotrypsin, the C$\bar{1}$s B chain exhibits three deletions at positions 441−442 (four residues), 560−561 (two residues), and 645−646 (three residues) and six minor insertions (Arg 466, Arg 481, Thr 538−Ser 539, Thr 598, Lys 608, Asn 627−Lys 629). There are three major modifications consisting of seven- and eight-residue insertions (Val 504−Asn 510 and Val 583−Ala 590) occurring in two loops, and a five-residue extension (Thr 669−Asp 673) at the C-terminal end of the chain. As expected, the core of the protease, containing the catalytic triad, is highly conserved, whereas all of the modifications are located on the surface. Thus, the two extended loops (Val 504−Asn 510 and Val 583−Ala 590) are located on one side of the entrance of the active site cleft (see Figure 9). Interestingly, the three deletions all occur within the same area, also in the vicinity of the active site cleft entrance. It should be noticed that the C-terminal extension (Thr 669−Asp 673) was constructed as the continuity of the C-terminal α-helix of chymotrypsin (see Figure 9). These three major modifications are the less constrained and therefore the less reliable parts of the model, due to the lack of good templates. From a functional point of view, it should be emphasized that the model is consistent with a classical activation mechanism involving formation of a salt bridge between the N-terminal Ile 423 and Asp 616 and that the S1 subsite (Asp 611), giving trypsin-like specificity, is located in a well-conserved region.

The 15-residue intermediary segment (Pro 408−Arg 422) which is C-terminal to the γ region in C$\bar{1}$s, and links CCP module V and the B chain moiety in proenzyme C1s (Figure 3), was constructed by using the activation peptide of bovine chymotrypsinogen, also comprising 15 residues, as a template structure. The reason for this choice is that the C-terminal end of this segment is structurally better defined in chymotrypsinogen than in chymotrypsin. As in chymotrypsinogen, this segment is covalently linked in C1s to the serine protease domain through a disulfide bridge involving Cys 410 and Cys 534 (Figure 3). As described below, this disulfide bridge is one of the constraints which were used to fit the intermediary segment to the B chain. In this respect, it should be mentioned that Cys 534 of C1s and its counterpart in chymotrypsinogen (Cys 122) are found to be exactly superimposed. In contrast, there is a local displacement between Cys 410 of C1s and the corresponding residue (Cys 1) of chymotrypsinogen, due to their respective location at positions 3 and 1 in the two homologous segments. However, the Cys 410−Cys 534 Cα−Cα distance in the C1s model appears to be compatible with the formation of a disulfide bridge.

A problem for modeling the CCP modules IV and V of C1s arose from the restricted number of homologous structures available, namely, those of the 5th, 15th, and 16th modules from human factor H (Norman et al., 1991; Barlow et al., 1992, 1993). The sequences of these three modules were aligned with the aid of the hydrophobic cluster analysis (HCA) method (Gaboriaud et al., 1987), on the basis of the amino acids considered to form the hydrophobic core, as defined from the three-dimensional structure of modules H5 and H16 (Norman et al., 1991; Barlow et al., 1992). Modeling of modules IV and V of C1s was realized on the basis of the structures of modules H16 and H5, respectively, using the alignments shown in Figure 8. The overall sequence homology (including identities and conservative replacements) is 32% for the C1sIV/H16 pair and 29% for the C1sV/H5 pair. In contrast, if one only takes into account the core residues, as defined in Figure 8, the homology increases to 57% and 77%, respectively, clearly indicating that both pairs of modules have a common framework. Indeed, the resulting three-dimensional models of modules IV and V of C1s both show a five-β-stranded globular structure typical of the CCP modules, with some modifications, located in the C-terminal region of the modules and mostly occurring at the surface. Thus, compared to H16, C1sIV exhibits a single one-residue deletion, two minor insertions (Asn 329, Asn 334), and a six-residue insertion (Val 313−Gly 318) forming a bulge within the third β-strand (Figure 8). In the same way, compared to H5, C1sV exhibits a one-residue deletion, three minor insertions (Glu 348, Gly 381−Gly 382, and Asn 391, which is the N-glycosylation site), and a six-residue insertion (Val 398−Leu 403), extending a preexisting bulge within the fifth β-strand (Figures 8 and 9).

Assembly of the C-Terminal Part of the C$\bar{1}$s γ-B Region. Since cross-linking experiments provided no evidence for an interaction between the N-terminal module IV and the remainder of the γ-B region, assembly was restricted to module V, the intermediary segment, and the B chain. As mentioned above, the intermediary segment (Pro 408−Arg 422) was initially fitted to the surface of the B chain by means of the Cys 410−Cys 534 disulfide bridge, using the same orientation as that of the activation peptide in chymotrypsinogen. In a second step, module V was connected to the intermediary segment through the Val 407−Pro 408 peptide bond. At this stage of the assembly process, the three peptides involved in the cross-linking (E379−R422, I602−M610, Q665−D673) were known, but the cross-linking sites had not been identified. Nevertheless, it was immediately obvious that segments I602−M610 and Q665−D673 were located at two opposite faces of the B chain, each in the vicinity of one of the extremities of segment E379−R422. Indeed, there was a strong indication that one of the cross-links was between I602−M610 (B chain) and the C-terminal end of E379−R422 (intermediary segment). As a consequence, the other cross-link was very likely between Q665−D673 (C-terminal extension of the B chain) and the N-terminal part of E379−R422 (module V). Identification of the cross-linking sites confirmed these hypotheses. Thus, the cross-link between Glu 418 and Lys 608 was consistent with the model (Figure 9), allowing a more precise orientation of the C-terminal end of the intermediary

P1

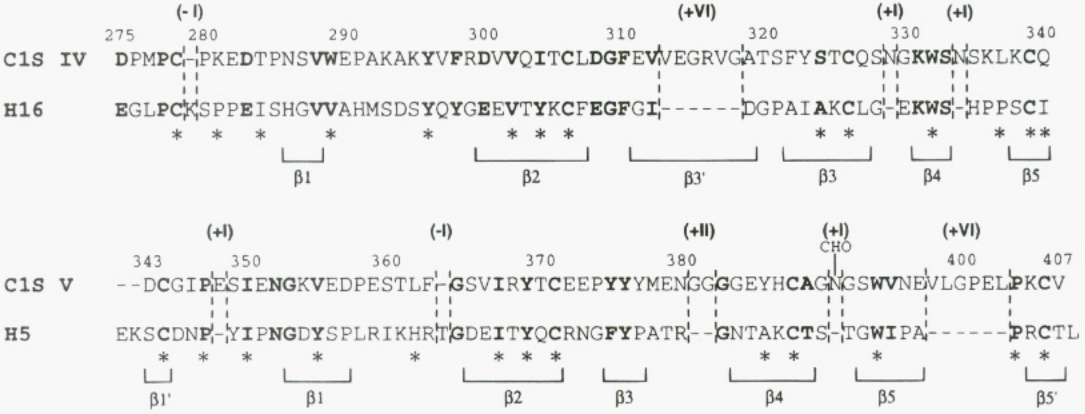

(- I)
(+VI)
(+I)
(+I)
275 280 290 300 310 320 330 340
C1S IV DPMPC-PKEDTPNSVWEPAKAKYVFRDVVQITCLDGFEVVEGRVGATSFYSTCQSNGKWSNSKLKCQ
H16 EGLPCKSPPEISHGVVAHMSDSYQYGEEVTYKCFEGFGI------DGPAIAKCLG-EKWS-HPPSCI
β1
β2
β3'
β3
β4
β5
(+I)
(-I)
(+II)
(+I)
CHO
(+VI)
343 350 360 370 380 400 407
C1S V --DCGIPESIENGKVEDPESTLF-GSVIRYTCEEPYYYMENGGGGEYHCAGNGSWVNEVLGPELPKCV
H5 EKSCDNP-YIPNGDYSPLRIKHRTGDEITYQCRNGFYPATR--GNTAKCTS-TGWIPA------PRCTL
β1'
β1
β2
β3
β4
β5
β5'

FIGURE 8: Sequence alignments used for modeling the CCP modules IV and V of C1s. The amino acid numbering used is that of C1s. Identical and conservatively replaced residues are shown in bold letters. $\beta1-\beta5$ are the β-strands of the 16th and 5th CCP modules of factor H, and residues considered to form the hydrophobic core of these modules are indicated by stars (Norman et al., 1991; Barlow et al., 1992). Insertions and deletions found in the CCP modules of C1s are shown by Roman numerals in parentheses. CHO indicates the N-linked oligosaccharide.

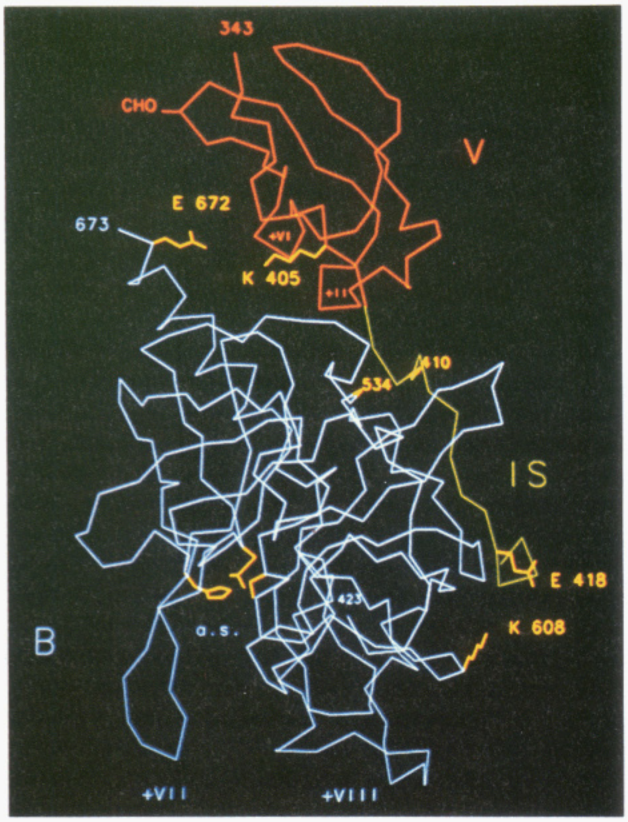

343
CHO
V
E 672
673
+VI
K 405
+II
534
410
IS
E 418
423
a.s.
K 608
B
+VII
+VIII

FIGURE 9: Three-dimensional Cα model of the C-terminal part of the catalytic region of C$\bar{1}$s. The CCP module V, the intermediary segment (IS), and the B chain are shown in orange, green, and blue, respectively. The side chains of the amino acids involved in the cross-links and of the active site amino acids (His 460, Asp 514, Ser 617), as well as the α carbons of the half-cystine residues Cys 410 and Cys 534, are shown in yellow. The major insertions of module V (Val 398–Leu 403 = +VI; Gly 381–Gly 382 = +II) and of the B chain (Val 504–Asn 510 = +VII; Val 583–Ala 590 = +VIII) are indicated. Abbreviations: CHO, N-linked oligosaccharide; a.s., active site.

segment, a region known to be flexible in active serine proteases. Identification of the other cross-link between Lys 405 and Glu 672 provided a distance constraint for the relative positioning of module V and the C-terminal extension of the B chain. In this case, some uncertainty arises from the fact that, contrary to Lys 405, located in a conserved area of module V (Figure 8), Glu 672 lies in the C-terminal extension of the B chain, arbitrarily modeled as the continuity of the preceding α-helix. Despite their relative inaccuracy, the constraints brought about by the two cross-links and the Cys 410–Cys 534 disulfide bridge allow the relative positioning of module V, the intermediary segment, and the catalytic B chain. They also provide evidence that module V interacts with the B chain on the side opposite to the active site cleft (Figure 9).

DISCUSSION

Chemical cross-linking with EDC is a frequently used method to identify ionic bonds involved in the interaction between two proteins or between individual domains within a protein. An advantage of this zero-length cross-linking agent is that any accessible carboxyl group can be activated and that, provided that it is close to the ϵ-amino group of a lysine residue, a cross-link may occur, thereby indicating that these groups form a salt bridge in the native protein(s). However, hydrophobic interactions, or even ionic interactions involving arginine and aspartic or glutamic acid, are not detected by this technique. The experiments reported in the present study also indicate that "abortive" reactions of EDC with carboxyl groups may occur, as shown by the identification, using mass spectrometry, of EDC-modified forms of peptides P278–M378 and V409–R422 and of the cross-linked peptides V398–P408/Q665–D673. Such reactions complicate analysis of the data, particularly peptide mapping and mass spectrometry analysis, and may, in some cases, interfere with identification of the cross-linked amino acids by N-terminal sequence analysis. Unambiguous identification of the amino acids involved in each cross-link could nevertheless be achieved from comparative sequence analysis of the corresponding unmodified peptide. Despite the limitations of the cross-linking technique, it should be emphasized that the experiments reported in this study were performed twice and that identical results were obtained in each case. As judged from SDS–PAGE analysis of the cross-linking between the γ fragment and the B chain, the overall yield of the reaction was 50–60%. It should be considered, however, that either of the two cross-links identified in this study (Lys 405–Glu 672 and Glu 418–Lys 608) is expected to yield a cross-linked γ-B molecule. It is likely, therefore, that the cross-linked γ-B species isolated by high-pressure gel permeation and hence the material contained in pool X1A each represent a mixture of species containing either one or the other cross-link, or both of them. Indeed, further analysis of the two cross-links after thermolytic digestion gives evidence, in both cases, for the presence of non-cross-linked peptides, indicating that both cross-linking reactions were incomplete.

Three-dimensional homology modeling of the serine protease domain of C$\bar{1}$s was realized according to the method defined by Greer (1990), using bovine chymotrypsin as a primary template. Although the resulting model is comparable to that obtained previously by Carter et al. (1984), there are noticeable differences with respect to the lengths and the precise boundaries of the major insertions, in particular those corresponding to the two extended loops located in the vicinity of the entrance of the active site cleft. The main reason for these discrepancies probably lies in the fact that the alignment used in the present study is based on a set of reference structures, whereas that used by Carter et al. was based on chymotrypsin alone. Despite the restricted number of reference structures available for CCP modules, the models obtained for modules IV and V of C1s, based respectively on the coordinates of the 16th and 5th modules of factor H, may be considered satisfactory, owing to the particularly high homology shown by each of these pairs of modules. Indeed, compared with the template structures, the models derived for modules IV and V each exhibit a single major insertion and likely provide, therefore, a reliable basis for the major part of their structure.

The information provided by both the chemical cross-linking experiments and the modeling studies allowed us to construct a model of the C-terminal part of the C$\bar{1}$s γ-B region, comprising module V, the 15-residue intermediary segment, and the serine protease B chain. In this regard, it should be emphasized that both approaches not only were complementary but also were consistent with each other. The three protein elements of the assembly are held together by means of three major constraints, of which two (the cross-link between Lys 405 and Glu 672 and the Cys 410−Cys 534 disulfide bridge) allow the relative positioning of module V and the B chain, indicating that these domains are likely closely associated to each other. Indeed, other interactions, not detectable by the cross-linking technique, could take place. For example, additional contacts may be provided by the six- and two-residue insertions of module V which, according to our model, form two bends possibly protruding toward the B domain (Figure 9). The occurrence of a tight association between module V and the B chain would be consistent with the neutron scattering studies performed on the γ-B fragment of C$\bar{1}$s (Zaccaï et al., 1990). It is also compatible with the calorimetric studies reported by Medved et al. (1989), indicating that modules IV and V of C1s unfold independently, at about the same temperature (60 °C), both in the intact protein and in the isolated γ fragment. Indeed, if one takes into account that the B chain melts at a lower temperature (49 °C), the proposed interaction between module V and the B chain cannot be expected to modify extensively the thermal stability of module V. This association between module V and the B chain represents the first example of an interaction between a CCP module and a serine protease domain. It is very likely also a characteristic of human C$\bar{1}$r, given the fact that this protease homologous to C$\bar{1}$s also contains the disulfide bridge (Cys 434−Cys 560) connecting the intermediary segment to the B chain and that EDC treatment yields evidence for a cross-link between Lys 426 of module V and the C-terminal Asp 688 of the B chain, homologous to the Lys 405−Glu 672 cross-link identified in C$\bar{1}$s (M. B. Lacroix, J. Ulrich, N. Scherrer, J. Gagnon, N. M. Thielens, and G. J. Arlaud, unpublished experiments).

The fact that module V of C$\bar{1}$s interacts with the B chain on the side opposite to both the activation site and the catalytic site (Figure 9) allows unimpeded access of the active site of C$\bar{1}$r to the Arg 422−Ile 423 bond cleaved upon activation, as well as of the active site of C$\bar{1}$s to its substrates C4 and C2. With respect to the precise functional role of module V, a likely hypothesis is that it participates in the recognition of one or both of these relatively large substrates (respective M_r 205 000 and 102 000) and allows their correct positioning relative to the active site. Indeed, previous studies based on the use of monoclonal antibodies (Matsumoto & Nagaki, 1986; Matsumoto et al., 1989) give evidence for the presence of a C4 binding site (presumably recognizing the C4b moiety) within the γ region of C$\bar{1}$s. Our model strongly suggests that this site is located within module V and provides a structural basis to test this hypothesis, e.g., by measuring the ability of peptides corresponding to selected portions of module V to inhibit C4 cleavage by C$\bar{1}$s. Other studies (Bing et al., 1978) were based on affinity labeling of C$\bar{1}$s by *m*-[[*o*-(2-chloro-5-(fluorosulfonyl)phenylureido)-phenoxy]butoxy]benzamidine, a reagent designed to bind reversibly through its benzamidine moiety to the S1 subsite (Asp 611) and to form covalent bonds, through its sulfonyl fluoride group, with residues adjacent to the binding site. As the label was found to be distributed on both the A and B chains, it was concluded that portions of the A chain are close to the active site. If one considers the length of the reagent, as calculated by the authors (17 Å), and the distance between Asp 611 and the proximal residues of module V, as estimated from our model (31 Å), it is highly unlikely that the observed labeling occurred in module V. In fact, according to the model shown in Figure 9, the only plausible hypothesis would be a reaction with the closest portion of the intermediary segment.

It is noteworthy that the cross-linking experiments performed in this study yielded no evidence for an interaction between CCP module IV and any of the other domains of the γ-B region. For the reasons mentioned above, this does not prove that there is no such interaction. However, this observation should be considered in light of the NMR structural studies performed on a pair of CCP modules (the 15th and 16th modules of human factor H), indicating the presence of a flexible "hinge" allowing a certain degree of freedom in the relative orientation of the modules (Barlow et al., 1993). If this feature also applies to C$\bar{1}$s, then module IV is expected to adopt a variety of orientations with respect to the remainder of the γ-B region, and therefore it is unlikely that cross-links will be detected. According to current models of the C1 complex (Schumaker et al., 1986; Weiss et al., 1986; Arlaud et al., 1987), activation of the catalytic region of C1s requires a contact with the corresponding regions of C$\bar{1}$r, located **inside** the complex, whereas, once activated, these same regions of C1s are expected to move **outside** the complex, in order to gain access to their substrates C4 and C2. It is tempting to speculate that the occurrence of a flexible hinge between modules IV and V of C1s is responsible for this movement.

ACKNOWLEDGMENT

We are indebted to Prof. I. D. Campbell and Dr. P. N. Barlow, who kindly sent us the coordinates of the 5th, 15th, and 16th CCP modules of human factor H. We thank Y. Petillot for mass spectrometry analysis and I. Bally for peptide synthesis.

REFERENCES

Abola, E. E., Bernstein, F. C., Bryant, S. H., Koetzle, T. F., & Weng, J. (1987) in *Crystallographic Database—Information Content, Software Systems, Scientific Applications* (Allen, F. H., Bergerhoff, G., & Sievers, R., Eds.) pp 107−132, Data Commission of the International Union of Crystallography, Bonn/Cambridge/Chester.

Arlaud, G. J., Sim, R. B., Duplaa, A. M., & Colomb, M. G. (1979) *Mol. Immunol. 16*, 445−450.

Arlaud, G. J., Colomb, M. G., & Gagnon, J. (1987) *Immunol. Today 8*, 106−111.

Barany, G., & Merrifield, R. B. (1980) in *The Peptides* (Gross, E., & Meienhoffer, J., Eds.) Vol. 2, pp 1−284, Academic Press, New York.

Barlow, P. N., Steinkasserer, A., Horne, T. J.; Pearce, J., Driscoll, P. C., Sim, R. B., & Campbell, I. D. (1992) *Biochemistry 31*, 3626−3634.

Barlow, P. N., Steinkasserer, A., Norman, D. G., Kieffer, B., Wiles, A. P., Sim, R. B., & Campbell, I. D. (1993) *J. Mol. Biol. 232*, 268−284.

Bartunik, H. D., Summers, L. J., & Bartsch, H. H. (1989) *J. Mol. Biol. 210*, 813−828.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shomanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−542.

Bing, D. H., Laura, R., Andrews, J. M., & Cory, M. (1978) *Biochemistry 17*, 5713−5718.

Birktoft, J. J., & Blow, D. M. (1972) *J. Mol. Biol. 68*, 187−240.

Bode, W., Chen, Z., Bartels, K., Kutzbach, C., Schmidt, G., & Bartunik, H. (1983) *J. Mol. Biol. 164*, 237−282.

Carter, P. E., Dunbar, B., & Fothergill, J. E. (1983) *Biochem. J. 215*, 565−571.

Carter, P. E., Dunbar, B., & Fothergill, J. E. (1984) *Philos. Trans. R. Soc. London, B 306*, 293−299.

Cooper, N. R. (1985) *Adv. Immunol. 37*, 151−216.

Fujinaga, M., & James, M. N. G. (1987) *J. Mol. Biol. 195*, 373−396.

Gaboriaud, C., Bissery, V., Benchetrit, T., & Mornon, J. P. (1987) *FEBS Lett. 224*, 149−155.

Greer, J. (1990) *Proteins 7*, 317−334.

Hess, D., Schaller, J., & Rickli, E. E. (1991) *Biochemistry 30*, 2827−2833.

Jones, T. A., & Thirup, S. (1986) *EMBO J. 5*, 819−822.

Jones, T. A., Zou, J. Y., Lowan, S. W., & Kjeldgaard, M. (1991) *Acta Crystallogr., Sect. A 47*, 110−119.

Laemmli, U. K. (1970) *Nature 227*, 680−685.

Mackinnon, C. M., Carter, P. E., Smyth, S. J., Dunbar, B., & Fothergill, J. E. (1987) *Eur. J. Biochem. 169*, 547−553.

Matsudaira, P. (1987) *J. Biol. Chem. 262*, 10035−10038.

Matsumoto, M., & Nagaki, K. (1986) *J. Immunol. 137*, 2907−2912.

Matsumoto, M., Nagaki, K., Kitamura, H., Kuramitsu, S., Nagasawa, S., & Seya, T. (1989) *J. Immunol. 142*, 2743−2750.

Means, G. E., & Feeney, R. E. (1971) in *Chemical Modification of Proteins*, pp 144−148, Holden-Day, Inc., San Francisco.

Medved, L. V., Busby, T. F., & Ingham, K. C. (1989) *Biochemistry 28*, 5408−5414.

Navia, M. A., Mc Keever, B. M., Springer, J. P., Lin, T.-Y., Williams, H. R., Fluder, E. M., Dorn, C. P., & Hoogsteen, K. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7−11.

Norman, D. G., Barlow, P. N., Baron, M., Day, A. J., Sim, R. B., & Campbell, I. D. (1991) *J. Mol. Biol. 219*, 717−725.

Read, R. J., & James, M. N. G. (1988) *J. Mol. Biol. 200*, 523−551.

Reid, K. B. M., Bentley, D. R., Campbell, R. D., Chung, L. P., Sim, R. B., Kristensen, T., & Tack, B. F. (1986) *Immunol. Today 7*, 230−234.

Remington, S. J., Woodbury, R. G., Reynolds, R. A., Matthews, B. W., & Neurath, H. (1988) *Biochemistry 27*, 8097−8105.

Sawyer, L., Shotton, D. M., Campbell, J. W., Wendell, P. L., Muirhead, H., Watson, H. C., Diamond, R., & Ladner, R. C. (1978) *J. Mol. Biol. 118*, 137−208.

Schumaker, V. N., Hanson, D. C., Kilchherr, E., Phillips, M. L., & Poon, P. H. (1986) *Mol. Immunol. 23*, 556−565.

Schumaker, V. N., Zavodszky, P., & Poon, P. H. (1987) *Annu. Rev. Immunol. 5*, 21−42.

Spycher, S. E., Nick, H., & Rickli, E. E. (1986) *Eur. J. Biochem. 156*, 49−57.

Tam, J. P., Heath, W. F., & Merrifield, R. B. (1983) *Int. J. Pept. Protein Res. 23*, 2939−2942.

Thielens, N. M., Van Dorsselaer, A., Gagnon, J., & Arlaud, G. J. (1990) *Biochemistry 29*, 3570−3578.

Tosi, M., Duponchel, C., Meo, T., & Julier, C. (1987) *Biochemistry 26*, 8516−8524.

Villiers, C. L., Arlaud, G. J., & Colomb, M. G. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 4477−4481.

Wang, D., Bode, W., & Huber, R. (1985) *J. Mol. Biol. 185*, 595−624.

Weiss, V., Fauser, C., & Engel, J. (1986) *J. Mol. Biol. 189*, 573−581.

Zaccaï, G., Aude, C. A., Thielens, N. M., & Arlaud, G. J. (1990) *FEBS Lett. 269*, 19−22.

BI9427781

P2

FEBS 15042 FEBS Letters 358 (1995) 323–328

Analysis of the N-linked oligosaccharides of human C1s using electrospray ionisation mass spectrometry

Yves Petillot[a], Pierre Thibault[b], Nicole M. Thielens[a], Véronique Rossi[a], Monique Lacroix[a], Bernadette Coddeville[c], Geneviève Spik[c], Verne N. Schumaker[d], Jean Gagnon[a], Gérard J. Arlaud[a],*

[a]*Laboratoire d'Enzymologie Moléculaire, Institut de Biologie Structurale Jean-Pierre Ebel (CEA/CNRS),*
41 avenue des Martyrs, 38027 Grenoble Cedex, France
[b]*Institute for Marine Biosciences, Halifax, Canada*
[c]*Laboratoire de Chimie Biologique, Université des Sciences et Technologies de Lille, Lille, France*
[d]*Department of Chemistry and Biochemistry, University of California, Los Angeles, CA, USA*

Received 30 November 1994; revised version received 14 December 1994

Abstract Information on the structures of the oligosaccharides linked to Asn residues 159 and 391 of the human complement protease $\overline{\mathrm{C1}}$s was obtained using mass spectrometric and monosaccharide analyses. Asn^{159} is linked to a complex-type biantennary, bisialylated oligosaccharide NeuAc2 Gal2 GlcNAc4 Man3 (molecular mass = 2206 ± 1). Asn^{391} is occupied by either a biantennary, bisialylated oligosaccharide, or a triantennary, trisialylated species NeuAc3 Gal3 GlcNAc5 Man3 (molecular mass = 2861 ± 1), or a fucosylated triantennary, trisialylated species NeuAc3 Gal3 GlcNAc5 Man3 Fuc1 (molecular mass = 3007 ± 1), in relative proportions of approximately 1:1:1. The carbohydrate heterogeneity at Asn^{391} gives rise to three major types of $\overline{\mathrm{C1}}$s molecules of molecular masses 79,318 ± 8 (A), 79,971 ± 8 (B), and 80,131 ± 8 (C), with an average mass of 79,807 ± 8. A minor modification, yielding an extra mass of 132 ± 2, is also detected within positions 1–153.

Key words: Complement; C1; N-Linked oligosaccharide; Electrospray ionisation mass spectrometry

---

1. Introduction

The first component of human complement, C1, is a complex enzyme comprising two serine proteases, C1r and C1s, that are sequentially activated upon binding of C1 to various activators (for reviews, see [1–3]). Proenzyme C1s, a single-chain protein, is converted through cleavage of a single Arg–Ile bond into $\overline{\mathrm{C1}}$s, the active protease responsible for the enzymic activity of $\overline{\mathrm{C1}}$, which contains two disulfide-linked chains A and B. The C-terminal B chain is a trypsin-like serine protease domain, whereas the N-terminal A chain is a mosaic-like polypeptide composed of five structural modules, including two pairs of internal repeats (I/III and IV/V) and one epidermal growth factor-like motif II [4,5]. The latter contains, at position 134, an Asn residue that undergoes partial (approx. 50%) β-hydroxylation [6]. The native $\overline{\mathrm{C1}}$s molecule can be split into two regions, the α region (comprising motifs I, II, and part of motif III) and the γ-B region (comprising motifs IV, V, and the catalytic B chain) [6,7]. The amino acid sequence of human C1s has been determined [4,5], indicating that the protein (673 amino acid residues) contains two N-glycosylation consensus sequences (-Asn159-Cys-Ser- and -Asn391-Gly-Ser-). Further studies [8] have shown that both positions are occupied by an oligosaccharide chain, but that only the carbohydrate linked to Asn^{159} is susceptible to cleavage by PNGase F. Since early carbohydrate analyses performed on intact $\overline{\mathrm{C1}}$s [9], no precise information on the nature of the two N-linked oligosaccharides has been obtained.

Electrospray ionisation MS has proved to be a valuable technique for precise mass determination of biopolymers [10,11]. The present work is based on the combined use of this technique and of monosaccharide analyses performed on $\overline{\mathrm{C1}}$s or various of its fragments and peptides containing one of the two oligosaccharides. Our data provide precise information on the composition of the oligosaccharides linked to Asn^{159} and Asn^{391}, indicating both to be of the complex glycotype, with the latter showing three major glycoforms. They also provide indirect evidence for a minor modification occurring within the N-terminal region of C1s.

2. Materials and methods

2.1. Materials

Human plasmin was obtained from Kabi Vitrum, Stockholm, Sweden. Trypsin (treated with 1-chloro-4-phenyl-3-(L-tosylamido)-butan-2-one) was from Sigma. PNGase F was purified from cultures of *Flavobacterium meningosepticum* according to the method of Tarentino et al. [12], modified as described in [8]. DFP was obtained from Sigma. ConA-Sepharose was purchased from Pharmacia, Uppsala, Sweden. α-Methyl-D-mannopyranoside was from Janssen Chimica (Noisy le Grand, France).

2.2. Proteins, fragments, and peptides

$\overline{\mathrm{C1}}$s was isolated from pooled human plasma as described previously [13], and its concentration was estimated from absorbance at 280 nm by using E(1%,1 cm) = 14.5 [7] and an M_r value of 79,800, derived from the present study. Recombinant human C1s was obtained from a baculovirus/insect cells expression system, and activated by incubation with human $\overline{\mathrm{C1}}$r as described previously [14]. The γ-B fragment from both human plasma $\overline{\mathrm{C1}}$s and recombinant $\overline{\mathrm{C1}}$s was obtained by limited proteolysis with plasmin, and the reaction was terminated in each case by blocking the active site serine residues of both γ-B and plasmin with DFP, as described in [6]. Both γ-B fragments were purified as described in [7]. The C1sα fragment was obtained by limited proteolysis with trypsin, and purified as described in [7]. Deglycosylation of C1sα with PNGase F was carried out as described previously [7]. Peptide C1sA 154–195 was obtained from CNBr cleavage of fragment C1s α2 [6].

---

*Corresponding author. Fax: (33) 76 88 54 94.
E-mail: arlaud@ibs.ibs.fr

Abbreviations: The nomenclature of complement components is that recommended by the World Health Organization; activated components are indicated by an overhead bar, e.g. $\overline{\mathrm{C1}}$s. ConA, concanavalin A; DFP, di-isopropyl phosphorofluoridate; MS, mass spectrometry; PNGase F, peptide *N*-glycosidase F.

0014-5793/95/$9.50 © 1995 Federation of European Biochemical Societies. All rights reserved.
SSDI 0014-5793(94)01429-9

Biochemistry **1997,** *36,* 6270−6282

P3

Structure and Assembly of the Catalytic Region of Human Complement Protease C$\bar{1}$r: A Three-Dimensional Model Based on Chemical Cross-Linking and Homology Modeling[†]

Monique Lacroix,[‡] Véronique Rossi,[‡] Christine Gaboriaud,[§] Sylvie Chevallier,[‡] Michel Jaquinod,[||]
Nicole M. Thielens,[‡] Jean Gagnon,[‡] and Gérard J. Arlaud*,[‡]

Laboratoire d'Enzymologie Moléculaire, Laboratoire de Cristallogénèse et Cristallographie des Protéines, and Laboratoire de Spectrométrie de Masse des Protéines, Institut de Biologie Structurale Jean-Pierre Ebel (CEA-CNRS), 41, avenue des Martyrs, 38027 Grenoble Cedex 1, France

Received October 30, 1996; Revised Manuscript Received February 18, 1997[⊗]

ABSTRACT: C1r is the modular serine protease responsible for autocatalytic activation of C1, the first component of the complement classical pathway. Its catalytic region is a noncovalent homodimer of two γ-B monomers, each comprising two contiguous complement control protein (CCP) modules, IV and V [also known as short consensus repeats (SCRs)], a 15-residue intermediary segment, and the serine protease B domain. With a view to gain insight into domain−domain interactions within this region, fragment C$\bar{1}$r (γ-B)$_2$, obtained by autolytic proteolysis of the active protease, was cross-linked with the water-soluble reagent 1-ethyl-3-[3-(dimethylamino)propyl]carbodiimide. Cross-linked species γ-B *intra* and γ-B *inter*, containing intra- and intermonomer cross-links, respectively, were isolated and then fragmented by CNBr cleavage and trypsin digestion. N-Terminal sequence and mass spectrometry analyses of the resulting cross-linked peptides allowed us to identify one intramonomer cross-link between Lys$_{426}$ of module V and the C-terminal Asp$_{688}$ of the serine protease B domain and one intermonomer cross-link between the N-terminal Gly$_{280}$ of fragment γ and Glu$_{493}$ of the B domain. Three-dimensional homology modeling of the CCP modules IV and V and of the B domain was also performed. The complementary information provided by chemical cross-linking and homology modeling studies was used to construct a three-dimensional model of the γ-B monomer, in which module V interacts with the serine protease on the side opposite to both the active site and the Arg$_{446}$−Ile$_{447}$ activation site. Also, a tentative three-dimensional model of the (γ-B)$_2$ dimer was built, indicating a loose "head to tail" association of the monomers, with the active sites facing opposite directions toward the outside of the dimer. The latter model is compared with available low-resolution structural data, and its functional implications are discussed in terms of the conformational changes occurring during C1r activation.

Human C1 is the complex modular protease that initiates the classical pathway of complement in response to the formation of immune complexes and to infection by various nonimmune activators, such as certain bacteria, parasites, and retroviruses, including HIV. Initial recognition of the target by C1 is mediated by C1q and triggers activation of its catalytic subunit C1s−C1r−C1r−C1s, a calcium-dependent assembly of two serine proteases.[1] This two-step activation process involves autolytic activation of C1r and then C$\bar{1}$r-mediated conversion of proenzyme C1s into the active enzyme responsible for limited proteolysis of C4 and C2, the protein substrates of C$\bar{1}$ [see reviews by Cooper (1985), Schumaker et al. (1987), and Arlaud et al. (1987a)].

Human C1r is a noncovalent homodimer of M_r 172 600, each monomer comprising a series of protein modules, including two CUB modules, a single EGF-like module, two CCP (or SCR) modules, and a serine protease domain. This same type of modular organization is also shared by C1s and by MASP, a protease which, in association with mannan binding lectin, forms a complex responsible for the initiation of the "lectin pathway" of complement activation (Takayama et al., 1994; Volanakis & Arlaud, 1997). Both C1r autoactivation and subsequent activation of C1s by C$\bar{1}$r are mediated by its catalytic (γ-B)$_2$ region (Lacroix et al., 1989), a noncovalent dimer that forms the core of the C1r−C1r dimer and C1s−C1r−C1r−C1s tetramer (Villiers et al., 1985). The various studies performed on C1r autoactivation are consistent with the occurrence of an intramolecular cross-

---

[†] This is Publication No. 424 from the Institut de Biologie Structurale Jean-Pierre Ebel. This work was supported by the Commissariat à l'Energie Atomique and the Centre National de la Recherche Scientifique. A preliminary report of this study was presented at the XVIth International Complement Workshop in Boston, MA, June 1996.

* To whom correspondence should be addressed.

[‡] Laboratoire d'Enzymologie Moléculaire.

[§] Laboratoire de Cristallogénèse et Cristallographie des Protéines.

[||] Laboratoire de Spectrométrie de Masse de Protéines.

[⊗] Abstract published in *Advance ACS Abstracts*, May 1, 1997.

[1] Abbreviations: CCP modules, complement control protein modules [also known as short consensus repeats (SCRs) or sushi modules]; CUB, *c*omplement C1r, C1s, *u*egf, *b*one morphogenetic protein; EDC, 1-ethyl-3-[3-(dimethylamino)propyl]carbodiimide; EDTA, ethylenediaminetetraacetic acid; HPLC, high-pressure liquid chromatography; MALDI, matrix-assisted laser desorption ionization; MASP, mannan binding lectin-associated serine protease; PTH, phenylthiohydantoin; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis. The nomenclature of complement components is that recommended by the World Health Organization; activated components are indicated by an overbar, e.g., C$\bar{1}$r. The nomenclature of protein modules is that defined by Bork and Bairoch (1995) following the International Workshop on Extracellular Protein Modules held in September 1994 in Margretetorp, Sweden.

S0006-2960(96)02719-5 CCC: $14.00 © 1997 American Chemical Society

P4

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 273, No. 2, Issue of January 9, pp. 1232–1239, 1998
Printed in U.S.A.

© 1998 by The American Society for Biochemistry and Molecular Biology, Inc.

Baculovirus-mediated Expression of Truncated Modular Fragments from the Catalytic Region of Human Complement Serine Protease C1s

EVIDENCE FOR THE INVOLVEMENT OF BOTH COMPLEMENT CONTROL PROTEIN MODULES IN THE RECOGNITION OF THE C4 PROTEIN SUBSTRATE*

(Received for publication, September 19, 1997)

Véronique Rossi‡§, Isabelle Bally‡, Nicole M. Thielens‡, Alfred F. Esser¶, and Gérard J. Arlaud‡‖

From the ‡Laboratoire d'Enzymologie Moléculaire, Institut de Biologie Structurale Jean-Pierre Ebel (CEA-CNRS), 41 avenue des Martyrs, 38027 Grenoble Cedex 1, France and the ¶University of Missouri-Kansas City School of Biological Sciences, Kansas City, Missouri 64110-2499

C1s is the modular serine protease responsible for cleavage of C4 and C2, the protein substrates of the first component of complement. Its catalytic region (γ-B) comprises two complement control protein (CCP) modules, a short activation peptide (ap), and a serine protease domain (SP). A baculovirus-mediated expression system was used to produce recombinant truncated fragments from this region, deleted either from the first CCP module (CCP$_2$-ap-SP) or from both CCP modules (ap-SP). The aglycosylated fragment CCP$_2$-ap-SP*ag* was also expressed by using tunicamycin. The fragments were produced at yields of 0.6–3 mg/liter of culture, isolated, and characterized chemically and then tested functionally by comparison with intact C1s and its proteolytic γ-B fragment. All recombinant fragments were expressed in a proenzyme form and cleaved by C$\bar{1}$r to generate active enzymes expressing esterolytic activity and reactivity toward C1 inhibitor comparable to those of intact C$\bar{1}$s. Likewise, the activated fragments γ-B, CCP$_2$-ap-SP, and ap-SP retained C$\bar{1}$s ability to cleave C2 in the fluid phase. In contrast, whereas fragment γ-B cleaved C4 as efficiently as C$\bar{1}$s, the C4-cleaving activity of CCP$_2$-ap-SP was greatly reduced (about 70-fold) and that of ap-SP was abolished. It is concluded that C4 cleavage involves substrate recognition sites located in both CCP modules of C$\bar{1}$s, whereas C2 cleavage is affected mainly by the serine protease domain. Evidence is also provided that the carbohydrate moiety linked to the second CCP module of C$\bar{1}$s has no significant effect on catalytic activity.

The first component of complement comprises two homologous, yet distinct modular serine proteases C1r[1] and C1s assembled in the form of a Ca^{2+}-dependent tetramer C1s-C1r-C1r-C1s. C1r is responsible for intrinsic C1 activation, a two-step process first involving autolytic C1r activation and then C$\bar{1}$r-mediated cleavage of proenzyme C1s. The active enzyme C$\bar{1}$s is a highly specific protease with trypsin-like specificity that mediates limited proteolysis of C4 and C2, the two protein substrates of C$\bar{1}$, thereby triggering activation of the classical pathway of complement in response to infection by various microorganisms (2–5). C4 cleavage occurs in the fluid phase and generates fragment C4b, which has the transient ability to bind covalently to the surface of the target. Cleavage of the second substrate C2, in contrast, takes place after prior formation of a C4b-C2 complex and yields C4b-C2a, the protease responsible for cleavage of complement protein C3 (5). This double proteolytic activity of C$\bar{1}$ is controlled by C1 inhibitor, a member of the serine protease inhibitor (serpin) family, which reacts stoichiometrically with both C$\bar{1}$r and C$\bar{1}$s to form covalent protease-inhibitor complexes, resulting in disruption of the C$\bar{1}$s-C$\bar{1}$r-C$\bar{1}$r-C$\bar{1}$s tetramer (2).

Human C1s is synthesized as a 673-residue single chain zymogen which, upon activation by C$\bar{1}$r, is cleaved between Arg^{422} and Ile^{423} to yield two disulfide-linked polypeptides, the N-terminal A chain comprising a series of five protein modules and the serine protease B domain (6–8). Limited proteolysis of C$\bar{1}$s with plasmin yields a C-terminal fragment γ-B, comprising two contiguous complement control protein (CCP) modules, a 15-residue intermediary segment homologous to the activation peptide in chymotrypsinogen, and the serine protease domain (9). The CCP modules (originally known as short consensus repeats) are structural motifs of about 60 residues homologous to those mostly found in various complement regulatory proteins known to interact with components C3b and/or C4b (10). The C-terminal CCP module of C1s bears a heterogeneous, complex-type *N*-linked oligosaccharide with both biantennary and triantennary species (11). Based on chemical cross-linking and three-dimensional homology modeling, it was shown recently that this module closely interacts with the serine protease domain on the side opposite to both the active site and the Arg^{422}-Ile^{423} bond cleaved upon activation (12).

From a functional point of view, the fact that the proteolytic fragment γ-B retains a functional active site (9) and studies based on the use of monoclonal antibodies indicating that the γ segment may be involved in C4 binding (13, 14) together sug-

---

* This work was supported in part by the Commissariat à l'Energie Atomique (CEA), the Centre National de la Recherche Scientifique, and National Institutes of Health Grant RO1-AI 19478 (to A. F. E.). This is publication number 493 from the Institut de Biologie Structurale Jean-Pierre Ebel. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Present address: University of Missouri-Kansas City School of Biological Sciences, 5100 Rockhill Rd., Kansas City, MO 64110-2499.

‖ To whom correspondence should be addressed. Tel.: (33) 4 76 88 49 81; Fax: (33) 4 76 54 94; E-mail: arlaud@ibs.ibs.fr.

[1] The abbreviations used are: the nomenclature of complement proteins is that recommended by the World Health Organization; activated components are indicated by an overbar, *e.g.* C$\bar{1}$s; the nomenclature of protein modules is that defined by Bork and Bairoch (1); CCP$_2$-ap-SP, ap-SP, recombinant fragments from the C1s catalytic region lacking the N-terminal CCP module and both CCP modules, respectively; CCP$_2$-ap-SP*ag*, aglycosylated fragment CCP$_2$-ap-SP; CCP, complement control protein; MALDI, matrix-assisted laser desorption ionization; PAGE, polyacrylamide gel electrophoresis.

This paper is available on line at http://www.jbc.org

Downloaded from http://www.jbc.org/ by guest on March 9, 2015

P4

gest that γ-B may represent the catalytic region of C1̄s. However, no precise information is available yet on the respective role of the serine protease domain and CCP modules in the various aspects of C1̄s catalytic activity. The present study provides identification of the domains of C1s involved in its activation by C1̄r, proteolytic activity toward C4 and C2, and reactivity toward C1 inhibitor, based on baculovirus-mediated expression of truncated recombinant fragments from the γ-B region.

EXPERIMENTAL PROCEDURES

Materials—Diisopropylphosphorofluoridate, neuraminidase (type X), and the synthetic substrate N^{α}-benzoyl-L-arginine ethyl ester were from Sigma. Human plasmin was obtained from Chromogenix AB (Mölndal, Sweden). Polyclonal anti-C1̄s antiserum was raised in rabbits according to standard procedures. Restriction enzymes were from Boehringer Mannheim. Vent$_R$ polymerase was from New England Biolabs. The pHC1s46 vector encoding human C1s (6) was kindly provided by Dr. Mario Tosi (Laboratoire d'Immunogénétique, Institut Pasteur, Paris). The pVT-Bac vector was kindly provided by Dr. Thierry Vernet (Institut de Biologie Structurale, Grenoble). Antibiotics and molecular biology reagents were from Appligen Oncor.

Proteins—Proenzyme C1s and activated C1r and C1s were purified from human plasma as described previously (15, 16). The C1̄s γ-B fragment was obtained by limited proteolysis with plasmin (17) and purified by ion-exchange chromatography on a Mono Q HR 5/5 column (Pharmacia Biotech Inc.) as described previously (12). The two glycoforms of the C1̄s γ-B fragment, containing either a biantennary or a triantennary oligosaccharide, were isolated by affinity chromatography on concanavalin A-Sepharose (Pharmacia) as described previously (11). The desialylated C1̄s γ-B fragment was obtained by treatment with neuraminidase (10 units/mg protein) for 5 h at 25 °C. Complement proteins C2, C4, and C1 inhibitor were isolated from human plasma by means of published procedures (18–20). For the determination of protein concentrations the following absorption coefficients (*A* (1%, 1 cm) at 280 nm) and molecular weights were used: C1r, 12.4 and 86,300 (21); C1s, 14.5 and 79,800 (21, 11); C1̄s γ-B, 18.3 and 47,520 (22, 11); C2, 10.0 and 100,000 (18); C4, 8.2 and 205,000 (23, 24); C1 inhibitor, 3.86 and 105,000 (25). The absorption coefficients (*A* (1%, 1 cm) at 280 nm) used for the recombinant fragments ap-SP (17.0) and CCP$_2$-ap-SP (16.4) were calculated from their amino acid composition by the method of Edelhoch (26), and their molecular weights were determined by mass spectrometry analysis (see "Results").

Construction of C1s Fragments CCP$_2$-ap-SP and ap-SP-containing Expression Plasmids—The following primers were obtained from Eurogentec (Seraing, Belgium) and used to amplify the desired human C1s cDNA sequences in a polymerase chain reaction using Vent$_R$ polymerase, according to established procedures: GAAGATCTTGACTGTGGCATTCCT (sense for CCP$_2$-ap-SP); GGAATTCTTAGTCCTCACGGGGGGT (antisense for CCP$_2$-ap-SP and ap-SP); and GAAGATCTTCCAGTCTGTGGAGTC (sense for ap-SP). The underlined sequences represent bases introduced either at the 5′ end of the C1s cDNA to create a *Bgl*II site for in-frame cloning into the pNT-Bac expression vector or at the 3′ end to create an *Eco*RI site. The pNT-Bac vector was constructed from the baculovirus expression vectors pVT-Bac (27) and pFast Bac1™ (Life Technologies Inc.). Briefly, the segment comprised between the *Eco*RV and *Kpn*I sites was excised from plasmid pVT-Bac and introduced between the *Sna*BI and *Kpn*I sites in pFast Bac1. The resulting vector contains the polyhedrin promoter and the multiple cloning site region of pVT-Bac (Fig. 1). It also encodes the melittin signal peptide previously used to enhance secretion of the papain precursor (27) and is therefore adapted to the production in a secreted form of fragments derived from the C-terminal end of proteins. The amplified DNAs were designated CCP$_2$-ap-SP, encoding residues 343–673 in mature human C1s, and ap-SP, encoding residues 408–673. These were purified and cloned into the pCR-Script Amp SK(+) intermediate vector (Stratagene) according to the manufacturer's instruction. Both fragments were excised with *Bgl*II and *Eco*RI and introduced into the *Bam*HI and *Eco*RI sites of the pNT-Bac vector by means of the T4 DNA ligase to create the plasmids pNT-Bac/CCP$_2$-ap-SP and pNT-Bac/ap-SP. Clones containing the insert were identified by restriction enzyme analysis and checked for the absence of mutations introduced by polymerase chain reaction by double-stranded DNA sequencing using the ^{T7}Sequencing™ kit (Pharmacia Biotech Inc.).

Cells and Viruses—The insect cells, *Spodoptera frugiperda* (Sf21), were provided by Dr. Jadwiga Chroboczek (Institut de Biologie Structurale, Grenoble). The cells were maintained in TC100 medium (Life Technologies, Inc.) containing 5% fetal calf serum (JRH Biosciences) supplemented with 50 IU/ml penicillin and 50 mg/ml streptomycin (Sigma). Recombinant baculoviruses were generated by using the Bac-to-Bac™ system (Life Technologies, Inc.) based on site-specific transposition of an expression cassette carried by the pNT-Bac-based recombinant plasmids into a baculovirus shuttle vector (bacmid) propagated in *Escherichia coli* (28). Transposition of the pNT-Bac/CCP$_2$-ap-SP and pNT-Bac/ap-SP plasmids into DH10Bac *E. coli* cells and selection of recombinant bacteria were performed as recommended by the manufacturer (Life Technologies, Inc.). The recombinant bacmids were purified using the Qiagen midiprep DNA purification system (Qiagen S.A., Courtaboeuf, France) and transfected into Sf21 cells with cellfectin in Sf900 II SFM medium (Life Technologies, Inc.) as recommended by the manufacturer. The transfection supernatant was decanted 4 days later, centrifuged, and stored at 4 °C until use. The recombinant viruses, designated *r*CCP$_2$-ap-SP and *r*ap-SP, were titered by virus plaque assay and amplified as described by King and Possee (29).

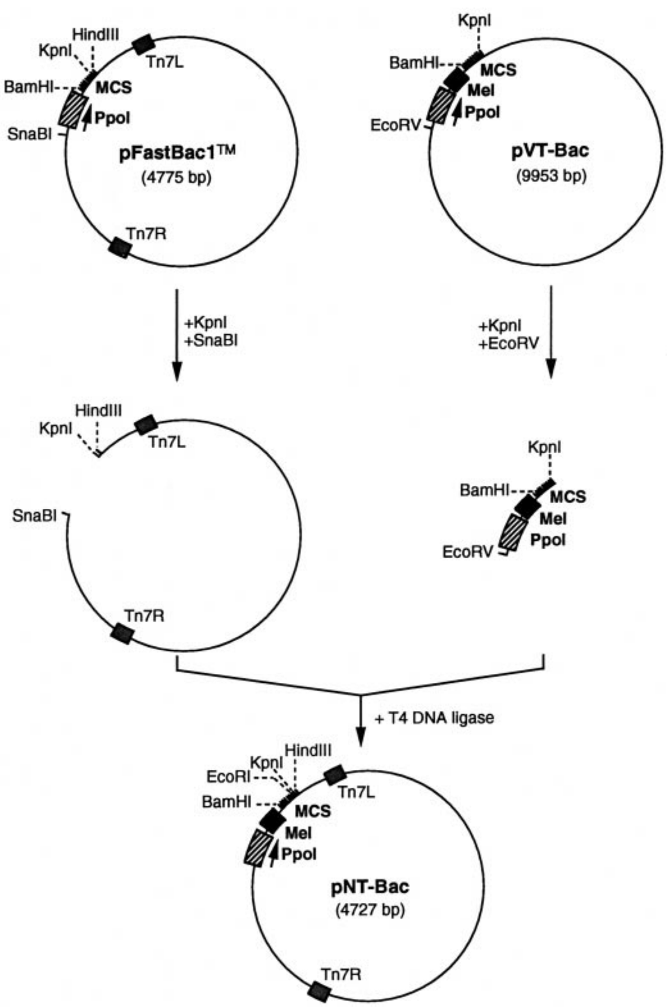


FIG. 1. **Schematic representation of the construction of the pNT-Bac vector.** The pNT-Bac vector was constructed from the baculovirus expression vectors pVT-Bac and pFast Bac1™, as described under "Experimental Procedures." *Mel*, melittin signal sequence; *MCS*, multiple cloning site region; *Ppol*, polyhedrin promoter. *Tn7L* and *Tn7R*, *left* and *right* hands of the bacterial transposon Tn7.

Protein Production and Purification—Sf21 cells (5×10^5 cells/ml in spinner flasks or 2×10^7 cells/175-cm^2 tissue culture flask) were infected with the recombinant viruses at a multiplicity of infection of 2–3 in Sf900 II SFM medium containing 50 IU/ml penicillin and 50 μg/ml streptomycin. Tunicamycin (5 μg/ml) was added for the production of the aglycosylated form of the CCP$_2$-ap-SP fragment. After 96 h at 28 °C, the supernatant was collected by centrifugation, and diisopropylphosphorofluoridate was added to a final concentration of 2 mM. The culture supernatant was dialyzed against 5 mM EDTA, 20 mM sodium phosphate, pH 8.6, and loaded at 2 ml/min onto a Mono Q HR 10/10 column (Pharmacia) equilibrated in the same buffer containing 1 mM diisopro-

Downloaded from http://www.jbc.org/ by guest on March 9, 2015

P4

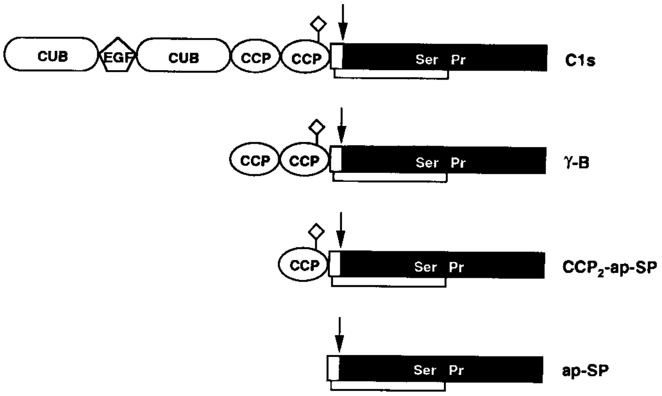


FIG. 2. **Comparative representation of the modular structures of human C1s and of the truncated fragments used in this study.** The nomenclature and symbols used for protein modules are those defined by Bork and Bairoch (1). *CUB*, module found in complement C1r/C1s, Uegf, and bone morphogenetic protein-1; *EGF*, epidermal growth factor-like module; *Ser Pr*, serine protease domain; the unlabeled segment corresponds to the activation peptide. The *arrow* indicates the Arg^{422}-Ile^{423} bond cleaved on activation. The only disulfide bridge shown is that connecting the activation peptide to the serine protease domain. ◇, oligosaccharide linked to Asn^{391}.

pylphosphorofluoridate. Elution was carried out with a linear NaCl gradient from 0 to 350 mM in 30 min. Fractions containing the recombinant proteins were identified by Western blot analysis and dialyzed against 1.5 M ammonium sulfate, 0.1 M sodium phosphate, pH 7.4. Further purification was achieved by high pressure hydrophobic interaction chromatography on a TSK-Phenyl 5PW column (Beckman) equilibrated in the same buffer containing 1 mM diisopropylphosphorofluoridate. Elution was carried out by decreasing the ammonium sulfate concentration from 1.5 M to 0 in 30 min at a flow rate of 1 ml/min. Both recombinant fragments were dialyzed against 50 mM triethanolamine HCl, 145 mM NaCl, pH 7.4, concentrated to 0.2–0.5 mg/ml by ultrafiltration on Microsep^{TM} microconcentrators (molecular weight cut-off = 10,000) (Filtron), and stored at −20 °C.

Chemical Characterization of the Recombinant Proteins—N-terminal sequence analyses were performed either on soluble proteins or after SDS-PAGE analysis and electrotransfer, using an Applied Biosystems model 477A protein sequencer as described previously (12). The matrix-assisted laser desorption ionization (MALDI) technique was used for mass spectrometry analysis of the recombinant proteins, using a Voyager Elite XL instrument (PerSeptive Biosystems, Cambridge, MA), under conditions described previously (30).

Functional Tests—Activation of proenzyme C1s and the recombinant fragments CCP$_2$-ap-SP and ap-SP was tested by incubation of the proteins (6.5 μM) in the presence of active C$\bar{1}$r (C$\bar{1}$r/protein molar ratio = 0.045) for various periods at 37 °C, in 50 mM triethanolamine HCl, 145 mM NaCl, pH 7.4. In the case of C1s and CCP$_2$-ap-SP, activation was monitored by SDS-PAGE analysis under reducing conditions and subsequent quantification by gel scanning of the two-chain structure characteristic of the active state of the proteins (see Fig. 2). Activation of the smaller fragment ap-SP was measured by its ability to cleave the synthetic substrate N^{α}-benzoyl-L-arginine ethyl ester.

The proteolytic activity of C$\bar{1}$s and its activated fragments γ-B, CCP$_2$-ap-SP, and ap-SP toward C2 and C4 was measured by incubation of the protein substrates in the presence of the enzymes for various periods at 37 °C followed by SDS-PAGE analysis under reducing conditions and quantification by gel scanning of the fragments characteristic of the C$\bar{1}$s-mediated cleavage (respectively, C2a and C2b, and the C4α' chain of the C4b fragment) (31, 32). For determination of the kinetic parameters, the C2 and C4 concentrations ranged from 0.5×10^{-6} M to 4×10^{-6} M, with fixed enzyme concentrations of 2×10^{-9} M and 1.2×10^{-9} M, respectively. All experiments were conducted in 20 mM sodium phosphate, 150 mM NaCl, 5 mM EDTA, pH 7.4, whereas enzyme dilutions were performed in 50 mM triethanolamine HCl, 145 mM NaCl, pH 7.4, containing 1 mg/ml ovalbumin.

The esterolytic activity of C$\bar{1}$s and its various activated fragments was measured against N^{α}-benzoyl-L-arginine ethyl ester, using an alcohol dehydrogenase/NAD-coupled spectrophotometric test, as described previously (33). The reactivity of the activated proteases toward C1 inhibitor was assessed by incubation in the presence of equimolar amounts of C1 inhibitor for 15 min at 37 °C in 50 mM triethanolamine HCl, 145 mM NaCl, pH 7.4, followed by SDS-PAGE analysis under reducing conditions of the formation of the covalent C1 inhibitor-serine protease B chain complexes (20).

Other Methods—SDS-PAGE analysis was performed as described previously (12). Western blot analysis of the recombinant proteins after SDS-PAGE was performed by electrotransfer to a nitrocellulose membrane and blocking of unoccupied sites with 5% non-fat dry milk in 10 mM Tris-HCl, 200 mM NaCl, pH 7.2. The immobilized proteins were visualized by using a rabbit polyclonal anti-C$\bar{1}$s antiserum, followed by a goat anti-rabbit immunoglobulin fraction conjugated to alkaline phosphatase (Sigma) and staining according to the manufacturer's instructions.

RESULTS

Expression and Purification of Recombinant Proteins—The modular structures of human C1s and of the various truncated fragments used in the present study are depicted in Fig. 2. Construction of the recombinant baculoviruses, transposition of the pNT-Bac/CCP$_2$-ap-SP and pNT-Bac/ap-SP plasmids into DH10Bac *E. coli* cells, and transfection of the corresponding recombinant bacmids into Sf21 cells were performed as described under "Experimental Procedures." The virus titer of the transfection supernatants after amplification was about 10^7 plaque forming units/ml in both cases. Infection of Sf21 cells by the supernatants was conducted for various periods at 28 °C, and the amount of recombinant proteins secreted into the culture medium was monitored by SDS-PAGE and Western blot analysis. Under normal culture conditions, protein bands with the expected apparent molecular weights (see below) became detectable at 48 h, their intensity increasing progressively to reach a plateau at 96 h. In contrast, secretion of the CCP$_2$-ap-SP fragment in the presence of tunicamycin reached a maximum at 72 h and decreased considerably after 96 h. Analysis of protein production in the cell pellets showed kinetic profiles similar to those observed for the supernatants. Under optimal production conditions, the amounts of recombinant protein secreted into the culture medium were estimated to be approximately 0.6 μg/ml for ap-SP, 3 μg/ml for CCP$_2$-ap-SP, and 1.5 μg/ml for the aglycosylated fragment CCP2-ap-SP (CCP$_2$-ap-SP*ag*) produced in the presence of tunicamycin.

The recombinant fragments were initially purified by fractionation of the culture supernatants by ion-exchange chromatography on a Mono Q column (Pharmacia). Both fragment CCP$_2$-ap-SP and CCP$_2$-ap-SP*ag* yielded visible peaks at 280 nm eluting during the ascending salt gradient, slightly later than the C$\bar{1}$s γ-B fragment produced by limited proteolysis. Fragment ap-SP, in contrast, was not retained on the column and eluted in the large flow-through fraction. Final purification was achieved by hydrophobic interaction chromatography. As illustrated in Fig. 3, fragment CCP$_2$-ap-SP eluted as a single peak at the end of the descending ammonium sulfate gradient, slightly later than the whole C$\bar{1}$s γ-B fragment. Its aglycosylated form eluted essentially at the same position, whereas the shorter fragment ap-SP eluted at the middle of the gradient and was therefore significantly more hydrophilic. The amounts of purified proteins produced were typically about 0.4, 1.5, and 0.5 mg per liter of culture for fragments ap-SP, CCP$_2$-ap-SP, and CCP$_2$-ap-SP*ag*, respectively.

The recombinant fragments were essentially pure as judged from SDS-PAGE analysis (Fig. 4). Each yielded a single band under both non-reducing and reducing conditions, indicating that they were in the proenzyme form, with apparent molecular weights of about 30,000 (ap-SP), 40,000 (CCP$_2$-ap-SP), and 38,000 (CCP$_2$-ap-SP*ag*), consistent with the predicted values. No sign of proteolytic degradation was observed on the gels upon storage of the recombinant proteins for several months at −20 °C.

Chemical Characterization of Recombinant Proteins—Edman degradation of the purified fragments CCP$_2$-ap-SP and

Downloaded from http://www.jbc.org/ by guest on March 9, 2015

P4

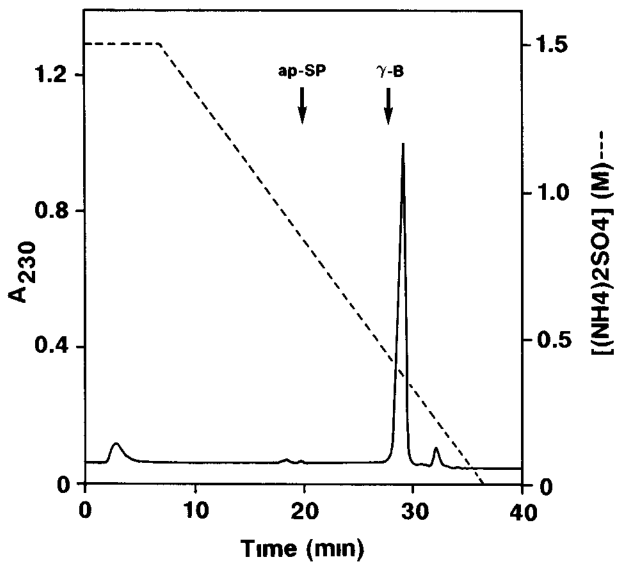


FIG. 3. **Purification of recombinant fragment CCP$_2$-ap-SP by high pressure hydrophobic interaction chromatography.** The fraction obtained from ion-exchange chromatography of the culture supernatant was further purified by hydrophobic interaction chromatography on a TSK-Phenyl 5PW column, and elution was realized by a descending ammonium sulfate gradient as described under "Experimental Procedures." Protein was detected by absorbance at 230 nm. The elution positions of fragments γ-B and ap-SP are indicated by *arrows*.

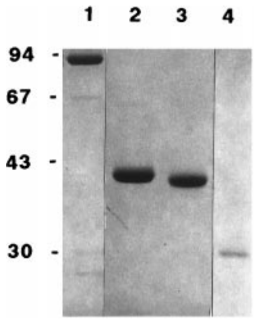


FIG. 4. **SDS-PAGE analysis of the recombinant proteins.** *Lane 1*, proenzyme C1s isolated from human serum; *lane 2*, fragment CCP$_2$-ap-SP; *lane 3*, aglycosylated fragment CCP$_2$-ap-SP; *lane 4*, fragment ap-SP. All samples were analyzed under reducing conditions. Molecular masses of marker proteins (expressed in kDa) are shown on the *left side* of the gel.

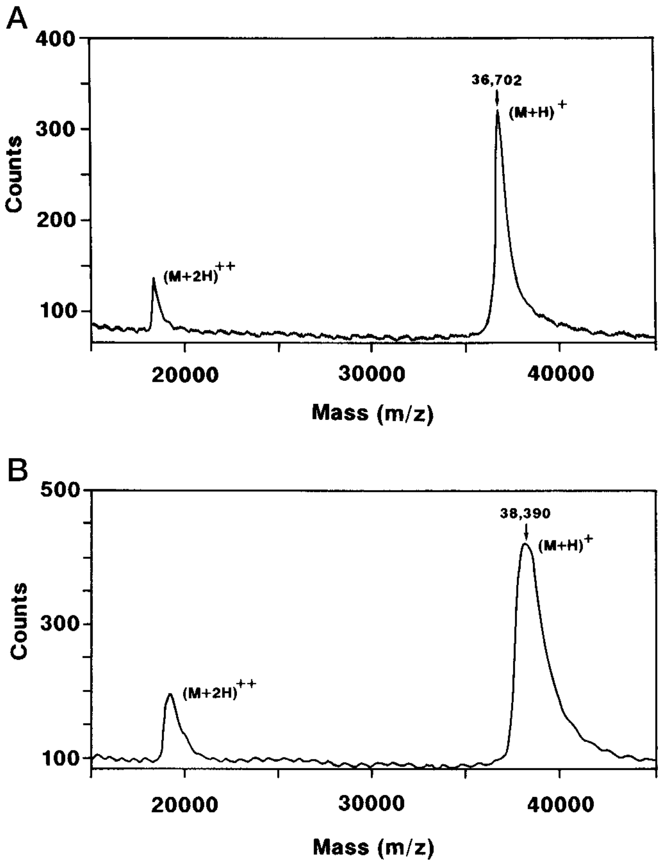


FIG. 5. **Analysis by MALDI mass spectrometry of the recombinant fragments CCP$_2$-ap-SP*ag* (*A*) and CCP$_2$-ap-SP (*B*).** Each spectrum exhibits two peaks corresponding to the $(M + H)^+$ and $(M + 2H)^{2+}$ protonated forms of the fragments. The mass values given in the text are derived from the $(M + H)^+$ forms.

CCP$_2$-ap-SP*ag* yielded a single sequence *Asp-Leu*-Asp-(Cys)-Gly-Ile-Pro-Glu-Ser-Ile . . . in both cases, corresponding to the segment Asp^{343}-Ile^{350} of human C1s preceded by the two residues Asp-Leu expected to be added at the N terminus, due to the introduction of bases at the 5′ end of the cDNA to create a *Bgl*II site (see "Experimental Procedures"). In the same way, fragment ap-SP yielded a single sequence *Asp-Leu*-Pro-Val-(Cys)-Gly-Val-Pro-Arg-Glu . . . derived from the segment Pro^{408}-Glu^{415} of C1s, corresponding to the N-terminal end of the activation peptide.

Analysis of fragment ap-SP by MALDI mass spectrometry yielded a molecular mass of 29,627 ± 15 Da, consistent with a sequence comprising residues Asp-Leu followed by the C-terminal segment Pro^{408}-Asp^{673} of C1s (calculated mass = 29,634 Da). In the same way, as shown in Fig. 5*A*, fragment CCP$_2$-ap-SP*ag* yielded a mass value of 36,701 ± 18 Da, in full agreement with that predicted for the sequence Asp-Leu-(Asp^{343} . . . Asp^{673}) (36,707 Da). Analysis of the glycosylated fragment CCP$_2$-ap-SP consistently yielded a wide, heterogeneous peak, centered on mass values ranging from 37,824 ± 19 Da to 38,389 ± 19 Da (as illustrated in the mass spectrum shown in Fig. 5*B*), depending on the preparation. Thus, the corresponding deduced masses for the oligosaccharide attached to Asn^{391} ranged from 1,114 ± 19 Da to 1,682 ± 19 Da. Considering that glycoproteins expressed in baculovirus/insect cells systems mainly bear high mannose oligosaccharides (34), and in keeping with our previous studies (11), the above values are consistent with the occurrence of short, heterogeneous oligosaccharides comprising two *N*-acetylglucosamine residues and a variable number of mannose residues ranging from 4 to 8 (calculated masses of the carbohydrate moiety = 1,055 and 1,704, respectively). The significantly high differences between the experimental and calculated values likely arise from the fact that the former are average figures which do not necessarily correspond to a round number of mannose residues (see Fig. 5*B*). In summary, N-terminal sequence and mass spectrometry analyses clearly showed that the recombinant proteins had the expected sequences and that a heterogeneous high mannose oligosaccharide was attached to Asn^{391} of fragment CCP$_2$-ap-SP.

Functional Characterization of Recombinant Proteins—As judged from SDS-PAGE analysis under reducing conditions, fragments CCP$_2$-ap-SP, CCP$_2$-ap-SP*ag*, and ap-SP retained a monochain structure upon prolonged incubation at 37 °C, indicating that the preparations were free of contaminating proteases. To test their activation by C$\bar{1}$r, *i.e.* their ability to be split at the Arg^{422}-Ile^{423} bond cleaved upon activation of intact C1s, each fragment was incubated for various periods at 37 °C in the presence of C$\bar{1}$r. SDS-PAGE analysis under reducing conditions indicated that all three fragments were cleaved by

Downloaded from http://www.jbc.org/ by guest on March 9, 2015

P4

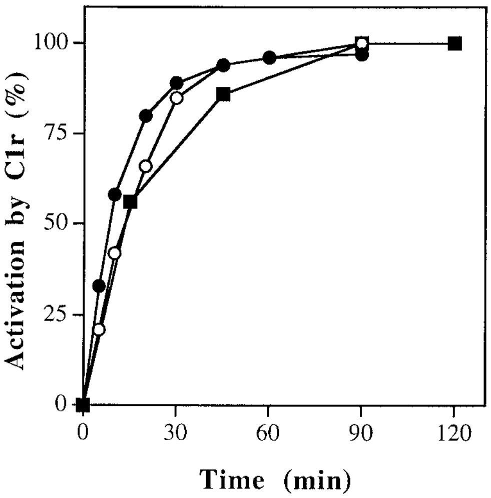


FIG. 6. **Kinetic analysis of the activation of C1s and the recombinant fragments CCP$_2$-ap-SP and ap-SP by C1̄r.** Proenzyme C1s (●) and its recombinant fragments CCP$_2$-ap-SP (○) and ap-SP (■) (each 6.5 μM) were incubated with activated C1r (C1̄r/protein molar ratio = 0.045) for various periods at 37 °C, in 50 mM triethanolamine HCl, 145 mM NaCl, pH 7.4. In the case of C1s and CCP$_2$-ap-SP, activation was monitored by SDS-PAGE quantification of the two-chain structure characteristic of the active state of the proteins. Activation of fragment ap-SP was measured by esterolytic activity toward N^{α}-benzoyl-L-arginine ethyl ester (see "Experimental Procedures").

C1̄r: (i) in the case of CCP$_2$-ap-SP and CCP$_2$-ap-SP*ag*, a two-chain structure became clearly visible, as shown by the appearance of bands corresponding to the serine protease domain and to the CCP$_2$-ap moieties; (ii) the apparent molecular weight of ap-SP was slightly decreased, suggesting removal of the short N-terminal activation peptide. Edman degradation of the C1̄r-treated fragments gave evidence for a second sequence ^{423}Ile-Ile-Gly-Gly-Ser-Asp-Ala . . . , in addition to the single N-terminal sequence yielded upon analysis of the monochain fragments (see above), confirming C1̄r-mediated cleavage of the ^{422}Arg-Ile^{423} bond in all cases. A comparative kinetic analysis of the activation of proenzyme C1s and the recombinant fragments CCP$_2$-ap-SP and ap-SP by C1̄r showed that all three proteins were activated in a comparable manner (Fig. 6). It was concluded, therefore, that the structural determinants involved in the cleavage of proenzyme C1s by C1̄r are contained solely in the C-terminal part of the protein, comprising the activation peptide and the serine protease domain.

The activated forms of the recombinant proteins CCP$_2$-ap-SP and ap-SP were prepared by incubation with C1̄r for 90 min at 37 °C under the conditions described in the legend to Fig. 6 and tested for their enzymic activity as described below. Esterolytic activity was tested against N^{α}-benzoyl-L-arginine ethyl ester, which is not among the best synthetic substrates of C1̄s, and was chosen because of its lack of cleavage by C1̄r (33). Activated C1̄s and its fragments CCP$_2$-ap-SP and ap-SP all cleaved the synthetic substrate, with values of k_{cat} and K_m ranging from 0.015 ± 0.003 to 0.020 ± 0.003 s^{-1} and from 2.0 ± 0.2 to 2.2 ± 0.2 mM, not significantly different from each other. It became clear, therefore, that C1̄r-mediated cleavage of the ^{422}Arg-Ile^{423} bond induced formation of a functional active site in all cases.

To check whether the activated recombinant fragments also displayed proteolytic activity, their ability to cleave the C1̄s protein substrates C2 and C4 was tested and compared with that of intact C1̄s and its proteolytic γ-B fragment. The two glycoforms of C1̄s γ-B, containing either a biantennary or a triantennary, complex-type oligosaccharide (11) were also tested. Kinetic analyses of C2 cleavage as a function of enzyme concentration clearly showed that all C1̄s fragments had the ability to cleave C2 and yielded the C2a and C2b fragments characteristic of cleavage by intact C1̄s (not shown). As summarized in Table I, the larger fragment γ-B and both fragments CCP$_2$-ap-SP (either glycosylated or aglycosylated) displayed kinetic constants very similar to those of intact C1̄s. That the carbohydrate linked to the second CCP module had no significant influence on C2 cleavage was confirmed by the fact that the two glycoforms of C1̄s γ-B, as well as the desialylated protein, all exhibited comparable kinetic constants (not shown). The shorter fragment ap-SP also cleaved C2, although with a k_{cat} value slightly less (about 50%) than those determined for the other fragments. Nevertheless, all of the fragments cleaved C2 efficiently, indicating that the first CCP module of C1̄s and very likely also the second CCP module do not play a major role in C2 cleavage.

TABLE I
Kinetic constants for proteolytic cleavage of C2 by intact C1̄s and various C1̄s fragments

Enzyme	k_{cat}	K_m	k_{cat}/K_m
	s^{-1}	μM	$M^{-1} \cdot s^{-1}$
C1̄s	7.4 ± 1	10.0 ± 3.0	740,000
γ-B	5.8 ± 1	5.6 ± 2.0	1,030,000
CCP$_2$-ap-SP	7.4 ± 1	7.7 ± 2.5	960,000
CCP$_2$-ap-SP*ag*	6.7 ± 1	7.7 ± 2.5	870,000
ap-SP	3.2 ± 1	8.6 ± 2.5	370,000

In the next step, we compared the ability of the various C1̄s fragments to cleave C4. Kinetic analyses as a function of enzyme concentration were performed initially, indicating that the larger fragment γ-B cleaved C4 to yield the expected C4α′ chain, with an efficiency comparable to that of intact C1̄s (not shown). In contrast, as illustrated in Fig. 7, a dramatic decrease in cleaving efficiency was observed for fragments CCP$_2$-ap-SP and CCP$_2$-ap-SP*ag*, whereas C4-cleaving activity was abolished in the case of the shorter fragment ap-SP. No proteolytic activity toward C4 was detectable after incubation for 30 min at 37 °C using ap-SP/C4 molar ratios up to 0.07 (not shown). Determination of the kinetic constants for C4 cleavage confirmed the above findings, indicating that C1̄s and its γ-B fragment had comparable k_{cat} and K_m values (Table II). The two glycoforms and the asialylated form of C1̄s γ-B all exhibited similar catalytic constants (not shown). In the case of both fragments CCP$_2$-ap-SP, in contrast, the k_{cat} decreased about 15–20-fold, and the K_m increased about 4-fold, leading to a mean decrease of the k_{cat}/K_m value of about 70, in agreement with the data shown in Fig. 7. Thus, C4-cleaving activity was strongly reduced upon deletion of the N-terminal CCP module and abolished when both CCP modules were removed, indicating that both modules participate in C4 recognition by C1̄s.

Finally, experiments aimed at checking the ability of the activated recombinant C1s fragments to react with the physiological protein inhibitor of C1̄s, C1 inhibitor, were performed. After incubation with equimolar amounts of C1 inhibitor followed by SDS-PAGE analysis under reducing conditions (not shown), C1̄s and all three fragments CCP$_2$-ap-SP, CCP$_2$-ap-SP*ag*, and ap-SP yielded a band of apparent molecular weight >120,000, corresponding to the covalent complex formed between C1 inhibitor and the serine protease domain. In all cases, including C1̄s, formation of the covalent complex was not complete under the conditions used, as shown by the presence of significant amounts of C1 inhibitor and the free form of the serine protease domain. However, the relative proportion of complex formed was comparable in all cases. As judged from these experiments, all fragments exhibited a reactivity toward C1 inhibitor comparable to that of intact C1̄s, indicating that the C-terminal end of the protease, comprising the activation

Downloaded from http://www.jbc.org/ by guest on March 9, 2015

P4

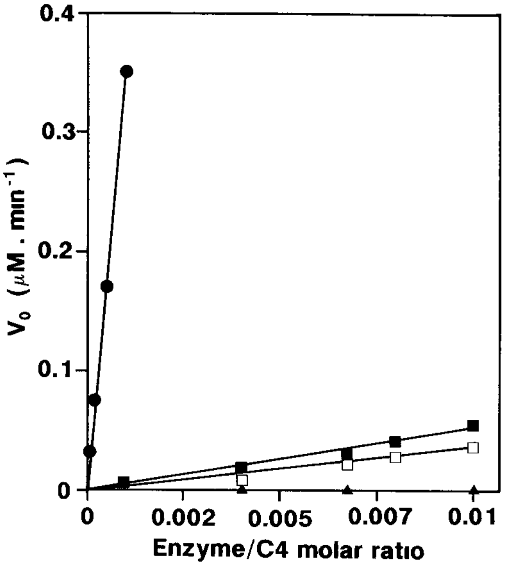


FIG. 7. **Comparative analysis of C4 cleavage by increasing amounts of activated C1s and its recombinant fragments CCP$_2$-ap-SP, CCP$_2$-ap-SP*ag*, and ap-SP.** C4 (2.25 μM) was incubated at 37 °C in 20 mM sodium phosphate, 150 mM NaCl, 5 mM EDTA, pH 7.4, in the presence of increasing concentrations of C$\bar{1}$s (●) or the activated recombinant fragments CCP$_2$-ap-SP (■), CCP$_2$-ap-SP*ag* (□), and ap-SP (▲). The initial rate of C4 cleavage V_0 was measured by kinetic analysis by SDS-PAGE of the production of the C4α′ chain of fragment C4b (see "Experimental Procedures").

TABLE II
Kinetic constants for proteolytic cleavage of C4 by intact C$\bar{1}$s and various C$\bar{1}$s fragments

Enzyme	k_{cat}	K_m	k_{cat}/K_m
	s^{-1}	μM	$M^{-1} \cdot s^{-1}$
C$\bar{1}$s	7.6 ± 1	5.0 ± 2.0	1,520,000
γ-B	7.6 ± 1	5.7 ± 2.0	1,330,000
CCP$_2$-ap-SP	0.4 ± 0.1	20.0 ± 5.0	20,000
CCP$_2$-ap-SP*ag*	0.5 ± 0.1	22.0 ± 5.0	23,000
ap-SP	[a]	[a]	[a]

[a] Values not measurable due to the lack of C4 cleaving activity.

peptide and the serine protease domain, contains the structural determinants required for reaction with this inhibitor.

DISCUSSION

The baculovirus/insect cells expression system has been used successfully to produce a number of intact multidomain proteins, including human complement proteases C1r and C1s (35, 36). In the present study, the recently developed Bac-to-Bac™ system (28) was used to express recombinant modular fragments from the catalytic region of human C1s. Although the secretion yields (0.6–3.0 mg/liter cell culture) were low compared with some of those reported for intact proteins, they were sufficient to produce the purified fragments for chemical and functional characterization. In this respect, the significantly lower level of secretion of fragment ap-SP is probably related, at least in part, to a decreased stability and/or solubility compared with the larger fragment CCP$_2$-ap-SP. Indeed, we have provided evidence that the second CCP module of C$\bar{1}$s is closely associated with the serine protease domain (12). Removal of this module may therefore be expected to unmask hydrophobic areas of the serine protease domain and thereby to decrease its solubility. In contrast, the observed slight decrease in the secretion yield of the aglycosylated fragment CCP$_2$-ap-SP*ag* is probably not linked to a reduced stability of the protein in solution due to the lack of the oligosaccharide chain but rather is an indirect consequence of the deleterious effect of tunicamycin on insect cells, which is known to affect secretion of some proteins (37). As shown by Edman degradation and mass spectrometry analyses, the recombinant C1s fragments produced in this study were homogeneous at the polypeptide level and had the expected N-terminal sequences and amino acid compositions. In contrast, the analyses performed on the glycosylated fragment CCP$_2$-ap-SP are consistent with the occurrence at Asn^{391} of heterogeneous high mannose type oligosaccharides containing a varying number of mannose residues ranging from 4 to 8, depending on the preparation, and even within a given preparation (see Fig. 5). Such a heterogeneity of the carbohydrate moiety of the glycoproteins expressed in baculovirus/insect cells systems could be a major disadvantage if these are produced for crystallographic purposes.

Several lines of evidence indicate that the carbohydrate borne by the second CCP module of C$\bar{1}$s is not significantly involved in the functional activities of this protease. Thus, the two glycoforms and the asialylated form of fragment C$\bar{1}$s γ-B all display comparable proteolytic activities toward C2 and C4. Moreover, the aglycosylated form of fragment CCP$_2$-ap-SP exhibits proteolytic kinetic constants similar to those of its glycosylated counterpart. It is very likely therefore that the two naturally occurring glycoforms of human C$\bar{1}$s display identical catalytic properties *in vivo*. The fact that complete removal of the carbohydrate from fragment CCP$_2$-ap-SP does not impair its functional activity also suggests that this does not significantly affect the folding of the protein. Thus, despite the location of Asn^{391} in the vicinity of the CCP module/serine protease domain interface (12), the oligosaccharide chain itself is probably not directly involved in the interaction between the two domains. This observation lends further support to the above-mentioned hypothesis that the decreased secretion of fragment CCP$_2$-ap-SP*ag* results from an indirect effect of tunicamycin on insect cells.

All of the recombinant fragments were produced and remained stable in a proenzyme form, indicating that, like the parent C1s molecule, none of them undergoes self-activation. In contrast, both fragments CCP$_2$-ap-SP and ap-SP were readily activated in solution by C$\bar{1}$r, with an efficiency comparable to intact C1s. The shorter C1s fragment ap-SP therefore contains the structural information required for C$\bar{1}$r-mediated cleavage. Indeed, this appears consistent with the fact that C1s cleavage by C$\bar{1}$r normally does not occur in the fluid phase, but inside the C1 complex, in which the activation site of the former protease and the active site of the latter are expected to be pre-positioned with respect to each other (4). It is likely, therefore, that C1s recognition by C$\bar{1}$r mainly involves a restricted number of residues on either side of the C1s Arg^{422}-Ile^{423} bond cleaved upon activation. In any case, it is clear that efficient cleavage does not require accessory binding sites outside the serine protease domain of C1s. It is noteworthy that C$\bar{1}$r-mediated cleavage of fragments CCP$_2$-ap-SP and ap-SP leads to the development of esterolytic activity in both cases, which again implies that the latter is self-sufficient for the formation of a fully active site, in agreement with current knowledge on active site formation in other serine proteases (38).

Our observation that the short C$\bar{1}$s fragment ap-SP retains the ability to form a stable complex with C1 inhibitor indicates that this contains the structural elements required for reaction with C1 inhibitor. It should be mentioned, however, that the reaction of C$\bar{1}$s with C1 inhibitor is a complex and still not fully understood process. Like other members of the serpin family, C1 inhibitor acts as a pseudosubstrate: C$\bar{1}$s cleaves C1 inhibitor at the Arg^{444}-Thr^{445} "reactive center," then the cleavage products do not dissociate, but a tight, SDS-stable enzyme-inhibitor complex is formed (25). It has been recently proposed

Downloaded from http://www.jbc.org/ by guest on March 9, 2015

that formation of this stable complex involves a secondary interaction site on C1 inhibitor (39). Obviously, a precise kinetic analysis of the reaction of C1 inhibitor with native C$\bar{1}$s and its ap-SP fragment would be necessary to check whether the latter behaves exactly as the parent protein at the different steps of the reaction. Nevertheless, our observation that C$\bar{1}$s and its fragment ap-SP yield comparable amounts of stable complexes with C1 inhibitor suggests that the proposed secondary interaction site on C1 inhibitor binds to the serine protease domain of C$\bar{1}$s.

With regard to C2 cleaving activity, all C$\bar{1}$s fragments were found to cleave this substrate with an efficiency comparable with that of the intact protease. The fact that the shorter fragment ap-SP exhibits a slightly decreased k_{cat} value can be rationalized by the following observations: (i) as discussed above, removal of the second CCP module may be expected to uncover hydrophobic patches of the serine protease domain and possibly to induce local denaturation resulting in partial loss of substrate binding site(s) at the surface of the domain; (ii) due to its lower solubility, the uncertainty on concentration measurements and hence on k_{cat} was probably higher in the case of fragment ap-SP. Indeed, the observation that ap-SP exhibits a K_m value for C2 comparable with that of the other fragments tends to favor the latter explanation. Therefore, our conclusion is that C2 recognition and cleavage by C$\bar{1}$s only involve structural motifs located within the ap-SP fragment and probably confined to the serine protease domain. In any case, our data provide no indication of a significant involvement of the remainder of the C$\bar{1}$s molecule, particularly the CCP modules.

In contrast, our data clearly show that, while fragment γ-B exhibits a C4 cleaving activity comparable with that of intact C$\bar{1}$s, this activity is greatly reduced upon removal of the first CCP module and abolished when both CCP modules are deleted. Therefore, we conclude that both of these modules contribute accessory recognition sites for C4, which likely mediate binding and positioning of the substrate in such a way that the scissile bond fits into the C$\bar{1}$s active site, thereby allowing efficient cleavage of the protein. This hypothesis is consistent with the observation that fragment CCP$_2$-ap-SP shows both an increased K_m for C4 and a decreased k_{cat} value (see Table II). Thus, removal of the first CCP module would result in partial loss of the C4 binding site, resulting in improper positioning of the protein, and thereby in decreased cleavage efficiency. Our finding that both CCP modules of C$\bar{1}$s participate in C4 binding is in full agreement with previous studies (13, 14) indicating that monoclonal antibodies directed to the γ segment of the protease inhibit its ability to cleave C4. From a more general point of view, it should be emphasized that all complement regulatory proteins known to bind the C4b fragment of C4 (including complement receptor type 1 (CR1), membrane cofactor protein, decay accelerating factor, and C4b-binding protein) are composed mainly or entirely of CCP modules (10). In the case of CR1, it has been shown by deletion experiments and site-directed mutagenesis that both the first and second CCP modules are required for C4b binding (40). By analogy, the most likely hypothesis is therefore that the two CCP modules of C$\bar{1}$s bind to the C4b moiety of C4. This does not preclude, however, that other binding sites may be involved in the C$\bar{1}$s-C4 interaction. In this respect, it should be mentioned that plasmin degradation of the serine protease domain of C$\bar{1}$s was reported to result in the loss of C4 binding ability (14). In the same way, it has been suggested that C$\bar{1}$s may interact with the C4a moiety of C4 through acidic residues close to the active site (41). It may be hypothesized, therefore, that efficient positioning of C4 with respect to C$\bar{1}$s requires two types of interactions: (i) between the C4b moiety and the CCP modules and (ii) between the C4a moiety and the serine protease domain.

It is not surprising that C4 and C2 recognition by C$\bar{1}$s involves distinct structural requirements, since these protein substrates are quite different both in terms of size and modular organization. Also, it is well established that, *in vivo*, C4 cleavage occurs in the fluid phase, whereas C2 is cleaved within the C2-C4b complex that forms after covalent attachment of C4b in the vicinity of the C$\bar{1}$ complex (5). It may be expected, therefore, that appropriate "presentation" of C2 with respect to the C$\bar{1}$s active site is effected, at least in part, by C4b. In this respect, it is noteworthy that initial binding of C2 to C4b has recently been shown to involve the CCP modules of C2 (42). Thus, the C4b moiety of native C4 would bind first to the CCP modules of C$\bar{1}$s, then the C4b fragment released upon C$\bar{1}$s cleavage would acquire preferential affinity for the homologous modules of C2.

The present study provides the first precise identification of the domains of C$\bar{1}$s involved in the various aspects of its catalytic activity, and this study also shows that its very restricted specificity and its ability to discriminate its two protein substrates arise for a large part from the occurrence of accessory substrate binding sites located outside the serine protease domain. Further studies are in progress to identify the residues of the CCP modules of C$\bar{1}$s that participate in C4 binding.

Acknowledgments—We thank Dr. Mario Tosi for providing the human C1s cDNA; Dr. Thierry Vernet for providing the pVT-Bac expression vector; Dr. Jean Gagnon for determining N-terminal sequences; and Dr. Michel Jaquinod for performing mass spectrometry analyses. The help provided by Dr. Thierry Vernet and Dr. Jadviga Chroboczek during the initial stages of this work is gratefully acknowledged.

REFERENCES

1. Bork, P., and Bairoch, A. (1995) *Trends Biochem. Sci.* **20,** (suppl.) C3
2. Cooper, N. R. (1985) *Adv. Immunol.* **37,** 151–216
3. Schumaker, V. N., Zavodszky, P., and Poon, P. H. (1987) *Annu. Rev. Immunol.* **5,** 21–42
4. Arlaud, G. J., Colomb, M. G., and Gagnon, J. (1987) *Immunol. Today* **8,** 106–111
5. Müller-Eberhard, H. J. (1988) *Annu. Rev. Immunol.* **5,** 21–42
6. Tosi, M., Duponchel, C., Méo, T., and Julier, C. (1987) *Biochemistry* **26,** 8516–8524
7. Mackinnon, C. M., Carter, P. E., Smyth, S. J., Dunbar, B., and Fothergill, J. E. (1987) *Eur. J. Biochem.* **169,** 547–553
8. Spycher, S. E., Nick, H., and Rickli, E. E. (1986) *Eur. J. Biochem.* **156,** 49–57
9. Villiers, C. L., Arlaud, G. J., and Colomb, M. G. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82,** 4477–4481
10. Reid, K. B. M., Bentley, D. R., Campbell, D. R., Chung, L. P., Sim, R. B., Kristensen, T., and Tack, B. F. (1986) *Immunol. Today* **7,** 230–234
11. Pétillot, Y., Thibault, P., Thielens, N. M., Rossi, V., Lacroix, M., Coddeville, B., Spik, G., Schumaker, V. N., Gagnon, J., and Arlaud, G. J. (1995) *FEBS Lett.* **358,** 323–328
12. Rossi, V., Gaboriaud, C., Lacroix, M., Ulrich, J., Fontecilla-Camps, J. C., Gagnon, J., and Arlaud, G. J. (1995) *Biochemistry* **34,** 7311–7321
13. Matsumoto, M., and Nagaki, K. (1986) *J. Immunol.* **137,** 2907–2912
14. Matsumoto, M., Nagaki, K., Kitamura, H., Kuramitsu, S., Nagasawa, S., and Seya, T. (1989) *J. Immunol.* **142,** 2743–2750
15. Arlaud, G. J., Villiers, C. L., Chesne, S., and Colomb, M. G. (1980) *Biochim. Biophys. Acta* **616,** 116–129
16. Arlaud, G. J., Sim, R. B., Duplaa, A. M., and Colomb, M. G. (1979) *Mol. Immunol.* **16,** 445–450
17. Thielens, N. M., Van Dorsselaer, A., Gagnon, J., and Arlaud, G. J. (1990) *Biochemistry* **29,** 3570–3578
18. Thielens, N. M., Villiers, M. B., Reboul, A., Villiers, C. L., and Colomb, M. G. (1982) *FEBS Lett.* **141,** 19–24
19. Dodds, A. W. (1993) *Methods Enzymol.* **223,** 46–61
20. Reboul, A., Arlaud, G. J., Sim, R. B., and Colomb, M. G. (1977) *FEBS Lett.* **79,** 45–50
21. Thielens, N. M., Aude, C. A., Lacroix, M. B., Gagnon, J., and Arlaud, G. J. (1990) *J. Biol. Chem.* **265,** 14469–14475
22. Zaccaï, G., Aude, C. A., Thielens, N. M., and Arlaud, G. J. (1990) *FEBS Lett.* **269,** 19–22
23. Isenman, D. E., and Kells, D. I. C. (1982) *Biochemistry* **21,** 1109–1117
24. Budzko, D. B., and Müller-Eberhard, H. J. (1970) *Immunochemistry* **7,** 227–234
25. Davis, A. E., Aulak, K. S., Zahedi, K., Bissler, J. J., and Harrison, R. A. (1993) *Methods Enzymol.* **223,** 97–12
26. Edelhoch, H. (1967) *Biochemistry* **6,** 1948–1954
27. Tessier, D. C., Thomas, D. Y., Laliberté, F., Khouri, H. E., and Vernet, T. (1991) *Gene* (*Amst.*) **98,** 177–183
28. Luckow, V. A., Lee, S. C., Barry, G. F., and Olins, P. O. (1993) *J. Virol.* **67,** 4566–4579

Downloaded from http://www.jbc.org/ by guest on March 9, 2015

P4

29. King, L. A., and Possee, R. D. (1992) *The Baculovirus Expression System: A Laboratory Guide,* pp. 111–114, Chapman and Hall, Ltd., London
30. Lacroix, M., Rossi, V., Gaboriaud, C., Chevallier, S., Jaquinod, M., Thielens, N. M., Gagnon, J., and Arlaud, G. J. (1997) *Biochemistry* **36,** 6270–6282
31. Kerr, M. A., and Porter, R. R. (1987) *Biochem. J.* **171,** 99–107
32. Schreiber, R. D., and Müller-Eberhard, H. J. (1974) *J. Exp. Med.* **140,** 1324–1335
33. Arlaud, G. J., and Thielens, N. M. (1993) *Methods Enzymol.* **223,** 61–82
34. Yeh, J.-C., Seals, J. R., Murphy, C. I., van Halbeek, H., and Cummings, R. D. (1993) *Biochemistry* **32,** 11087–11099
35. Gal, P., Sarvari, M., Szilagyi, K., Zavodszky, P., and Schumaker, V. N. (1989) *Complement Inflammation* **6,** 433–441
36. Luo, C., Thielens, N. M., Gagnon, J., Gal, P., Sarvari, M., Tseng, Y., Tosi, M., Zavodszky, P., Arlaud, G. J., and Schumaker, V. N. (1992) *Biochemistry* **31,** 4254–4262
37. O'Reilly, D. R., Miller, L. K., and Luckow, V. A. (1992) *Baculovirus Expression Vectors: A Laboratory Manual,* pp. 221–224, W. H. Freeman & Co., New York
38. Perona, J. J., and Craik, C. S. (1995) *Protein Sci.* **4,** 337–360
39. He, S., Sim, R. B., and Whaley, K. (1997) *FEBS Lett.* **405,** 42–46
40. Krych, M., Hourcade, D., and Atkinson, J. P. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 4353–4357
41. Carter, P. E., Dunbar, B., and Fothergill, J. E. (1983) *Biochem. J.* **215,** 565–571
42. Xu, Y., and Volanakis, J. E. (1997) *J. Immunol.* **158,** 5958–5965

Downloaded from http://www.jbc.org/ by guest on March 9, 2015

P5

CMLS, Cell. Mol. Life Sci. 54 (1998) 171–178
1420-682X/98/020171-08 $ 1.50 + 0.20/0
© Birkhäuser Verlag, Basel, 1998

CMLS Cellular and Molecular Life Sciences

NMR structures of the C-terminal end of human complement serine protease C1s

P. Gans[a], V. Rossi[b], C. Gaboriaud[c], I. Bally[b], J.-F. Hernandez[b], M. J. Blackledge[a] and G. J. Arlaud[b], *

[a]Laboratoire de Résonance Magnétique Nucléaire, Institut de Biologie Structurale Jean-Pierre Ebel (CEA/CNRS), 41 Avenue des Martyrs, F-38027 Grenoble Cedex 1 (France)
[b]Laboratoire d'Enzymologie Moléculaire, Institut de Biologie Structurale Jean-Pierre Ebel (CEA/CNRS), 41 Avenue des Martyrs, F-38027 Grenoble Cedex 1 (France), Fax +33 476885494, e-mail: arlaud@ibs.ibs.fr
[c]Laboratoire de Cristallogénèse et Cristallographie des Protéines, Institut de Biologie Structurale Jean-Pierre Ebel (CEA/CNRS), 41 Avenue des Martyrs, F-38027 Grenoble Cedex 1 (France)

Received 19 November 1997; accepted 24 November 1997

Abstract. Synthetic peptides derived from the C-terminal end of the human complement serine protease C1s were analysed by circular dichroism and nuclear magnetic resonance (NMR) spectroscopy. Circular dichroism indicates that peptides 656-673 and 653-673 are essentially unstructured in water and undergo a coil-to-helix transition in the presence of increasing concentrations of trifluoroethanol. Two-dimensional NMR analyses performed in water/trifluoroethanol solutions provide evidence for the occurrence of a regular α-helix extending from Trp659 to Ser668 (peptide 656-673), and from Tyr656 to Ser668 (peptide 653-673), the C-terminal segment of both peptides remaining unstructured under the conditions used. Based on these and other observations, we propose that the serine protease domain of C1s ends in a 13-residue α-helix (656Tyr-Ser668) followed by a five-residue C-terminal extension. The latter appears to be flexible and is probably locked within C1s through a salt bridge involving Glu672.

Key words. Complement; C1s; serine proteases; peptide synthesis; NMR structure.

Human C1s is a highly specific modular protease that is responsible for the enzymic activity of C1, the initial component of the classical pathway of the complement system [1, 2]. Its catalytic region, involved in the recognition and limited proteolysis of complement components C4 and C2, comprises two contiguous 'complement control protein' (CCP) modules connected, through a short intermediary segment, to a serine protease domain. Although no information about the crystallographic structure of C1s is presently available, recent studies based on chemical cross-linking and homology modelling [3] have led to a three-dimensional model of the assembly of the C-terminal part of the catalytic region, comprising the second CCP module, the intermediary segment, and the serine protease domain. This model is based, in large part, on the occurrence of an ionic bond between Glu672, the penultimate residue of the serine protease domain, and Lys405, in the second CCP module.

A major uncertainty of this model arises from the structure of the C-terminal end of the serine protease domain which, based upon homology with serine proteases of known structure [4], was initially simply assumed to form a continuous α-helix. However, compared with the corresponding region of classical serine proteases such as trypsin, chymotrypsin and elastase, the C-terminal end of C1s is prolonged by a six-residue extension which contains a proline residue and is therefore unlikely to adopt an α-helical conformation. Indeed, recent structural studies on blood coagulation

---

* Corresponding author.

Article No. mb982008 *J. Mol. Biol.* (1998) **282**, 459–470

JMB

Evolutionary Conserved Rigid Module-domain Interactions can be Detected at the Sequence Level: The Examples of Complement and Blood Coagulation Proteases

Christine Gaboriaud[1]*, Véronique Rossi[2]
Juan Carlos Fontecilla-Camps[1] and Gérard J. Arlaud[2]

[1]*Laboratoire de Cristallogenèse et Cristallographie des Protéines and* [2]*Laboratoire d'Enzymologies Moléculair Institut de Biologie Structurale Jean-Pierre Ebel, 41 avenue des Martyrs, 38027 Grenoble Cedex 1, France*

**Corresponding author*

Several extracellular modular proteins, including proteases of the complement and blood coagulation cascades, are shown here to exhibit conserved sequence patterns specific for a particular module-domain association. This was detected by comparative analysis of sequence variability in different multiple sequence alignments, which provides a new tool to investigate the evolution of modular proteins. A first example deals with the proteins featuring a common complement control protein (CCP) module-serine protease (SP) domain pattern at their C-terminal end, defined here as the CCP-SP sub-family. These proteins include the complement proteases C1r, C1s and MASPs, the *Limulus* clotting factor C, and the proteins of the haptoglobin family. A second example deals with blood coagulation factors VII, IX and X and protein C, all featuring a common epidermal growth factor (EGF)-SP C-terminal assembly. Highly specific motifs are found at the connection between the CCP or EGF module and the activation peptide of the SP domain: [P/A]-x-C-x-[P/A]-[I/V]-C-G-x-[P/S/K] in the case of the CCP-SP proteins, and C-x-[P/S]-x-x-x-[Y/F]-P-C-G in the case of the EGF-SP proteins. Each motif is strictly conserved in the whole sub-family and it is detected in no more than one other known protein sequence. Strikingly, most of the conserved residues specific to each sub-family appear to be clustered at the interface between the SP domain and the CCP or EGF module. We propose that a rigid module-domain interaction occurs in these proteins and has been conserved through evolution. The functional implications of these assemblies, underlined by such evolutionary constraints, are discussed.

© 1998 Academic Press

Keywords: complement; blood coagulation; domain-domain interaction; protein module; sequence homology

Introduction

Comparative analysis of protein sequences has revealed that many extracellular proteins are composed from a limited repertoire of sequence patterns or modules. These "mosaic" or "modular" proteins may thus be described as the linear juxtaposition of contiguous modules and/or domains. Modules may be defined as a subset of domains that are used repeatedly as "building blocks" in diverse proteins and have probably appeared through genetic shuffling (Patthy, 1994; Doolittle, 1995; Hegyi & Bork, 1997). Several mosaic proteins play an essential role in extracellular biological pathways. The two best-studied cases are the blood coagulation and the complement cascades, two major host defense systems (Patthy, 1993). The former system is initiated after blood vessel injury, the latter after recognition of infectious microor-

---

Present address: V. Rossi, University of Missouri-Kansas City School of Biological Sciences, 407 Biological Sciences Building, 5100 Rockhill Road, Kansas City, MO 64110-2499, USA

Abbreviations used: CCP, complement control protein module; SP, serine protease; EGF, epidermal growth factor module; HSSP, homology derived secondary structure of proteins; PDB, Brookhaven Protein Data Bank; M-D, module-domain.

E-mail address of the corresponding author: chris@lccp.ibs.fr

© 1998 Academic Press

The EMBO Journal Vol. 19 No. 8 pp. 1755–1765, 2000

P7

Crystal structure of the catalytic domain of human complement C1s: a serine protease with a handle

Christine Gaboriaud[1], Véronique Rossi[2], Isabelle Bally[2], Gérard J.Arlaud[2] and Juan Carlos Fontecilla-Camps

LCCP and [2]LEM, Institut de Biologie Structurale J.-P.Ebel
CEA-CNRS, 41, rue Jules Horowitz, 38027 Grenoble Cedex 1, France

[1]Corresponding author
e-mail: chris@lccp.ibs.fr

C1s is the highly specific modular serine protease that mediates the proteolytic activity of the C1 complex and thereby triggers activation of the complement cascade. The crystal structure of a catalytic fragment from human C1s comprising the second complement control protein (CCP2) module and the chymotrypsin-like serine protease (SP) domain has been determined and refined to 1.7 Å resolution. In the areas surrounding the active site, the SP structure reveals a restricted access to subsidiary substrate binding sites that could be responsible for the narrow specificity of C1s. The ellipsoidal CCP2 module is oriented perpendicularly to the surface of the SP domain. This arrangement is maintained through a rigid module–domain interface involving intertwined proline- and tyrosine-rich polypeptide segments. The relative orientation of SP and CCP2 is consistent with the fact that the latter provides additional substrate recognition sites for the C4 substrate. This structure provides a first example of a CCP–SP assembly that is conserved in diverse extracellular proteins. Its implications in the activation mechanism of C1 are discussed.

Keywords: complement/modular structure/molecular recognition/serine protease/substrate specificity

Introduction

The complement cascade is a major system of innate immunity against pathogenic micro-organisms in mammals and in other vertebrate species. Activation of the classical pathway of complement is mediated by C1, a 790 000 Da macromolecular complex comprising three components: C1q, C1r and C1s. C1q is a hexamer of collagen-rich heterotrimers that has the overall shape of a bunch of tulips and internally binds two modular proteases C1r and C1s. The binding of C1q to the target micro-organism leads to autolytic activation of C1r, which, in turn, activates C1s by cleaving the Arg426–Ile427 bond (Cooper, 1985; Arlaud *et al*., 1987, 1998). In contrast with the low specificity of the initial recognition step mediated by C1q, C1s restrictively cleaves at a single arginyl peptide bond in both complement C4 and C2 proteins, the natural substrates of C1. A cascade of subsequent proteolytic reactions results in various biological activities designed to provide a first line of defense against microbial infection. Another facet of the protective action of complement lies in its role in immune tolerance, as exemplified by its involvement in the clearance of apoptotic cells (Mevorach *et al*., 1998) and in graft rejection (Dalmasso, 1992). Uncontrolled side-effects of complement may also result in various autoimmune pathologies, including Alzheimer's disease (Rogers *et al*., 1992).

C1s comprises, starting from the N-terminus, a CUB module (Bork and Beckmann, 1993), an epidermal growth factor (EGF)-like module, a second CUB module, two complement control protein (CCP, also known as sushi and SCR) modules (Mackinnon *et al*., 1987; Tosi *et al*., 1987) and a C-terminal chymotrypsin-like serine protease (SP) domain. This modular architecture is also found in C1r (Arlaud *et al*., 1998) and in the mannan-binding lectin-associated serine proteases (MASPs), a family of proteolytic enzymes involved in the activation of the recently discovered 'lectin pathway' of complement activation (Sato *et al*., 1994; Thiel *et al*., 1997; Endo *et al*., 1998).

Insights into the molecular architecture and functional role of the catalytic region of C1s have been obtained only in the past few years. Studies based on chemical cross-linking and homology modeling strongly suggested the occurrence of a close interaction between the second CCP module and the SP domain in both C1s and C1r (Rossi *et al*., 1995; Lacroix *et al*., 1997). A similar conclusion (Gaboriaud *et al*., 1998) was reached from a comparative analysis of amino acid sequence variability in the family of proteins containing the CCP–SP motif found in species ranging from invertebrates to mammals (Tosi *et al*., 1989; Muta *et al*., 1991; Matsushita *et al*., 1998). Studies carried out with recombinant modular fragments from the C1s catalytic region indicated that although C2 cleavage and reactivity towards C1 inhibitor are mediated mainly by the SP domain, efficient C4 cleavage requires substrate binding sites located in both CCP modules (Rossi *et al*., 1998).

CCP modules are known to be involved in many recognition processes, including the binding of several complement factors to fragments C3b and C4b (Reid and Day, 1989). Five CCP module structures have been determined by NMR spectroscopy: H5, the H15–H16 pair from human complement factor H (Barlow *et al*., 1992, 1993) and the VC3–VC4 module pair from the vaccinia virus complement control protein (VCP) (Wiles *et al*., 1997). More recently, the 3 Å resolution X-ray structure of the N-terminal pair of CCP modules from the complement regulatory protein CD46 has also been reported (Casasnovas *et al*., 1999). Structural studies performed on pairs of CCP modules indicate that these may exhibit quite different inter-modular orientations and show varying degrees of flexibility at the interface

© European Molecular Biology Organization

P7

Table I. Final refinement and model geometry statistics

Protein residues	303
Protein atoms	2284
Carbohydrate residues	3
Waters/ions/buffer	320/2 sulfates/1TES
Resolution (Å)	15–1.7
R_{work}/No. of observations	0.185/36847
R_{free}/No. of observations	0.220/1969
Residues in dual conformations	10[a]
Residues with disordered side chains after C_β	14[b]
R.m.s. bond lengths (Å)	0.011
R.m.s. bond angle associated distances (Å)	0.024
R.m.s. planarity (Å)	0.016

[a] Ser349, Ser365, Cys371, Val395, Cys410, Thr469, Ser474, Cys580, Cys613, Lys614.

[b] Glu351, Glu356, Asp357, Glu359, Glu373, Glu385, Glu397, Glu402, Arg414, Glu506, Asp541, Met545, Lys629, Gln665.

between the two modules (Barlow *et al*., 1993; Wiles *et al*., 1997; Casasnovas *et al*., 1999; Kirkitadze *et al*., 1999). As discussed below, this flexibility may be relevant for the activation and activity of the proteases of C1.

There is relatively little structural information concerning the proteases of the complement system. Pioneering studies in this field have been performed on factor D, providing a structural basis for the unique activation and control mechanisms of this chymotrypsin-like serine protease (Volanakis and Narayana, 1996; Jing *et al*., 1999). The present study describes the first three-dimensional (3D) structure of an activated catalytic fragment from a modular protease of the complement system and provides information that can be used as a basis for further studies of the structure–function relationships of C1s as well as other modular proteins of the CCP–SP family.

Results and discussion

We have determined the structure of a recombinant fragment from the catalytic domain of human C1s, consisting of the CCP2 module (residues 342–406) linked to the C-terminal chymotrypsin-like SP domain (residues 410–668). The structure was solved with the combined use of molecular replacement (MR) and automatic model building (WARP) and refined at 1.7 Å resolution (see Materials and methods). The final R_{work} and R_{free} factors are 0.185 and 0.220, respectively, and the stereochemistry of the model is of good quality (Table I).

Structure of the SP domain

As in other trypsin-like serine proteinases, the core of the C-terminal domain of C1s folds into two six-stranded β-barrels connected by three *trans*-segments, several surface loops and one C-terminal α-helix (Figures 1A and 2A). The catalytic residues Ser617(195), His460(57) and Asp514(102) (C1s numbering, followed by chymotrypsinogen numbering in brackets) are located at the junction of both barrels in a manner virtually identical to the one observed in trypsin. C1s also presents all the other key structural features of the active conformation (for a review, see Perona and Craik, 1997). Asp611(189), the well-established determinant of the specificity of cleavage after Arg and/or Lys residues, lies at the bottom of the primary specificity pocket (S1). A putative sulfate ion is found at the active site, where three of its oxygen atoms form hydrogen bonds with the active Ser617(195) Oγ (2.68 Å), the His460(57) Nε2 (2.84 Å) and the Lys614(192) Nζ (2.9 Å), respectively.

The surface segments of C1s that differ in conformation relative to the other SPs of known 3D structure are listed in Table II. In the mammalian SPs, fine substrate specificity is determined by surface loops 1–3 and A–E (Perona and Craik, 1997). C1s displays unique structural features in all of these loops (Table IIB and Figure 2A). The two major insertions in loops 3 and C are grouped on one side of the entrance of the active site, whereas the deletions in loops 1, 2 and A are on the opposite side (Figures 2A and 3). Disordered conformations are observed in loops 3 and E at both ends of the specific substrate binding region.

The C-terminal segment 669–673 is flexible, as observed previously in the NMR structural analysis of the C-terminal segments 653–673 and 656–673 of C1s (Gans *et al*., 1998). The disorder observed at the C-terminus of the cleaved activation peptide (418–422) and in the loop homologous to the trypsin calcium binding loop (478–486) are common features of SP domain structures. The salt bridges between Glu418 and Lys608, and between Glu672 and Lys405, previously identified by chemical cross-linking (Rossi *et al*., 1995), are not observed in the crystal structure. This is because these glutamic acid residues are located at the two disordered C-termini. This suggests that these segments have alternative configurations, one of which was trapped by the chemical reaction. Four cysteines (371, 410, 580, 613) are observed with dual conformations, and the inter-chain disulfide bridge (Cys410–Cys534) is not formed, most likely because of radiation damage (Burmeister, 2000). However, no conformational change is required to bridge cysteines 410 and 534.

Structure of the CCP2 module

As with other CCP modules, CCP2 folds into a small and compact hydrophobic core enveloped by six β-strands and stabilized by four conserved cysteines forming disulfide bridges in a 1–3, 2–4 pattern (Figures 1A, 2B and C). There are overall lower structural similarities among the CCP modules (Table III) than among the SP domains (Table II). Unexpectedly, the VCP CCP3 module has the least similar structure relative to C1s CCP2, despite it having the most homologous amino acid sequence (Table III). Structurally, the CD46 CCP2 module is the closest to C1s CCP2 (Table III; Figure 2B). As shown in Figure 2C, the relative structural orientation of the B2 and B4 strands is shared by all the CCP structures, whereas the topology of the other strands relative to this central conserved core is variable, especially at the regions that form the interfaces with the preceding and following module or domain.

The most prominent structural features specific to the C1s CCP module are: (i) the presence of tyrosines 375, 376 and 377, which form a surface patch near the C-terminal extremity of the module; and (ii) the insertion loops located between B3 and B4 and between B5 and B6 (Figure 2B and C). The large insertion between B5 and B6 is found in only 23 of the 1109 known CCP module amino acid sequences. The B1 to B2 loop, previously called the 'hypervariable loop' (Wiles *et al*., 1997), lies on the side opposite to these insertions (Figure 2B) and is shorter in C1s CCP2 than in most other CCP modules.

The glycosylation site of C1s is located in the loop connecting B4 to B5 (Figure 2B), on the side of the protein that contains the insertion loops. Two *N*-acetyl-glucosamines (NAG) and a fucose bound to the proximal NAG (Pétillot *et al*., 1995) (Figure 2B) were built into well-defined electron density (not shown).

Structure of the rigid CCP2 module–SP domain interface

The overall shape of the CCP2–SP assembly is that of a bludgeon with the ellipsoidal CCP2 module tightly anchored on the more globular SP domain, on the side opposite to the active site (Figures 1A and 4A). The perpendicular orientation of the long axis of CCP2 occurs in the absence of crystallographic packing constraints and results from a very rigid interface. This rigidity arises from interactions among a proline- and tyrosine-rich hydrophobic framework involving residues within the 372–377 stretch from CCP2, the 526(114)–534(122) region from SP and the short intermediary segment spanning residues 407–410 (Figure 1B and C).

The connecting 407–410 segment is clamped on one side by the CCP module β-stranded tyrosine 375–377 patch and, on the opposite side, by residues Pro532(120) to Cys534(122) of the SP domain (Figure 1C). This clamp is strengthened on both sides by other direct interactions between the SP domain and the CCP module: Lys525(113) with Asn352 and Glu372, and the 526(114)–528(116) segment with Tyr375 and Tyr376 (Figure 1B).

In the crystal structure, there are no covalent bonds connecting the regions of SP and CCP involved in the interface. On the one hand, cleavage of the Arg426(15)–Ile427(16) bond has occurred at the C-terminal end of the chymotrypsinogen-like activation peptide [residues 410(1)–426(15)], generating two polypeptide chains. On the other hand, as indicated above, the inter-subunit Cys410(1)–Cys534(122) disulfide bridge is not formed. The following non-covalent interactions are observed: 23 van der Waals contacts, seven direct and four water-mediated hydrogen bonds and one salt bridge.

A template structure for other CCP module–SP domain assemblies involved in several humoral defense systems

In addition to the C1r and MASPs complement proteases, which share the modular architecture of C1s, the CCP–SP family includes several proteins belonging to humoral defense systems: *Limulus* clotting factor C, which initiates a small defensive coagulation cascade in the hemolymph of the primitive horseshoe crab in the presence of traces of endotoxin lipopolysaccharides (Muta *et al*., 1991); haptoglobin (Hp), which is an acute-phase plasma glycoprotein (Dobryszycka, 1997) that binds free hemoglobin with extremely high affinity in order to remove it from plasma (Hwang and Greer, 1980); and the haptoglobin-related protein (Hpr), which is the toxic component of the trypanosome lytic factors directed against the African cattle parasite *Trypanosoma brucei brucei* (Smith *et al*., 1995). As detailed in the legend to Figure 1C, the residues maintaining the interface framework, which are clustered around the connecting segment, are highly conserved in the various sequences of the CCP–SP family. This strongly suggests that the rigid CCP2–SP assembly observed in C1s is also present in these proteins (Gaboriaud *et al*., 1998). Consequently, the C1s structure constitutes a reliable template for the study of proteins of this family.

Obstruction of C1s binding sub-sites: a striking similarity with blood clotting enzymes

In the absence of inhibitor in the CCP2–SP structure, the substrate binding sub-sites are unoccupied and cannot be unambiguously identified. However, canonical 'substrate binding-like' conformations of the P4 to P3′ binding sites for the inhibitors have been observed in numerous serine protease–inhibitor complexes (Bode and Huber, 1992). We have superimposed such complexes onto the structure of the C1s SP domain in order to identify S4 to S2 and S1′ to S3′ C1s sub-sites (P1, P2, etc. and P1′, P2′, etc. refer to residue positions on the N- and C-terminal sides of the substrate scissile peptide bond, respectively, whereas S1, S2, etc. and S1′, S2′, etc. represent the corresponding binding sites on the protease). It appears from this analysis that the access to the C1s binding sub-sites is severely restricted, especially when compared with the open canyon found in digestive enzymes such as trypsin. This situation is reminiscent of the occlusion of the binding sites observed in some highly specific proteases of the coagulation cascade (Bode *et al*., 1997).

The major structural features of the C1s sub-sites (Figure 3) may be summarized as follows. (i) Access to S1 may be affected by the partially disordered Lys614(192) that lies immediately above its entrance. (ii) The major insertion loop C is found just above the entrance to the active site [see Glu506(97D) in Figure 3], in a position similar to that of loop B in thrombin where it restricts the access to S2 (Figure 2A). (iii) The cluster formed by Phe511(99), Tyr595(174)and Trp640(215) is reminiscent of the hydrophobic box of coagulation factors IXa and Xa (Bode *et al*., 1997), although the first two residues are significantly displaced in C1s: Phe511(99) fills up part of the space normally assigned to S2 and Tyr595(174) occupies S4. (iv) On the opposite side of the active site, the accessibility of S2′ is restricted by Trp444(41) and Arg563(151), which are slightly displaced relative to the coagulation factors because of the conformation of loops A and D (Table II).

Implications for the specific inhibition and substrate recognition of C1s

The steric constraints observed in the sub-sites of C1s SP are likely to contribute to its high substrate specificity. They may explain (i) why there is only one plasma inhibitor of C1s and (ii) why C1s is not inhibited by the archetypal bovine pancreatic trypsin inhibitor (BPTI) (Arlaud and Thielens, 1993). In the case of thrombin, the poor inhibition by BPTI has been shown to be due to both the insertion at loop B and the presence of Glu(192) (instead of Gln) at the entrance of the active site (van de Locht *et al*., 1997). By analogy, in C1s the presence of both the loop C insertion and Lys614(192) above the active site entrance may explain the lack of inhibition by BPTI (Figure 3). The restricted access to S2 may also

P7

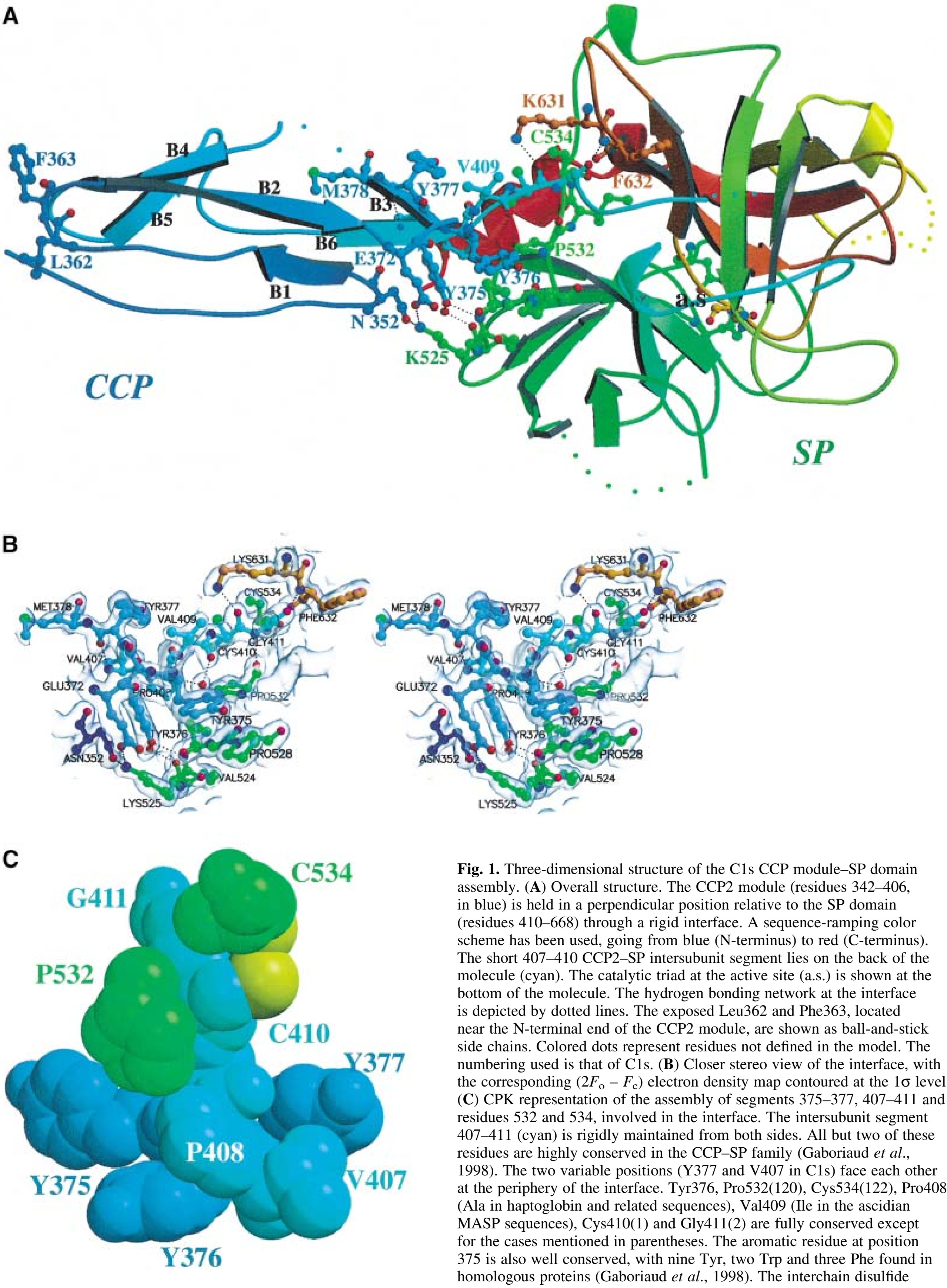

A
F363
B4
B2
M378
B3
Y377
V409
K631
C534
F632
B5
B6
E372
P532
L362
Y376
Y375
B1
a.s
N 352
K525
CCP
SP
B
LYS631
CYS534
MET378
TYR377
VAL409
PHE632
GLY411
CYS410
VAL407
GLU372
PRO408
PRO532
TYR375
TYR376
PRO528
ASN352
VAL524
LYS525
C
C534
G411
P532
C410
Y377
P408
V407
Y375
Y376

Fig. 1. Three-dimensional structure of the C1s CCP module–SP domain assembly. (**A**) Overall structure. The CCP2 module (residues 342–406, in blue) is held in a perpendicular position relative to the SP domain (residues 410–668) through a rigid interface. A sequence-ramping color scheme has been used, going from blue (N-terminus) to red (C-terminus). The short 407–410 CCP2–SP intersubunit segment lies on the back of the molecule (cyan). The catalytic triad at the active site (a.s.) is shown at the bottom of the molecule. The hydrogen bonding network at the interface is depicted by dotted lines. The exposed Leu362 and Phe363, located near the N-terminal end of the CCP2 module, are shown as ball-and-stick side chains. Colored dots represent residues not defined in the model. The numbering used is that of C1s. (**B**) Closer stereo view of the interface, with the corresponding ($2F_o - F_c$) electron density map contoured at the 1σ level. (**C**) CPK representation of the assembly of segments 375–377, 407–411 and residues 532 and 534, involved in the interface. The intersubunit segment 407–411 (cyan) is rigidly maintained from both sides. All but two of these residues are highly conserved in the CCP–SP family (Gaboriaud *et al.*, 1998). The two variable positions (Y377 and V407 in C1s) face each other at the periphery of the interface. Tyr376, Pro532(120), Cys534(122), Pro408 (Ala in haptoglobin and related sequences), Val409 (Ile in the ascidian MASP sequences), Cys410(1) and Gly411(2) are fully conserved except for the cases mentioned in parentheses. The aromatic residue at position 375 is also well conserved, with nine Tyr, two Trp and three Phe found in homologous proteins (Gaboriaud *et al.*, 1998). The interchain disulfide bridge Cys410(1)–Cys534(122) is not observed in the X-ray structure probably because of radiation damage; however, the cysteine side chains are displayed here in a conformation compatible with disulfide bridge formation (sulfur atoms are shown in yellow).

explain why either a Ala→Val or a Ala→Asp change at the P2 position of the human C1 inhibitor leads to a dysfunctional molecule with diminished inhibitory activity towards both C1s and C1r [a condition resulting in systemic lupus erythematosus (Zahedi *et al.*, 1997)]. It should be noted, however, that although in C2 the P2 sub-site is occupied by Gly, it corresponds to a Gln in C4. The fitting of the bulky Gln to a not very accessible S2 sub-site may be compensated by additional binding sub-sites. Indeed, the presence of both C1s CCP modules is required for efficient proteolytic cleavage of C4, indicating an extended recognition surface (Rossi *et al.*, 1998). In addition, the occurrence in C1s of a cluster containing Phe511(99), Tyr595(174) and Trp640(215) (Figure 3) may explain its affinity for hydrophobic P3 residues (Leu, Val) and small P4 residues (Gly, Ser), as found in C4, C2 and C1 inhibitor.

Functional implications of SP rigid modular extensions in C1s and blood-clotting enzymes

As the proteases of the complement cascade, the blood-clotting enzymes are under tight regulation to prevent potentially harmful effects. These highly selective serum enzymes all bear a modular N-terminal extension and act in the context of macromolecular complexes, attached to their site of action. Their activity is limited in time and space by inhibitors and modulated by several effectors.

A major difference between C1s and the vitamin K-dependent blood-clotting enzymes is the orientation of the module at the N-terminal side of the SP domain. In the X-ray structures of coagulation factors Xa (Padmanabhan *et al.*, 1993), IXa (Brandstetter *et al.*, 1995), VIIa (Banner *et al.*, 1996) and of protein C (Mather *et al.*, 1996), the epidermal growth factor-like module (EGF2) lies tangent and binds extensively to the chymotrypsin-like SP domain (Figure 4B). The mean surface buried at the EGF module–SP domain interface in the vitamin K-dependent enzymes goes from 562 Å^2 in factor IXa [code 1pfx in the Protein Data Bank (PDB; Bernstein *et al.*, 1977)] to 742 Å^2 in factor VIIa (PDB code 1dan). Furthermore, EGF2 interacts with the SP domain on the side opposite to the substrate binding site, whereas C1s CCP2 extends in the same direction as this site (Figures 3 and 4A). As in C1s, only two residues link the C-terminal end of the EGF2 to the N-terminus of the SP domain [the conserved inter-subunit Cys(1)–Cys(122) disulfide bond]. Because of the perpendicular orientation of C1s CCP2, the surface buried by the module/domain assembly represents only 8.3% of the total surface of the CCP module and 3.6% of the total surface of the SP domain (378 Å^2 out of 4528 Å^2 and 10526 Å^2, respectively). Thus, the buried surface in the coagulation factors represents a 150–200% increase relative to C1s.

Although both the EGF and CCP modules are involved in the control of protease activity, the different module–SP domain interaction implies a different kind of modulation. In the case of coagulation factors VIIa, IXa and Xa, the pair of EGF modules and the preceding GLA domain are known to bind the respective cofactors (Banner *et al.*, 1996; Bode *et al.*, 1997). In C1s, the CCP modules probably provide additional binding sites for C4 (Rossi *et al.*, 1998). The catalytic domain of C1s represents the first example of a rigid modular extension of the substrate binding site of an SP domain. An extension of the recognition site of an SP has also been reported in the microplasmin–staphylokinase–microplasmin ternary complex (Parry *et al.*, 1998). However, in that case the interaction takes place between a bacterial cofactor and a plasma SP.

A trypsin-like, closed S1 pocket

A further difference between C1s and the coagulation enzymes lies in the Na$^+$-induced allosteric regulation of activity observed in the latter. This regulation has been structurally correlated to the architecture of a water

Table II. Structural comparisons of the C1s SP domain with homologous X-ray structures

A Overall structure

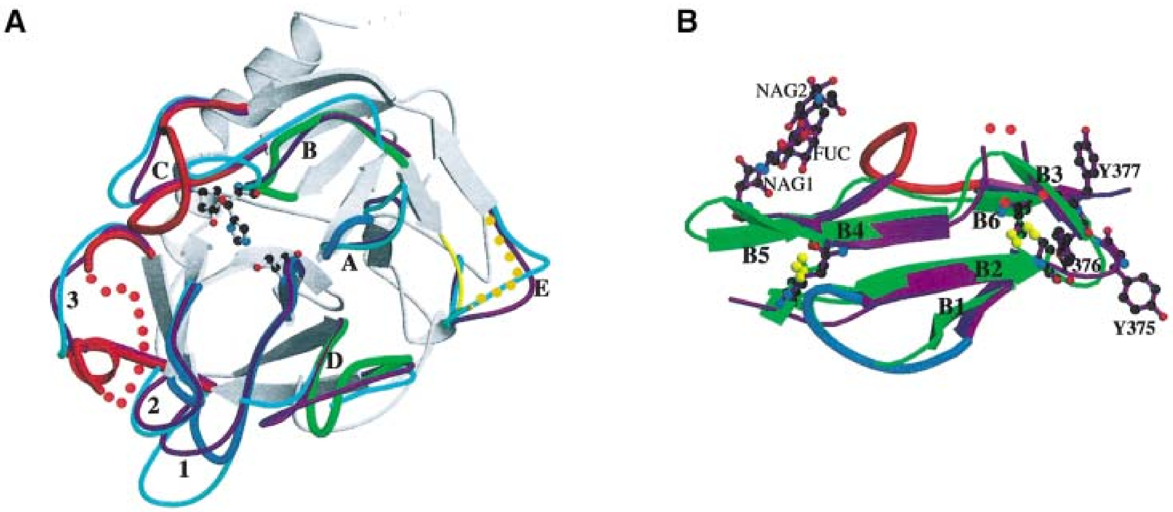
	Chymotrypsin	Thrombin	Factor Xa	Trypsin	Protein C
Global r.m.s. (Å)/nb Ca	1.0/194	0.8/182	1.0/188	0.8/183	1.0/194
Sequence identity (%)	37	38	35	36	37
PDB code	3gch	1fpc	1fax	1trn	1aut

B Surface segments of the C1s SP domain significantly different (d >1.5 Å) from the homologous segments of other SP domains (chymotrypsin, trypsin, thrombin, factor VII, factor D, factor Xa, and factor IXa)

C1s numbering	Chymotrypsin numbering	Loop label[a]	Type of modification
430–431	23–24		same length
441–443	34–40	A	deletion
466–467	61–62	B	variable
501–509	96–98	C	major insertion
538–543	126–129		minor insertion
560–562	145–146	D	variable
577–596	165–175	3	major insertion
607–608	185–187	1	deletion
625–630	203–205		minor insertion
646	221–224	2	deletion

[a] Loop labels as defined by Perona and Craik (1997) and detailed in Figure 2A. Loop E is not defined in C1s structure.

P7



C

```
                 B1                B2       B3  nd    B4  g    B5 insertion B6
     b          bbbb         b  bbbbbbb    bbb      bbbbb |    bbbb        b-bbb
C1s LDCGIPESI-ENGKVEdp----ESTLFGSVIRYTCEepYYYMEnggGGEYHCAGN---GSWVNevlgpELPK-CVP
CD1   CEEPpTF-EamELIgkpk--pyYEIGERVDYKCKkgyfyip-plaTHTICDRN---HTWL-----pvsddacyr
CD2 ETCpyirDP-LNGQAVpan---GTYEFGYQMHFICNe-GYYLI--geEILYCELkgsvAIWS-----gKPPI-CEK
VC3 vkcqsPPSisnGRHNgye----dfYTDGSVVTYscns-GYSli--gnsGVLCSg----gEWSD------ppt-cqi
VC4 vkcphpt-i-sngylssgfk--rsysYNDNVDFKCKy-gykls--gsSSSTCSPG---NTWKP-----elpk-cvr
H15 IPCSQPPQI-EHGTINssrssqesYAHGTKLSYTCEg-GFRIs--eENETTCY-m---gkws------spPQ-CEg
H16  pCKSPPEi-shGVVahms---dsYQYGEEVTYKCFe-gfgid--gpaIAKCL----geKWSH------pps-cik
      *         *           *     * *    * *         *        *            *
     bbb        bbbb Hyp_V bbb  bbbbbbb    bbbb     bbbbb      bbb         b-bbb
Cons  C  P  I  NG           F  G  I Y C   GY I   G     C      G WS        P  C
            L  H            Y     L F    F V                   T
            V                     V          L
```

Fig. 2. Comparison of the C1s SP domain (A) and the C1s CCP2 module (B and C) with the 3D structures of homologous proteins. (**A**) Superimposition of the structures of the SP domains of C1s, human trypsin (magenta), and human thrombin (cyan). The structures are shown in the 'standard' orientation, i.e. with the active-site cleft facing the viewer and the substrate binding site running from left to right. The catalytic triad residues of C1s are shown with ball-and-sticks (distances: Ser Oγ–His Nε2 = 2.86 Å; His Nδ1–Asp Oδ2 = 2.74 Å). The common core is shown in white, and only the variable surface loops around the active site are highlighted. The C1s loops are thicker and color-coded red for major insertions, blue for deletions, and green for other significant modifications. Dots represent disordered segments. The loops are labeled as in Perona and Craik (1997): A, 34–41; B, 56–64; C, 97–103; D, 143–149; E, 74–86; 1, 185–188; 2, 217–225; 3, 169–175. (**B**) Superimposed structures of the CCP modules from C1s (magenta) and CD46-2 (green). The two insertions in the C1s structure are shown in red, the deletion in blue. Red dots represent the disordered Gly–Gly stretch. The three carbohydrate residues observed in the structure are also displayed. The disulfide bridges from the two CCPs are superimposed. (**C**) Structural alignment of the CCP modules. The secondary structure (bbb) and strand numbering in the C1s structure are shown above. Pairs of residues with C_α–C_α distances of <2.0 Å are shown in upper case, the other residues in lower case. The deletion–insertion color code is the one used in (B). The consensus sequence (Cons), and the positions of the eight commonly observed β-strands (bbb) depicted above it, are taken from Wiles *et al.* (1997). The 11 consensus residues used for the core superposition (see Table III) are indicated by *. The carbohydrate attachment site is indicated by *g*.

Table III. Structural comparisons of the C1s CCP2 module with homologous structures

	VC3[a]	VC4[a]	H15[b]	H16[b]	CD1[c]	CD2[c]
PDB code	1vvc	1vvc	1hfi	1hcc	1ckl	1ckl
R.m.s on cysteines (Å)	1.6	1.0	1.1	0.9	1.2	0.4
Core r.m.s[d] (Å)	1.6	1.3	1.0	1.3	1.8	0.7
Global r.m.s. (Å)/nb Ca	2.2/56	2.0/54	1.7/56	2.0/55	1.7/51	1.5/57
Sequence identity[e] (%)	30.3	24.1	28.6	27.3	23.5	24.5

[a] Modules 3 and 4 from the vaccinia virus complement control protein (Wiles *et al.*, 1997).

[b] Modules 15 and 16 from the complement factor H (Barlow *et al.*, 1993).

[c] Modules 1 and 2 from the complement regulatory protein CD46 (Casasnovas *et al.*, 1999).

[d] The superposition was based on the C_α of the 11 consensus residues comprised in conserved strands, as indicated by * in Figure 2C.

[e] The sequence identity was measured on the homologous residues defined by the global superposition.

channel where Na^+ binds, which connects the bottom of the S1 binding site to an aperture at the bottom surface of the molecule (Dang and Dicera, 1996; Guinto *et al.*, 1999). Although it was postulated that this channel should be

P7

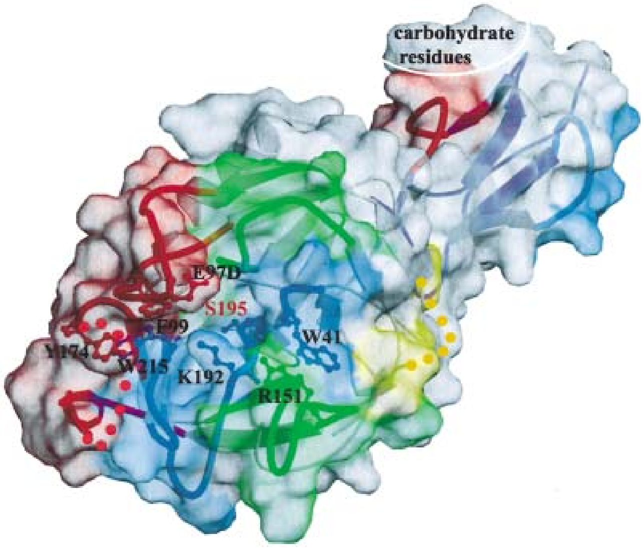


Fig. 3. Specific surface features around the catalytic site and substrate binding sub-sites of the C1s catalytic domain. The surface and the underlying loops have been colored using the same code. Also shown are the side chains of residues restricting the access of the substrate (see text) and of the active site Ser195. The CCP module constitutes an extension of the SP domain substrate binding site. The molecule is shown in the same orientation as in Figure 2A. The chymotrypsinogen residue numbering has been used throughout.

common to all proteases exhibiting either Tyr or Phe at position 225 (Guinto *et al*., 1999), the position of Tyr(225) in C1s is closer to Pro225 of trypsin (Ca–Ca distance = 0.4 Å) than to Tyr225 of thrombin (distance = 1.4 Å), casting doubts about the generality of the above proposition. This significant difference between C1s and the coagulation enzymes may result from the three-residue deletion in loop 2, which appears to be a unique feature of C1s. This deletion induces a narrowing of the bottom of the S1 specificity pocket, which is closed as in the case of trypsin.

The CCP module extension may act as a spacer and a handle

Current C1 models (Schumaker *et al*., 1986; Weiss *et al*., 1986; Arlaud *et al*., 1987) propose that the SP domain of C1s, once activated by C1r (an event expected to take place inside the C1 complex), becomes exposed to the solvent in order to be able to cleave C4 and C2. Both the orientation of CCP2 and its very rigid interaction with SP suggest that it could be involved, as both a spacer and a handle, in the movement required to expose the active site of C1s. Figure 5 depicts a very simple model that illustrates this hypothesis. In the initial state, the SP domains of C1r and C1s are buried inside C1 and oriented in a position compatible with the activating cleavage of the Arg446–Ile447 bond of C1s by C1r. This position would be similar to the one observed in the microplasmin–staphylokinase–microplasmin ternary complex (Parry *et al*., 1998). In this configuration, the C1s catalytic domain cannot cleave its protein substrates, because the essential S′ binding sub-sites are buried and face the C1r catalytic domain (Figure 5).

In the next step, we propose that the region connecting the CCP1 and CCP2 modules acts as a hinge allowing the SP domain to move to a position outside the C1 complex, a likely requirement for the cleavage of C4 and C2 (Figure 5). The proposed initial and final positions of SP are based on the comparison between C1s CCP2 and the structures of the CCP1–CCP2 module pair of CD46 (Casasnovas *et al*., 1999) and the CCP3–CCP4 module pair of VCP (Wiles *et al*., 1997). In the buried conformation, Leu362 (Figures 1A and 5), could help stabilize the CCP1–CCP2 interactions as observed for the corresponding modules of CD46 (Casasnovas *et al*., 1999; Figure 5).

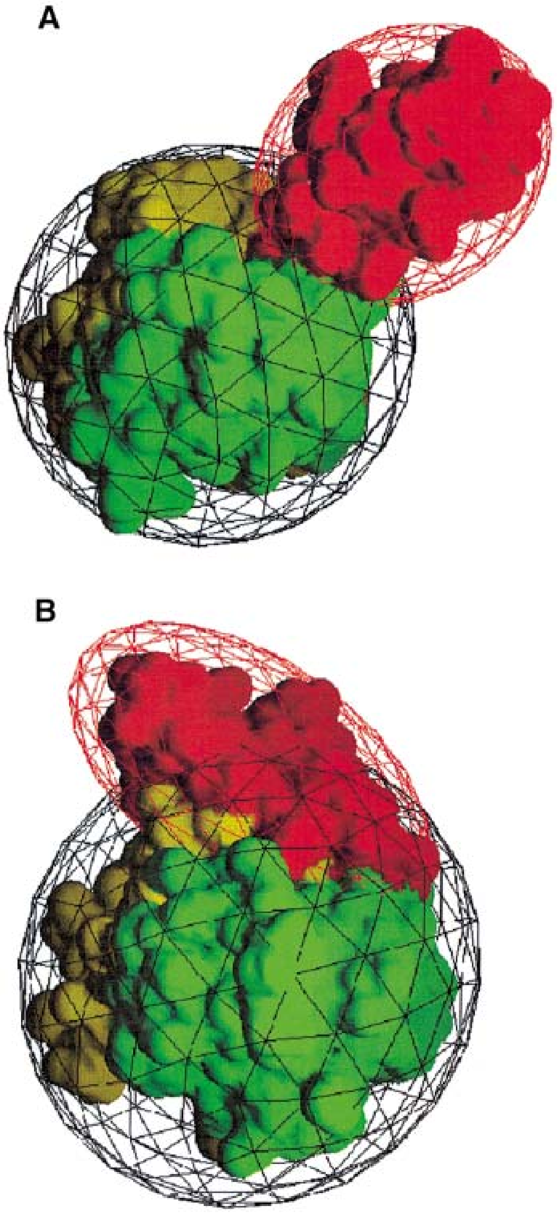


Fig. 4. The C1s and the vitamin K-dependent blood-clotting enzyme SP domains bear rigid modular N-terminal extensions, but they have strikingly different orientations. (**A**) The CCP modular extension of C1s (depicted in red) sits perpendicular to the SP domain surface (towards the viewer). The two SP subdomains are colored yellow and green. The molecule is seen in a perpendicular top view relative to Figures 3 and 2A, the binding site lying at the bottom of the figure. (**B**) The EGF modular extension (red) of blood coagulation factor VIIa (PDB code 1dan) lies tangentially to the surface of the SP domain. The SP domain in viewed in the same orientation as in (A) and the same color code has been used.

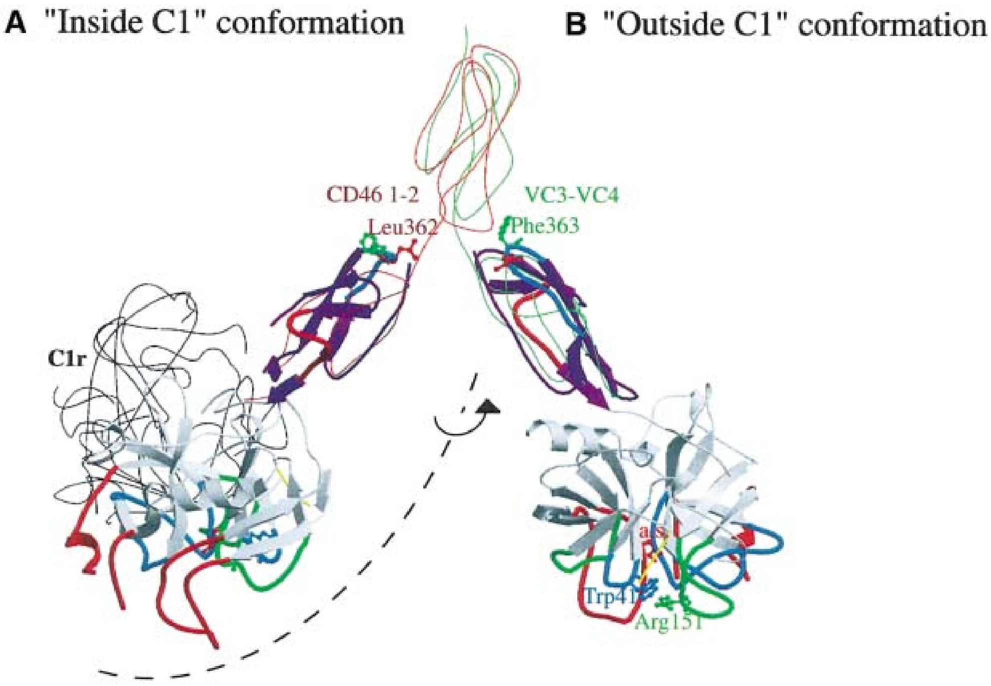


Fig. 5. Illustration of the potential role of the C1s CCP2 module handle in the putative displacement of the SP active site. The activation of C1s by C1r inside the C1 complex (**A**) requires an orientation of the C1s SP that is not compatible with its ability to cleave C4 and C2: the side chains of Trp41 (blue), Arg151 (green) and the active site Ser195 (red) face the C1r molecule inside the C1 complex. The thin black line indicates the position of the homologous C1r catalytic domain as required for C1s activation. In this orientation, the homologous CCP2 modular extension of C1r lies behind the SP domain. Correct positioning of C1s for proteolysis would require a conformational change in the CCP1–CCP2 hinge region (**B**). C1s colors are coded as in Figure 3. The thin red and green lines show the conformation of the CCP module pairs (CD46 CCP 1–2 and VCP CCP 3–4, respectively) used as references in the construction of this model, as detailed in the text. The exposed hydrophobic side chains of Leu362 (red) and Phe363 (green) observed in the structure are proposed to stabilize each of the two alternative conformations of the C1s CCP1–CCP2 module pair.

In the exposed orientation, Phe363 would stabilize the C1s CCP1–CCP2 interface as does the homologous Tyr81 in VCP CCP4 (Wiles *et al*., 1997). A similar movement may take place in the homologous C1r catalytic domains, which are expected to switch from a position suitable for their own activation to a position appropriate for C1s activation.

Conclusions and perspectives

Here we report the first 3D structure of a modular complement protease. The extreme specificity of C1s seems to arise from a combination of restriction elements in the vicinity of the active site, and additional specific recognition elements on the CCP modules for C4. Analogous exclusion mechanisms have been reported in the case of the blood-clotting enzymes. However, the module–SP interactions are very different in the two cases. The rigid perpendicular orientation of CCP2 relative to the SP domain in C1s is very likely to exist in other members of the CCP–SP family. In addition to their postulated substrate recognition function, the CCP2 modules of both C1s and C1r could play the role of rigid spacers and handles helping to position the SP domain so that it can either be activated or exert its catalytic function. This certainly represents a key feature of C1r and C1s, as these proteases are held together through their N-terminal ends (Thielens *et al*., 1990) as parts of a macromolecular complex, instead of acting freely in solution. Similar conclusions may also be extrapolated to the homologous MASPs proteases, which also participate in multimolecular complexes (Sato *et al*., 1994; Thiel *et al*., 1997; Endo *et al*., 1998).

We plan to use site-directed mutagenesis in order to further increase our understanding of the role played by the different structural determinants on tuning the proteolytic activity of C1s.

Materials and methods

Protein expression and purification

A recombinant fragment corresponding to the second CCP module and the SP domain of C1s was expressed and purified as described previously (Rossi *et al*., 1998). The protein comprised the human C1s segment Asp343–Asp673 preceded by an Asp–Leu sequence added at the N-terminal end due to the introduction of a restriction site at the 5′ end of the cDNA. Briefly, recombinant baculoviruses containing the plasmid pNT-Bac/CCP$_2$-ap-SP were generated by using the Bac-to-Bac system (Life Technologies, Inc.) and used to infect Sf21 insect cells in serum-free medium. The recombinant protein was isolated from the cell culture supernatant by anion-exchange chromatography on a Mono Q column followed by hydrophobic interaction chromatography on a TSK-Phenyl column.

Crystallization and data collection

Pooled fractions were concentrated to ~6 mg/ml in 50 mM triethanolamine–hydrochloride, 145 mM NaCl pH 7.4. Crystals were obtained by mixing 2 µl of protein solution with 2 µl of reservoir solution on a coverslip and then sealing over the reservoir using vacuum grease. The reservoirs contained 1 ml of 32–26% PEG 4000 at pH 5.6, 7.4 or 8.4, with or without 100 mM ammonium sulfate. Crystals were allowed to grow in the cold room (4°C) for periods of 1 to 3 months. Crystals were initially characterized in the laboratory (space group $P2_1$, cell parameters listed in Table IV) and data were collected subsequently on the BM02 and ID14-EH4 synchrotron beamlines at the European Synchrotron Radiation

Table IV. Data collection and processing statistics relevant to the two distinct steps of refinement of the model

	Native 1	Native 2
Crystallization parameters	pH 5.7 [PEG4K] 32%	pH 7.4 [PEG4K] 34% $[(NH_4)_2SO_4]$ = 100 mM
Unit cell parameters	a = 39.08 Å b = 79.65 Å c = 60.21 Å β = 91.24°	a = 39.54 Å b = 80.39 Å c = 61.31 Å β = 91.21°
Wavelength (Å)	0.98	0.93
Resolution (Å)	2.2	1.7
No. of measurements	64464	115710
No. of unique reflections	18464	39489
Wilson B factor ($Å^2$)	20.8	21.7
Coverage overall (%)	98.0 (95.3)[a]	96.7 (75.8)[a]
I/Sigma(I) all data	11.6 (3.5)[a]	16.9 (3.4)[a]
R_{sym} overall (%)[b]	6.7 (14.8)[a]	3.7 (24.7)[a]

[a] Numbers in parentheses correspond to the highest resolution shell: 2.4–2.2 Å for native 1, 1.75–1.7 Å for native 2.

[b] $R_{sym} = \Sigma[I - \langle I \rangle]/\Sigma I$, where the summation is over all symmetry equivalent reflections.

Facility (ESRF), Grenoble. Crystals were cooled at cryo-temperature and stored in solid propane as described by Vernède and Fontecilla-Camps (1999). The first native dataset was measured to a resolution of 2.2 Å. An additional 180° of data to 1.7 Å resolution were recorded later from a different cryo-cooled crystal on the beamline ID14-EH4. All data were processed, reduced and scaled using the XDS package (Kabsch, 1993). Details are given in Table IV.

Structure determination, refinement and analysis

The structure was solved in two steps. First, molecular replacement was applied using the native 1 dataset (Table IV). Rotational and translational searches were performed using the AMoRe package (Navaza, 1994) with data from 15 to 3 Å resolution and the bovine thrombin structure (1fpc from the PDB; Bernstein *et al*., 1977) as a search model. The rotational search showed a unique solution with a correlation value on intensities of 0.21. Translational search and rigid body fitting resulted in a contrasted solution with a correlation value of 0.28 and an R value of 50.7% between 15 and 3 Å resolution. Similar results were obtained with other SP models [chymotrypsin (2cga), protein C (1aut)]. Correction and refinement of the model were carried out in the areas where electron density maps clearly showed up relevant information. Rounds of refinement using X-PLOR (Brünger, 1992) were interspersed with interactive computer graphics model building using O (Jones *et al*., 1991). At this stage, the R_{work} and R_{free} were 0.413 and 0.516, respectively, using data between 15 and 2.2 Å resolution. However, at this stage it was not possible to build unambiguously the surface insertion loops, the activation peptide and the CCP module.

The high-resolution native 2 dataset (Table IV) combined with the use of the powerful WARP procedure (Perrakis *et al*., 1999) allowed the complete structure of the C1s CCP2–SP to be solved in a second step. The main chain was automatically built for 279 residues whereas the side chains were constructed using O (Jones *et al*., 1991) into a good electron density map. This latter step was restricted in most cases to the manual choice of the correct rotamer. Twenty additional residues were built manually, along with three carbohydrate residues and two sulfate ions that were clearly defined in the electron density map. Refinement was then carried out using Refmac (Murshudov *et al*., 1997), starting from an R_{work} of 0.296 and an R_{free} of 0.309 for data between 15 and 1.7 Å resolution. ARP (Lamzin and Wilson, 1993) was used for solvent building. A number of residues displayed alternative conformations and were modeled as such. Some residues poorly defined in the electron density maps were truncated to alanines (Table I). The stereochemistry of the structure was assessed with PROCHECK (Laskowski *et al*., 1993). Details and statistics of the final model are presented in Table I. The structure factors of the high-resolution dataset 2 and the atomic coordinates have been deposited in the Protein Data Bank with accession code 1elv.

The detailed analysis of the contacts between the CCP module and the SP domain was carried out using Hbplus (McDonald and Thornton, 1994) and LIGPLOT (Wallace *et al*., 1995). The surface areas were calculated with NACCESS (Hubbard and Thornton, 1993). Figures were generated with several combined uses of MOLSCRIPT (Kraulis, 1991), BOBSCRIPT, GRASP (Nicholls *et al*., 1991) and Raster3D (Merritt and Bacon, 1997). The sequence alignments were the same as described previously (Gaboriaud *et al*., 1998) or extracted from the recent version of the pfam database (Bateman *et al*., 1999).

Quality of the model

Most of the structure was well-defined in a good high resolution electron density map. The final refined model comprises residues 342–379, 382–417, 423–477, 487–582 and 595–668. Of the 303 amino acid residues defined in the structure, 99.6% are in the favorable or additionally allowed regions in the Ramachandran plot. Only Thr646 is in the generously allowed region. Five segments (380–381, 418–422, 478–486, 583–594 and 669–673) display high flexibility or disordered conformations.

Acknowledgements

The authors are grateful to M.Roth and W.Burmeister as local contacts at the ESRF. The quality of the beam-lines they developed was critical for this study. The authors are grateful to D.Housset for the collection of the high resolution dataset and for numerous stimulating discussions. We thank A.Perrakis for his advice and help in the installation of the WARP software. The work of M.Qian and C.Olivier at early stages of the C1s project is also acknowledged. This work was supported by the Commissariat à l'Energie Atomique and the Centre National de la Recherche Scientifique.

References

Arlaud,G.J. and Thielens,N. (1993) Human complement serine proteases C1r and C1s and their proenzymes. *Methods Enzymol*., **223**, 61–82.

Arlaud,G.J., Colomb,M.G. and Gagnon,J. (1987) A functional model of the human C1 complex. *Immunol. Today*, **8**, 106–111.

Arlaud,G.J., Volanakis,J.E., Thielens,N.M., Narayana,S.V.L., Rossi,V. and Xu,Y. (1998) The atypical serine proteases of the complement system. *Adv. Immunol*., **69**, 249–307.

Banner,D.W., D'Arcy,A., Chene,C., Winkler,F.K., Guha,A., Konigsberg,W.H., Nemerson,Y. and Kirchhofer,D. (1996) The crystal structure of the complex of blood coagulation factor VIIa with soluble tissue factor. *Nature*, **380**, 41–46.

Barlow,P.N., Norman,D.G., Steinkasserer,A., Horne,T.J., Pearce,J., Driscoll,P.C., Sim,R.B. and Campbell,I.D. (1992) Solution structure of the fifth repeat of factor H: a second example of the complement control protein module. *Biochemistry*, **31**, 3626–3634.

Barlow,P.N., Steinkasserer,A., Norman,D.G., Kieffer,B., Wiles,A.P., Sim,R.B. and Campbell,I.D. (1993) Solution structure of a pair of complement modules by nuclear magnetic resonance. *J. Mol. Biol*., **232**, 268–284.

Bateman,A., Birney,E., Durbin,R., Eddy,S.R., Finn,R.D. and Sonnhammer,E.L. (1999) Pfam 3.1: 1313 multiple alignments and profile HMMs match the majority of proteins. *Nucleic Acids Res*., **27**, 260–262.

Bernstein,F.C., Koetzle,T.F., Williams,G.J., Meyer,E.E.,Jr, Brice,M.D., Rodgers,J.R., Kennard,O., Shimanouchi,T. and Tasumi,M. (1977) The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol*., **112**, 535–542.

Bode,W. and Huber,R. (1992) Natural protein proteinase inhibitors and their interaction with proteinases. *Eur. J. Biochem*., **204**, 433–451.

Bode,W., Brandstetter,H., Mather,T. and Stubbs,M.T. (1997) Comparative analysis of haemostatic proteinases: structural aspects of thrombin, factor Xa, factor IXa and protein C. *Thromb. Haemost*., **78**, 501–511.

Bork,P. and Beckmann,G. (1993) The CUB domain. A widespread module in developmentally regulated proteins. *J. Mol. Biol*., **231**, 539–545.

Brandstetter,H., Bauer,M., Huber,R., Lollar,P. and Bode,W. (1995) X-ray structure of clotting factor IXa: active site and module structure related to Xase activity and hemophilia B. *Proc. Natl Acad. Sci. USA*, **92**, 9796–9800.

P7

Brünger,A.T. (1992) *X-PLOR version 3.1, a System for X-ray Crystallography and NMR*. Yale University Press, New Haven, CT.

Burmeister,W.P. (2000) Structural changes in a cryo-cooled protein crystal due to radiation damage. *Acta Crystallogr. D*, **56**, 328–341.

Casasnovas,J.M., Larvie,M. and Stehle,T. (1999) Crystal structure of two CD46 domains reveals an extended measles virus-binding surface. *EMBO J.*, **18**, 2911–2922.

Cooper,N.R. (1985) The classical complement pathway: activation and regulation of the first complement component. *Adv. Immunol.*, **37**, 151–216.

Dalmasso,A.P. (1992) The complement system in xenotransplantation. *Immunopharmacology*, **24**, 149–160.

Dang,Q.D. and Di Cera,E. (1996) Residue 225 determines the Na$^+$-induced allosteric regulation of catalytic activity in serine proteases. *Proc. Natl Acad. Sci. USA*, **93**, 10653–10656.

Dobryszycka,W. (1997) Biological functions of haptoglobin—new pieces to an old puzzle. *Eur. J. Clin. Chem. Clin. Biochem.*, **35**, 647–654.

Endo,Y., Takahashi,M., Nakao,M., Saiga,H., Sekine,H., Matsushita,M., Nonaka,M. and Fujita,T. (1998) Two lineages of mannose-binding lectin-associated serine protease (MASP) in vertebrates. *J. Immunol.*, **161**, 4924–4930.

Gaboriaud,C., Rossi,V., Fontecilla-Camps,J.C. and Arlaud,G.J. (1998) Evolutionary conserved rigid module–domain interactions can be detected at the sequence level: the examples of complement and blood coagulation proteases. *J. Mol. Biol.*, **282**, 459–470.

Gans,P., Rossi,V., Gaboriaud,C., Bally,I., Hernandez,J.F., Blackledge,M.J. and Arlaud,G.J. (1998) NMR structures of the C-terminal end of human complement serine protease C1s. *Cell. Mol. Life Sci.*, **54**, 171–178.

Guinto,E.R., Caccia,S., Rose,T., Fütterer,K., Waksman,G. and Di Cera,E. (1999) Unexpected crucial role of residue 225 in serine proteases. *Proc. Natl Acad. Sci. USA*, **96**, 1852–1857.

Hubbard,S.J. and Thornton,J.M. (1993) 'NACCESS', Computer Program, Department of Biochemistry and Molecular Biology, University College London, UK.

Hwang,P.K. and Greer,J. (1980) Interaction between hemoglobin subunits in the hemoglobin.haptoglobin complex. *J. Biol. Chem.*, **255**, 3038–3041.

Jing,H., Macon,K., Moore,D., DeLucas,L.J., Volanakis,J.E. and Narayana,S.V.L. (1999) Structural basis of profactor D activation: from a highly flexible zymogen to a novel self-inhibited serine protease, complement factor D. *EMBO J.*, **18**, 804–814.

Jones,T.A, Zou,J.-Y., Cowan,S.W. and Kjeldgaard,M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A*, **47**, 110–119.

Kabsch,W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.*, **26**, 795–800.

Kirkitadze,M.D. *et al.* (1999) Independently melting modules and highly structured intermodular junctions within complement receptor type 1. *Biochemistry*, **38**, 7019–7031.

Kraulis,P.J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946–950.

Lacroix,M., Rossi,V., Gaboriaud,C., Chevallier,S., Jaquinod,M., Thielens,N.M., Gagnon,J. and Arlaud,G.J. (1997) Structure and assembly of the catalytic region of human complement protease C1r: a three-dimensional model based on chemical cross-linking and homology modeling. *Biochemistry*, **36**, 6270–6282.

Lamzin,V. and Wilson,K. (1993) Automated refinement of protein model. *Acta Crystallogr. D*, **49**, 129–147.

Laskowski,R.A., MacArthur,M.W., Moss,D.S. and Thornton,J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.*, **26**, 283–291.

Mackinnon,C.M., Carter,P.E., Smyth,S.J., Dunbar,B. and Fothergill,J.E. (1987) Molecular cloning of cDNA for human complement component C1s. The complete amino acid sequence. *Eur. J. Biochem.*, **169**, 547–553.

Mather,T., Oganessyan,V., Hof,P., Huber,R., Foundling,S., Esmon,C. and Bode,W. (1996) The 2.8 Å crystal structure of Gla-domainless activated protein C. *EMBO J.*, **15**, 6822–6831.

Matsushita,M., Endo,Y., Nonaka,M. and Fujita,T. (1998) Complement-related serine proteases in tunicates and vertebrates. *Curr. Opin. Immunol.*, **10**, 29–35.

McDonald,I.K. and Thornton,J.M. (1994) Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.*, **238**, 777–793.

Merritt,E.A. and Bacon,D.J. (1997) Raster3D photorealistic molecular graphics. *Methods Enzymol.*, **277**, 505–524.

Mevorach,D., Mascarenhas,J.O., Gershov,D. and Elkon,K.B. (1998) Complement-dependent clearance of apoptotic cells by human macrophages. *J. Exp. Med.*, **188**, 2313–2320.

Murshudov,G.N., Vagin,A.A. and Dodson,E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D*, **53**, 240–255.

Muta,T., Miyata,T., Misumi,Y., Tokunaga,F., Nakamura,T., Toh,Y., Ikehara,Y. and Iwanaga,S. (1991) *Limulus* factor C. An endotoxin-sensitive serine protease zymogen with a mosaic structure of complement-like, epidermal growth factor-like and lectin-like domains. *J. Biol. Chem.*, **26**, 6554–6561.

Navaza,J. (1994) AMoRe: an automated package for molecular replacement. *Acta Crystallogr. A*, **50**, 157–163.

Nicholls,A., Sharp,K.A. and Honig,B. (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins*, **11**, 281–296.

Padmanabhan,K., Padmanabhan,K.P., Tulinsky,A., Park,C.H., Bode,W., Huber,R., Blankenship,D.T., Cardin,A.D. and Kisiel,W. (1993) Structure of human Des (1–45) factor Xa at 2.2 Å resolution. *J. Mol. Biol.*, **232**, 947–966.

Parry,M.A, Fernandez-Catalan,C., Bergner,A., Huber,R., Hopfner,K.P., Schlott,B., Guhrs,K.H. and Bode,W. (1998) The ternary microplasmin–staphylokinase–microplasmin complex is a proteinase–cofactor–substrate complex in action. *Nature Struct. Biol.*, **10**, 917–923.

Perona,J.J. and Craik,C.S. (1997) Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. *J. Biol. Chem.*, **272**, 29987–29990.

Perrakis,A., Morris,R. and Lamzin,V.S. (1999) Automated protein model building combined with iterative structure refinement. *Nature Struct. Biol.*, **6**, 458–463.

Pétillot,Y. *et al.* (1995) Analysis of the N-linked oligosaccharides of human C1s using electrospray ionisation mass spectrometry. *FEBS Lett.*, **358**, 323–328.

Reid,K.B.M. and Day,A.J. (1989) Structure–function relationships of the complement components. *Immunol. Today*, **10**, 177–180.

Rogers,J. *et al.* (1992) Complement activation by β-amyloid in Alzheimer disease. *Proc. Natl Acad. Sci. USA*, **89**, 10016–10020.

Rossi,V., Gaboriaud,C., Lacroix,M., Ulrich,J., Fontecilla-Camps,J.C., Gagnon,J. and Arlaud,G.J. (1995) Structure of the catalytic region of human complement protease C1s: study by chemical cross-linking and three-dimensional homology modeling. *Biochemistry*, **34**, 7311–7321.

Rossi,V., Bally,I., Thielens,N.M., Esser,A.F. and Arlaud,G.J. (1998) Baculovirus-mediated expression of truncated modular fragments from the catalytic region of human complement serine protease C1s. Evidence for the involvement of both complement control protein modules in the recognition of the C4 protein substrate. *J. Biol. Chem.*, **273**, 1232–1239.

Sato,T., Endo,Y., Matsushita,M. and Fujita,T. (1994) Molecular characterization of a novel serine protease involved in activation of the complement system by mannose-binding protein. *Int. Immunol.*, **6**, 665–669.

Schumaker,V.N., Hanson,D.C., Kilchherr,E., Phillips,M.L. and Poon,P.H. (1986) A molecular mechanism for the activation of the first component of complement by immune complexes. *Mol. Immunol.*, **23**, 557–565.

Smith,A.B., Esko,J.D. and Hajduk,S.L. (1995) Killing of trypanosomes by the human haptoglobin related protein. *Science*, **268**, 284–286.

Thiel,S. *et al.* (1997) A second serine protease associated with mannan-binding lectin that activates complement. *Nature*, **386**, 506–510.

Thielens,N.M., Aude,C.A., Lacroix,M.B., Gagnon,J. and Arlaud,G.J. (1990) Ca^{2+} binding properties and Ca^{2+}-dependent interactions of the isolated NH$_2$-terminal α fragments of human complement proteases C1r and C1s. *J. Biol. Chem.*, **265**, 14469–14475.

Tosi,M., Duponchel,C., Meo,T. and Julier,C. (1987) Complete cDNA sequence of human complement C1s and close physical linkage of the homologous genes *C1s* and *C1r*. *Biochemistry*, **26**, 8516–8524.

Tosi,M., Duponchel,C., Meo,T. and Couture-Tosi,E. (1989) Complement genes *C1r* and *C1s* feature an intronless serine protein domain closely related to haptoglobin. *J. Mol. Biol.*, **208**, 709–714.

van de Locht,A., Bode,W., Huber,R., Le Bonniec,B.F., Stone,S.R., Esmon,C.T. and Stubbs,M.T. (1997) The thrombin E192Q–BPTI complex reveals gross structural rearrangements: implications for the interaction with antithrombin and thrombomodulin. *EMBO J.*, **16**, 2977–2984.

P7

Vernède,X. and Fontecilla-Camps,J.C. (1999) A method to stabilize reduced and/or gas-treated protein crystals by flash-cooling under a controlled atmosphere. *J. Appl. Crystallogr.*, **32**, 505–509.

Volanakis,J.E. and Narayana,S.V.L. (1996) Complement factor D, a novel serine protease. *Protein Sci.*, **5**, 553–564.

Wallace,A.C., Laskowski,R.A. and Thornton,J.M. (1995) LIGPLOT: a program to generate schematic diagrams of protein–ligand interactions. *Protein Eng.*, **8**, 127–134.

Weiss,V., Fauser,C. and Engel,J. (1986) Functional model of subcomponent C1 of human complement. *J. Mol. Biol.*, **189**, 573–581.

Wiles,A.P., Shaw,G., Bright,J., Perczel,A., Campbell,I.D. and Barlow,P.N. (1997) NMR studies of a viral protein that mimics the regulators of complement activation. *J. Mol. Biol.*, **272**, 253–265.

Zahedi,R., Wisnieski,J. and Davis,A.E.,III (1997) Role of the P2 residue of complement 1 inhibitor (Ala443) in determination of target protease specificity: inhibition of complement and contact system proteases. *J. Immunol.*, **159**, 983–988.

Received January 26, 2000; accepted February 14, 2000

P8

Interaction Properties of Human Mannan-Binding Lectin (MBL)-Associated Serine Proteases-1 and -2, MBL-Associated Protein 19, and MBL[1,2]

Nicole M. Thielens,[3]* Sándor Cseh,* Steffen Thiel,† Thomas Vorup-Jensen,† Véronique Rossi,* Jens C. Jensenius,† and Gérard J. Arlaud*

The mannan-binding lectin (MBL) activation pathway of complement plays an important role in the innate immune defense against pathogenic microorganisms. In human serum, two MBL-associated serine proteases (MASP-1, MASP-2) and MBL-associated protein 19 (MAp19) were found to be associated with MBL. With a view to investigate the interaction properties of these proteins, human MASP-1, MASP-2, MAp19, as well as the N-terminal complement subcomponents C1r/C1s, Uegf, and bone morphogenetic protein-1-epidermal growth factor (CUB-EGF) segments of MASP-1 and MASP-2, were expressed in insect or human kidney cells, and MBL was isolated from human serum. Sedimentation velocity analysis indicated that the MASP-1 and MASP-2 CUB-EGF segments and the homologous protein MAp19 all behaved as homodimers (2.8–3.2 S) in the presence of Ca^{2+}. Although the latter two dimers were not dissociated by EDTA, their physical properties were affected. In contrast, the MASP-1 CUB-EGF homodimer was not sensitive to EDTA. The three proteins and full-length MASP-1 and MASP-2 showed no interaction with each other as judged by gel filtration and surface plasmon resonance spectroscopy. Using the latter technique, MASP-1, MASP-2, their CUB-EGF segments, and MAp19 were each shown to bind to immobilized MBL, with K_D values of 0.8 nM (MASP-2), 1.4 nM (MASP-1), 13.0 nM (MAp19 and MASP-2 CUB-EGF), and 25.7 nM (MASP-1 CUB-EGF). The binding was Ca^{2+}-dependent and fully sensitive to EDTA in all cases. These data indicate that MASP-1, MASP-2, and MAp19 each associate as homodimers, and individually form Ca^{2+}-dependent complexes with MBL through the CUB-EGF pair of each protein. This suggests that distinct MBL/MASP complexes may be involved in the activation or regulation of the MBL pathway. ***The Journal of Immunology,*** **2001, 166: 5068–5077.**

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

Studies performed over the past decade have led to the discovery of a novel complement activation pathway named the mannan-binding lectin (MBL)[4] or lectin pathway. This pathway may play an important role in innate immunity, especially against encapsulated microorganisms that have the ability to evade activation of the classical and alternative pathways of complement (2–4). The recognition component of this pathway is MBL, a member of the collectin family (5) that binds through C-type lectin domains to neutral carbohydrates on the surface of various pathogenic microorganisms (6). Two modular MBL-associated serine proteases, MASP-1 (2) and MASP-2 (3), were initially found to be associated to MBL. A third protein component MBL-associated protein 19 (MAp19), generated by alternative splicing of the *MASP-2* gene (7) and, more recently, a further protease MASP-3 (8) were also shown to be associated with MBL.

MASP-1 and MASP-2 exhibit a modular structure homologous to that of C1r and C1s, the proteases of the C1 complex of complement (9), with an N-terminal module originally found in complement subcomponents C1r/C1s, Uegf, and bone morphogenetic protein-1 (CUB module) (10) followed by an epidermal growth factor (EGF)-like module, a second CUB module, two contiguous complement control protein modules (11), and a C-terminal chymotrypsin-like serine protease domain (see Fig. 1). MAp19 is a nonenzymic protein comprising the N-terminal CUB-EGF segment of MASP-2 followed by four unique residues (7, 12). As yet, very little is known about the stoichiometry, the assembly, and the activation mechanisms of the MBL-MASP complex(es). An analysis of the interaction between recombinant fragments of MASP-1 and MASP-2, and MBL from the rat was published very recently, providing evidence that both MASP-1 and MASP-2 associate with MBL in a Ca^{2+}-dependent manner through interactions involving the N-terminal CUB-EGF-CUB moiety of the proteases (13).

This study was undertaken to investigate the interaction properties of human MASP-1, MASP-2, and MAp19, as well as the ability of each protein to associate with MBL. For this purpose, MASP-1, MASP-2, the CUB-EGF segments of each protease, and MAp19 were expressed using insect or human kidney cells, and MBL was purified from human serum. Sedimentation velocity analysis and surface plasmon resonance spectroscopy demonstrated that MASP-1, MASP-2, and MAp19 each associate as homodimers

---

*Laboratoire d'Enzymologie Moléculaire, Institut de Biologie Structurale Jean-Pierre Ebel (Commissariat à l'Energie Atomique-Centre National de la Recherche Scientifique), Grenoble, France; and †Department of Medical Microbiology and Immunology, University of Aarhus, Aarhus, Denmark

Received for publication November 22, 2000. Accepted for publication February 14, 2001.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] This work was supported in part by the Commissariat à l'Energie Atomique, the Center National de la Recherche Scientifique, and the Danish Medical Research Council. S.C. is the recipient of a FEBS long-term fellowship.

[2] A preliminary report of this study was presented at the XVIIIth International Complement Workshop in Salt Lake City, Utah, July 23–27, 2000.

[3] Address correspondence and reprint requests to Dr. Nicole Thielens, Laboratoire d'Enzymologie Moléculaire, Institut de Biologie Structurale Jean-Pierre Ebel, 41 rue Jules Horowitz, 38027 Grenoble Cedex 1, France. E-mail address: thielens@ibs.fr

[4] Abbreviations used in this paper: MBL, mannan-binding lectin; the nomenclature of protein modules defined by Bork and Bairoch (1); CUB module, module originally found in complement subcomponents C1r/C1s, Uegf, and bone morphogenetic protein-1; EGF, epidermal growth factor; MAp19, 19-kDa MBL-associated protein; MASP, MBL-associated serine protease; RU, resonance unit(s); S, sedimentation coefficient; D, diffusion constant.

Copyright © 2001 by The American Association of Immunologists 0022-1767/01/$02.00

P9

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 44, Issue of November 2, pp. 40880–40887, 2001
Printed in U.S.A.
© 2001 by The American Society for Biochemistry and Molecular Biology, Inc.

Substrate Specificities of Recombinant Mannan-binding Lectin-associated Serine Proteases-1 and -2*

Received for publication, June 26, 2001, and in revised form, August 27, 2001
Published, JBC Papers in Press, August 29, 2001, DOI 10.1074/jbc.M105934200

Véronique Rossi‡, Sandor Cseh‡§, Isabelle Bally‡, Nicole M. Thielens‡, Jens C. Jensenius¶, and Gérard J. Arlaud‡‖

From the ‡Laboratoire d'Enzymologie Moléculaire, Institut de Biologie Structurale Jean-Pierre Ebel (CEA-CNRS), 41 rue Jules Horowitz, 38027 Grenoble Cedex 1, France and the ¶Department of Medical Microbiology and Immunology, University of Aarhus, DK 8000, Aarhus, Denmark

Mannan-binding lectin (MBL)-associated serine proteases-1 and 2 (MASP-1 and MASP-2) are homologous modular proteases that each interact with MBL, an oligomeric serum lectin involved in innate immunity. To precisely determine their substrate specificity, human MASP-1 and MASP-2, and fragments from their catalytic regions were expressed using a baculovirus/insect cells system. Recombinant MASP-2 displayed a rather wide, C1s-like esterolytic activity, and specifically cleaved complement proteins C2 and C4, with relative efficiencies 3- and 23-fold higher, respectively, than human C1s. MASP-2 also showed very weak C3 cleaving activity. Recombinant MASP-1 had a lower and more restricted esterolytic activity. It showed marginal activity toward C2 and C3, and no activity on C4. The enzymic activity of both MASP-1 and MASP-2 was specifically titrated by C1 inhibitor, and abolished at a 1:1 C1 inhibitor:protease ratio. Taken together with previous findings, these and other data strongly support the hypothesis that MASP-2 is the protease that, in association with MBL, triggers complement activation via the MBL pathway, through combined self-activation and proteolytic properties devoted to C1r and C1s in the C1 complex. In view of the very low activity of MASP-1 on C3 and C2, our data raise questions about the implication of this protease in complement activation.

Mannan-binding lectin (MBL)[1] is an oligomeric C-type lectin that recognizes arrays of neutral carbohydrates such as mannose and *N*-acetylglucosamine on the surface of pathogenic microorganisms (2). This selectivity endows MBL with the ability to discriminate self from infectious non-self, and confers this "ante-antibody" a major role in innate immunity, as underlined by numerous clinical reports indicating that MBL deficiency is linked with increased susceptibility to infectious diseases (3–5). In addition to its role as an opsonin (3), MBL has devised the ability to associate to several modular proteases termed MASPs (MBL-associated serine proteases) (6–9). A single MASP entity was initially identified, and characterized as a protease with the ability to cleave complement proteins C4, C2, and C3 (7, 10). Further studies by Thiel *et al.* (6) revealed that MASP was indeed a mixture of two related but distinct proteases, MASP-1 and MASP-2, and that only the latter had the ability to cleave C4. A third protein component MAp19, arising from alternative splicing of the MASP-2 gene (11, 12), and very recently a further protease MASP-3 (13) were also shown to be associated with MBL.

MASP-1 and MASP-2 show a domain organization identical to that of C1r and C1s, the enzymatic components of the C1 complex of complement (14), with an N-terminal CUB module (15) followed by an epidermal growth factor-like module, a second CUB module, two contiguous CCP modules (16), and a C-terminal chymotrypsin-like serine protease domain (see Fig. 1). By analogy with human C1s, it may be anticipated that the proteolytic activity and specificity of the MASPs is defined by the two CCP modules together with the serine protease domain (17), and that the latter forms a rigid association with the preceding, second CCP module (18). Comparative analysis of the cDNAs of C1r, C1s, and the MASPs in different animal species reveal that these fall into two groups (19, 20). In the smaller and probably more ancient group comprising MASP-1 and the ascidian MASPs, the active site serine is encoded by a "TCN" codon (where N is A, T, G, or C), and the histidine-loop disulfide bridge is present. In the larger group encompassing C1r, C1s, MASP-2, and most of the known animal MASPs, the active site serine is encoded by an "AGY" codon (where Y is T or C), and the histidine-loop is missing. The only available information dealing with the substrate specificity of MASP-1 and MASP-2 has been obtained on proteases isolated from human serum, and is somewhat controversial. Thus, whereas it is now widely accepted that MASP-2 cleaves both C4 and C2 (6, 21, 22), the observation that MASP-1 can cleave C3 and hence directly trigger complement activation (21, 23) is debated (22, 24). In addition, the relative efficiency of MASP-2 with respect to C4 and C2 cleavage has not been assessed.

The objective of the present work was to produce recombinant human MASP-1 and MASP-2 and catalytic fragments thereof, to precisely determine their substrate specificity on both protein substrates and synthetic esters, and to measure the kinetic parameters of their activity. Our data reveal that

---

* This work was supported in part by the Commissariat à l'Energie Atomique, the Centre National de la Recherche Scientifique, and the Danish Medical Research Council. A preliminary report of this study was presented at the Innate Immunity Workshop in Santorini, Greece, October 11–15, 2000. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Recipient of a FEBS long-term fellowship.

‖ To whom correspondence should be addressed. Tel.: 33-4-38-78-49-81; Fax: 33-4-38-78-54-94; E-mail: arlaud@ibs.fr.

[1] The abbreviations used are: MBL, mannan-binding lectin; MASP, MBL-associated serine protease; SP, serine protease domain; Tos-Arg-OMe; *p*-tosyl-L-arginine methyl ester; Z-Gly-Arg-S-Bzl, *N*-carboxybenzyloxyglycine-L-arginine thiobenzyl ester; Ac-Gly-Lys-OMe, *N*-acetylglycine-L-lysine methyl ester; Bz-Arg-OEt, N^{α}-benzoyl-L-arginine ethyl ester; CCP module, complement control protein module; CUB module, protein module originally found in complement subcomponents C1r/C1s, Uegf, and bone morphogenetic protein-1; DFP, diisopropyl fluorophosphate; PAGE, polyacrylamide gel electrophoresis; α_2M, α_2-macroglobulin; the nomenclature of protein modules is that defined by Bork and Bairoch (1).

Downloaded from http://www.jbc.org/ by guest on March 23, 2015

P9

MASP-2 cleaves C4 much more efficiently than does C1s, emphasizing the physiological relevance of MASP-2 with respect to complement activation. In contrast, recombinant MASP-1 shows only marginal C3 cleaving activity, raising questions about its involvement in complement activation.

EXPERIMENTAL PROCEDURES

Materials—Diisopropyl phosphorofluoridate was from Acros Organics, Noisy le Grand, France. The plasmids containing the full-length MASP-1 and MASP-2 cDNAs were obtained as described previously (6, 25). Oligonucleotides were obtained from Oligoexpress, Paris, France. Trypsin was obtained from Sigma, Saint Quentin Fallavier, France. Z-Gly-Arg-S-Bzl was from Enzyme Systems Products, Livermore, CA. Ac-Gly-Lys-OMe, Bz-Arg-OEt, and Tos-Arg-OMe were obtained from Sigma.

Proteins—MBL was isolated from human plasma according to the procedure described by Tan *et al.* (26), with a further purification step using ion-exchange chromatography, as described in Thielens *et al.* (27). Recombinant full-length MASP-1 was expressed using a baculovirus insect cell system and purified by ion-exchange chromatography and affinity chromatography on an UltraLink™-MBL column, as described previously (27). Activated C1s and complement proteins C4, C2, and C3 were purified from human plasma according to published procedures (28–31). Partially purified C5 was prepared as described by Al Salihi *et al.* (31), and partially purified factor B was obtained from the flow-through fraction of the last C2 purification step on C4b-Sepharose (30). C1 inhibitor was purified from human plasma essentially as described in Ref. 32, except that concanavalin A-agarose (Sigma) was used instead of concanavalin A-Sepharose. Purified human α_2-macroglobulin (α_2M) was kindly provided by L. Sottrup-Jensen (University of Aarhus). The concentrations of purified proteins were determined using the following absorption coefficients ($A^{1\%}_{1\,cm}$ at 280 nm) and molecular weights: C1s, 14.5 and 79,800; C4, 8.3 and 205,000; C2, 8.9 and 102,000; C3, 10.0 and 185,000 (29, 30, 33). The absorption coefficients ($A^{1\%}_{1\,cm}$ at 280 nm) used for the recombinant fragments CCP1/2-SP and CCP2-SP of MASP-2 (18.3 and 19.2, respectively) were calculated from the number of Trp, Tyr, and disulfides according to the method of Edelhoch (34), using molecular weight values of 42,710 and 35,330, calculated from the sequence (6). Because of the low amounts of material recovered, the concentrations of full-length MASP-1 and MASP-2 and of the CCP1/2-SP fragment of MASP-1 were estimated on the basis of Coomassie Blue staining after SDS-PAGE analysis using appropriate internal standards and respective molecular weights of 82,000, 74,200, and 45,200 (6, 9).

Construction of Expression Plasmids Containing the CCP1/2-SP Fragment of MASP-1—A DNA fragment encoding residues 299–699 of human MASP-1 was amplified by polymerase chain reaction using the Vent$_R$ polymerase and the MASP-1b/pDR$_2\Delta$EF1α vector (25) as a template, according to established procedures. The sequence of the sense primer (5′-GAAGATCTCAATGAGTGCCCAGAGCT-3′) introduced a *Bgl*II restriction site (underlined) and allowed in-frame cloning with the melittin signal peptide of the pNT-Bac expression vector (17). The antisense primer (5′-GGAATTC**TCA**GTTCCTCACTCC-3′) introduced a stop codon (bold face) followed by an *Eco*RI site (underlined) at the 3′ end of the fragment. The polymerase chain reaction product was digested with *Bgl*II and *Eco*RI, purified, and cloned into the *Bam*HI/*Eco*RI sites of the pNT-Bac vector.

Construction of Expression Plasmids Containing Full-length MASP-2 and MASP-2-derived Fragments—The DNA fragment encoding the MASP-2 signal peptide plus the mature protein (amino acids 1–671) was excised from the pBS-MASP-2 vector (6) by digestion with *Xho*I and *Eco*RI and cloned into the corresponding restriction sites of the pFast-Bac1 baculovirus transfer vector (Life Technologies, Inc.). DNA fragments encoding residues 298–686 (CCP1/2-SP construct) and residues 365–686 (CCP2-SP construct) of human MASP-2 were amplified by polymerase chain reaction using the Vent$_R$ polymerase and the pFast-Bac MASP-2 vector as a template. The sequences of the sense primers (5′-CGGGATCCCCAGCCTTGCCCTTATCC-3′ for the CCP1/2-SP construct and 5′-CGGGATCCCGACTGTGGCCCTCCTG-3′ for the CCP2-SP construct) introduced a *Bam*HI restriction site (underlined) and allowed in-frame cloning with melittin signal peptide of the pNT-Bac expression vector. The antisense primer (5′-GGAATTC**TTA**AAAATCACTAATTATGTTCTC-3′) was identical for both constructs and introduced a stop codon (bold face) followed by an *Eco*RI site (underlined). Both polymerase chain reaction fragments were digested with *Eco*RI and *Bam*HI, purified, and cloned into the corresponding sites of the pNT-Bac vector. All DNA constructs used in this study were checked for the absence of mutations by double-stranded DNA sequencing (Genome Express, Grenoble).

Cells and Viruses—The insect cells *Spodoptera frugiperda* (Ready-Plaque Sf9 cells from Novagen) and *Trichoplusia ni* (High Five™) were routinely grown and maintained as described previously (27). Recombinant baculoviruses were generated using the Bac-to-Bac™ system (Life Technologies, Inc.). The bacmid DNA was purified using the Qiagen midiprep purification system (Qiagen S.A., Courtaboeuf, France) and used to transfect Sf9 insect cells utilizing cellfectin in Sf900 II SFM medium (Life Technologies, Inc.) as described in the manufacturer's protocol. Recombinant virus particles were collected 4 days later, titrated by virus plaque assay, and amplified as described by King and Possee (35).

Protein Production and Purification—High Five cells (1.75×10^7 cells/175-cm^2 tissue culture flask) were infected with the recombinant viruses at a multiplicity of infection of 2–3 in Sf900 II SFM medium at 28 °C for 48 h (MASP-2 fragments) or 72 h (full-length MASP-2 and CCP1/2-SP fragment of MASP-1). Culture supernatants were collected by centrifugation.

The supernatant containing MASP-2 was dialyzed against 50 mM NaCl, 1 mM EDTA, 50 mM triethanolamine hydrochloride, pH 8.5, and loaded onto a Q-Sepharose Fast Flow column (Amersham Pharmacia Biotech) (2.8 × 10 cm) equilibrated in the same buffer. Elution was carried out by applying a 800-ml linear gradient from 50 to 500 mM NaCl in the same buffer. Fractions containing the recombinant protein were identified by Western blot analysis, and dialyzed against 50 mM NaCl, 1 mM EDTA, 50 mM triethanolamine hydrochloride, pH 8.0. Further purification was achieved by ion-exchange chromatography on a Mono-Q HR 5/5 column (Amersham Pharmacia Biotech) equilibrated in the same buffer. Elution was carried out with a linear NaCl gradient from 50 to 500 mM in 90 min.

The supernatants containing the CCP1/2-SP and CCP2-SP fragments of MASP-2 were dialyzed against 5 mM EDTA, 20 mM Na$_2$HPO$_4$, pH 8.6, and loaded onto a Q-Sepharose Fast Flow column (2.8 × 10 cm) equilibrated in the same buffer. Elution was carried out by applying a 700-ml linear gradient from 0 to 350 mM NaCl in the same buffer. Fractions containing the recombinant proteins were identified by Western blot analysis, dialyzed against 1.5 M (NH$_4$)$_2$SO$_4$, 0.1 M Na$_2$HPO$_4$, pH 7.4, and further purified by high-pressure hydrophobic interaction chromatography on a TSK-Phenyl 5PW column (Beckman) equilibrated in the same buffer. Elution was carried out by decreasing the (NH$_4$)$_2$SO$_4$ concentration from 1.5 M to 0 in 30 min. Both purified recombinant fragments were dialyzed against 145 mM NaCl, 50 mM triethanolamine hydrochloride, pH 7.4, concentrated up to 0.2 mg/ml by ultrafiltration, and stored at −20 °C.

The supernatant containing the CCP1/2-SP fragment of MASP-1 was dialyzed against 50 mM Na acetate, pH 5.1, and loaded on a SP-Sepharose column (Amersham Pharmacia Biotech) (2.8 × 8 cm) equilibrated in the same buffer. Elution was performed with a 800-ml linear gradient from 0 to 600 mM NaCl. Recombinant proteins were dialyzed against 50 mM triethanolamine hydrochloride, 145 mM NaCl, pH 7.4, concentrated by ultrafiltration to 0.05–0.5 mg/ml, and stored at 0 °C.

Polyacrylamide Gel Electrophoresis and Immunoblotting—SDS-PAGE analysis was performed as described previously (36). Western blot analysis and immunodetection of the recombinant proteins were carried out as described by Rossi *et al.* (17), or using the ECL detection procedure of Amersham Pharmacia Biotech. The antibodies used were the mouse monoclonal anti-MASP-2 antibody 1.3B7, a rabbit polyclonal anti-MASP-2 antibody (37), and a rabbit anti-peptide antibody directed against the serine protease domain of MASP-1 (6).

N-terminal Sequence Determination—N-terminal sequence analyses were performed after SDS-PAGE and electrotransfer, using an Applied Biosystems model 477A protein sequencer as described previously (38).

Proteolytic Assays—The proteolytic activity of MASP-1, MASP-2, and C1s toward C2, C3, C4, C5, and factor B was measured by incubation at different enzyme:protein ratios for varying periods at 37 °C, as indicated in the text. The extent of reaction was determined after SDS-PAGE analysis under reducing conditions by gel scanning of the cleavage fragments, as described previously (17). For determination of the kinetic parameters, the concentrations of C2 ranged from 1.0 to 4 μM, and those of C4 ranged from 50 to 500 nM (MASP-2 cleavage) or from 0.5 to 4 μM (C1s cleavage). Fixed enzyme concentrations of 2 nM were used in all cases, and enzyme dilutions were performed in 50 mM triethanolamine HCl, 145 mM NaCl, pH 7.4, containing 1 mg/ml ovalbumin. The kinetic constants were determined by the Lineweaver-Burk method using linear regression analysis and are based on duplicate or triplicate measurements of initial rates at 6 separate substrate concentrations.

Downloaded from http://www.jbc.org/ by guest on March 23, 2015

P9

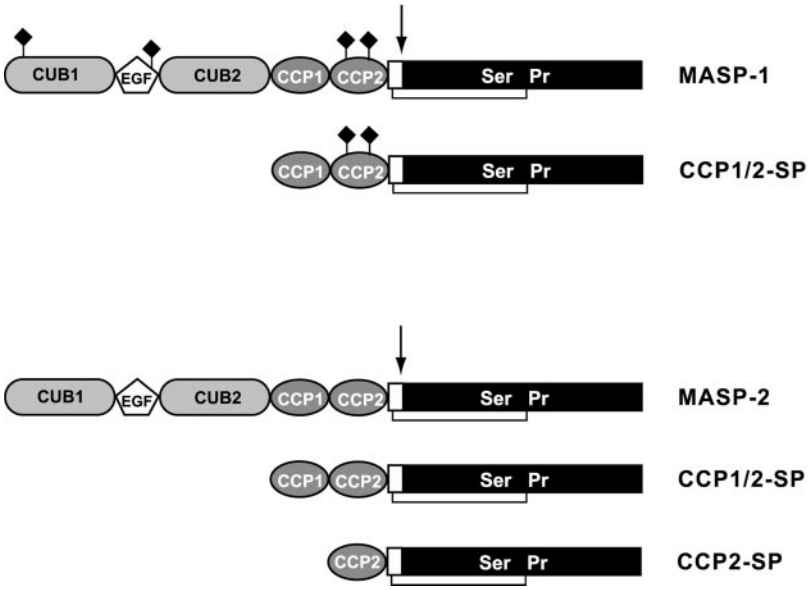


FIG. 1. **Modular structure of the recombinant proteins and fragments used in this study.** The nomenclature and symbols used for protein modules are those defined by Bork and Bairoch (1). *Ser Pr*, serine protease domain. The *arrow* indicates the Arg-Ile bond cleaved upon activation. The only disulfide bridge shown is that connecting the activation peptide to the serine protease domain. ♦, *N*-linked oligosaccharides.

All kinetic analyses were conducted in 150 mM NaCl, 5 mM EDTA, 20 mM sodium phosphate, pH 7.4.

Reactivity Toward Inhibitors—Complex formation between C1 inhibitor and MASP-1 or MASP-2 was assessed by incubation of the proteases with excess molar ratios of C1 inhibitor for 15 min at 37 °C in 50 mM triethanolamine-HCl, 145 mM NaCl, pH 7.4, followed by SDS-PAGE analysis under reducing conditions. C1 inhibitor-protease complexes were revealed by Western blot analysis using a rabbit polyclonal antibody directed against full-length MASP-2, and a rabbit anti-peptide antibody directed against the serine protease domain of MASP-1. Titration of the enzymatic activity of C1s, MASP-1, and MASP-2 by C1 inhibitor was performed by preincubating these proteases (at concentrations of 0.25 μM, 20 nM, and 0.25 μM, respectively) with increasing concentrations of C1 inhibitor (10–500 nM) for 15 min at 37 °C. The residual C2 cleaving activity of MASP-2 and C1s was then measured by incubating the enzymes (3.75 and 7.5 nM, respectively) for 15 min at 37 °C in the presence of 2.25 μM C2, followed by SDS-PAGE analysis, as described above. In the case of MASP-1, the residual esterolytic activity was measured on Bz-Arg-OEt, as described above.

The reactivity of C1s and MASP-2 toward α_2M was measured by preincubating the enzymes (0.25 μM each) for 30 min at 37 °C in the presence of increasing concentrations (3.7–300 μM) of α_2M. Residual C4 cleaving activity was then measured by incubating the enzymes (1.25 nM each) for 15 min at 37 °C in the presence of 2.25 μM C4, followed by SDS-PAGE analysis.

Esterolytic Assays—Esterolytic activities were measured on the synthetic esters Ac-Gly-Lys-OMe and Tos-Arg-OMe using a spectrophotometric assay based on the measurement of methanol released upon hydrolysis (39). Assays on the thioester Z-Gly-Arg-S-Bzl and on Bz-Arg-OEt were carried out as described by McRae *et al.* (40), and Arlaud and Thielens (39). All assays were conducted at 30 °C, at substrate concentrations of 0.1–3 mM (Ac-Gly-Lys-OMe and Tos-Arg-OMe), 0.5–3 mM (Bz-Arg-OEt), or 0.075–0.2 mM (Z-Gly-Arg-S-Bzl).

RESULTS

Production and Characterization of Recombinant MASP-2 and C-terminal Fragments from MASP-1 and MASP-2—The modular structures of MASP-1, MASP-2, and of the truncated fragments used in the present study are depicted in Fig. 1. The recombinant baculoviruses used for expression of each construct were obtained as described under "Experimental Procedures" and used to infect High Five insect cells for various periods at 28 °C. The amount of recombinant material recovered in the culture medium, as estimated by SDS-PAGE and Western blot analysis, ranged from 0.15 (MASP-2) to 2 μg/ml (CCP1/2-SP fragment of MASP-1). Protein purification was performed as described under "Experimental Procedures," using an initial ion-exchange fractionation step in all cases. Because of the low amounts of material recovered, protein detection and analysis was performed routinely using Western blot analysis rather than Coomassie Blue staining.

Due to its very low recovery, recombinant MASP-2 could not be purified to homogeneity since attempts to completely remove contaminant proteins resulted in a nearly complete loss of material. SDS-PAGE analysis of the partially purified MASP-2 fraction under nonreducing conditions revealed two bands reactive with specific antibodies: (i) a major, 80-kDa species corresponding to the full-length protease, yielding two sequences, Thr-Pro-Leu-Gly-Pro-Lys-Trp-Pro-Glu-Pro . . . and Ile-Tyr-Gly-Gly-Gln-Lys-Ala-Lys-Pro-Gly . . . , corresponding to the N-terminal ends of the mature protein and of the serine protease domain, respectively; (ii) a 45-kDa species yielding only the first sequence above, and corresponding to a truncated fragment derived from the N-terminal end of the protein. Analysis under reducing conditions indicated that full-length MASP-2 was recovered in a partially activated form, since only 20–30% of the protein migrated as a single chain, proenzyme species, whereas the remainder yielded two bands at 45 and 28 kDa, corresponding to the N-terminal A chain and the serine protease domain, respectively. Based on Coomassie Blue staining after SDS-PAGE analysis, the relative amount of full-length MASP-2 in the partially purified fraction averaged 10% of the total protein contents.

Both the CCP1/2-SP and CCP2-SP fragments of MASP-2 could be purified to homogeneity. On SDS-PAGE analysis, the CCP1/2-SP fragment migrated under nonreducing conditions as a 44-kDa band, which upon reduction split into two bands corresponding to the serine protease domain (27 kDa) and the CCP1/2 segment (17 kDa) (Fig. 2, *lanes 1* and *2*), indicating complete activation of the protease. In contrast, the shorter CCP2-SP fragment migrated as a single band of about 35 kDa both under nonreducing and reducing conditions (Fig. 2, *lanes 3* and *4*), indicating that this species had retained a single chain, proenzyme structure. N-terminal sequence analysis of the latter fragment yielded a single sequence Asp-Pro-Asp-Cys-Gly-Pro-Pro . . . corresponding to the expected N-terminal end of the fragment, whereas the activated CCP1/2 SP fragment yielded equivalent amounts of two sequences: Asp-Pro-Gln-Pro-Cys-Pro-Tyr . . . (N-terminal end) and Ile-Tyr-Gly-Gly-Gln-Lys-Ala . . . (serine protease domain). Protein staining with Coomassie Blue (not shown) yielded the same pictures as observed by Western blot analysis, indicating that no major fragment or contaminant was present in the purified preparations.

Downloaded from http://www.jbc.org/ by guest on March 23, 2015

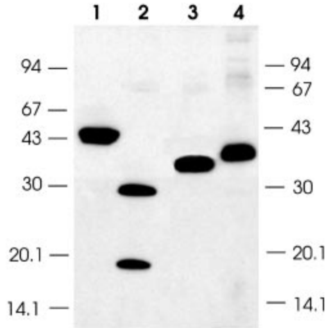


FIG. 2. **SDS-PAGE analysis of the recombinant MASP-2 fragments used in this study.** *Lanes 1* and *2*, MASP-2 CCP1/2-SP fragment (unreduced, and reduced); *lanes 3* and *4*, MASP-2 CCP2-SP fragment (unreduced, and reduced). Protein was revealed by Western blot analysis using an anti-MASP-2 polyclonal antibody. Molecular masses of unreduced and reduced standard proteins (expressed in kDa) are shown, respectively, on the *left* and *right* sides of the gel.

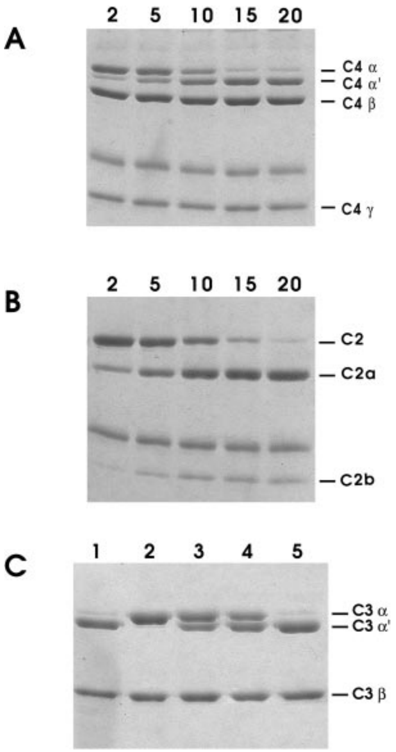


FIG. 3. **SDS-PAGE analysis of the C4, C2, and C3 cleaving activities of recombinant MASP-2.** C4 (2.25 μM) (*A*) and C2 (2.25 μM) (*B*) were incubated in the presence of the recombinant MASP-2 CCP1/2-SP fragment (1 and 2 nM, respectively) for periods of 2–20 min at 37 °C, as indicated. *C*, C3 (2 μM) was incubated for 5 min at 37 °C in the presence of trypsin (0.15 μM) (*lane 1*), or overnight at 37 °C, either alone (*lane 2*) or in the presence of MASP-2 concentrations of 16 (*lane 3*), 33 (*lane 4*), and 330 nM (*lane 5*). Cleavage reactions were monitored by SDS-PAGE analysis of the samples under reducing conditions. The unlabeled band on gels *A* and *B* corresponds to the ovalbumin contained in the buffer used for enzyme dilution (see "Experimental Procedures").

As in the case of full-length MASP-2, the CCP1/2-SP fragment of MASP-1 could not be purified to homogeneity, mainly because of the low amounts of recombinant material available. SDS-PAGE analysis indicated that the fragment migrated under nonreducing conditions as a single band of about 50 kDa. Upon reduction, part of the material (40%) still migrated as a single chain, proenzyme species, whereas the remainder yielded a doublet at 32 and 34 kDa, reactive with antibodies directed to the C-terminal end of the molecule, and corresponding to the serine protease domain. Sequence analysis confirmed these findings, as the proenzyme species yielded a single sequence Asp-Leu-Val-Glu-Leu-Pro-Glu-Leu-Gln-Pro . . . corresponding to the N-terminal end of the fragment, the doublet at 32–34 kDa yielding the expected sequence Ile-Phe-Asn-Gly-Arg-Pro-Ala-Gln-Lys-Gly . . . characteristic of the serine protease domain. Thus, as previously observed in the case of full-length MASP-1 expressed in the same system, the recombinant CCP1/2-SP fragment was recovered in a partially activated form. As judged from Coomassie Blue staining, the MASP-1 CCP1/2-SP fragment was estimated at 20–30% of the total protein contents of the partially purified fraction, depending on the preparation.

Proteolytic Activity of Recombinant MASP-2—Recombinant MASP-2 and its CCP1/2-SP fragment both cleaved C4 very efficiently, as shown by their ability to convert the C4α chain into the smaller C4-α' species (Fig. 3*A*). In both cases, almost complete cleavage of C4 was achieved upon incubation for about 15 min at 37 °C at an enzyme:protein molar ratio of 1:2000. Both recombinant proteases also readily and specifically cleaved C2, as shown by their ability to split the protein into its characteristic C2a and C2b fragments (Fig. 3*B*). Again, C2 cleavage was essentially complete after 20 min at 37 °C at an enzyme:protein molar ratio of 1:1000. Under these conditions, no significant C4 or C2 cleavage was observed when these proteins were incubated in the presence of the proenzyme CCP2-SP fragment of MASP-2. The ability of MASP-2 to cleave C3 was also tested using the same methodology. As illustrated in Fig. 3*C*, incubation of C3 with increasing amounts of either full-length MASP-2 or its catalytic fragment CCP1/2-SP led to an increased production of the C3α' fragment. However, the cleavage reaction was very inefficient, as complete conversion of C3α into C3α' required overnight incubation at 37 °C at an enzyme:protein molar ratio of 1:6 (Fig. 3*C, lane 5*). For comparison, complete cleavage of C3α by trypsin was achieved at an enzyme:protein molar ratio of 1:13 in only 5 min at 37 °C (Fig. 3*C, lane 1*). Using partially purified preparations of C5 and factor B as substrates, no significant cleavage of these complement proteins by either MASP-2 or its CCP1/2-SP fragment could be detected after overnight incubation at 37 °C at enzyme:protein molar ratios up to 1:10 (not shown).

The kinetic parameters for C2 and C4 cleavage by MASP-2 were determined using the full-length protease and its CCP1/2-SP fragment, and the values were compared with those obtained with active C1s purified from human serum. In the case of C2, all three enzymes showed comparable k_{cat} values, whereas both MASP-2 species exhibited K_m values slightly, but consistently lower than C1s (Table I and Fig. 4). As a result, the C2 cleavage efficiency of both MASP-2 species, as measured by the k_{cat}/K_m ratio, was slightly higher than that of C1s. With respect to C4 cleavage, again, all three enzymes exhibited similar k_{cat} values. In contrast, both MASP-2 and its CCP1/2-SP fragment showed K_m values in the nanomolar range, *i.e.* 26–32 times lower than the value of 1.92 ± 0.5 μM determined for C1s. As a result, the k_{cat}/K_m ratios for MASP-2 and its CCP1/2-SP fragment were 12- and 23-fold higher, respectively, than that for C1s, indicating a much higher efficiency in the case of C4 cleavage. In this respect, it should be mentioned that the k_{cat} values for MASP-2 were significantly lower than those determined for the CCP1/2-SP fragment. This is likely because of the fact that the concentration of MASP-2 could only be roughly estimated (see "Experimental Procedures") and, we believe, was probably overestimated rather than the opposite. As a consequence, the k_{cat} values determined for the CCP1/2-SP fragment should be regarded as the most representative. It may be estimated therefore that MASP-2 exhibits C2 and C4 cleaving efficiencies about 3 and 23 times higher, respectively, than C1s (see Table I).

Downloaded from http://www.jbc.org/ by guest on March 23, 2015

P9

TABLE I

Kinetic constants for proteolytic cleavage of C2 and C4 by Cls and MASP-2

Enzyme	C2 cleavage			C4 cleavage		
	k_{cat}	K_m	k_{cat}/K_m	k_{cat}	K_m	k_{cat}/K_m
	s^{-1}	μM	$M^{-1} \times s^{-1}$	s^{-1}	nM	$M^{-1} \times s^{-1}$
Cls	5.3 ± 1.3	12.3 ± 3.0	4.3×10^5	4.7 ± 1.2	1920 ± 500	2.45×10^6
MASP-2 CCP1/2-SP fragment	7.4 ± 1.8	6.6 ± 1.5	11.2×10^5	3.4 ± 0.9	60 ± 15	56.7×10^6
MASP-2	4.9 ± 1.2[a]	6.5 ± 1.5	7.5×10^{5a}	2.2 ± 0.6[a]	74 ± 19	29.7×10^{6a}

[a] k_{cat} values determined for MASP-2 are likely underestimated, due to overestimation of the enzyme concentration.

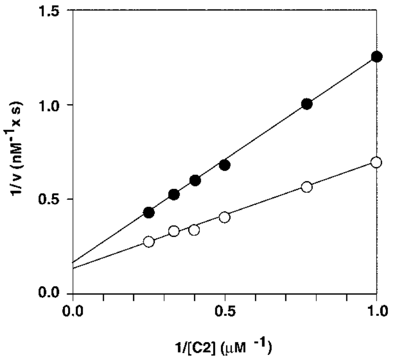


FIG. 4. **Lineweaver-Burk plots of C2 cleavage by C1s and the CCP1/2-SP fragment of MASP-2.** Kinetic analyses were performed as described under "Experimental Procedures," using C2 concentrations ranging from 1 to 4 μM, and fixed enzyme concentrations of 2 nM. ●, cleavage by C1s; ○, cleavage by the CCP1/2-SP fragment of MASP-2.

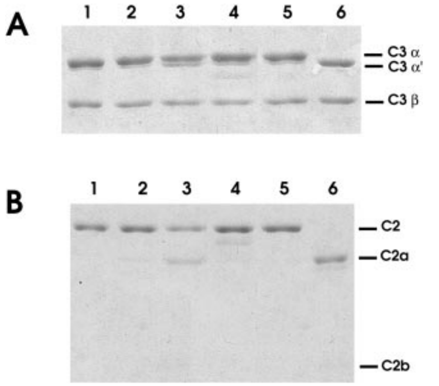


FIG. 5. **SDS-PAGE analysis of the C3 and C2 cleaving activities of recombinant MASP-1.** *A*, C3 (1 μM) was incubated overnight at 37 °C alone (*lane 2*), in the presence of 0.1 μM full-length MASP-1 (*lane 3*), in the presence of MASP-1 plus 0.5 μM C1 inhibitor (*lane 4*), or in the presence of MASP-1 plus 5 mM DFP (*lane 5*). In a control experiment, C3 (1 μM) was incubated for 10 min at 37 °C in the presence of 0.1 μM trypsin (*lane 6*). *Lane 1*, control, nonincubated C3. *B*, C2 (2 μM) was incubated overnight at 37 °C alone (*lane 2*), in the presence of 0.2 μM MASP-1 (*lane 3*), in the presence of MASP-1 plus 1 μM C1 inhibitor (*lane 4*), or in the presence of MASP-1 plus 5 mM DFP (*lane 5*). In a control experiment, C2 (1 μM) was incubated for 1 h at 37 °C in the presence of 0.1 μM C1s (*lane 6*). *Lane 1*, control, nonincubated C2 sample. Cleavage reactions were monitored by SDS-PAGE analysis of the samples under reducing conditions.

Proteolytic Activity of Recombinant MASP-1—The proteolytic activity of recombinant MASP-1 was tested as described above for MASP-2, using both the full-length protease and its C-terminal CCP1/2-SP fragment. For either species, no detectable activity was observed on human C4, C5, and factor B, even after overnight incubation at 37 °C using enzyme:substrate molar ratios up to 1:10 (not shown). Using C3 as a substrate, a low but significant activity was consistently detected using either full-length MASP-1 or the CCP1/2-SP fragment. However, partial C3 cleavage (about 13%) required overnight incubation at 37 °C at an enzyme:substrate molar ratio as high as 1:10 (Fig. 5*A, lane 3*). This faint activity was nevertheless clearly attributable to MASP-1, since it was blocked by pretreatment with C1 inhibitor, and by DFP (Fig. 5*A, lanes 4* and *5*). In a control experiment, complete C3 cleavage was achieved upon incubation for 10 min at 37 °C with trypsin using a 1:10 enzyme:substrate molar ratio (Fig. 5*A, lane 6*). As shown in Fig. 5*B*, C2 was also found to be cleaved by recombinant MASP-1. Again, however, this reaction was quite inefficient since significant cleavage (about 28%) required overnight incubation at 37 °C in the presence of either full-length MASP-1 or its CCP1/2-SP fragment at a 1:10 enzyme:substrate ratio (Fig. 5*B, lane 3*). Pretreatment of MASP-1 with C1 inhibitor or DFP reduced the extent of cleavage to about 5% (Fig. 5*B, lanes 4* and *5*), *i.e.* a value comparable with that observed when C2 alone was incubated overnight at 37 °C (Fig. 5*B, lane 2*). The observed C2 cleavage was therefore clearly mediated by MASP-1, but was quite inefficient compared with the C2 cleaving activity of MASP-2 or C1s.

Esterolytic Activities of MASP-1 and MASP-2—The esterolytic activity of recombinant MASP-1 and MASP-2 was measured on various synthetic substrates and compared with that of human C1r and C1s. For both MASP-1 and MASP-2, comparable results were obtained using either the full-length proteases or their catalytic CCP1/2-SP fragment. Full-length MASP-1 and the MASP-2 CCP1/2-SP fragment were routinely used for determination of the kinetic constants shown in Table II. MASP-2 and C1s efficiently cleaved the four substrates tested in this study, and both enzymes exhibited similar k_{cat} and K_m values. However, C1s showed a significantly better efficiency on Ac-Gly-Lys-OMe, Bz-Arg-Oet, and Tos-Arg-OMe, whereas MASP-2 was more efficient on the thioester Z-Gly-Arg-S-Bzl, because of a remarkably low K_m value (Table II). In keeping with previous reports (39), C1r hydrolyzed Ac-Gly-Lys-OMe and Z-Gly-Arg-S-Bzl, but had no activity on Bz-Arg-OEt and Tos-Arg-OMe. Likewise, MASP-1 also exhibited a restricted esterolytic activity, as it was not reactive toward Ac-Gly-Lys-OMe and Tos-Arg-OMe, and cleaved Z-Gly-Arg-S-Bzl and Bz-Arg-OEt to significant extents. Interestingly, of the four proteases used, MASP-1 showed the least K_m value toward Bz-Arg-OEt (Table II). All of the esterolytic activities measured with recombinant MASP-1 and MASP-2 were blocked after pretreatment of the proteases with excess C1 inhibitor.

Reactivity of MASP-1 and MASP-2 Toward Inhibitors—After incubation with a molar excess of C1 inhibitor, full-length MASP-1 and MASP-2, as well as their CCP1/2-SP fragments, all specifically reacted with this inhibitor, as shown by the formation of stable complexes migrating on SDS-PAGE analysis under reducing conditions as species of over 100 kDa, corresponding to C1 inhibitor-serine protease domain complexes (not shown). To determine the stoichiometry of the reaction between MASP-2 and C1 inhibitor, the CCP1/2-SP fragment of MASP-2 was preincubated with increasing molar ratios of C1 inhibitor and then the residual protease activity was measured using the C2 cleavage assay. As shown in Fig. 6*B*, increasing the C1 inhibitor:MASP-2 ratio led to a linear decrease of enzymatic activity, with complete inhibition at a molar ratio of about 1:1. In a comparative experiment, the same stoichiome-

Downloaded from http://www.jbc.org/ by guest on March 23, 2015

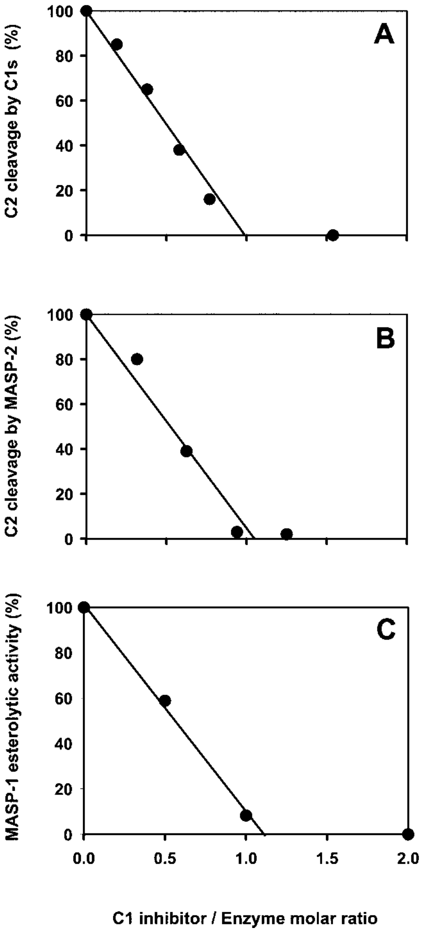


FIG. 6. **Stoichiometry of the reaction of MASP-1, MASP-2, and C1s with C1 inhibitor.** MASP-1 (20 nM), MASP-2 (0.25 μM), and C1s (0.25 μM) were preincubated for 15 min at 37 °C in the presence of increasing concentrations of C1 inhibitor to achieve C1 inhibitor:protease ratios up to 2.0, as indicated. The residual C2 cleaving activity of C1s (*A*) and MASP-2 (*B*), and the residual esterolytic activity of MASP-1 on Bz-Arg-OEt (*C*) were measured as described under "Experimental Procedures."

try was determined in the case of C1s, in agreement with previous reports (32). The stoichiometry of the reaction between MASP-1 and C1 inhibitor was determined in a similar way, by measuring the esterolytic activity of the protease on Bz-Arg-OEt. As shown in Fig. 6*C*, preincubation with increasing concentrations of C1 inhibitor led to a linear decrease of MASP-1 esterolytic activity, with again complete inhibition at a C1 inhibitor:MASP-1 molar ratio of about 1:1. Thus, as determined earlier in the case of C1r and C1s (32, 41), both MASP-1 and MASP-2 reacted with C1 inhibitor in a 1:1 stoichiometry.

As mentioned above, the proteolytic activity of MASP-1 toward C3 and C2 was prevented in the presence of 5 mM DFP. Because of the very low activity of MASP-1 on these substrates, it was not possible to further analyze its sensitivity with respect to DFP. In the case of MASP-2, preincubation of the protease with 2 mM DFP for 30 min at 37 °C resulted in 86% inhibition of its C4 cleaving ability, and nearly complete inhibition was obtained when this treatment was performed twice in a row. Thus, MASP-2 exhibited a DFP sensitivity comparable with that previously determined in the case of C1r and C1s (42). We also checked the effect of α_2M on the C4 cleaving activity of MASP-2 and C1s. Both proteases were used at a concentration of 0.5 μM in the presence of increasing molar ratios of α_2M up to 120:1. Under these conditions, this protein exerted no detectable inhibitory effect on the proteolytic activ-

TABLE II

Kinetic constants for the esterolytic activity of MASP-1, MASP-2, Clr and Cls on various synthetic substrates

Enzyme	Ac-Gly-Lys-OMe			Bz-Arg-OEt			Z-Gly-Arg-S-Bzl			Tos-Arg-OMe		
	k_{cat}	K_m	k_{cat}/K_m	k_{cat}	K_m	k_{cat}/K_m	k_{cat}	K_m	k_{cat}/K_m	k_{cat}	K_m	k_{cat}/K_m
	s^{-1}	mM	$M^{-1} \times s^{-1}$	s^{-1}	mM	$M^{-1} \times s^{-1}$	s^{-1}	μM	$M^{-1} \times s^{-1}$	s^{-1}	mM	$M^{-1} \times s^{-1}$
MASP-1	—[a]	—[a]	—[a]	0.35 ± 0.03	0.18 ± 0.02	1.94×10^3	1.23 ± 0.1	310 ± 30	4.0×10^3	—[a]	—[a]	—[a]
MASP-2	133 ± 12	6.7 ± 0.6	19.8×10^3	2.4 ± 0.2	4.4 ± 0.5	0.54×10^3	12.7 ± 1.2	40 ± 5	318×10^3	7.1 ± 0.7	2.8 ± 0.03	2.5×10^3
Cls	167 ± 15	2.5 ± 0.2	66.8×10^3	9.5 ± 0.9	4.0 ± 0.4	2.37×10^3	24.3 ± 2.3	190 ± 20	128×10^3	11.9 ± 1.2	1.2 ± 0.1	9.9×10^3
Clr	8.5 ± 0.8	1.6 ± 0.2	5.3×10^3	—[a]	—[a]	—[a]	8.9 ± 0.8	190 ± 20	47×10^3	—[a]	—[a]	—[a]

[a] Values not measurable due to the lack of activity.

Downloaded from http://www.jbc.org/ by guest on March 23, 2015

ity of either protease, indicating that, unlike MASP-1 (43), and like C1s (39), MASP-2 is not sensitive to α_2M.

DISCUSSION

Full-length MASP-2, as well as fragments from the C-terminal catalytic regions of MASP-1 and MASP-2 were expressed using a baculovirus/insect cell system. As observed previously in the case of full-length MASP-1 expressed in the same system (27), the recombinant material was produced at very low yields (0.15–2.0 μg/ml). This rendered isolation of the recombinant material difficult, but nevertheless allowed us to achieve complete purification of MASP-2 fragments CCP1/2-SP and CCP2-SP, and partial purification of the MASP-1 CCP1/2-SP fragment and of full-length MASP-2. As judged from all functional assays performed in this study, the latter two preparations were functionally pure, *i.e.* devoid of detectable contaminant esterolytic or proteolytic activities. Thus, full-length MASP-1 on the one hand, and full-length MASP-2 on the other hand, exhibited esterolytic and proteolytic activities comparable with those of their CCP1/2-SP fragments, both in terms of efficiency and specificity. In addition, these activities were all blocked by C1 inhibitor, providing further evidence that they were not due to contaminant proteases arising from the insect cells. The recombinant material used in this study was therefore appropriate for a precise assessment of the enzymic properties and specificity of MASP-1 and MASP-2.

As previously observed for the full-length recombinant protease (27), the CCP1/2-SP fragment of MASP-1 was recovered in a partially (about 60%) activated form, suggesting that either autolytic activation or extrinsic proteolytic cleavage occurred during the synthesis process. Similarly, in the case of MASP-2, the full-length species was also recovered in a partially activated form. Interestingly, the short catalytic fragment CCP2-SP of MASP-2 was recovered in a fully proenzyme state, whereas the larger fragment CCP1/2-SP was secreted by the insect cells as a partially activated form, and underwent complete activation during the purification process. These latter findings appear not consistent with an activation process mediated by an extrinsic protease, as access to the Arg^{444}-Ile^{445} activation site of MASP-2 is not expected to be made easier in the larger CCP1/2-SP fragment than in the shorter CCP2-SP fragment. On the other hand, this strikingly different behavior of the two fragments would be consistent with an autolytic activation mechanism of MASP-2 involving interaction between two molecules through their C-terminal catalytic region, if this process requires the first CCP module. This latter hypothesis appears plausible, in light of the observation that dimerization of the catalytic domains of the homologous protease C1r involves its first CCP module (44). In the same way, the fact that recombinant full-length MASP-2 only undergoes partial activation may be explained by analogy with previous findings on C1r (48, 49) indicating that the N-terminal, Ca^{2+}-binding CUB-EGF-CUB region of C1r exerts a negative control on the activation of the C-terminal catalytic region of the protease. The above observations support the hypothesis of an intrinsic ability of MASP-2 to self-activate, in full agreement with previous reports by Vorup-Jensen *et al.* (45) indicating that reconstitution of MBL with MASP-2 alone is sufficient to trigger complement activation.

Although recombinant MASP-2 and its CCP1/2-SP fragment displayed a detectable C3 cleaving activity, this was only observed upon overnight incubation at elevated enzyme:substrate ratios. It is likely therefore that this faint activity is nonspecific and has no biological relevance. In contrast, in agreement with previous reports (6, 21, 45), this work confirms that MASP-2 specifically cleaves and activates C4 and C2 and provides the first detailed analysis of the kinetic parameters of these reactions. These data indicate that (i) MASP-2 cleaves C2 slightly more efficiently (2–3 times) than does C1s; (ii) MASP-2 exhibits a C4-cleaving efficiency that is 20–25 times higher than that of C1s, arising mainly from a much lower K_m value (about 60 nM *versus* about 2 μM). These findings have direct functional implications, since C4 cleavage, the initial event in the formation of the classical pathway C3-convertase C4b,2a, occurs in the fluid phase and is the limiting step of this process. Thus, the relative efficiency of C4 cleavage by MASP-2 and C1s *in vivo* will directly depend on their respective K_m values for C4 (about 60 nM and 2 μM) relative to the C4 concentration in serum (about 3 μM) (46). It can be deduced from these figures that, whereas C1s will cleave C4 in serum at a rate well below the V_m value, MASP-2 will always function at or close to maximal velocity. C2 cleavage, the second step of the assembly of the C4b,2a-convertase, is postulated to take place after prior binding of C2 to membrane-bound C4b, in close vicinity of either C1 or the MBL-MASP complex. Therefore the C2 cleavage efficiency of MASP-2 or C1s is not directly related to their K_m values for this substrate, which in both cases are well above the C2 concentration in serum, *i.e.* 0.13–0.43 μM (47).

According to the above figures, MASP-2 is expected to be about 20–25 times more efficient than C1s with respect to C4 cleavage. As a consequence, despite the low amounts of MASP-2 in serum, a complex utilizing MBL as a recognition unit and MASP-2 as a catalytic unit is probably as active as C1 in terms of complement activation. These considerations provide strong support to the biological relevance of the lectin pathway of complement activation as a major innate defense mechanism. Together with experiments suggesting that MASP-2 has the ability to self-activate (Ref. 45 and this study), and binds individually to MBL with high affinity (27), the above results strongly suggest that the entity responsible for triggering the MBL-mediated lectin pathway of complement is an MBL·MASP-2 complex, in which MASP-2 combines the ability to self-activate as well as to cleave C4 and C2, and hence fulfills the roles of both C1r and C1s in C1. In keeping with previous data (21), our data also demonstrate that the proteolytic activity of MASP-2 is specifically titrated by C1-inhibitor, with a 1:1 C1 inhibitor:MASP-2 stoichiometry. As well documented in the case of the C1 enzymes (41), it may be anticipated therefore that the proteolytic activity of MASP-2 in serum is tightly and specifically controlled by C1 inhibitor.

This work also provides the first analysis of the esterolytic activities and specificities of MASP-1 and MASP-2. In keeping with the fact that MASP-2 and C1s specifically cleave the same protein substrates, they exhibit similar esterolytic activities on the various synthetic substrates tested in this study, which they cleave with comparable k_{cat} and K_m values. However, as measured by the k_{cat}/K_m ratio, the efficiency of MASP-2 was found consistently lower than that of C1s, with the exception of the thioester Z-Gly-Arg-S-Bzl, which, because of a remarkably low K_m value, is the best known substrate of MASP-2. Thus, the shared ability of MASP-2 and C1s to cleave C4 and C2 likely arises in part from common structural features at or in the vicinity of their active sites. On the other hand, the fact that MASP-2 has a much lower K_m value for C4 is probably related to structural determinants located outside the active site area, very likely in the CCP modules, as previously demonstrated in the case of C1s (17). MASP-1, in contrast, exhibits a lower esterolytic activity, that is restricted to two of the substrates tested in this study, Bz-Arg-OEt and Z-Gly-Arg-S-Bzl. Nevertheless, it is noteworthy that MASP-1 cleaves Bz-Arg-OEt with an efficiency comparable with that of C1s, because of a rather low K_m value (see Table II). Also, the fact that MASP-1 does not hydrolyze Ac-Gly-Lys-OMe suggests that its

Downloaded from http://www.jbc.org/ by guest on March 23, 2015

P9

activity may be restricted to arginyl bonds only. As judged from its rather low and narrow esterolytic activity, MASP-1 appears functionally much closer to C1r than to MASP-2 and C1s, which display broader activities and are more efficient enzymes.

In keeping with data obtained by Matsushita *et al.* (21) using material purified from human serum, both recombinant full-length MASP-1 and its catalytic CCP1/2-SP fragment showed detectable C2 and C3 cleaving activities. However, these reactions were quite inefficient, since in both cases only partial cleavage could be obtained after overnight incubation at elevated enzyme:substrate ratios. In this respect, it should be stressed that the C2 cleaving activity of recombinant MASP-1 is quite negligible compared with that of recombinant MASP-2 (~20,000-fold less). In the same way, the C3 cleaving activity of MASP-1 was consistently found to be less than that of MASP-2, which itself is extremely low compared the MASP-2 activities on C4 and C2. With respect to C2 cleavage by MASP-1, it is difficult to compare quantitatively our data with those of Matsushita *et al.* (21), since these authors used a hemolytic assay to test C2 activation. As for C3 cleavage, as judged from the data presented by Matsushita *et al.* (21), the activity exhibited by the MASP-1 fraction purified from serum appears comparable with that obtained with our recombinant material, since only partial cleavage was obtained under the conditions used. In this respect, it should also be mentioned that other studies have reported no detectable C3 cleavage by the total MASP fraction (22, 24). In summary, although our data certainly validate previous findings by Matsushita *et al.* (21) that MASP-1 exhibits detectable C2 and C3 cleaving activities, they also clearly show that these activities are quite negligible compared with the very efficient C4 and C2 cleaving activities of MASP-2. It appears very unlikely, therefore, that the physiological function of MASP-1 is to cleave C3 (which would be a very inefficient means of triggering complement, compared with C1s or MASP-2), and/or to cleave C2 (which, alone, would be physiologically not relevant). If this hypothesis is correct, then the physiological substrate(s) of MASP-1 in serum, and hence its physiological role, remain to be established.

Acknowledgments—We thank Jean-Pierre Andrieu and Jean Gagnon for determining N-terminal sequences.

REFERENCES

1. Bork, P., and Bairoch, A. (1995) *Trends Biochem. Sci.* **20,** (suppl.) C03
2. Weis, W. I., Drickamer, K., and Hendrickson, W. A. (1992) *Nature* **360,** 127–134
3. Turner, M. W. (1996) *Immunol. Today* **17,** 532–540
4. Hoffmann, J. A., Kafatos, F. C., Janeway, C. A., and Ezekowitz, R. A. (1999) *Science* **284,** 1313–1318
5. Turner, M. W., and Hamvas, R. M. J. (2000) *Rev. Immunogenetics* **2,** 305–322
6. Thiel, S., Vorup-Jensen, T., Stover, C. M., Schwaeble, W., Laursen, S. B., Poulsen, K., Willis, A. C., Eggleton, P., Hansen, S., Holmskov, U., Reid, K. B. M., and Jensenius, J. C. (1997) *Nature* **386,** 506–510
7. Matsushita, M., and Fujita, T. (1992) *J. Exp. Med.* **176,** 1497–1502
8. Takada, F., Takayama, Y., Hatsuse, H., and Kawakami, M. (1993) *Biochem. Biophys. Res. Commun.* **196,** 1003–1009
9. Sato, T., Endo, Y., Matsushita, M., and Fujita, T. (1994) *Int. Immunol.* **6,** 665–669
10. Matsushita, M., and Fujita, T. (1995) *Immunobiology* **194,** 443–448
11. Stover, C. M., Thiel, S., Thelen, M., Lynch, N. J., Vorup-Jensen, T., Jensenius, J. C., and Schwaeble, J. W. (1999) *J. Immunol.* **162,** 3481–3490
12. Takahashi, M., Endo, Y., Fujita, T., and Matsushita, M. (1999) *Int. Immunol.* **11,** 859–864
13. Dahl, M. R., Thiel, S., Willis, A. C., Vorup-Jensen, T., Christensen, T., Petersen, S. V., and Jensenius, J. C. (2001) *Immunity,* **15,** 127–135
14. Arlaud, G. J., Volanakis, J. E., Thielens, N. M., Narayana, S. V. L., Rossi, V., and Xu, Y. (1998) *Adv. Immunol.* **69,** 249–307
15. Bork, P., and Beckmann, G. (1993) *J. Mol. Biol.* **231,** 539–545
16. Reid, K. B. M., Bentley, D. R., Campbell, R. D., Chung, L. P., Sim, R. B., Kristensen, T., and Tack, B. F. (1986) *Immunol. Today* **7,** 230–234
17. Rossi, V., Bally, I., Thielens, N. M., Esser, A. F., and Arlaud, G. J. (1998) *J. Biol. Chem.* **273,** 1232–1239
18. Gaboriaud, C., Rossi, V., Bally, I., Arlaud, G. J., and Fontecilla-Camps, J. C. (2000) *EMBO J.* **19,** 1755–1765
19. Ji, X., Azumi, K., Sasaki, M., and Nonaka, M. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 6340–6345
20. Endo, Y., Takahashi, M., Nakao, M., Saiga, H., Sekine, H., Matsushita, M., Nonaka, M., and Fujita, T. (1998) *J. Immunol.* **161,** 4924–4930
21. Matsushita, M., Thiel, S., Jensenius, J. C., Terai, I., and Fujita, T. (2000) *J. Immunol.* **165,** 2637–2642
22. Wong, N. K. H., Kojima, M., Dobo, J., Ambrus, G., and Sim, R. B. (1999) *Mol. Immunol.* **36,** 853–861
23. Matsushita, M., Endo, Y., and Fujita, T. (1998) *Immunobiol.* **199,** 340–347
24. Zhang, Y., Suankatray, C., Jones, D. R., Zhang, X. H., Lint, T. F., and Gewurz, H. (1998) *Mol. Immunol.* **35,** 390 (abstr.)
25. Vorup-Jensen, T. (2000) *Ph. D. Thesis*, University of Aarhus
26. Tan, S. M., Chung, M. C. M., Kon, O. L., Thiel, S., Lee, S. H., and Lu, J. (1996) *Biochem. J.* **319,** 329–332
27. Thielens, N. M., Cseh, S., Thiel, S., Vorup-Jensen, T., Rossi, V., Jensenius, J. C., and Arlaud, G. J. (2001) *J. Immunol.* **166,** 5068–5077
28. Arlaud, G. J., Reboul, A., Sim, R. B., and Colomb, M. G. (1979) *Biochim. Biophys. Acta* **576,** 151–162
29. Dodds, A. W. (1993) *Methods Enzymol.* **223,** 46–61
30. Thielens, N. M., Villiers, M. B., Reboul, A., Villiers, C. L., and Colomb, M. G. (1982) *FEBS Lett.* **141,** 19–24
31. Al Salihi, A., Ripoche, J., Pruvost, L., and Fontaine, M. (1982) *FEBS Lett.* **150,** 238–242
32. Reboul, A., Arlaud, G. J., Sim, R. B., and Colomb, M. G. (1977) *FEBS Lett.* **79,** 45–50
33. Pétillot, Y., Thibault, P., Thielens, N. M., Rossi, V., Lacroix, M., Coddeville, B., Spik, G., Schumaker, V. N., Gagnon, J., and Arlaud, G. J. (1995) *FEBS Lett.* **358,** 323–328
34. Edelhoch, H. (1967) *Biochemistry* **6,** 1948–1954
35. King, L. A., and Possee, R. D. (1992) in *The Baculovirus Expression System: A Laboratory Guide*, pp. 111–114, Chapman and Hall Ltd., London
36. Thielens, N. M., Aude, C. A., Lacroix, M. B., Gagnon, J., and Arlaud, G. J. (1990) *J. Biol. Chem.* **265,** 14469–14475
37. Petersen, S. V., Poulsen, K., Stover, C. M., Koch, C., Vorup-Jensen, T., and Thiel, S. (1998) *Mol. Immunol.* **35,** 409 (abstr.)
38. Rossi, V., Gaboriaud, C., Lacroix, M., Ulrich, J., Fontecilla-Camps, J. C., Gagnon, J., and Arlaud, G. J. (1995) *Biochemistry* **34,** 7311–7321
39. Arlaud, G. J., and Thielens, N. M. (1993) *Methods Enzymol.* **223,** 61–82
40. McRae, B. J., Lin, T.-Y.,., and Powers, J. C. (1981) *J. Biol. Chem.* **256,** 12362–12366
41. Sim, R. B., Arlaud, G. J., and Colomb, M. G. (1979) *Biochem. J.* **179,** 449–457
42. Arlaud, G. J., Sim, R. B., Duplaa, A. M., and Colomb, M. G. (1979) *Mol. Immunol.* **16,** 445–450
43. Terai, I., Kobayashi, K., Matsushita, M., Fujita, T., and Matsuno, K. (1995) *Int. Immunol.* **7,** 1579–1584
44. Lacroix, M., Rossi, V., Gaboriaud, C., Chevallier, S., Jaquinod, M., Thielens, N. M., Gagnon, J., and Arlaud, G. J. (1997) *Biochemistry* **36,** 6270–6282
45. Vorup-Jensen, T., Petersen, S. V., Hansen, A. G., Poulsen, K., Schwaeble, W., Sim, R. B., Reid, K. B. M., Davis, S. J., Thiel, S., and Jensenius, J. C. (2000) *J. Immunol.* **165,** 2093–2100
46. Isenman, D. (2000) in *The Complement Factsbook* (Morley, B. J., and Walport, M. J., eds) pp. 95–103, Academic Press, New York
47. Xu, Y., and Volanakis, J. E. (2000) in *The Complement Factsbook* (Morley, B. J., and Walport, M. J., eds) pp. 73–77, Academic Press, New York
48. Lacroix, M. B., Aude, C. A., Arlaud, G. J., and Colomb, M. G. (1989) *Biochem. J.* **257,** 885–891
49. Thielens, N. M., Illy, C., Bally, I. M., and Arlaud, G. J. (1994) *Biochem. J.* **301,** 509–516

Downloaded from http://www.jbc.org/ by guest on March 23, 2015

P11

Functional Role of the Linker between the Complement Control Protein Modules of Complement Protease C1s[1]

Isabelle Bally,* Véronique Rossi,* Nicole M. Thielens,* Christine Gaboriaud,† and Gérard J. Arlaud[2]*

C1s is the modular serine protease responsible for cleavage of C4 and C2, the protein substrates of the first component of C (C1). Its catalytic domain comprises two complement control protein (CCP) modules connected by a four-residue linker Gln^{340}-Pro-Val-Asp^{343} and a serine protease domain. To assess the functional role of the linker, a series of mutations were performed at positions 340–343 of human C1s, and the resulting mutants were produced using a baculovirus-mediated expression system and characterized functionally. All mutants were secreted in a proenzyme form and had a mass of 77,203–77,716 Da comparable to that of wild-type C1s, except Q340E, which had a mass of 82,008 Da, due to overglycosylation at Asn^{391}. None of the mutations significantly altered C1s ability to assemble with C1r and C1q within C1. Whereas the other mutations had no effect on C1s activation, the Q340E mutant was totally resistant to C1r-mediated activation, both in the fluid phase and within the C1 complex. Once activated, all mutants cleaved C2 with an efficiency comparable to that of wild-type C1s. In contrast, most of the mutations resulted in a decreased C4-cleaving activity, with particularly pronounced inhibitory effects for point mutants Q340K, P341I, V342K, and D343N. Comparable effects were observed when the C4-cleaving activity of the mutants was measured inside C1. Thus, flexibility of the C1s CCP_1-CCP_2 linker plays no significant role in C1 assembly or C1s activation by C1r inside C1 but plays a critical role in C4 cleavage by adjusting positioning of this substrate for optimal cleavage by the C1s active site. *The Journal of Immunology,* **2005, 175: 4536–4542.**

The C1 complex is a 790-kDa multimolecular assembly comprised of three components, C1q, C1r, and C1s, which triggers the classical pathway of C, a major route of innate immunity against pathogenic microorganisms. Recognition of the target microorganism is mediated by C1q and triggers autolytic activation of C1r, a modular serine protease (SP),[3] which in turn activates a second protease C1s through cleavage of its Arg^{422}-Ile^{423} bond (1–4). Active C1s then specifically cleaves a single arginyl peptide bond in C proteins C4 and C2, the natural substrates of C1, thereby triggering a cascade of proteolytic reactions that elicit various biological activities designed to provide a first line of defense against infection.

Human C1s is synthesized as a 673-residue single-chain zymogen, which starting from the N terminus, comprises a CUB module (5), an epidermal growth factor (EGF) module, a second CUB module, two complement control protein (CCP) modules (6), and a chymotrypsin-like SP domain. This modular architecture is shared by C1r, and by the mannan-binding lectin-associated SPs (MASPs), a group of proteases involved in the lectin pathway of C (7). From a functional point of view, it is known that assembly of C1s-C1r-C1r-C1s, the tetrameric catalytic subunit of C1, involves Ca^{2+}-dependent heterodimeric C1r-C1s interactions mediated by the N-terminal CUB_1-EGF moiety of each protease (8–10). On the other hand, it is established that the enzymatic properties of C1r and C1s are mediated by their C-terminal CCP_1-CCP_2-SP regions (11, 12). In the case of C1s, expression of modular fragments has led to the conclusion that, whereas C2 cleavage only requires the SP domain, C4 cleavage involves substrate recognition sites located in the CCP modules (13). The structure of the CCP_2-SP fragment of C1s has been solved by x-ray crystallography, revealing that the ellipsoidal CCP_2 module is tightly anchored on the more globular SP domain by means of a rigid interface and may therefore act as a spacer and a handle (14). The structure of the C1s CCP_1 module is not known yet, and the flexibility at the CCP_1-CCP_2 junction has not been characterized. Nevertheless, a number of structural studies have been performed on CCP module pairs, indicating that these exhibit varying degrees of flexibility at their intermodular interface (15–18). Based on these observations, it was hypothesized that the CCP_1-CCP_2 linker of C1s exhibits substantial flexibility, allowing the SP domain to switch from a position inside C1, allowing activation by C1r, to a more external location, required for cleavage of C4 and C2 (14). From a general standpoint, current knowledge of the C1 structure strongly suggests that its ability to mediate its finely tuned function relies for a large part on the flexibility of its individual subunits (4).

The objective of this study was to assess the particular role of the C1s CCP_1-CCP_2 linker in the various facets of the interaction and catalytic properties of this enzyme. Using site-directed mutagenesis, we show that this segment is not involved in the C1s activation mechanism but plays a key role in modulating its activity toward C4.

Materials and Methods

Materials

Diisopropyl phosphorofluoridate was purchased from Sigma-Aldrich. The Pfu polymerase was obtained from Invitrogen Life Technologies. Restriction enzymes were purchased from New England Biolabs. The polyclonal

---

*Laboratoire d'Enzymologie Moléculaire and †Laboratoire de Cristallographie et Cristallogénèse des Protéines, Institut de Biologie Structurale Jean-Pierre Ebel, Grenoble, France

Received for publication April 29, 2005. Accepted for publication July 19, 2005.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] This work was supported by the Commissariat à l'Energie Atomique, Centre National de la Recherche Scientifique, and Université Joseph Fourier, Grenoble.

[2] Address correspondence and reprint requests to Dr. Gérard J. Arlaud, Institut de Biologie Structurale Jean-Pierre Ebel, 41 rue Jules Horowitz, 38027 Grenoble Cedex 1, France. E-mail address: arlaud@ibs.fr

[3] Abbreviations used in this paper: SP, serine protease; CUB, module originally identified in C1r/C1s, Uegf, and bone morphogenetic protein; EGF, epidermal growth factor; CCP, complement control protein; MASP, mannan-binding lectin-associated SP.

Copyright © 2005 by The American Association of Immunologists, Inc.

0022-1767/05/$02.00

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

P11

anti-C1s antiserum used for Western blot analysis was raised in rabbits according to standard procedures. The C1s-Sepharose column was prepared by coupling activated C1s purified from human serum (40 mg) to cyanogen bromide-activated Sepharose 4B (30 ml) according to the method of March et al. (19). The pBS-C1s plasmid containing the full-length human C1s cDNA (20) was kindly provided by Dr. M. Tosi (University of Rouen, Rouen, France). Oligonucleotides were purchased from MWG Biotec.

Proteins

C1q, proenzyme C1r, and activated C1s were purified from human plasma as described previously (21, 22). The activated CCP$_2$-SP fragment of human C1r was expressed in a baculovirus/insect cells system and purified by ion-exchange and hydrophobic interaction chromatography as described previously (23). C proteins C2 and C4 were isolated from human plasma by means of published procedures (24, 25). The concentrations of purified proteins were determined using the following absorption coefficients ($A_{1\%,\ 1\ \mathrm{cm}}$ at 280 nm) and molecular weights: C1q, 6.8 and 459,300; C1r, 12.4 and 86,300 (8); C1r fragment CCP2-SP, 14.9 and 40,400 (23); C2, 10.0 and 100,000 (24); and C4, 8.2 and 205,000 (26, 27).

Expression and purification of recombinant wild-type and mutant C1s

Recombinant wild-type C1s and its variants were expressed using a baculovirus/insect cells system. A DNA fragment encoding the C1s signal peptide plus the full-length mature protein (amino acid residues 1–673) was amplified by PCR using the Pfu polymerase and cloned into the *Bam*HI and *Eco*RI sites of the pFastBac1 vector (Invitrogen Life Technologies). The expression plasmids coding for all C1s mutants were generated using the QuickChange XL site-directed mutagenesis kit (Stratagene). The mutagenic oligonucleotides were designed according to the manufacturer's recommendations, and a silent restriction site was introduced in each case for screening of positive clones. The pFastBac1/C1s expression plasmid coding for wild-type C1s was used as a template. The sequences of all mutants were confirmed by dsDNA sequencing (Genome Express).

The insect cells *Spodoptera frugiperda* (Ready-Plaque Sf9 cells obtained from Novagen) and *Trichoplusia ni* (High Five) were routinely grown and maintained as described previously (28). The bacmid DNA was used to transfect Sf9 cells using cellfectin in Sf900 II SFM medium (Invitrogen Life Technologies) as described in the manufacturer's protocol. The recombinant baculoviruses were generated using the Bac-to-Bac system (Invitrogen Life Technologies), amplified, and titrated as described previously (28). High Five cells (1.75×10^7 cells/175-cm^2 tissue culture flask) were infected with the recombinant viruses at a multiplicity of infection of two to three in Sf900 II SFM medium at 28°C for 48 h. The recombinant C1s molecules were purified from the cell culture supernatants using a one-step affinity chromatography on a Sepharose-C1s column. Briefly, each supernatant was dialyzed against 145 mM NaCl, 2 mM CaCl$_2$, and 50 mM triethanolamine hydrochloride (pH 7.4) containing 2 mM diisopropyl phosphorofluoridate and loaded onto the C1s-Sepharose column equilibrated in the same buffer. Elution was conducted with the same buffer containing 5 mM EDTA instead of CaCl$_2$. In between two runs, the column was systematically washed with 50 mM triethanolamine hydrochloride and 0.5 M NaCl (pH 7.4). Each purified C1s variant was concentrated by ultrafiltration to 0.2–0.7 mg/ml, dialyzed against 145 mM NaCl, and 50 mM triethanolamine hydrochloride (pH 7.4) and stored at −20°C until use. The concentrations of the purified recombinant C1s variants were determined using an absorption coefficient ($A_{1\%,\ 1\ \mathrm{cm}}$ at 280 nm) of 14.5 (8) and the m.w. value determined in each case by mass spectrometry analysis (see *Results*).

Chemical characterization of the recombinant proteins

N-terminal sequence analysis was performed using an Applied Biosystems model 492 as described previously (11). Mass spectrometry analysis of the recombinant C1s molecules was performed using the MALDI technique on a Voyager Elite XL instrument (Applied Biosystems) under conditions described previously (29). The experimental error was approximately ±0.05% of the mass values, i.e., approximately ±40 Da for the values shown in Table I.

Fluid-phase activation of the C1s mutants

To check their ability to undergo activation in the fluid phase, the C1s mutants (0.14 mg/ml) were incubated with the activated C1r CCP$_2$-SP fragment (1.15 μg/ml) for 2 h at 37°C in 145 mM NaCl and 50 mM triethanolamine hydrochloride (pH 7.4), containing 1 mg/ml OVA. The activation extent was determined by SDS-PAGE analysis (30) under reducing conditions and measurement by gel scanning of the amount of the two chains characteristic of active C1s relative to the single-chain proenzyme species.

Analysis by surface plasmon resonance spectroscopy of the ability of the C1s mutants to associate with C1r and C1q

Analyses were performed using a BIAcore 3000 instrument (BIAcore AB). C1q was immobilized on the surface of a CM5 sensor chip (BIAcore AB) using the amine coupling chemistry. The C1s-C1r-C1r-C1s tetramer was reconstituted from equimolar amounts of proenzyme C1r and each individual C1s variant. Binding of each sample (50 nM) was measured over 16,500 resonance units of immobilized C1q, at a flow rate of 20 μl/min in 145 mM NaCl, 2 mM CaCl$_2$, and 50 mM triethanolamine hydrochloride (pH 7.4), containing 0.005% surfactant P20 (BIAcore AB). Each sample was injected over a surface with immobilized BSA for subtraction of the bulk refractive index background. Regeneration of the surfaces was achieved by injection of 10 μl of 145 mM NaCl, 5 mM EDTA, and 50 mM triethanolamine hydrochloride (pH 7.4). For the tetramer containing wild-type C1s, the apparent dissociation constant was determined from binding experiments performed at varying concentrations (1–5 nM) as described previously (31).

Table I. *Mass spectrometry analysis of the C1s variants*

C1s Variant	Experimental Mass (Da)[a]	Calculated Mass of the Polypeptide (Da)	Deduced Mass of the Carbohydrates (Da)[b]
Wild type	77,413	74,887	2,526
Q340A	77,338	74,830	2,508
Q340K	77,716	74,887	2,829
Q340R	77,317	74,915	2,402
Q340S	77,302	74,846	2,456
Q340D	77,366	74,874	2,492
Q340E	82,008	74,888	7,120
Q340E/N391Q	76,117	74,902	1,215
Q340E/D343N	77,203	74,887	2,316
Q340D/D343E	77,540	74,888	2,652
P341I	77,413	74,903	2,510
V342K	77,425	74,916	2,509
D343N	77,231	74,886	2,345
Q340K/D343N	77,345	74,886	2,459
Q340K/P341I	77,548	74,903	2,645
Q340S/P341I	77,339	74,862	2,477
Q340K/P341I/V342K	77,493	74,932	2,561

[a] The experimental error is approximately ±40 Da for all values.
[b] These values account for the carbohydrates attached to Asn^{159} and Asn^{391}.

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

Activation of the C1s mutants in the C1 complex

The C1 complex (0.25 μM) was reconstituted by mixing C1q, proenzyme C1r, and each individual C1s variant, in relative molar amounts 1:2:2, respectively, and incubated for 90 min at 37°C in 145 mM NaCl, 1 mM $CaCl_2$, and 50 mM triethanolamine hydrochloride (pH 7.4). Samples were submitted to SDS-PAGE analysis (30) under reducing conditions using 10% acrylamide gels. C1s was revealed by Western blotting after electrotransfer to a nitrocellulose membrane as described previously (13), and its extent of activation was determined by gel scanning from the relative proportion of the bands corresponding to the active and proenzyme forms.

Proteolytic assays

The proteolytic activity of the activated C1s variants toward C2 and C4 was measured by incubation of these protein substrates (2.25 μM each) at enzyme:protein molar ratios of 1:500 and 1:1125, respectively, for varying periods at 37°C. Enzyme dilutions were performed in 145 mM NaCl, 50 mM triethanolamine hydrochloride, and 1 mg/ml OVA (pH 7.4), containing 1 mM $CaCl_2$ in the case of the reconstituted C1 samples. The extent of reaction was determined after SDS-PAGE analysis under reducing conditions by gel scanning of the cleavage fragments, as described previously (13). For determination of the kinetic parameters of C4 cleavage, the C4 concentration ranged from 0.5 to 4 μM, and C1s was used at a fixed concentration of 2 nM. The kinetic constants were determined by the Lineweaver-Burk method using linear regression analysis and are based on triplicate measurements of initial rates at six different substrate concentrations.

Modeling of the C1s catalytic domain

A three-dimensional model of the CCP_1-CCP_2-SP fragment of C1s was assembled from the C1r CCP_1 module, as seen in the x-ray structure of the C1r CCP_1-CCP_2-SP fragment (12), and the C1s CCP_2-SP domain, as seen in the x-ray structure of the corresponding fragment (14). Relative positioning of CCP_1 with respect to CCP_2 was achieved by superimposing the C1r CCP_1-CCP_2 module pair onto the C1s CCP_2 module. The root mean square deviation between the two CCP2 modules was 0.67 Å, based on 56 Cα positions. In CCP_1, Lys^{355} (equivalent to $\text{Gln}^{340}$of C1s) was replaced by Gln.

Results

The objective of this study was to perform point mutations in the sequence stretch 340–343 located between the CCP_1 and CCP_2 modules of human C1s and to determine the functional properties of the resulting mutants to assess the role of this region in the function of C1s, particularly in the context of the C1 complex. Our strategy was based initially on a comparative analysis of the sequence of this segment in C1s and in other proteases of the same family (C1r, MASP-1, MASP-2, and MASP-3) (Fig. 1). The wild-type and mutant C1s molecules were expressed in a baculovirus/insect cells system and secreted in the culture supernatant, at concentrations ranging from 5 to 10 mg/L culture, except for one of the mutants (Q340S/P341I), which showed a significantly decreased production yield (1 mg/L). All recombinant proteins could be successfully purified by affinity chromatography on a C1s-Sepharose column in the presence of Ca^{2+} ions, providing strong indication that, as expected, the interaction properties characteristic of their N-terminal CUB_1-EGF moiety (10) were unaltered. All variants were essentially pure, as assessed by SDS-PAGE analysis, and exhibited a single-chain structure under reducing conditions, indicating that they were in the proenzyme state (Fig. 2).

	336 347
C1s_human	KLK**C**QPVD**C**GIP
C1s_hamster	RLR**C**QPVD**C**GIP
C1r_human	MPR**C**KIKD**C**GQP
C1r_mouse	MPR**C**KIKN**C**GQP
MASP1/3_human	IPT**C**KIVD**C**RAP
MASP1/3_mouse	IPT**C**KIVD**C**GAP
MASP1/3_rat	IPT**C**KIVD**C**GVP
MASP3_guinea pig	IPT**C**QIVD**C**GAP
MASP3_rabbit	IPT**C**KMVD**C**GVP
MASP2_human	MPA**C**SIVD**C**GPP
MASP2_rat	IPE**C**SIID**C**GPP
MASP2_mouse	MPE**C**SIID**C**GPP

FIGURE 1. Sequence alignment of the segments at the junction between the CCP_1 and CCP_2 modules of proteases of the C1r/C1s/MASP family. The last cysteine of modules CCP_1 and the first cysteine of modules CCP_2 are shown in bold. The amino acid numbering is that of human C1s.

*The Q340E mutation enhances glycosylation at Asn^391^*

A series of point mutations were performed initially at position 340, where the Gln residue was replaced by Lys (as in C1r and MASP-1/3), Ser (as in MASP-2), Ala, Arg, Asp, or Glu. As judged by SDS-PAGE analysis, all of the mutants had an electrophoretic behavior similar to that of the wild-type protein, except mutant Q340E, which surprisingly exhibited a much higher apparent m.w. (Fig. 2). This observation was confirmed by mass spectrometry analysis, which revealed that, whereas the other point mutants at position 340 had mass values of 77,302–77,716 Da comparable to that of wild-type C1s (77,413 Da), the Q340E mutant had a mass of 82,008 Da (Table I). The sequence of the plasmid corresponding to the Q340E mutant was confirmed by dsDNA sequencing, whereas N-terminal sequence analysis of the recombinant protein yielded the expected Glu-Pro-Thr-Met-Tyr. . . sequence. Therefore, it appeared likely that the observed increase in molecular mass was due to an increase in the size of one or both of the N-linked oligosaccharides attached to C1s, at positions 159 and 391 (32). Preliminary characterization of mutant Q340E based on treatment with peptide:*N*-glycosidase F and fragmentation with plasmin (data not shown) provided indication that the observed modification likely occurred at the carbohydrate linked to Asn^{391}. To confirm this hypothesis, a double mutant was expressed, carrying both the Q340E mutation and an Asn to Gln mutation at position 391. Analysis by mass spectrometry of the resulting Q340E/N391Q mutant yielded a mass of 76,117 Da, much lower than the value determined for Q340E, which is consistent with the presence at position 159 of a carbohydrate with a mass of 1,215 ± 40 Da (Table I). Therefore, it became clear that the extra mass of mutant Q340E was due to an increase in the size of the oligosaccharide chain linked to Asn^{391}, which, as deduced from the values determined for the Q340E and Q340E/N391Q mutants, had an estimated mass of ~5905 Da. This value is ~5-fold that of the oligosaccharide chains attached to the other C1s variants expressed in this study (Table I).

To understand why the Q340E mutation had an effect on the extent of glycosylation at Asn^{391}, a three-dimensional model of the C1s catalytic domain was constructed as described under *Materials and Methods*. Based on this model, it appeared plausible, considering the location of residues Gln^{340} and Asp^{343} on either side of the CCP_1-CCP_2 interface, that the Q340E mutation would generate repulsion between the carboxyl groups and thereby possibly modify the relative positioning of the CCP modules (see *Discussion* for more details). Strong support to this hypothesis came from

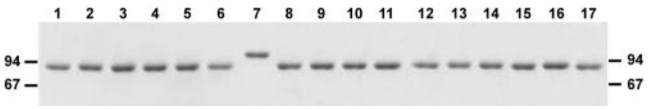


FIGURE 2. SDS-PAGE analysis of the recombinant C1s variants. *Lanes 1–17* correspond to wild-type C1s and mutants Q340A, Q340K, Q340R, Q340S, Q340D, Q340E, Q340E/N391Q, Q340E/D343N, Q340D/D343E, P341I, V342K, D343N, Q340K/D343N, Q340K/P341I, Q340S/P341I, and Q340K/P341I/V342K, respectively. All samples were analyzed under reducing conditions. Molecular masses of marker proteins (expressed in kDa) are indicated.

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

the observation that the double mutant Q340E/D343N had a mass of 77,203 Da, indicating that removal of the negative charge at position 343 restored a normal glycosylation pattern at position 391 (Table I). The double mutant Q340D/D343E also exhibited normal glycosylation at Asn^{391}, suggesting that the presence of Glu at position 340, and Asp at position 343 was an absolute prerequisite for modification of the carbohydrate size. As listed in Table I, a series of other mutations were performed within the C1s sequence stretch 340–343, and none of these other mutations had an effect on the size of the oligosaccharide attached to Asn^{391}.

The Q340E mutation impairs C1s activation

To check whether any of the mutations at position 340 could interfere with C1s ability to undergo activation in the fluid phase, the corresponding variants were incubated with the activated C1r CCP$_2$-SP fragment, and the extent of activation was measured by SDS-PAGE analysis as described under *Materials and Methods*. As shown in Fig. 3, replacement of Gln^{340} by Ala, Lys, Arg, Ser, or Asp had no significant effect on the activation process, but the Q340E mutant was virtually totally resistant to C1r-mediated cleavage. Activation was fully restored in the Q340E/N391Q mutant lacking glycosylation at Asn^{391} and in Q340E/D343N exhibiting a normal glycosylation pattern (Fig. 3*B*), demonstrating that the lack of activation of Q340E was due to the increased size of the oligosaccharide attached to Asn^{391}. Again, none of the other mutations performed at positions 340–343 had a significant effect on C1r-mediated activation of C1s (data not shown).

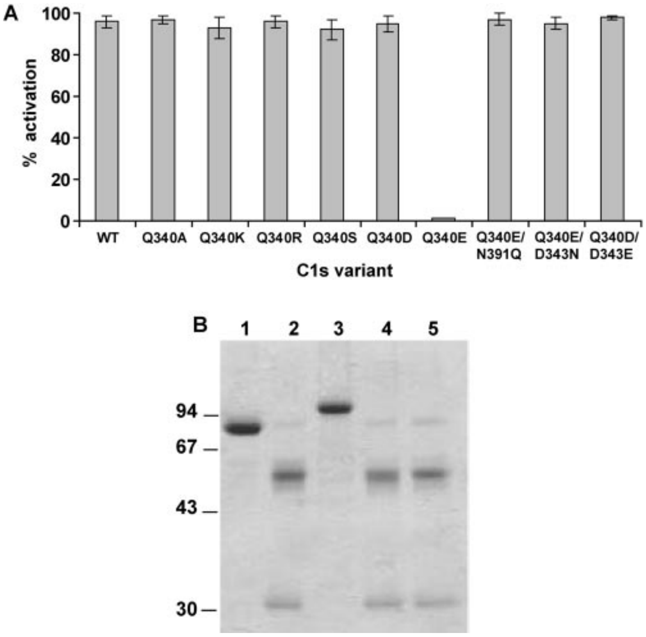


FIGURE 3. Fluid-phase C1s activation by C1r is impaired in the Q340E mutant. *A*, Different C1s variants comprising a mutation at position 340 were submitted to fluid-phase activation by the CCP$_2$-SP fragment of C1r as described under *Materials and Methods*, and the extent of activation was compared with that of wild-type C1s. Error bars correspond to the SD of duplicate experiments. *B*, SDS-PAGE analysis of C1r-mediated activation of selected C1s mutants. *Lane 1*, untreated wild-type C1s; *lanes 2–5*, wild-type C1s and mutants Q340E, Q340E/N391Q, and Q340E/D343N incubated with C1r CCP$_2$-SP for 2 h at 37°C. All samples were analyzed under reducing conditions. OVA was not incorporated in the samples for the sake of clarity. Molecular masses of marker proteins (expressed in kDa) are indicated.

Mutations in the CCP$_1$-CCP$_2$ linker have no significant effect on C1s ability to associate with C1r and C1q and to undergo activation in C1

The ability of the C1s mutants to associate with C1r and C1q to form the C1 complex was measured by incubating each variant with an equimolar amount of proenzyme C1r and then monitoring interaction of the resulting Ca^{2+}-dependent C1s-C1r-C1r-C1s tetramer with immobilized C1q using surface plasmon resonance spectroscopy. As illustrated by the representative examples shown in Fig. 4, most of the samples yielded binding curves very similar to those obtained in the case of wild-type C1s, although slight and likely not significant variations in the binding intensity were observed for some mutants, such as Q340E, P341I, or V342K. The kinetic parameters of the interaction were determined in the case of the tetramer containing wild-type C1s, yielding k_{on} and k_{off} values of $1.0 \pm 0.2 \times 10^6\ \mathrm{M}^{-1}\ \mathrm{s}^{-1}$ and $1.8 \pm 0.3 \times 10^{-3}\ \mathrm{s}^{-1}$, respectively, and a resulting K_{D} of $1.8 \pm 0.5 \times 10^{-9}$ M, indicative of high affinity.

To check whether the C1s mutants retained the ability to activate inside C1, each variant was mixed with appropriate amounts of C1q and proenzyme C1r, and the resulting C1 complex was allowed to activate as described under *Materials and Methods*. As previously observed in the fluid phase, the Q340E mutant was totally resistant to activation, whereas none of the other mutations had a significant effect on the activation process (data not shown). Again, activation of the Q340E/N391Q, Q340E/D343N, and Q340D/D343E mutants was similar to that of wild-type C1s.

Mutations in the CCP$_1$-CCP$_2$ linker alter C4 cleavage but not C2 cleavage

We next investigated the ability of the activated C1s mutants to cleave C4 and C2. For this purpose, all mutants were submitted to fluid-phase activation by the C1r CCP$_2$-SP fragment, and a comparative analysis of their proteolytic activity was performed using the assays described under *Materials and Methods*. As shown in Fig. 5, all mutants cleaved C2 with a relative efficiency comparable to that of wild-type C1s, indicating that none of the mutations had a significant effect on the recognition and/or cleavage of this substrate. In contrast, most of the mutations significantly altered C4 cleavage (Fig. 6). Thus, with the exception of Q340A and Q340E/N391Q, all mutations resulted in a decreased C4-cleaving activity, with particularly pronounced inhibitory effects for point mutants Q340K, P341I, V342K, and D343N. Interestingly, the multiple mutations Q340K/P341I, Q340S/P341I, and Q340K/P341I/V342K, yielding CCP$_1$-CCP$_2$ linkers identical to those of human

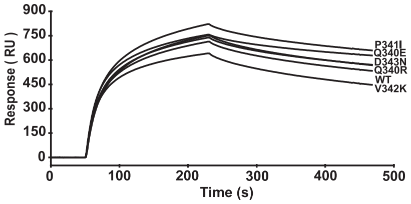


FIGURE 4. Analysis by surface plasmon resonance spectroscopy of the interaction between immobilized C1q and the C1s-C1r-C1r-C1s tetramer reconstituted from various C1s mutants. Wild-type C1s and its mutants were incubated individually with proenzyme C1r in the presence of Ca^{2+} ions, and the resulting C1s-C1r-C1r-C1s tetramers (50 nM) were allowed to bind to C1q (16,500 resonance units) immobilized on the surface of a sensor chip, as described under *Materials and Methods*. From top to bottom, the binding curves shown correspond to P341I, Q340E, D343N, Q340R, wild-type C1s, and V342K.

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

P11

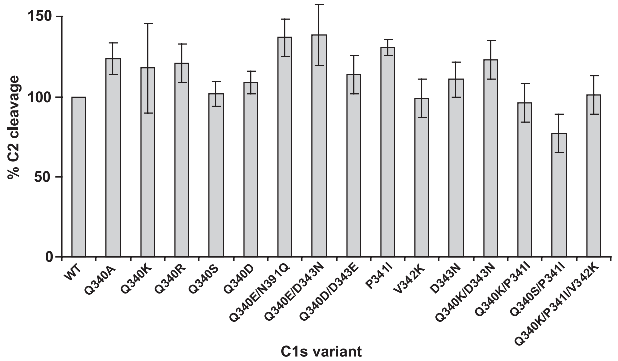


FIGURE 5. Comparative analysis of C2 cleavage by the C1s variants. Wild-type C1s and its mutants were activated in the fluid phase by the C1r CCP_2-SP fragment and tested for their ability to cleave C2, as described under *Materials and Methods*. Error bars correspond to the SD of triplicate experiments.

MASP-1/3, MASP-2, and C1r, respectively (see Fig. 1), all resulted in C4-cleaving activities $< 20\%$ compared with that of wild-type C1s.

Additional experiments were aimed at measuring the C4-cleaving activity of the activated C1s mutants in the context of the macromolecular C1 complex. For this purpose, representative mutants were incorporated in C1, allowed to activate, and their C4-cleaving activity was measured in the complex. As shown in Fig. 6, although differences were observed in the case of mutants Q340A and Q340K, most of the mutants tested (P341I, D343N, Q340K/D343N, Q340S/P341I, and Q340E/N391Q) exhibited a cleavage pattern similar to that obtained with isolated C1s.

To investigate why the mutations in the CCP_1-CCP_2 linker altered the C4-cleaving activity of C1s, the kinetic constants of the reaction were determined for some of the mutants in their isolated form. As listed in Table II, all mutants tested had K_m values similar to that determined for wild-type C1s. Consistent with the data shown in Fig. 6, mutants Q340A and Q340E/N391Q had k_{cat} values close to that of wild-type C1s. In contrast, all mutants showing decreased C4-cleaving activity in Fig. 6 (Q340K, P341I, D343N, and Q340S/P341I) had decreased k_{cat} values.

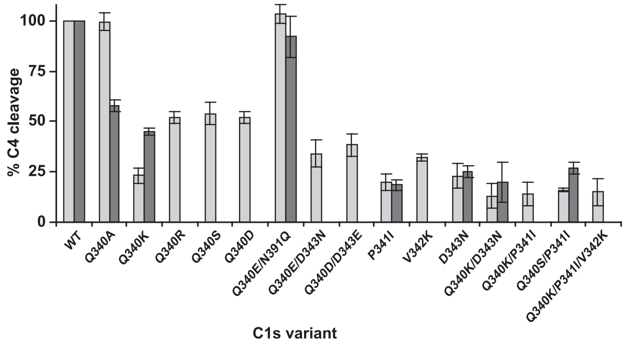


FIGURE 6. Comparative analysis of C4 cleavage by the C1s variants. Wild-type C1s and all mutants were activated in the fluid phase by the C1r CCP_2-SP fragment and tested for their ability to cleave C4, as described under *Materials and Methods*. The resulting C4-cleaving activity of each mutant is expressed relative to that of wild-type C1s and shown as light gray bars. In a second series of experiments, wild-type C1s and selected mutants (Q340A, Q340K, Q340E/N391Q, P341I, D343N, Q340K/D343N, and Q340S/P341I) were incorporated into C1, allowed to activate, and then tested for their ability to cleave C4 in the context of the C1 complex. The relative C4-cleaving activity of these mutants is shown as dark gray bars. Error bars correspond to the SD of triplicate experiments.

Table II. *Kinetic constants for proteolytic cleavage of C4 by different C1s variants*

C1s Variant	$k_{cat}{}^a$ (s^{-1})	$K_m{}^a$ (μM)	k_{cat}/K_m $(M^{-1} \times s^{-1})$
Wild type	6.7 ± 1.0	2.3 ± 0.9	2.95×10^6
Q340A	7.7 ± 0.7	2.6 ± 0.1	3.00×10^6
Q340E/N391Q	7.3 ± 0.5	2.0 ± 0.5	3.65×10^6
Q340K	2.5 ± 0.9	3.2 ± 1.1	0.80×10^6
P341I	1.8 ± 0.9	3.2 ± 0.7	0.60×10^6
D343N	2.6 ± 1.0	1.9 ± 0.5	1.40×10^6
Q340S/P341I	1.3 ± 1.0	2.9 ± 0.6	0.45×10^6

[a] Mean values determined from three separate experiments.

Discussion

The aim of this study was to explore the functional role of the segment 340–343 at the interface between the CCP_1 and CCP_2 modules of C1s. For this purpose, a series of mutants were generated by site-directed mutagenesis, expressed using a baculovirus/insect cells system, and tested for their ability to mediate the various interaction and catalytic properties of C1s, as measured outside and within the C1 complex.

A totally unexpected finding of these investigations is that substitution of Glu for Gln at position 340 results in a marked increase in the mass of the oligosaccharide chain attached to Asn^{391} from a value of ~1250 Da, corresponding to the average size of the high-mannose N-linked carbohydrates usually produced by insect cells (13), to a value of ~5905 Da. The presence of Glu at position 340 was found to be a specific requirement, as substitution with other residues, including Asp, was ineffective. In addition, none of the other mutants produced in this study had abnormal glycosylation at Asn^{391}. The fact that the double mutant Q340E/D343N exhibited a normal glycosylation pattern rules out the possibility of a specific recognition of Glu^{340} by one of the enzymes involved in the glycosylation process. On the other hand, this observation provides strong support to the hypothesis that introducing a Glu at position 340 generates repulsion between its negative charge and that of residue Asp^{343}. Thus, despite the fact that the model of the C1s catalytic domain at the origin of this hypothesis (Fig. 7) is partly based on the CCP_1-CCP_2 interface as seen in C1r (12), it is probably valid to some extent, at least with regard to the relative positioning of residues 340 and 343 on either side of the linker. Based on these observations, we suggest that repulsion between the carboxyl groups of Glu^{340} and Asp^{343} modifies the structure and/or flexibility of the CCP_1-CCP_2 linker of C1s in such a way that it facilitates access of the glycosylation enzymes to Asn^{391}. This

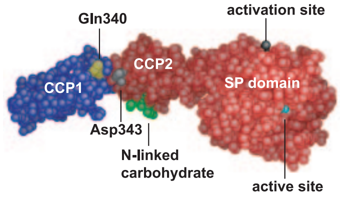


FIGURE 7. Three-dimensional space-filling model of the C1s catalytic domain. The model was built as described under *Materials and Methods*. The CCP_1 module is colored blue, whereas CCP_2 and the SP domain are colored red. The side chains of residues Gln^{340} and Asp^{343} are shown in yellow and gray, respectively. The position of the active site is shown by residue Ser^{617} (light blue), and the approximate location of the Arg^{422}-Ile^{423} activation site is indicated by residue Phe^{417} (black). The structure of the carbohydrate moiety attached to Asn^{391}, shown in green, is as defined in the x-ray structure of the C1s CCP_2-SP fragment (14).

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

hypothesis appears plausible in light of the location of the oligosaccharide chain relative to the CCP_1-CCP_2 linker (Fig. 7). It is interesting to note that, whereas Asn^{159} of C1s purified from human serum is linked to a homogeneous biantennary oligosaccharide, Asn^{391} is occupied by either a biantennary oligosaccharide, or a triantennary, trisialylated species, or a fucosylated triantennary, trisialylated species, in relative proportions of ~1:1:1 (32). It appears likely from our data that this natural heterogeneity arises from an intrinsic variability in the conformation of the C1s CCP_1-CCP_2 linker.

Another interesting observation is that the increased size of the carbohydrate attached to Asn^{391} totally prevents C1s activation by C1r, both in the fluid phase and in the context of the C1 complex. Therefore, it may be concluded that due to its abnormal size the carbohydrate chain linked to Asn^{391} extends over the C1s SP domain on the side opposite to the active site, hence creating steric hindrance and preventing access of the C1r SP domain to the C1s activation site (Fig. 7).

In contrast, overglycosylation at Asn^{391} has no effect on C1 reconstitution, as judged from the ability of the Q340E mutant to assemble with C1r and C1q (Fig. 4). In this respect, it is remarkable that none of the mutations at the CCP_1-CCP_2 linker had any significant effect on the ability of C1s to incorporate in C1 as shown by surface plasmon resonance spectroscopy, implying that these have no impact on the ability of C1s to associate with C1r, and of the resulting C1s-C1r-C1r-C1s tetramer to associate with C1q. In addition, once incorporated in C1, all mutants, except Q340E, retained activation properties similar to that of wild-type C1s. It may be concluded from these observations that flexibility of the C1s CCP_1-CCP_2 linker plays no significant role in C1 assembly or in the mechanism involved in C1s activation by C1r inside the C1 complex. Thus, contrary to what was initially hypothesized (14), it appears unlikely that C1s activation by C1r requires a movement of its SP domain toward the interior of the C1 complex to allow cleavage of its susceptible Arg-Ile bond by the C1r active site. As proposed in a recent model (4), a more likely hypothesis is that C1s SP domain is located at the periphery of C1 and that the C1r SP domain moves toward the outside to mediate cleavage.

Our observation that none of the mutations in the CCP_1-CCP_2 linker had a significant effect on the C2-cleaving activity of C1s is consistent with previous findings (13) indicating that the CCP_2-SP and SP fragments of C1s retain the ability of full-length C1s to cleave C2 and provide additional evidence that C2 recognition and cleavage by C1s only involve structural motifs located within its SP domain. In contrast, our data indicate that most of the mutations in the CCP_1-CCP_2 linker result in a significant decrease of the C4-cleaving activity of C1s. Again, these findings are in line with previous studies indicating that removal of the CCP_1 module of C1s greatly reduces (~70-fold) its ability to cleave C4 (13). Taken together, these observations strongly support the hypothesis that C4 recognition by C1s involves a major binding site located in CCP_1, allowing positioning of the protein substrate in such a way that its cleavage site fits properly into the C1s active site. In this hypothesis, flexibility at the CCP_1-CCP_2 junction appears as a critical factor because it would be used to finely adjust positioning of C4 for optimal cleavage by the C1s active site. This hypothesis is strongly supported by the observation that, for all mutants tested, the decreased C4-cleaving ability arises from a decreased k_{cat} value, the K_m of the reaction being unchanged. Indeed, modification at the CCP_1-CCP_2 junction would not significantly modify the affinity of the interaction between C4 and the CCP_1 module but would be expected to alter the flexibility of the linker and thereby, in most cases, would not allow optimal positioning of C4 with respect to C1s active site, hence the observed decreased k_{cat} values. In this respect, it is interesting to note that, whereas the Q340K mutant is poorly active on C4, the double mutant Q340E/N391Q shows optimal activity. A tempting hypothesis is that due to the repulsion effect discussed above the CCP_1-CCP_2 linker shows increased flexibility in the latter case, allowing optimal cleavage. In contrast, the linker could be "locked" in a more rigid configuration in the Q340K mutant, hence its decreased efficiency. It is also noteworthy that replacement of the C1s CCP_1-CCP_2 linker by that of other members of the C1r/C1s/MASP family results in all cases in a decreased C4-cleaving activity. The fact that this also applies to MASP-2, a protease that shares the ability to cleave C4, provides strong indication that C1s and MASP-2 use different mechanisms to recognize this substrate.

Finally, it is interesting to note that most of the mutations at the CCP_1-CCP_2 interface have very similar effects on the C4-cleaving activity of C1s, whether this is measured on the isolated protease or in the context of the C1 complex. In this respect, it appears unlikely that the discrepancies observed in the case of Q340A and Q340K reflect a particular behavior of these mutants. Thus, in line with other data discussed above, it may be deduced from these observations that the C1s catalytic domain is readily accessible in C1, in full agreement with the hypothesis that it is located at the periphery of the complex (4).

Acknowledgments

We thank B. Dublet and J.-P. Andrieu for performing mass spectrometry and N-terminal sequence analyses.

Disclosures

The authors have no financial conflict of interest.

References

1. Cooper, N. R. 1985. The classical complement pathway: activation and regulation of the first complement component. *Adv. Immunol.* 37: 151–216.
2. Schumaker, V. N., P. Zavodszky, and P. H. Poon. 1987. Activation of the first component of complement. *Annu. Rev. Immunol.* 5: 21–42.
3. Arlaud, G. J., M. G. Colomb, and J. Gagnon. 1987. A functional model of the human C1 complex. *Immunol. Today* 8: 106–111.
4. Gaboriaud, C., N. M. Thielens, L. A. Gregory, V. Rossi, J. C. Fontecilla-Camps, and G. J. Arlaud. 2004. Structure and activation of the C1 complex of complement: unraveling the puzzle. *Trends Immunol.* 25: 368–373.
5. Bork, P., and G. Beckmann. 1993. The CUB domain. A widespread module in developmentally regulated proteins. *J. Mol. Biol.* 231: 539–545.
6. Reid, K. B. M., D. R. Bentley, R. D. Campbell, L. P. Chung, R. B. Sim, T. Kristensen, and B. F. Tack. 1986. Complement system proteins that interact with C3b or C4b: a superfamily of structurally related proteins. *Immunol. Today* 7: 230–234.
7. Fujita, T. 2002. Evolution of the lectin-complement pathway and its role in innate immunity. *Nat. Rev. Immunol.* 2: 346–353.
8. Thielens, N. M., C. A. Aude, M. B. Lacroix, J. Gagnon, and G. J. Arlaud. 1990. Ca^{2+} binding properties and Ca^{2+}-dependent interactions of the isolated NH_2-terminal α fragments of human complement proteases C1r and C1s. *J. Biol. Chem.* 265: 14469–14475.
9. Busby, T. F., and K. C. Ingham. 1990. NH_2-terminal calcium-binding domain of human complement C1s mediates the interaction of C1r with C1q. *Biochemistry* 29: 4613–4618.
10. Gregory, L. A., N. M. Thielens, G. J. Arlaud, J. C. Fontecilla-Camps, and C. Gaboriaud. 2003. X-ray structure of the Ca^{2+}-binding interaction domain of C1s: insights into the assembly of the C1 complex of complement. *J. Biol. Chem.* 278: 32157–32164.
11. Rossi, V., C. Gaboriaud, M. Lacroix, J. Ulrich, J. C. Fontecilla-Camps, J. Gagnon, and G. J. Arlaud. 1995. Structure of the catalytic region of human complement protease C1s: study by chemical cross-linking and three-dimensional homology modeling. *Biochemistry* 34: 7311–7321.
12. Budayova-Spano, M., M. Lacroix, N. M. Thielens, G. J. Arlaud, J. C. Fontecilla-Camps, and C. Gaboriaud. 2002. The crystal structure of the zymogen catalytic domain of complement protease C1r reveals that a disruptive mechanical stress is required to trigger activation of the C1 complex. *EMBO J.* 21: 231–239.
13. Rossi, V., I. Bally, N. M. Thielens, A. F. Esser, and G. J. Arlaud. 1998. Baculovirus-mediated expression of truncated modular fragments from the catalytic region of human complement serine protease C1s: evidence for the involvement of both complement control protein modules in the recognition of the C4 protein substrate. *J. Biol. Chem.* 273: 1232–1239.

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

P11

14. Gaboriaud, C., V. Rossi, I. Bally, G. J. Arlaud, and J. C. Fontecilla-Camps. 2000. Crystal structure of the catalytic domain of human complement C1s: a serine protease with a handle. *EMBO J.* 19: 1755–1765.
15. Barlow, P. N., A. Steinkasserer, D. G. Norman, B. Kieffer, A. P. Wiles, R. B. Sim, and I. D. Campbell. 1993. Solution structure of a pair of complement modules by nuclear magnetic resonance. *J. Mol. Biol.* 232: 268–284.
16. Wiles, A. P., G. Shaw, J. Bright, A. Perczel, I. D. Campbell, and P. N. Barlow. 1997. NMR studies of a viral protein that mimics the regulators of complement activation. *J. Mol. Biol.* 272: 253–265.
17. Casasnovas, J. M., M. Larvie, and T. Stehle. 1999. Crystal structure of two CD46 domains reveals an extended measles virus-binding surface. *EMBO J.* 18: 2911–2922.
18. Kirkitadze, M. D., and P. N. Barlow. 2001. Structure and flexibility of the multiple domain proteins that regulate complement activation. *Immunol. Rev.* 180: 146–161.
19. March, S. C., I. Parikh, and P. Cuatrecasas. 1974. A simplified method for cyanogen bromide activation of agarose for affinity chromatography. *Anal. Biochem.* 60: 149–152.
20. Luo, C., N. M. Thielens, J. Gagnon, P. Gal, M. Sarvari, Y. Tseng, M. Tosi, P. Zavodszky, G. J. Arlaud, and V. N. Schumaker. 1992. Recombinant human complement subcomponent C1s lacking β-hydroxyasparagine, sialic acid, and one of its two carbohydrate chains still reassembles with C1q and C1r to form a functional C1 complex. *Biochemistry* 31: 4254–4262.
21. Arlaud, G. J., R. B. Sim, A. M. Duplaa, and M. G. Colomb. 1979. Differential elution of Clq, Clr and Cls from human Cl bound to immune aggregates: use in the rapid purification of Cl subcomponents. *Mol. Immunol.* 16: 445–450.
22. Arlaud, G. J., C. L. Villiers, S. Chesne, and M. G. Colomb. 1980. Purified proenzyme C1r: some characteristics of its activation and subsequent proteolytic cleavage. *Biochim. Biophys. Acta* 616: 116–129.
23. Lacroix, M., C. Ebel, J. Kardos, J. Dobo, P. Gal, P. Zavodszky, G. J. Arlaud, and N. M. Thielens. 2001. Assembly and enzymatic properties of the catalytic domain of human complement protease C1r. *J. Biol. Chem.* 276: 36233–36240.
24. Thielens, N. M., M. B. Villiers, A. Reboul, C. L. Villiers, and M. G. Colomb. 1982. Human complement subcomponent C2: purification and proteolytic cleavage in fluid phase by C1s, $C1r_2$-$C1s_2$ and C1. *FEBS Lett.* 141: 19–24.
25. Dodds, A. W. 1993. Small-scale preparation of complement components C3 and C4. *Methods Enzymol.* 223: 46–61.
26. Isenman, D. E., and D. I. C. Kells. 1982. Conformational and functional changes in the fourth component of human complement produced by nucleophilic modification and by proteolysis with C1s. *Biochemistry* 21: 1109–1117.
27. Budzko, D. B., and H. J. Müller-Eberhard. 1970. Cleavage of the fourth component of human complement (C4) by C1 esterase: isolation and characteristics of the low molecular weight product. *Immunochemistry* 7: 227–234.
28. Thielens, N. M., S. Cseh, S. Thiel, T. Vorup-Jensen, V. Rossi, J. C. Jensenius, and G. J. Arlaud. 2001. Interaction properties of human mannan-binding lectin (MBL)-associated serine proteases-1 and -2, MBL-associated protein 19, and MBL. *J. Immunol.* 166: 5068–5077.
29. Lacroix, M., V. Rossi, C. Gaboriaud, S. Chevallier, M. Jaquinod, N. M. Thielens, J. Gagnon, and G. J. Arlaud. 1997. Structure and assembly of the catalytic region of human complement protease C1r: a three-dimensional model based on chemical cross-linking and homology modeling. *Biochemistry* 36: 6270–6282.
30. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680–685.
31. Gregory, L. A., N. M. Thielens, M. Matsushita, R. Sorensen, G. J. Arlaud, J. C. Fontecilla-Camps, and C. Gaboriaud. 2004. The x-ray structure of human mannan-binding lectin-associated protein 19 (MAp19) and its interaction site with mannan-binding lectin and L-ficolin. *J. Biol. Chem.* 279: 29391–29397.
32. Pétillot, Y., P. Thibault, N. M. Thielens, V. Rossi, M. Lacroix, B. Coddeville, G. Spik, V. N. Schumaker, J. Gagnon, and G. J. Arlaud. 1995. Analysis of the N-linked oligosaccharides of human C1s using electrospray ionisation mass spectrometry. *FEBS Lett.* 358: 323–328.

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

P12

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 280, NO. 51, pp. 41811–41818, December 23, 2005
© 2005 by The American Society for Biochemistry and Molecular Biology, Inc. Printed in the U.S.A.

Functional Characterization of Complement Proteases C1s/Mannan-binding Lectin-associated Serine Protease-2 (MASP-2) Chimeras Reveals the Higher C4 Recognition Efficacy of the MASP-2 Complement Control Protein Modules*

Received for publication, April 8, 2005, and in revised form, October 11, 2005 Published, JBC Papers in Press, October 14, 2005, DOI 10.1074/jbc.M503813200

Véronique Rossi, Florence Teillet, Nicole M. Thielens, Isabelle Bally, and Gérard J. Arlaud[1]

From the Laboratoire d'Enzymologie Moléculaire, Institut de Biologie Structurale Jean-Pierre Ebel, 38027 Grenoble Cedex 1, France

C1s and mannan-binding lectin-associated serine protease-2 (MASP-2) are the proteases that trigger the classical and lectin pathways of complement, respectively. They have identical modular architectures and cleave the same substrates, C2 and C4, but show markedly different efficiencies toward C4. Multisite-directed mutagenesis was used to engineer hybrid C1s/MASP-2 molecules where either the complement control protein (CCP) modules or the serine protease (SP) domain of C1s were swapped for their MASP-2 counterparts. The resulting chimeras ($\text{C1s}_{(\text{MASP-2 CCP1/2})}$ and $\text{C1s}_{(\text{MASP-2 SP})}$, respectively) were expressed and characterized chemically and functionally. Whereas $\text{C1s}_{(\text{MASP-2 SP})}$ was recovered as an active enzyme, $\text{C1s}_{(\text{MASP-2 CCP1/2})}$ was produced in a proenzyme form and was susceptible to activation by C1r, indicating that the activation properties of the chimeras were dictated by the nature of their SP domain. Similarly, each activated chimera had an esterolytic activity characteristic of its own SP domain and cleaved C2 with an efficiency comparable with that of their parent C1s and MASP-2 proteases. Both chimeras cleaved C4, but whereas $\text{C1s}_{(\text{MASP-2 SP})}$ and C1s had K_m values in the micromolar range, $\text{C1s}_{(\text{MASP-2 CCP1/2})}$ and MASP-2 had K_m values in the nanomolar range, resulting in 21–27-fold higher k_{cat}/K_m ratios. Thus, the higher C4 cleavage efficiency of MASP-2 arises from a higher substrate recognition efficacy of its CCP modules. Remarkably, $\text{C1s}_{(\text{MASP-2 CCP1/2})}$ retained C1s ability to associate with C1r and C1q to form a pseudo-C1 complex and to undergo activation within this complex, indicating that the C1s-CCP modules have no direct implication in either function.

The classical and lectin pathways of complement activation, two major components of innate immunity against pathogenic microorganisms in vertebrates, elicit the same defense reactions in response to infection, and involve similar recognition and triggering mechanisms. C1, the complex that triggers the classical pathway, comprises a flower-like hexameric recognition subunit C1q and two copies of homologous, yet distinct modular serine proteases, C1r and C1s, which together form a Ca^{2+}-dependent tetramer C1s-C1r-C1r-C1s. C1r is responsible for internal C1 activation, a two-step process first involving autolytic C1r activation and then cleavage of proenzyme C1s. Activated C1s is a highly specific trypsin-like protease that mediates limited proteolysis of C4 and C2, the natural protein substrates of C1, thereby triggering activation of the pathway (1–3).

The lectin pathway can be triggered by three different oligomeric lectins (mannan-binding lectin, L-ficolin and H-ficolin), which exhibit an overall structure reminiscent of that of C1q and share the ability to form a complex with mannan-binding lectin-associated serine protease-2 (MASP-2)[2] (4–6). MASP-2 (7) has a modular structure homologous to that of C1r and C1s, with an N-terminal CUB module (8), a Ca^{2+}-binding epidermal growth factor-like module (9), a second CUB module, two contiguous CCP modules (10), and a C-terminal chymotrypsin-like serine protease (SP) domain. The catalytic subunit of each activating complex is thought to be a MASP-2 homodimer, which thus combines C1r property to self-activate and C1s ability to cleave C4 and C2 (11–13).

Whereas C1s and MASP-2 have comparable C2 cleaving activities, MASP-2 cleaves C4 more efficiently than does C1s, because of a much lower K_m for this substrate (12). Expression of recombinant modular segments from the C-terminal catalytic region of both proteases has revealed that, whereas C2 cleavage is mediated by their SP domain alone, C4 cleavage requires additional sites located in the CCP modules, likely involved in substrate recognition (14, 15). The present study was undertaken to assess the relative contribution of the SP domain and the CCP modules of C1s and MASP-2 to their C4 cleaving activity and investigate what determines the higher efficiency of MASP-2 toward this substrate. For this purpose, we took advantage of the fact that C1s and MASP-2 share homologous modular structures to engineer chimeric C1s molecules comprising either the CCP modules or the SP domain of MASP-2. This domain swapping strategy allows us to demonstrate that the higher C4-cleavage efficiency of MASP-2 arises from the higher C4 recognition efficacy of its CCP modules and reveals that the SP domain of each protease is able to cooperate with the CCP modules of the other.

EXPERIMENTAL PROCEDURES

Materials—Diisopropyl phosphorofluoridate and the synthetic substrate Ac-Gly-Lys-OMe were from Sigma. Oligonucleotides were obtained from Oligoexpress, Paris. Restriction enzymes were from New

---

* This work was supported by the Commissariat à l'Energie Atomique, the CNRS, and the Université Joseph Fourier, Grenoble. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] To whom correspondence should be addressed. Tel.: 33-4-38-78-49-81; Fax: 33-4-38-78-54-94; E-mail: arlaud@ibs.fr.

[2] The abbreviations used are: MASP, mannan-binding lectin-associated serine protease; CCP, complement control protein; CUB, module originally found in complement C1r/C1s, Uegf, and bone morphogenetic protein; SP, serine protease domain; $\text{C1s}_{(\text{MASP-2 SP})}$, chimeric C1s molecule containing the SP domain of MASP-2; $\text{C1s}_{(\text{MASP-2 CCP1/2})}$, chimeric C1s molecule containing the $\text{CCP}_{1/2}$ modules of MASP-2.

P12

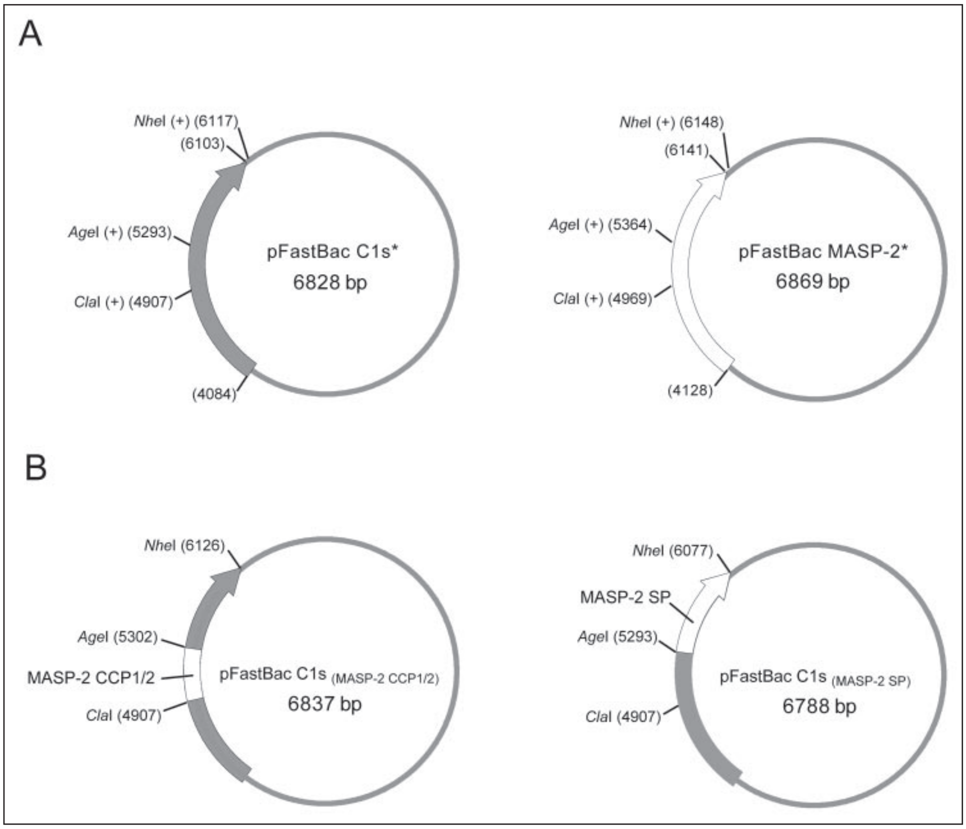


FIGURE 1. **Construction of the expression plasmids.** Three single restriction sites ClaI, AgeI, and NheI were introduced into the parental plasmids pFastBac C1s and pFastBac MASP-2. The resulting intermediate plasmids pFastBac C1s* and pFastBac MASP-2* (*A*) were used to generate the chimeric expression plasmids pFastBac C1s$_{\text{(MASP-2 CCP1/2)}}$ and pFastBac C1s$_{\text{(MASP-2 SP)}}$ (*B*), as described under "Experimental Procedures." The *numbers* in *parentheses* indicate positions of domain boundaries and restriction sites in the DNA sequence of the vectors.

England Biolabs, Beverly, MA. The C1s-Sepharose column was prepared by coupling activated C1s purified from human serum to CNBr-activated Sepharose 4B according to March *et al.* (16). Plasmids pBS-C1s, containing the full-length human C1s cDNA (17), and pBS-MASP-2, containing the full-length human MASP-2 cDNA (7), were kindly provided by Dr. Mario Tosi (University of Rouen, France) and Dr. Steffen Thiel (University of Aarhus, Denmark), respectively. Plasmid pFastBac1 was from Invitrogen.

Proteins—Complement proteins C2, C4, C1q, activated C1s, and proenzyme C1r were purified from human plasma according to published procedures (18–21). Wild-type C1s, the activated CCP$_2$-SP fragment of C1r, and the CCP$_{1/2}$-SP fragment of MASP-2 were produced in insect cells and purified as described previously (12, 22). C1 inhibitor was purified from human plasma essentially as described in Rossi *et al.* (12). The concentration of purified proteins was determined using the following absorption coefficients ($A_{1\%,1\ \text{cm}}$ at 280 nm) and the following molecular weights: C2, 10.0 and 100,000 (18); C4, 8.3 and 205,000 (19); C1q, 6.8 and 459,300 (20); C1r, 12.4 and 86,300 (23); wild-type C1s and C1s*, 14.5 and 77,400; C1r fragment CCP$_2$-SP, 14.9 and 40,400 (22), MASP-2 CCP$_{1/2}$-SP fragment, 18.3 and 42,700 (12), C1 inhibitor, 4.5 (24) and 104,000 (25). The absorption coefficients ($A_{1\%,1\ \text{cm}}$ at 280 nm) used for recombinant C1s$_{\text{(MASP-2 CCP1/2)}}$ and C1s$_{\text{(MASP-2 SP)}}$ (14.8 and 15.7, respectively) were calculated according to Edelhoch (26) using molecular weight values of 76,700 and 75,800, respectively, calculated from the amino acid sequence and the carbohydrate content. Because of the contaminating A chain fragment in the C1s$_{\text{(MASP-2 SP)}}$ preparation, the relative amount of the full-length, active enzyme was determined from SDS-PAGE and N-terminal sequence analysis.

Construction of Expression Plasmids—The parental plasmids used in this study, pFastBac C1s and pFastBac MASP-2, were derived from pBS-C1s and pBS-MASP-2 as described previously (12, 22). The plasmids encoding the chimeric proteins were generated by site-directed mutagenesis of the parental plasmids using the QuikChange$^{\text{TM}}$ multi site-directed mutagenesis kit from Stratagene (La Jolla, CA). Three single restriction sites, ClaI, AgeI, and NheI, were introduced into pFastBac C1s and pFastBac MASP-2 at the boundaries between modules CUB$_2$ and CCP$_1$, between CCP$_2$ and the SP domain, and after the stop codon, respectively (Fig. 1). The resulting mutated plasmids named pFastBac C1s* and pFastBac MASP-2* were subsequently used for domain swapping. The chimeric plasmid pFastBac C1s$_{\text{(MASP-2 CCP1/2)}}$ was obtained by inserting the ClaI (4969)/AgeI (5364) segment of pFastBac MASP-2* within the ClaI (4907)/AgeI (5293) sites of pFastBac C1s*. The chimeric plasmid pFastBac C1s$_{\text{(MASP-2 SP)}}$ was obtained in a similar way by insertion of the AgeI (5364)/NheI (6148) segment of pFastBac MASP-2* within the AgeI (5293)/NheI (6117) sites of pFastBac C1s*. All DNA constructs used in this study were checked by double-stranded DNA sequencing (Genome Express, Grenoble, France).

Cells and Viruses—The insect cells *Spodoptera frugiperda* (Ready-Plaque Sf9 cells from Novagen) and *Trichoplusia ni* (High Five$^{\text{TM}}$) were routinely grown and maintained as described previously (27). Recombinant baculoviruses were generated using the Bac-to-Bac$^{\text{TM}}$ system (Invitrogen). The bacmid DNA was purified using the Qiagen midiprep purification system (Qiagen S. A., Courtaboeuf, France) and used to transfect Sf9 insect cells utilizing cellfectin in Sf900 II SFM medium (Invitrogen) as recommended by the manufacturer. Recombinant virus particles were collected 4 days later and amplified as described by King and Possee (28).

Protein Production and Purification—All chimeric proteins were produced and purified using the same procedure, except that diisopropyl phosphorofluoridate was omitted at all steps in the case of C1s$_{\text{(MASP-2 SP)}}$. Briefly, High Five cells (1.75×10^{7} cells/175-cm^2 tissue culture flask) were infected with the recombinant viruses at a multiplicity of infection of 2–3 in Sf900 II SFM medium at 28 °C for 72 h. Culture supernatants were collected by centrifugation, dialyzed against 50 mM triethanolamine hydrochloride, 145 mM NaCl, pH 7.4, and loaded onto the C1s-Sepharose column (3 × 5 cm) after the addition of 2 mM $CaCl_2$ and 2 mM diisopropyl phosphorofluoridate. Elution of the bound material was carried out with the same buffer containing 5 mM EDTA instead of $CaCl_2$. The eluate was concentrated to 0.3–0.5 mg/ml by centrifugation on a PM-30 membrane (Amicon, Millipore), dialyzed against 50 mM triethanolamine hydrochloride, 145 mM NaCl, pH 7.4, and stored at −20 °C until use.

Activation of Wild-type C1s, C1s, and $C1s_{(MASP\text{-}2\ CCP1/2)}$*—Wild-type C1s and its C1s* and $C1s_{(MASP\text{-}2\ CCP1/2)}$ variants (3.6 μM each) were activated by incubation with a 2.2% molar ratio of the activated C1r CCP_2-SP fragment for 100 min at 37 °C in 50 mM triethanolamine hydrochloride, 145 mM NaCl, pH 7.4. Activation was monitored by SDS-PAGE analysis under reducing conditions and subsequent quantification by gel scanning of the A and B chains characteristic of activated C1s.

Esterolytic Assays—The esterolytic activity of wild-type C1s, C1s*, the $CCP_{1/2}$-SP fragment of MASP-2, and both C1s/MASP-2 chimeras was measured on the synthetic ester Ac-Gly-Lys-OMe using an assay described previously (29). Initial cleavage rates were determined at 30 °C at a substrate concentration of 1 mM and at protease concentrations ranging from 1 to 3 nM. The esterolytic activity of each variant (measured as mol of substrate cleaved/s/mol of enzyme) was expressed relative to that of wild-type C1s.

Proteolytic Assays—The kinetic parameters of C4 and C2 cleavage were determined using the SDS-PAGE analysis-based method described previously (12, 14). For C4 cleavage by the $CCP_{1/2}$-SP fragment of MASP-2 and the $C1s_{(MASP\text{-}2\ CCP1/2)}$ chimera, the enzyme concentration was 0.5 nM, and substrate concentrations ranged from 50 to 500 nM. C4 cleavage by wild-type C1s, C1s*, and the $C1s_{(MASP\text{-}2\ SP)}$ hybrid was performed at a 1 nM enzyme concentration, with substrate concentrations ranging from 0.5 to 4 μM. C2 cleavage was carried out under the same conditions in all instances, *i.e.* a 2 nM enzyme concentration and with C2 concentrations ranging from 1 to 4 μM. The kinetics constants were determined from initial rates analyses by the Lineweaver-Burk method using linear regression analysis and are based on four to five experiments at six different substrate concentrations.

Titration with C1 Inhibitor—Titration of the enzymatic activity of C1s and the $C1s_{(MASP\text{-}2\ CCP1/2)}$ and $C1s_{(MASP\text{-}2\ SP)}$ chimeras was performed essentially as described previously (12) by preincubating the proteases (0.25 μM each) with increasing concentrations of C1 inhibitor (50–500 nM) for 15 min at 37 °C in 145 mM NaCl, 50 mM triethanolamine-HCl, 1 mg/ml ovalbumin, pH 7.4. Residual C2 cleaving activity was then measured by incubating the enzymes (3.75 nM for $C1s_{(MASP\text{-}2\ SP)}$ and 7.5 nM for C1s and $C1s_{(MASP\text{-}2\ CCP1/2)}$) with 2.25 μM C2 for 10 min at 37 °C followed by SDS-PAGE analysis of the cleavage as described above.

Surface Plasmon Resonance Spectroscopy—Analyses were performed using a BIAcore 3000 instrument (BIAcore AB, Uppsala, Sweden). C1q was immobilized on the surface of a CM5 sensor chip (BIAcore AB) using the amine-coupling chemistry as described previously (30). The C1s-C1r-C1r-C1s tetramer (10 nM) was reconstituted from proenzyme C1r and each C1s variant in 145 mM NaCl, 2 mM $CaCl_2$, 50 mM triethanolamine hydrochloride, pH 7.4, containing 0.005% surfactant P20 (BIAcore AB). Binding to immobilized C1q (10,000 resonance units) was measured at a flow rate of 20 μl/min in the buffer used for tetramer reconstitution. Equivalent volumes of each tetramer were injected on a surface with immobilized bovine serum albumin to serve as blank sensorgrams. Regeneration of the surface was realized with injection of the same buffer containing 5 mM EDTA instead of 2 mM $CaCl_2$. Data were analyzed by global fitting to a 1:1 Langmuir binding model of the association and dissociation phases for several concentrations simultaneously using the BIAevaluation 3.1 software (BIAcore AB). The apparent equilibrium dissociation constants (K_D) were calculated from the ratio of the dissociation k_{off} and association k_{on} rate constants, and maximal binding capacities (R_{max}) were determined using the same model.

Activation of $C1s_{(MASP\text{-}2\ CCP1/2)}$ within the C1 Complex—C1 (0.25 μM) was reconstituted from equimolar amounts of C1q and the tetramer containing proenzyme C1r and either wild-type C1s, C1s*, or $C1s_{(MASP\text{-}2\ CCP1/2)}$ and incubated in 145 mM NaCl, 2 mM $CaCl_2$, 50 mM triethanolamine hydrochloride, pH 7.4, for various periods at 37 °C to allow spontaneous activation (31). The reaction mixtures were submitted to SDS-PAGE analysis under reducing conditions, and activation was measured by Western blot using a polyclonal anti-C1s antibody as described previously (14).

Other Methods—SDS-PAGE analysis was performed as described previously (23). Western blot analysis of the recombinant proteins was performed according to published procedures (14) using anti-C1s and anti-MASP-2 antibodies (12). N-terminal sequence analysis was performed as described previously (32).

RESULTS

The objective of this study was to evaluate the respective contribution of the SP domain and CCP modules of C1s and MASP-2 to their catalytic activity, notably with respect to their differential efficiency toward C4. For this purpose, two chimeric C1s molecules containing either the SP domain of MASP-2 ($C1s_{(MASP\text{-}2\ SP)}$) or its CCP_1 and CCP_2 modules ($C1s_{(MASP\text{-}2\ CCP1/2)}$) were engineered using multisite-directed mutagenesis. As detailed under "Experimental Procedures," the domain swapping procedure used required introduction of new restriction sites at module boundaries in the C1s cDNA, resulting in slight amino acid changes (Fig. 2). Thus, at the CUB_2-CCP_1 junction, the Asp^{275}-Pro^{276} segment of wild-type C1s was replaced by Ser-Met in $C1s_{(MASP\text{-}2\ CCP1/2)}$ and by Ser-Ile in $C1s_{(MASP\text{-}2\ SP)}$ and the intermediate species C1s* containing only the restriction sites (Fig. 2*B*). At the CCP_2-SP boundary, Lys^{405} of wild-type C1s was replaced by Val in C1s* and $C1s_{(MASP\text{-}2\ SP)}$. As detailed later, the intermediate C1s* species was used as a control to evaluate the possible effects of these amino acid changes.

Chemical Characterization of the C1s/MASP-2 Chimeras—Wild-type C1s, C1s*, $C1s_{(MASP\text{-}2\ CCP1/2)}$, and $C1s_{(MASP\text{-}2\ SP)}$ were expressed in a baculovirus/insect cell system and secreted in the culture supernatant at concentrations of about 10, 6, 2, and 0.15 mg/liter, respectively. All C1s variants could be purified by affinity chromatography on a C1s-Sepharose column in the presence of Ca^{2+} ions, indicating that the interaction properties characteristic of their N-terminal CUB_1-epidermal growth factor moiety (33) were unaltered. As assessed by SDS-PAGE analysis, C1s, C1s*, and $C1s_{(MASP\text{-}2\ CCP1/2)}$ were essentially pure and exhibited a single-chain structure under reducing conditions, indicating that they were in the proenzyme state (Fig. 3*A*). Whereas C1s* and wild-type C1s had similar apparent molecular masses (83 kDa), that of $C1s_{(MASP\text{-}2\ CCP1/2)}$ was significantly lower (78 kDa), consistent with the fact that the MASP-2 CCP_2 module lacks *N*-linked glycosylation, contrary to its counterpart in C1s (Fig. 2*A*). As expected, $C1s_{(MASP\text{-}2\ CCP1/2)}$ was reactive to both anti-C1s and anti-MASP-2 antibodies (data not shown).

SDS-PAGE analysis of the $C1s_{(MASP\text{-}2\ SP)}$ chimera under nonreducing conditions (Fig. 3*B*) revealed two bands generally present in equivalent amounts, (i) a 79-kDa species corresponding to the full-length protease, yielding two sequences, Glu-Pro-Thr-Met-Tyr-Gly-Glu-Ile-Leu-Ser

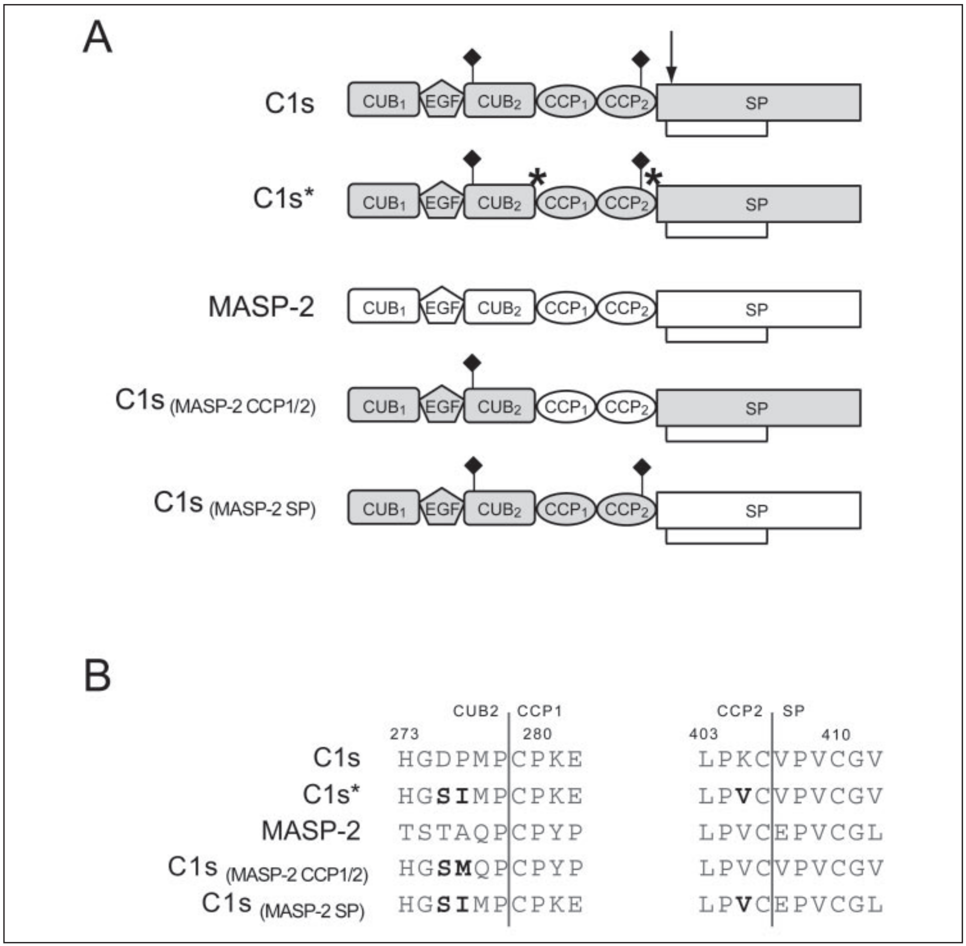

A
C1s
C1s*
MASP-2
C1s (MASP-2 CCP1/2)
C1s (MASP-2 SP)
CUB1
EGF
CUB2
CCP1
CCP2
SP
B
CUB2 | CCP1
273
280
CCP2 | SP
403
410
C1s HGDPMP|CPKE LPKC|VPVCGV
C1s* HGSIMP|CPKE LPVC|VPVCGV
MASP-2 TSTAQP|CPYP LPVC|EPVCGL
C1s (MASP-2 CCP1/2) HGSMQP|CPYP LPVC|VPVCGV
C1s (MASP-2 SP) HGSIMP|CPKE LPVC|EPVCGL

FIGURE 2. **Structures of the C1s/MASP-2 chimeras.** *A*, modular structures of C1s, C1s*, MASP-2, and the $C1s_{(MASP\text{-}2\ CCP1/2)}$ and $C1s_{(MASP\text{-}2\ SP)}$ chimeras. The only disulfide bridge shown is that connecting the N-terminal moiety of the proteases to their SP domain after cleavage of the susceptible Arg-Ile bond (indicated by the *arrow*) upon activation. *N*-Linked oligosaccharides are represented as *black diamonds*. *Stars* indicate the approximate position of the amino acid changes introduced at domain boundaries by insertion of the restriction sites. *B*, sequence alignment *highlighting* the amino acid changes introduced at module boundaries in C1s*, $C1s_{(MASP\text{-}2\ CCP1/2)}$, and $C1s_{(MASP\text{-}2\ SP)}$. Modified residues are shown in *bold*. The residue numbering is that of C1s.

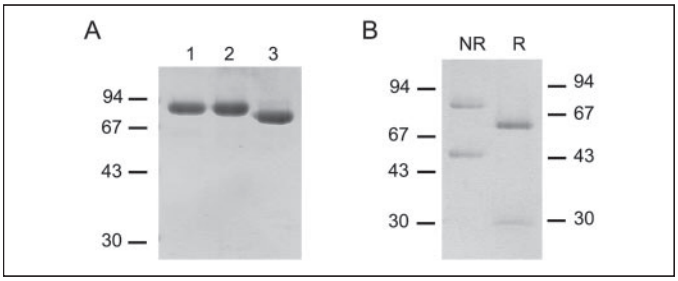


FIGURE 3. **SDS-PAGE analysis of the recombinant C1s variants used in this study.** *A*, analysis under reducing conditions of C1s (*lane 1*), C1s* (*lane 2*), and $C1s_{(MASP\text{-}2\ CCP1/2)}$ (*lane 3*). Molecular masses of reduced standard proteins (expressed in kDa) are indicated. *B*, analysis of the $C1s_{(MASP\text{-}2\ SP)}$ chimera under nonreducing (*NR*) and reducing (*R*) conditions. Molecular masses of unreduced and reduced standard proteins are shown on the *left* and *right sides* of the gel, respectively.

and Ile-Tyr-Gly-Gly-Gln-Lys-Ala-Lys-Pro-Gly, corresponding to the N-terminal ends of C1s and of the MASP-2 SP serine protease domain, respectively and (ii) a 50-kDa species yielding only the first sequence above and reacting only with anti-C1s antibodies, corresponding to a truncated fragment derived from the N-terminal end of the chimeric protein. Analysis under reducing conditions indicated that full-length $C1s_{(MASP\text{-}2\ SP)}$ was totally activated, as it yielded two bands at 29 and 60 kDa corresponding to the serine protease domain and the N-terminal chain expected to be released upon activation (see Fig. 2). The former chain reacted only with anti-MASP-2 antibodies (data not shown), whereas the latter co-migrated with the truncated fragment mentioned above.

Functional Characterization of the C1s/MASP-2 Chimeras—Activation of C1s* and the $C1s_{(MASP\text{-}2\ CCP1/2)}$ chimera was readily achieved by incubation with the CCP_2-SP fragment of C1r, as assessed by the release of two chains upon SDS-PAGE analysis under reducing conditions (data not shown). As expected, the light chains, corresponding to the SP domains (30 kDa), co-migrated with that of wild-type C1s. In contrast, the N-terminal heavy chain of $C1s_{(MASP\text{-}2\ CCP1/2)}$ had an apparent mass of 58 kDa, significantly lower than those of C1s and C1s* (61 kDa). Compared with wild-type C1s, no significant difference was observed in the activation kinetics of C1s* and $C1s_{(MASP\text{-}2\ CCP1/2)}$, and >95% activation was achieved in all cases upon incubation with C1r CCP_2-SP under the conditions described under "Experimental Procedures."

The esterolytic activity of the various activated C1s variants was next measured using Ac-Gly-Lys-OMe as a substrate and compared with those of wild-type C1s and of the $CCP_{1/2}$-SP fragment of MASP-2, which retains the esterolytic and proteolytic activities of full-length MASP-2 (12). As illustrated in Fig. 4, C1s* and $C1s_{(MASP\text{-}2\ CCP1/2)}$ cleaved this substrate with an efficiency very similar to that of wild-type C1s. In contrast, $C1s_{(MASP\text{-}2\ SP)}$ had a significantly lower activity comparable with that of the $CCP_{1/2}$-SP fragment of MASP-2. This clearly indicated that each activated variant was catalytically active and had an esterolytic activity characteristic of its own SP domain (12).

Both activated C1s/MASP-2 chimeras specifically cleaved C2, as assessed by their ability to split this protein into its characteristic C2a and C2b fragments (data not shown). In both cases, C2 cleavage was essentially complete after incubation for 20 min at 37 °C at an enzyme: substrate molar ratio of 1:1000. The kinetic parameters of C2 cleavage were determined for both chimeras and compared with those obtained

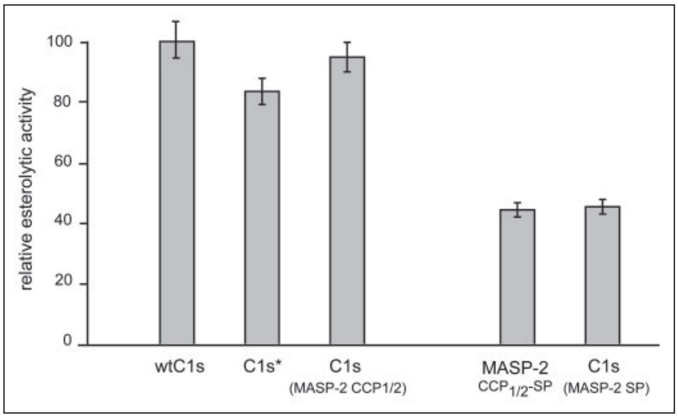


FIGURE 4. **Esterolytic activity of the C1s variants.** Wild-type (*wt*) C1s, the $CCP_{1/2}$-SP fragment of MASP-2, and the C1s variants were tested for their ability to cleave Ac-Gly-Lys-OMe as described under "Experimental Procedures." Esterolytic activities are expressed relative to that of wild-type C1s, which was 50.8 s^{-1}. *Error bars* correspond to the S.D. of duplicate experiments.

TABLE ONE

Kinetic constants for proteolytic cleavage of C2 by the C1s variants

Mean values were determined from four to five measurements ±S.D.

Enzyme	k_{cat}	K_m	k_{cat}/K_m
	s^{-1}	μM	$M^{-1}\,s^{-1}$
Wild-type C1s	5.1 ± 1.2	6.1 ± 0.3	8.4×10^5
C1s*	5.2 ± 1.0	9.2 ± 2.0	5.6×10^5
MASP-2[a]	5.6 ± 1.4	5.2 ± 1.0	10.8×10^5
$C1s_{(MASP\text{-}2\ CCP1/2)}$	5.1 ± 1.5	9.2 ± 0.8	5.5×10^5
$C1s_{(MASP\text{-}2\ SP)}$	5.2 ± 1.5	7.3 ± 1.8	7.1×10^5

[a] Values were determined using the MASP-2 CCP1/2 SP fragment.

with wild-type C1s, C1s*, and the $CCP_{1/2}$ fragment of MASP-2. As listed in TABLE ONE, all enzymes showed comparable k_{cat} values. Although slight differences were observed in the K_m values, the C2 cleavage efficiency of each chimera, as measured by the k_{cat}/K_m ratio, was not significantly different from the values determined for wild-type C1s and the $CCP_{1/2}$-SP fragment of MASP-2.

Both C1s/MASP-2 chimeras mediated specific cleavage of C4, as shown by their ability to convert its C4α chain into the smaller C4α' fragment (data not shown). N-terminal sequence analysis of the reaction mixtures indicated that both enzymes yielded a C4α' fragment with the sequence Ala-Leu-Glu-Ile-Leu-Gln-Glu-Glu-Asp-Leu, providing unambiguous evidence that cleavage occurred at the Arg^{737}-Ala^{738} bond normally recognized by wild-type C1s and MASP-2. Whereas $C1s_{(MASP\text{-}2\ CCP1/2)}$ was readily active at C4 concentrations of 50–500 nM, efficient cleavage by $C1s_{(MASP\text{-}2\ SP)}$ required substrate concentrations in the micromolar range. Determination of the kinetic parameters of C4 cleavage revealed that wild-type C1s, the $CCP_{1/2}$-SP fragment of MASP-2, and both C1s/MASP-2 chimeras had comparable k_{cat} values (TABLE TWO). In contrast, whereas $C1s_{(MASP\text{-}2\ SP)}$, wild-type C1s, and C1s* all showed K_m values in the micromolar range, $C1s_{(MASP\text{-}2\ CCP1/2)}$ and the $CCP_{1/2}$-SP fragment of MASP-2 both exhibited K_m values in the nanomolar range. As a result, the k_{cat}/K_m ratios for the latter two enzymes were 21–27-fold higher than those determined for the three proteases containing the C1s-CCP modules (TABLE TWO). Thus, introduction of the MASP-2 SP domain in C1s did not significantly modify the C4 cleaving activity of the latter, but introduction of the MASP-2 CCP modules resulted in a strong increase in C4 cleavage efficiency due to a dramatic decrease in K_m.

TABLE TWO

Kinetic constants for proteolytic cleavage of C4 by the C1s variants

Mean values were determined from four to five measurements ±S.D.

Enzyme	k_{cat}	K_m	k_{cat}/K_m
	s^{-1}	nM	$M^{-1}\,s^{-1}$
Wild-type C1s	5.4 ± 1.1	6600 ± 1300	8.2×10^5
C1s*	5.3 ± 1.3	5000 ± 900	10.6×10^5
MASP-2[a]	1.9 ± 0.2	85 ± 12	223×10^5
$C1s_{(MASP\text{-}2\ CCP1/2)}$	3.2 ± 0.3	147 ± 19	218×10^5
$C1s_{(MASP\text{-}2\ SP)}$	2.9 ± 0.7	2800 ± 700	10.4×10^5

[a] Values were determined using the MASP-2 CCP1/2 SP fragment.

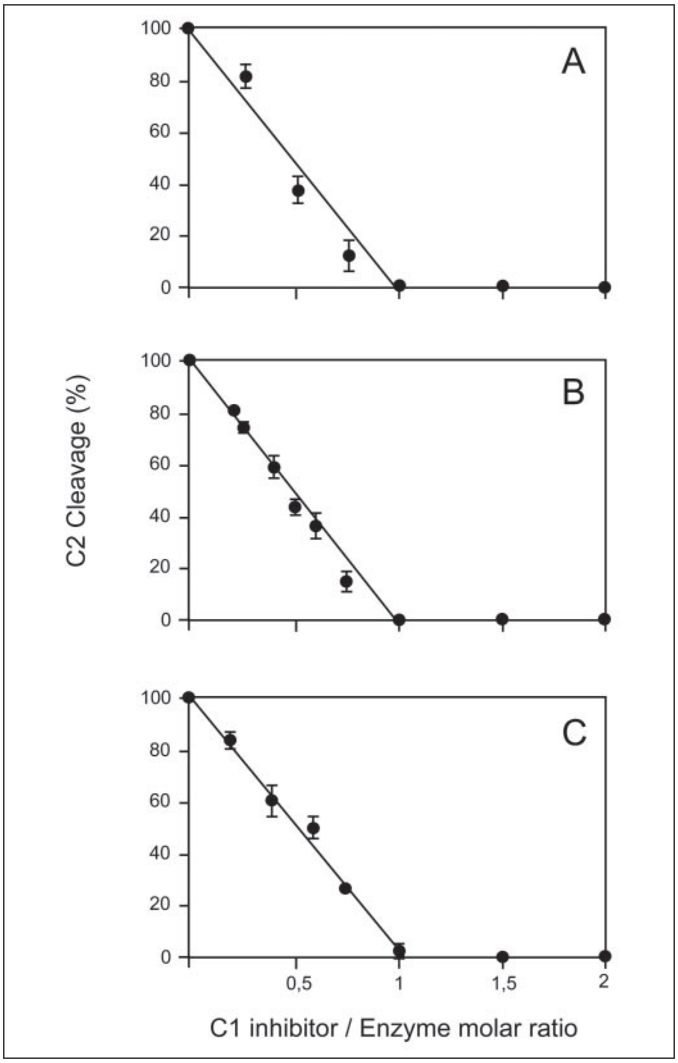


FIGURE 5. **Active site titration of C1s, $C1s_{(MASP\text{-}2\ CCP1/2)}$, and $C1s_{(MASP\text{-}2\ SP)}$ by C1 inhibitor.** C1s, $C1s_{(MASP\text{-}2\ CCP1/2)}$, and $C1s_{(MASP\text{-}2\ SP)}$ (0.25 μM each) were preincubated for 15 min at 37 °C in the presence of increasing concentrations of C1 inhibitor to achieve C1 inhibitor:protease ratios up to 2.0, as indicated. The residual C2 cleaving activity of C1s (*A*), $C1s_{(MASP\text{-}2\ CCP1/2)}$ (*B*), and $C1s_{(MASP\text{-}2\ SP)}$ (*C*) was measured as described under "Experimental Procedures." *Error bars* correspond to the S.D. of triplicate experiments.

To determine the active site concentration of the C1s/MASP-2 hybrids, these were titrated with C1 inhibitor. As observed for wild-type MASP-2 (12) and C1s (Fig. 5*A*), increasing the C1 inhibitor:protease ratio led to a linear decrease of the C2 cleaving activity of both chimeras, with complete inhibition at a molar ratio close to 1:1 (Fig. 5, *B* and *C*). Thus, the protein concentration estimates determined for these enzymes (see "Experimental Procedures") actually matched their active

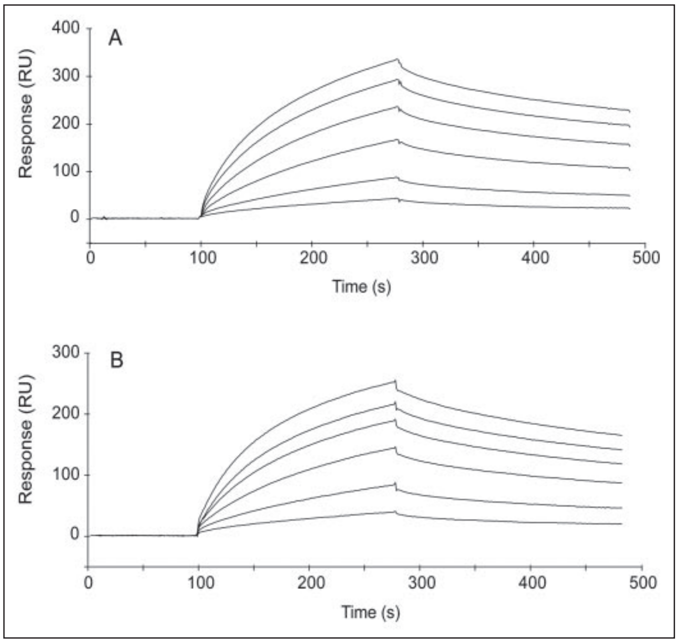


FIGURE 6. **Analysis by surface plasmon resonance spectroscopy of the ability of the C1s$_{\text{(MASP-2 CCP1/2)}}$ chimera to associate with C1r and C1q within C1.** The C1s-C1r-C1r-C1s tetramer was reconstituted from proenzyme C1r and either wild-type C1s (*A*) or C1s$_{\text{(MASP-2 CCP1/2)}}$ (*B*). Interaction with immobilized C1q was measured as described under "Experimental Procedures" using in each case analyte concentrations of 1, 2, 4, 6, 8, and 10 nM (from *bottom* to *top*). *RU*, resonance units.

TABLE THREE

Kinetic and dissociation constants for the interaction between immobilized C1q and the C1s-C1r-C1r-C1s tetramer assembled from C1r and different C1s variants

Mean values were determined from two separate experiments.

Protein	k_{on}	k_{off}	K_D
	$M^{-1}\,s^{-1}$	s^{-1}	nM
Wild-type C1s	1.09×10^6	1.68×10^{-3}	1.56
C1s*	0.94×10^6	1.84×10^{-3}	1.97
C1s$_{\text{(MASP-2 CCP1/2)}}$	0.97×10^6	2.89×10^{-3}	3.04

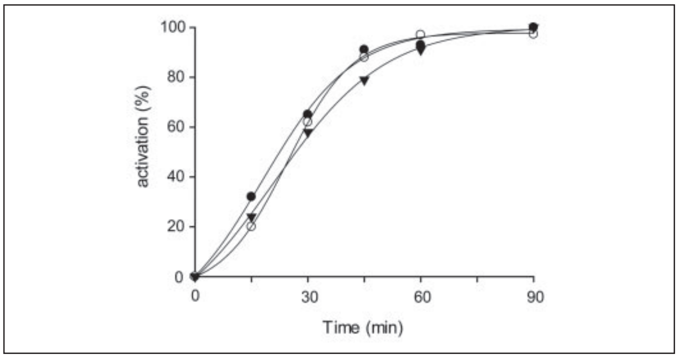


FIGURE 7. **Activation of wild-type C1s, C1s*, and C1s$_{\text{(MASP-2 CCP1/2)}}$ within C1.** C1 was reconstituted from appropriate amounts of C1q, C1r, and either wild-type C1s, C1s*, or C1s$_{\text{(MASP-2 CCP1/2)}}$. Activation was measured by SDS-PAGE and Western blot analyses as described under "Experimental Procedures." ●, wild-type C1s; ▼, C1s*; ○, C1s$_{\text{(MASP-2 CCP1/2)}}$.

site concentrations, yielding further indication that these chimeric molecules were properly folded and retained full enzymatic activity.

Functional Behavior of C1s$_{\text{(MASP-2 CCP1/2)}}$ within the C1 Complex—The ability of the C1s/MASP-2 hybrids to associate with C1r and C1q to form a pseudo-C1 complex was measured by incubating each variant with an equimolar amount of proenzyme C1r and then monitoring the interaction of the resulting pseudo-C1s-C1r-C1r-C1s tetramer with immobilized C1q using surface plasmon resonance spectroscopy. Interaction was observed in the case of C1s$_{\text{(MASP-2 SP)}}$, but likely due to the presence of the N-terminal fragment, the binding curves did not fit a simple 1:1 Langmuir interaction model, and therefore, the kinetic constants could not be determined. The other chimera C1s$_{\text{(MASP-2 CCP1/2)}}$ yielded binding curves strikingly similar to those obtained in the case of wild-type C1s (Fig. 6) and C1s* (not shown). Determination of the kinetic parameters yielded comparable k_{on} values (TABLE THREE). Only a slight increase in k_{off} and in the resulting K_D value was observed in the case of C1s$_{\text{(MASP-2 CCP1/2)}}$, indicating that this chimeric molecule essentially retained the C1s ability to associate with C1r and C1q within C1.

To check its ability to undergo activation inside C1, C1s$_{\text{(MASP-2 CCP1/2)}}$ was mixed with appropriate amounts of C1q and proenzyme C1r, and the resulting pseudo C1 complex was allowed to activate as described under "Experimental Procedures." As illustrated in Fig. 7, C1s$_{\text{(MASP-2 CCP1/2)}}$ yielded an activation curve very similar to those obtained with wild-type C1s and C1s*, demonstrating that replacement of the CCP$_1$-CCP$_2$ segment of C1s by its MASP-2 counterpart did not significantly alter its ability to undergo activation in C1.

DISCUSSION

The major aim of this study was to investigate why MASP-2 and C1s, the homologous modular proteases that trigger the lectin and classical pathways of complement, respectively, exhibit differential catalytic efficiencies toward their common substrate C4. For this purpose, chimeric C1s molecules containing either the CCP$_1$-CCP$_2$ module pair or the SP domain of MASP-2 were engineered by domain swapping and expressed in insect cells. In this respect, it is worthy to note that, whereas wild-type C1s and the control molecule C1s* were produced at comparable yields (6–10 mg/liter), the expression level of the C1s$_{\text{(MASP-2 SP)}}$ chimera was much lower (0.15 mg/liter) and similar to those observed for full-length MASP-2 (12) or its CCP$_{1/2}$-SP segment. In contrast, the other chimera C1s$_{\text{(MASP-2 CCP1/2)}}$ was consistently produced at a higher yield (2 mg/liter). As judged from these figures, proteins harboring the SP domain of MASP-2 were produced in significantly smaller amounts, suggesting that the presence of this particular domain is a limiting factor for expression in a baculovirus/insect cells system. Indeed, comparative analysis of the cDNA sequences of the various MASP-2 segments indicates that the SP domain contains a significantly higher proportion of unfavorable codons, as defined for *Drosophila melanogaster* (34), suggesting that the observed low expression level may be linked to codon usage.

Another conclusion arising from this study is that the activation characteristics of the C1s/MASP-2 chimeras are dictated by the nature of their SP domain. Thus, as also observed for wild-type C1s, the C1s$_{\text{(MASP-2 CCP1/2)}}$ chimera containing the C1s SP domain was secreted in a proenzyme form and did not undergo activation upon prolonged storage. Conversely, the chimeric molecule containing the MASP-2 SP domain had a behavior strikingly similar to that of wild-type MASP-2 expressed in a baculovirus/insect cells system (12). Thus, C1s$_{\text{(MASP-2 SP)}}$ was recovered as a two-chain, active form, indicating that it underwent spontaneous activation during or after the synthesis and secretion processes. In addition, as described previously for MASP-2 (12), C1s$_{\text{(MASP-2 SP)}}$ contained a significant amount of a 50-kDa N-terminal fragment, suggesting autolytic cleavage at a site located at the junction between the CCP$_2$ module and the SP domain. It cannot be excluded, however, that this fragment resulted from other processes, such as abortive synthesis due to particular features of the cDNA coding for the MASP-2 SP domain.

With respect to activation, it should also be emphasized that the variants containing the C1s SP domain (C1s$_{\text{(MASP-2 CCP1/2)}}$ and C1s*) were both susceptible to activation by C1r in a manner comparable with wild-type C1s. The fact that this property is retained by C1s$_{\text{(MASP-2 CCP1/2)}}$ chimera provides further indication that recognition by C1r is restricted to the C1s SP domain and does not involve determinants in the upstream part of the protein, in full agreement with previous findings (14).

Each C1s/MASP-2 chimera retained an esterolytic activity characteristic of its own SP domain, providing a first indication that the functionality of this domain was not significantly altered by the swapping process. The fact that each chimera reacted with C1 inhibitor in a 1:1 stoichiometry fully supported this conclusion. With respect to C2 cleavage, both chimeras had an efficiency comparable with that of their parent proteases C1s and MASP-2. This was expected considering that C2 cleavage by C1s and MASP-2 only requires their SP domain (14, 15) and that both proteases have comparable efficiencies toward C2 (12). Thus, introduction of the MASP-2 CCP modules in C1s was not expected to modify its C2 cleaving activity, whereas introduction of the MASP-2 SP domain was expected to have little effect.

Analysis of C4 cleavage by the C1s/MASP-2 chimeras yields precise information about the relative contribution of the SP domain and CCP modules of C1s and MASP-2 to their activity toward this substrate and provides an explanation for the observed higher efficiency of MASP-2. Swapping the C1s-CCP modules for those of MASP-2 had no significant effect on k_{cat} but resulted in a considerable increase in C4 cleavage efficiency by decreasing K_m to a value similar to that determined for MASP-2 (12) and its CCP$_{1/2}$-SP fragment. This observation is fully consistent with previous data indicating that C4 cleavage by C1s and MASP-2 requires substrate recognition by sites located in their CCP modules (12, 15) and clearly demonstrates that the CCP modules of MASP-2 are much more efficient in terms of C4 recognition than are those of C1s. On the other hand, the fact that replacing the C1s SP domain by its MASP-2 counterpart had no significant impact on k_{cat} provides strong indication that both domains are equally catalytically active toward this substrate. It may be concluded from these observations that the higher C4 cleavage efficiency of MASP-2 compared with C1s arises mainly, if not only, from a better efficacy of its CCP modules in terms of C4 recognition. In this respect it should be noted that, whereas C4 binding is considered to involve the CCP$_1$ module of C1s (12, 35), the CCP$_2$ module has been implicated in the case of MASP-2 (15). In light of these observations, it appears likely that the markedly different ability of these proteases to recognize C4 arises from the fact that they make use of distinct binding sites rather than homologous sites with differential binding properties.

It is remarkable that the C1s$_{\text{(MASP-2 CCP1/2)}}$ chimera is virtually as efficient toward C4 as is MASP-2 itself despite the unnatural junction between the MASP-2 CCP$_2$ module and the C1s SP domain. The structure of the CCP$_2$-SP interface has been determined for both C1s and MASP-2 (36, 37), indicating that the domains interact in a similar fashion in each case, through a network of interactions involving hydrogen bonds and van der Waals contacts. Nevertheless, although many of the interacting residues are conserved in both proteins, there are significant differences (37). It appears likely, therefore, that despite these expected local differences, each domain retains the ability to accommodate its partner in such a way that the stability of the interface is maintained, at least to a certain extent. Indeed, the intrinsic flexibility of the CCP$_2$-SP interface, as observed in the case of MASP-2 (37), possibly allows this type of adaptation. From a general standpoint, the fact that both C1s$_{\text{(MASP-2 CCP1/2)}}$ and C1s$_{\text{(MASP-2 SP)}}$ are readily active on C4 demonstrates that the SP domain of each protease is able to cooperate efficiently with the CCP modules of the other, implying that the overall C4 cleavage mechanism does not involve stringent structural constraints between catalytic and recognition elements. An alternative possibility is that both chimeras retain flexibility at their CCP$_2$-SP junction, allowing them to retain efficient C4 cleaving ability.

As shown by surface plasmon resonance spectroscopy and activation measurements, the C1s$_{\text{(MASP-2 CCP1/2)}}$ chimera also retains C1s ability to associate with C1q and C1r and to undergo activation in the context of the resulting pseudo C1 complex, indicating that, remarkably, swapping the C1s-CCP modules for those of MASP-2 has no significant effect on either function. In full agreement with recent studies based on site-directed mutagenesis (35), this clearly demonstrates that, in terms of interaction or recognition, the CCP$_1$-CCP$_2$ segment of C1s plays no direct role in C1 assembly and in the C1s activation process but is probably only used as a linker in both cases. In contrast, the flexibility at the CCP$_1$-CCP$_2$ interface (35) and the recognition properties of the CCP modules are key factors of the C4 cleavage reaction.

Acknowledgement—We thank J.-P. Andrieu for performing N-terminal sequence analyses.

REFERENCES

1. Cooper, N. R. (1985) *Adv. Immunol.* **37,** 151–216
2. Schumaker, V. N., Zavodszky, P., and Poon, P. H. (1987) *Annu. Rev. Immunol.* **5,** 21–42
3. Gaboriaud, C., Thielens, N. M., Gregory, L. A., Rossi, V., Fontecilla-Camps, J. C., and Arlaud, G. J. (2004) *Trends Immunol.* **25,** 368–373
4. Matsushita, M., and Fujita, T. (1992) *J. Exp. Med.* **176,** 1497–1502
5. Matsushita, M., Endo, Y., and Fujita, T. (2000) *J. Immunol.* **164,** 2281–2284
6. Matsushita, M., Kuraya, M., Hamasaki, N., Tsujimura, M., Shiraki, H., and Fujita, T. (2002) *J. Immunol.* **168,** 3502–3506
7. Thiel, S., Vorup-Jensen, T., Stover, C. M., Schwaeble, W., Laursen, S. B., Poulsen, K., Willis, A. C., Eggleton, P., Hansen, S., Holmskov, U., Reid, K. B. M., and Jensenius, J. C. (1997) *Nature* **386,** 506–510
8. Bork, P., and Beckmann, G. (1993) *J. Mol. Biol.* **231,** 539–545
9. Gregory, L. A., Thielens, N. M., Matsushita, M., Sorensen, R., Arlaud, G. J., Fontecilla-Camps, J. C., and Gaboriaud, C. (2004) *J. Biol. Chem.* **279,** 29391–29397
10. Reid, K. B., Bentley, D. R., Campbell, R. D., Chung, L. P., Sim, R. B., Kristensen, T., and Tack, B. F. (1986) *Immunol. Today* **7,** 230–234
11. Vorup-Jensen, T., Petersen, S. V., Hansen, A. G., Poulsen, K., Schwaeble, W., Sim, R. B., Reid, K. B. M., Davis, S. J., Thiel, S., and Jensenius, J. C. (2000) *J. Immunol.* **165,** 2093–2100
12. Rossi, V., Cseh, S., Bally, I., Thielens, N. M., Jensenius, J. C., and Arlaud, G. J. (2001) *J. Biol. Chem.* **276,** 40880–40887
13. Teillet, F., Dublet, B., Andrieu, J. P., Gaboriaud, C., Arlaud, G. J., and Thielens, N. M. (2005) *J. Immunol.* **174,** 2870–2877
14. Rossi, V., Bally, I., Thielens, N. M., Esser, A. F., and Arlaud, G. J. (1998) *J. Biol. Chem.* **273,** 1232–1239
15. Ambrus, G., Gal, P., Kojima, M., Szilagyi, K., Balczer, J., Antal, J., Graf, L., Laich, A., Moffatt, B. E., Schwaeble, W., Sim, R. B., and Zavodszky, P. (2003) *J. Immunol.* **170,** 1374–1382
16. March, S. C., Parikh, I., and Cuatrecasas, P. (1974) *Anal. Biochem.* **60,** 149–152
17. Luo, C., Thielens, N. M., Gagnon, J., Gal, P., Sarvari, M., Tseng, Y., Tosi, M., Zavodszky, P., Arlaud, G. J., and Schumaker, V. N. (1992) *Biochemistry* **31,** 4254–4262
18. Thielens, N. M., Villiers, M. B., Reboul, A., Villiers, C. L., and Colomb, M. G. (1982) *FEBS Lett.* **141,** 19–24
19. Dodds, A. W. (1993) *Methods Enzymol.* **223,** 46–61
20. Arlaud, G. J., Sim, R. B., Duplaa, A.-M., and Colomb, M. G. (1979) *Mol. Immunol.* **16,** 445–450
21. Arlaud, G. J., Villiers, C. L., Chesne, S., and Colomb, M. G. (1980) *Biochim. Biophys. Acta* **616,** 116–129
22. Lacroix, M., Ebel, C., Kardos, J., Dobo, J., Gal, P., Zavodszky, P., Arlaud, G. J., and Thielens, N. M. (2001) *J. Biol. Chem.* **276,** 36233–36240
23. Thielens, N. M., Aude, C. A., Lacroix, M. B., Gagnon, J., and Arlaud, G. J. (1990) *J. Biol. Chem.* **265,** 14469–14475
24. Harpel, P. C. (1976) *Methods Enzymol.* **45,** 751–760
25. Bock, S. C., Skriver, K., Nielsen, E., Thogersen, H.-C., Wiman, B., Donaldson, V. H., Eddy, R. L., Marrinan, J., Radziejewska, E., Huber, R., Shows, T. B., and Magnusson, S. (1986) *Biochemistry* **25,** 4292–4301

26. Edelhoch, H. (1967) *Biochemistry* **6,** 1948–1954
27. Thielens, N. M., Cseh, S., Thiel, S., Vorup-Jensen, T., Rossi, V., Jensenius, J. C., and Arlaud, G. J. (2001) *J. Immunol.* **166,** 5068–5077
28. King, L. A., and Possee, R. D. (1992) The *Baculovirus Expression System: A Laboratory Guide*, pp. 111–114, Chapman and Hall, Ltd., London
29. Arlaud, G. J., and Thielens, N. M. (1993) *Methods Enzymol.* **223,** 61–82
30. Zundel, S., Cseh, S., Lacroix, M., Dahl, M. R., Matsushita, M., Andrieu, J. P., Schwaeble, W. J., Jensenius, J. C., Fujita, T., Arlaud, G. J., and Thielens, N. M. (2004) *J. Immunol.* **172,** 4342–4350
31. Tacnet-Delorme, P., Chevallier, S., and Arlaud, G. J. (2001) *J. Immunol.* **167,** 6374–6381
32. Rossi, V., Gaboriaud, C., Lacroix, M., Ulrich, J., Fontecilla-Camps, J. C., Gagnon, J., and Arlaud, G. J. (1995) *Biochemistry* **34,** 7311–7321
33. Gregory, L. A., Thielens, N. M., Arlaud, G. J., Fontecilla-Camps, J. C., and Gaboriaud, C. (2003) *J. Biol. Chem.* **278,** 32157–32164
34. Duret, L., and Mouchiroud, D. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 4482–4487
35. Bally, I., Rossi, V., Thielens, N. M., Gaboriaud, C., and Arlaud, G. J. (2005) *J. Immunol.* **175,** 4536–4542
36. Gaboriaud, C., Rossi, V., Bally, I., Arlaud, G. J., and Fontecilla-Camps, J. C. (2000) *EMBO J.* **19,** 1755–1765
37. Harmat, V., Gal, P., Kardos, J., Szilagyi, K., Ambrus, G., Vegh, B., Naray-Szabo, G., and Zavodszky, P. (2004) *J. Mol. Biol.* **342,** 1533–1546

P13

ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Molecular Immunology 45 (2008) 670–677

Molecular Immunology

www.elsevier.com/locate/molimm

Elucidation of the substrate specificity of the MASP-2 protease of the lectin complement pathway and identification of the enzyme as a major physiological target of the serpin, C1-inhibitor☆

Felicity K. Kerr[a], Adele R. Thomas[a], Lakshmi C. Wijeyewickrema[a], James C. Whisstock[a,b], Sarah E. Boyd[c], Dion Kaiserman[a], Antony Y. Matthews[a], Phillip I. Bird[a], Nicole M. Thielens[d], Véronique Rossi[d,1], Robert N. Pike[a,*,1]

[a] *Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Australia*
[b] *ARC Centre for Structural and Functional Microbial Genomics, Monash University, Clayton, Victoria 3800, Australia*
[c] *Faculty of Information Technology, Monash University, Clayton, Victoria 3800, Australia*
[d] *Laboratoire d'Enzymologie Moléculaire, Institut de Biologie Structurale Jean-Pierre Ebel, Grenoble, France*

Received 25 May 2007; received in revised form 9 July 2007; accepted 10 July 2007
Available online 20 August 2007

Abstract

Complement is a central component of host defence, but unregulated activation can contribute to disease. The system can be initiated by three pathways: classical, alternative and lectin. The classical and lectin pathways are initiated by the C1 and mannose-binding lectin (MBL) or ficolin complexes, respectively, with C1s the executioner protease of the C1 complex and MASP-2 its counterpart in the lectin complexes. These proteases in turn cleave the C4 and C2 components of the system. Here we have elucidated the cleavage specificity of MASP-2 using a randomised substrate phage display library. Apart from the crucial P_1 position, the MASP-2 S_2 and S_3 subsites (in that order) play the greatest role in determining specificity, with Gly residues preferred at P_2 and Leu or hydrophobic residues at P_3. Cleavage of peptide substrates representing the known physiological cleavage sequences in C2, C4 or the serpin C1-inhibitor (a likely regulator of MASP-2) revealed that MASP-2 is up to 1000 times more catalytically active than C1s. C1-inhibitor inhibited MASP-2 50-fold faster than C1s and much faster than any other protease tested to date, implying that MASP-2 is a major physiological target of C1-inhibitor.
© 2007 Elsevier Ltd. All rights reserved.

Keywords: Mannose-binding lectin-associated serine protease-2; Complement; Active site specificity; C1s; C1-inhibitor; Serpin; Protease

---

1. Introduction

The complement system can be activated in three ways: the classical, lectin and alternative pathways (Goldsby et al., 2003). The former two pathways are both activated following binding of a recognition molecule to pathogenic surfaces, these being C1q for the classical pathway and mannose-binding lectin (MBL) or ficolin for the lectin pathway (Fujita et al., 2004). Following binding of the recognition molecule, associated serine proteases are activated, which in turn cleave the complement C4 and C2 proteins. The C3 convertase (C4bC2a) complex then activates C3 and the rest of the complement system in turn, resulting in the formation of the membrane attack complex and release of small pro-inflammatory protein fragments, such as the potent anaphylotoxin, C5a. The initiating proteases of both the classical and lectin pathways are regulated by the serpin, C1-inhibitor

---

Abbreviations: NHMec, 7-amino-4-methyl coumarin group; Abz, 2-aminobenzoyl; Lys(Dnp), lysine dinitrophenyl; DMF, dimethyl formamide; PoPS, prediction of protease specificity program; MBL, mannose-binding lectin; MASP, mannose-binding lectin-associated serine protease; C1-Inh, C1-inhibitor; FQS, fluorescence-quenched substrate

☆ The nomenclature for residues in substrates is based on that outlined by Schechter and Berger (1967) for the substrates of proteases. The residues are numbered from the cleaved bond (P_1–P_1') as follows: P_n–···–P_4–P_3–P_2–P_1–P_1'–P_2'–P_3'–P_4'–···–P_n'; the corresponding subsites of the enzyme are denoted as S_n–···–S_4–S_3–S_2–S_1–S_1'–S_2'–S_3'–S_4'–···–S_n'.

* Corresponding author. Tel.: +61 3 99053923; fax: +61 3 99054699.
E-mail address: rob.pike@med.monash.edu.au (R.N. Pike).

[1] These authors contributed equally to this study.

0161-5890/$ – see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.molimm.2007.07.008

Functional Characterization of the Recombinant Human C1 Inhibitor Serpin Domain: Insights into Heparin Binding

Véronique Rossi,* Isabelle Bally,* Sarah Ancelet,* Yuanyuan Xu,† Véronique Frémeaux-Bacchi,‡ Romain R. Vivès,* Rabia Sadir,* Nicole Thielens,* and Gérard J. Arlaud*

Variants of the human C1 inhibitor serpin domain containing three *N*-linked carbohydrates at positions 216, 231, and 330 (C1inhΔ97), a single carbohydrate at position 330 (C1inhΔ97DM), or no carbohydrate were produced in a baculovirus/insect cells system. An N-terminally His-tagged C1inhΔ97 variant was also produced. Removal of the oligosaccharide at position 330 dramatically decreased expression, precluding further analysis. All other variants were characterized chemically and shown to inhibit C1s activity and C1 activation in the same way as native C1 inhibitor. Likewise, they formed covalent complexes with C1s as shown by SDS-PAGE analysis. C1 inhibitor and its variants inhibited the ability of C1r-like protease to activate C1s, but did not form covalent complexes with this protease. The interaction of C1 inhibitor and its variants with heparin was investigated by surface plasmon resonance, yielding K_D values of 16.7×10^{-8} M (C1 inhibitor), 2.3×10^{-8} M (C1inhΔ97), and 3.6×10^{-8} M (C1inhΔ97DM). C1s also bound to heparin, with lower affinity ($K_D = 108 \times 10^{-8}$ M). Using the same technique, 50% inhibition of the binding of C1 inhibitor and C1s to heparin was achieved using heparin oligomers containing eight and six saccharide units, respectively. These values roughly correlate with the size of 10 saccharide units yielding half-maximal potentiation of the inhibition of C1s activity by C1 inhibitor, consistent with a "sandwich" mechanism. Using a thermal shift assay, heparin was shown to interact with the C1s serine protease domain and the C1 inhibitor serpin domain, increasing and decreasing their thermal stability, respectively. ***The Journal of Immunology*, 2010, 184: 4982–4989.**

The C1 inhibitor is the only regulator of the early proteases of the classical and lectin pathways of complement (C1r and C1s and mannan-binding lectin-associated serine protease [SP]-2, respectively) (1–3). Other physiologically relevant target proteases of C1 inhibitor include mannan-binding lectin-associated SP-1 (2, 4), plasma kallikrein and activated factor XII of the contact system (5, 6), and activated factor XI of the intrinsic coagulation system (7). More recently, reactivity toward the C1r-like protease (C1r-LP) has been reported (8). Over the past few years, several potential anti-inflammatory effects of C1 inhibitor, independent of its protease inhibitory activity, and involving noncovalent interactions with various proteins, cell surfaces, or lipids have been revealed (9).

C1 inhibitor is a heavily glycosylated single-chain glycoprotein of 478 aa belonging to the SP inhibitor (serpin) superfamily (10). It has a two-domain structure, with a classical C-terminal serpin domain and a nonconserved, exceptionally large (~100 aa) N-terminal extension. Of the 13–20 potential *O*- and *N*-glycosylation sites identified in C1 inhibitor, 10–17 lie within this N-terminal domain, three *N*-linked oligosaccharides being attached to the serpin domain through residues Asn^{216}, Asn^{231}, and Asn^{330} (10) (Fig. 1). The N-terminal extension of C1 inhibitor does not seem to affect protease inhibition (11, 12), and its functional role remains elusive. However, interaction of C1 inhibitor with selectins on endothelial cells is known to be mediated by tetrasaccharides located in this area (9). Like other serpins, C1 inhibitor neutralizes its target proteases by presenting them a scissile peptide bond that matches their substrate specificity. This bond, identified as Arg^{444}-Thr^{445}, lies in the reactive center loop, an exposed segment with a stressed conformation located near the C-terminal end of the serpin domain. Reaction with C1 inhibitor results in the formation of a 1:1 covalent complex between the protease and the bulk of C1 inhibitor, observable by SDS-PAGE analysis (13). As for other serpins, the inhibitory activity of C1 inhibitor toward proteases, such as C1s or activated factor XI, can be greatly enhanced by heparin and other glycosaminoglycans (14), and the precise mechanism of this potentiation remains to be established.

Several groups have attempted to produce rC1 inhibitor or its serpin domain using various heterologous expression systems. Functional full-length C1 inhibitor was produced by transient expression in COS-1 cells (15) and using a baculovirus/insect cell system (16). Active C1 inhibitor was also expressed at high yield in *Pichia pastoris* and purified (17). Various constructs corresponding to the serpin domain of C1 inhibitor were also produced in COS-1 cells (10), *Escherichia coli* (18), and recently in *P. pastoris* (19). Using the latter recombinant protein, the crystal structure of the C1 inhibitor serpin domain has been solved, revealing a latent conformation with the uncleaved reactive center loop inserted into a seven-stranded antiparallel β-sheet (19). Based on this structure, a model for interaction with heparin has been proposed.

---

*Institut de Biologie Structurale Jean-Pierre Ebel, Unité Mixte de Recherche 5075, Centre National de la Recherche Scientifique-Commissariat à l'Energie Atomique-Université Joseph Fourier, Grenoble; ‡Hôpital Européen Georges Pompidou, Paris, France; and †University of Alabama at Birmingham, Birmingham, AL 35294

Received for publication June 24, 2009. Accepted for publication February 15, 2010.

This work was supported by the Commissariat à l'Energie Atomique, the Centre National de la Recherche Scientifique, and the Université Joseph Fourier (Grenoble, France).

Address correspondence and reprint requests to Dr. Véronique Rossi, Laboratoire d'Enzymologie Moléculaire, Institut de Biologie Structurale Jean-Pierre Ebel, 41 Rue Jules Horowitz, 38027 Grenoble Cedex 1, France. E-mail address: veronique.rossi@ibs.fr

Abbreviations used in this paper: CCP, complement control protein; C1r-LP, C1r-like protease; RU, resonance unit; SP, serine protease.

Copyright © 2010 by The American Association of Immunologists, Inc. 0022-1767/10/$16.00

www.jimmunol.org/cgi/doi/10.4049/jimmunol.0902016

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

P16

In the current study, we used a baculovirus/insect cells system to express and purify different variants of the serpin domain of C1 inhibitor, with the aim to further investigate its reactivity with proteases, the implication of *N*-linked carbohydrates, and the interaction with heparin.

Materials and Methods

Reagents and proteins

Diisopropyl phosphorofluoridate and *N*-acetylglycine-L-lysine methyl ester were obtained from Sigma-Aldrich (Saint-Quentin-Fallavier, France). Oligonucleotides were purchased from Eurogentec (Liège, Belgium). Restriction enzymes were from New England Biolabs (Beverly, MA). The pcDNA3C1inh plasmid containing the full-length human C1-inhibitor cDNA (Met^{458} allele) was provided by Véronique Frémeaux-Bacchi (Hôpital Européen Georges Pompidou). C1 inhibitor and activated C1s were purified from human plasma, according to published procedures (2, 20). The C1q recognition subunit of C1 and the proenzyme form of the C1s-C1r-C1r-C1s tetramer were purified from human plasma, as described previously (20, 21). Human recombinant proenzyme C1s and its complement control protein (CCP)$_2$-SP segment were expressed in a baculovirus/insect cells system (22, 23). Human C1r-LP was expressed in CHO-K1 cells and purified as described previously (8). The concentration of purified proteins was determined using the absorption coefficient ($A_{1\%,\ 1cm}$ at 280 nm) and m.w.: activated C1s, 14.5 and 79,800 (24, 25); proenzyme C1s and C1s Ser617Ala, 14.5 and 77,400 (22); C1s CCP$_2$-SP segment, 16.4 and 36,700 (23); C1 inhibitor, 4.5 and 104,000 (10, 26); C1q, 6.8 and 459,300 (20); and C1s-C1r-C1r-C1s, 13.5 and 330,000 (21). For the rC1 inhibitor variants, absorption coefficients were calculated by the method of Gill and von Hippel (27), and m.w. was determined by mass spectrometry: C1inhΔ97, 6.6 and 46,228; C1inhΔ97-ht, 6.6 and 46,740; and C1inhΔ97DM, 6.2 and 43,824.

Expression and purification of C1s Ser617Ala

The C1s mutant Ser617Ala was expressed using a baculovirus/insect cells system, essentially as described previously for other C1s variants (22). Briefly, a DNA fragment encoding the human C1s signal peptide plus the mature protein was amplified by PCR and cloned into the pFastBac1 vector (Invitrogen, San Diego, CA). The expression plasmid coding for the Ser617Ala mutant was generated using the QuickChange XL site-directed mutagenesis kit (Stratagene, La Jolla, CA). The mutagenic oligonucleotides were designed according to the manufacturer's recommendations, and the pFastBac1/C1s expression plasmid coding for wild-type C1s was used as a template. The recombinant baculovirus was generated using the Bac-to-Bac system (Invitrogen) and used to infect High Five cells. C1s Ser617Ala was purified from the cell culture supernatant using a single-step affinity chromatography on a Sepharose-C1s column, as described earlier (22).

Construction of the expression plasmid pNT-BacC1inhΔ97

The C1inhΔ97 construct encoding residues 98–478 of human C1 inhibitor was amplified by PCR using the Vent$_R$ polymerase and the full-length pcDNA3C1inh plasmid. The sequence of the sense primer (5′-GGGCCCAGCCGGATCCCGGGTCCTTCTGCCCA-3′) introduced a BamHI restriction site (underlined) at the 5′ end of the PCR product and allowed in-frame cloning with the melittin signal peptide of the pNT-Bac expression vector (23). The antisense primer (5′-GACCCCAGGGCC**TGA**GAAGCTTAGGATCAGGT-3′) introduced a stop codon (bold type) followed by a HindIII site (underlined) at the 3′ end. The amplified PCR product was purified and cloned into the BamHI and HindIII restriction sites of pNT-Bac.

Plasmids coding for the double mutant C1inhΔ97DM lacking glycosylation at Asn^{216} and Asn^{231} and the triple mutant C1inhΔ97TM lacking glycosylation at Asn^{216}, Asn^{231}, and Asn^{330} were generated, using the QuickChange multisite-directed mutagenesis kit (Stratagene), by mutating the corresponding asparagine residues to glutamine.

Construction of the expression plasmid pNT-BacC1inhΔ97-ht

A 6-His tag was inserted by site-directed mutagenesis in the pNT-BacC1inhΔ97 plasmid between the coding regions of the melittin signal peptide and of residue Gly^{98}. Two complementary 86-bp primers (5′-GCCCTTGTTTTTATGGTCGTGTACATTTCTTACAT-3′ and 5′-CTATGCGCATCACCATCACCATCACGGGTCCTTCTGCCCAGGACCTGTTAC-3′) containing the 6-His coding region (underlined) were used for this purpose. The QuickChange XL site-directed mutagenesis kit (Stratagene) was used, with a two-step amplification as recommended for large insertions.

Protein expression

The insect cells *Spodoptera frugiperda* (Sf21) and *Trichoplusia ni* (High Five) were routinely grown and maintained as described previously (28), and recombinant baculoviruses were generated using the Bac-to-Bac system (Invitrogen). The bacmid DNA was purified using the Qiagen midiprep purification system (Qiagen, Courtaboeuf, France) and used to transfect Sf21 insect cells with cellfectin in Sf900 II SFM medium (Invitrogen), as recommended by the manufacturer. For all C1 inhibitor variants, recombinant virus particles were collected 4 d later and amplified as described by King and Possee (29). High Five cells (1.75×10^7 cells/175 cm^2 tissue culture flask) were infected with the recombinant viruses at a multiplicity of infection of 2 in Sf900 II SFM medium at 28°C for 72 h. The culture supernatants containing the C1 inhibitor constructs were collected by centrifugation, supplemented with 1 mM diisopropyl phosphorofluoridate, and stored frozen at −20°C until use.

Purification of the C1inhΔ97 variants

The same procedure was used for all variants, except C1inhΔ97-ht. Culture supernatants (0.5 l) were dialyzed against 50 mM triethanolamine-HCl (pH 8.6) and loaded onto a Q-Sepharose Fast Flow column (2.8 × 10 cm) (GE Healthcare, Orsay, France) equilibrated in the same buffer. Elution was carried out by applying a 1.2-l linear gradient from 0–350 mM NaCl. Fractions containing the recombinant material were identified by Western blot analysis, dialyzed against 1.5 M (NH$_4$)$_2$SO$_4$, 0.1 M Na$_2$HPO$_4$ (pH 7.4), and further purified by high-pressure hydrophobic interaction chromatography on a TSK-Phenyl 5PW column (Beckman Coulter, Fullerton, CA) equilibrated in the same buffer. Elution was carried out by decreasing the (NH$_4$)$_2$SO$_4$ concentration from 1.5 M to 0 in 30 min. The purified proteins were dialyzed against 145 mM NaCl, 50 mM triethanolamine-HCl (pH 7.4), concentrated to ~1 mg/ml by ultrafiltration, and stored at −20°C.

Purification of C1inhΔ97-ht

The culture supernatant containing C1inhΔ97-ht was dialyzed against 145 mM NaCl, 50 mM triethanolamine-HCl (pH 7.4) and loaded onto a 2-ml His-select HF-Nickel affinity gel column (Sigma-Aldrich). Elution was performed by stepwise addition of 6-ml buffer fractions containing increasing imidazole concentrations of 10, 30, 50, and 250 mM. Fractions containing C1inhΔ97-ht were identified by SDS-PAGE analysis and Coomassie blue staining, dialyzed against 145 mM NaCl, 50 mM triethanolamine-HCl (pH 7.4), and stored at −20°C.

SDS-PAGE analysis of covalent complex formation between truncated C1 inhibitor variants and target proteases

Activated C1s (40 pmol) and C1r-LP (73 pmol) were incubated with equimolar amounts of native C1 inhibitor and different truncated variants, as indicated, for 90 min at 37°C in 145 mM NaCl, 50 mM triethanolamine-HCl (pH 7.4). Samples were analyzed by SDS-PAGE on 10% acrylamide gels under nonreducing conditions.

Inhibition assays

Inhibition of C1s esterolytic activity was measured by preincubating activated C1s (0.5 μM) with increasing concentrations of C1 inhibitor or its variants (0.1–1 μM) for 90 min at 37°C in 145 mM NaCl, 50 mM triethanolamine-HCl (pH 7.4) and then measuring residual C1s activity using *N*-acetylglycine-L-lysine methyl ester, as described previously (21).

The association rate constant of the inhibition of C1s esterolytic activity was determined as follows. Plasma C1s (6 nM) was diluted in 150 mM NaCl, 5 mM Na phosphate (pH 7.4) containing 1 mg/ml OVA and incubated for various periods at 20°C with excess C1 inhibitor or C1inhΔ97 (120 nM) in the presence of a large excess of each heparin oligomer (26 μg/ml). The residual C1s esterolytic activity [*E*] was measured at 30°C, as described above. The pseudo first-order rate constant, *k*ass, was calculated according to the equation $k\text{ass} = k\text{obs}/[I]_0$, where $[I]_0$ is the initial inhibitor concentration and *k*obs is determined from the plot $\ln [E] = -k\text{obs} \cdot t$ (30).

Inhibition of C1 activation was assessed by incubating the C1 complex (0.25 μM) reconstituted from equimolar amounts of C1q and the proenzyme C1s-C1r-C1r-C1s tetramer in the presence of increasing concentrations of C1 inhibitor or its variants (0.5–1.5 μM) for 90 min at 37°C in 145 mM NaCl, 50 mM triethanolamine-HCl, 1 mM CaCl$_2$ (pH 7.4). The reaction mixtures were analyzed by SDS-PAGE under reducing conditions, and C1s activation was measured by Western blot using a polyclonal anti-C1s Ab, as described previously (31).

Inhibition of C1r-LP–mediated C1s activation was measured by preincubating C1r-LP (1 μM) with equimolar amounts of C1 inhibitor or C1inhΔ97 for 30 min at 37°C in 145 mM NaCl, 50 mM triethanolamine-

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

HCl (pH 7.4). Samples were then incubated with proenzyme C1s (2 μM) for 30 min at 30°C in the same buffer, and the extent of C1s activation was assessed by SDS-PAGE under reducing conditions.

Preparation of heparin-derived oligomers

Fifteen-kDa porcine mucosal heparin (Sigma-Aldrich) (10 g) was digested with heparinase I (Grampian Enzymes, Orkney, U.K.) (8 mU/ml) in 150 ml 0.1 mg/ml BSA, 2 mM CaCl$_2$, 50 mM NaCl, and 5 mM Tris-HCl (pH 7.5) for 54 h at 25°C. The reaction was stopped by heating at 100°C for 5 min, and the digest was fractionated on a Bio-Gel P-10 column (4.4 × 150 cm) (Bio-Rad, Hercules, CA) equilibrated with 0.25 M NaCl and run at 1 ml/min. The eluted material, detected by absorbance at 232 nm, consisted of a graded series of oligomers, ranging from 2–18 saccharide units. Only the top fractions of each peak were pooled, and each pool was resubmitted to gel filtration. Samples were dialyzed against water and quantified.

Heparin biotinylation

Six-kDa heparin (Sigma-Aldrich) at 1 mM in PBS was reacted for 24 h at room temperature with 10 mM biotin-LC-hydrazine (Pierce, Rockford, IL), as described by Vivès et al. (32). The mixture was extensively dialyzed against water to remove excess biotin and freeze dried.

Thermal shift assay

Analyses were performed on an IQ5 96-well real-time PCR instrument (Bio-Rad), as described previously (33). Proteins (23 μl, 7–26 μM) in 145 mM NaCl, 50 mM triethanolamine-HCl (pH 7.4), in the presence or absence of a 2:1 molar excess of 15-kDa porcine mucosal heparin, were mixed with 2 μl of 5000× SYPRO Orange (Molecular Probes, Eugene, OR) diluted 1:100 in water. Samples were heat-denatured from 20°C to 95°C at 1°C/min. Protein unfolding was monitored by recording changes in the fluorescence of SYPRO Orange. The inflection point of the curves was determined by plotting the first derivative, and the melting temperatures were assessed from the minima. The fluorescence of buffers was checked as a control.

Surface plasmon resonance spectroscopy

Heparin-binding analyses were performed on a BIAcore 3000 instrument (GE Healthcare). Streptavidin (200 μg/ml in 10 mM Na acetate [pH 4.2]) was immobilized on a CM4 sensor chip (GE Healthcare), using the amine-coupling chemistry, in 10 mM HEPES 150 mM NaCl (pH 7.4) until a coupling level of 2500 resonance units (RUs) was reached. Biotinylated 6-kDa heparin was then captured on the streptavidin surface in 10 mM HEPES, 300 mM NaCl, 0.005% surfactant P20 (pH 7.4) to reach a final coupling value of 30–50 RUs. A flow cell with 2500 RU streptavidin was used as a negative control.

Heparin-binding analyses. Native C1 inhibitor, C1inhΔ97, C1inhΔ97DM, and activated C1s (15 μl) were injected over the heparin surface at 10 μl/min in 50 mM triethanolamine-HCl, 145 mM NaCl, 0.005% surfactant P20 (pH 7.4) at six concentrations, ranging from 0.05–1 μM. Surfaces were regenerated with 15 μl 2 M NaCl, with additional injections of 5 μl 0.05% SDS and 10 μl 2 M NaCl for the C1inhΔ97 variants. Data were analyzed by global fitting to a 1:1 Langmuir binding model of the association and dissociation phases simultaneously, using BIAevaluation 3.1 software (GE Healthcare). The apparent equilibrium dissociation constants (K_D) were calculated from the k_{off}/k_{on} ratios of the dissociation and association constants.

Competition by heparin oligomers. C1 inhibitor or C1s (each 0.53 μM) was injected on the heparin sensor chip at 50 μl/min in 50 mM triethanolamine-HCl, 145 mM NaCl, 0.005% surfactant P20 (pH 7.4) in the presence of various heparin oligomers (2–20 saccharide units) at concentrations of 12 mg/l for C1 inhibitor and 30 mg/l for C1s. The number of RUs determined in the presence of each oligomer was compared with the value measured in the absence of heparin.

The interaction between C1inhΔ97 and C1s Ser617Ala was monitored using a BIAcore X instrument (GE Healthcare). C1s was immobilized on the surface of a CM5 sensor chip (GE Healthcare) using the amine-coupling chemistry. Binding of C1inhΔ97 was measured over 10,000 RUs immobilized C1s Ser617Ala, at a flow rate of 20 μl/min in 10 mM HEPES, 145 mM NaCl (pH 7.4), containing 0.005% surfactant P20. Each sample was injected over a surface with immobilized BSA for subtraction of the bulk refractive index background.

Other methods

SDS-PAGE, N-terminal sequence, and MALDI mass spectrometry analyses were carried out as described previously (23, 24, 34). Western blot analysis was performed according to published procedures (23) using a polyclonal anti-C1 inhibitor antiserum.

Results

The objective of this study was to further functionally characterize the C-terminal serpin domain of human C1 inhibitor. For this purpose, the C1inhΔ97 DNA construct coding for residues 98–478 of human C1 inhibitor and two variants lacking *N*-linked carbohydrates at positions 216, 231, and 330 (C1inhΔ97TM) or solely containing the glycosylation site at Asn^{330} (C1inhΔ97DM) were expressed using a baculovirus/insect cells system (Fig. 1). All C1 inhibitor variants were secreted in the culture supernatants, at concentrations ~10 mg/l (C1inhΔ97 and C1inhΔ97-ht), 3 mg/l (C1inhΔ97DM), and 0.2 mg/l (C1inhΔ97TM). Thus, removal of the oligosaccharide linked to Asn^{330} resulted in a dramatic decrease of the expression yield, precluding purification of material in sufficient amounts to perform further analysis. *N*-terminal sequence analysis of the purified C1inhΔ97 variants yielded a major sequence *Asp-Pro*-Gly-Ser-Phe-(Cys)-Pro-Gly-Pro-Val… in all instances, corresponding to the segment Gly^{98}–Val^{105} of human C1 inhibitor preceded by the two residues Asp-Pro expected to be added at the N terminus, due to insertion of the BamHI restriction site at the cDNA 5′ end (see *Materials and Methods*). Occasionally, a minor (<10%) sequence Gly-Ser-Phe-(Cys)-Pro-Gly-Pro-Val… was also observed, indicating some heterogeneity in the maturation process. Sequence analysis of the C1inhΔ97-ht variant yielded the expected six-His cluster followed by the Gly^{98}-Ser-Phe-(Cys)-Pro-Gly-Pro-Val… sequence. Analysis of C1inhΔ97 and C1inhΔ97-ht by MALDI mass spectrometry consistently yielded rather homogeneous peaks centered on mass values of 46,228 ± 50 Da and 46,741 ± 50 Da, respectively, accounting for the polypeptide chains plus three high-mannose *N*-linked oligosaccharides with average deduced masses of 1,085–1,117 ± 50 Da. Similarly, analysis of the C1inhΔ97DM variant yielded a mean mass value of 43,824 ± 50 Da, accounting for the polypeptide chain plus the remaining oligosaccharide linked to Asn^{330} (deduced mass = 920 ± 50 Da).

Complex formation with target proteases

Analysis of the rC1 inhibitor variants by SDS-PAGE revealed major bands with apparent m.w. ~49,500 (C1inhΔ97-ht), 44,000 (C1inhΔ97), and 39,500 (C1inhΔ97DM) (Fig. 2). C1inhΔ97-ht and C1inhΔ97 were essentially pure as judged from SDS-PAGE analysis, whereas minor contaminants were still present in the

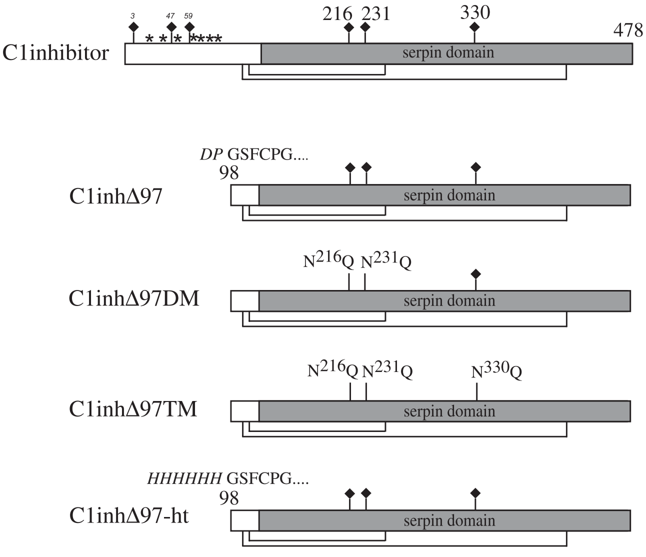


FIGURE 1. Schematic representation of the structure of human C1 inhibitor and of the C1inhΔ97 variants expressed in this study. The Cys^{101}–Cys^{406} and Cys^{108}–Cys^{183} disulfide bridges are represented. *, *O*-linked carbohydrates; ◆, *N*-linked carbohydrates.

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

purified C1inhΔ97DM preparation (Fig. 2*A*, *lane 4*). As observed for full-length C1 inhibitor (Fig. 2*A*, *lane 5*), all truncated variants reacted with active C1s to form SDS- and urea-stable complexes of apparent m.w. ~120,000 (C1inhΔ97), 116,000 (C1inhΔ97DM), and 120,000 (C1inhΔ97-ht). Covalent complex formation was almost complete when native C1 inhibitor was allowed to react with an equimolar amount of C1s for 90 min at 37°C, but it was only partial for the recombinant variants under the same conditions, suggesting a decreased efficiency of the truncated forms. Nevertheless, as illustrated in the case of C1inhΔ97-ht, increasing the C1 inhibitor variant/C1s molar ratio to 2:1 readily increased formation of the complex in all instances (Fig. 2*B*, *lane 4*).

Using C1r-LP as a second target protease, we checked its ability to form stable complexes with native C1 inhibitor and C1inhΔ97. As illustrated in Fig. 2*C*, C1r-LP alone migrated as a diffuse band of apparent m.w. 73,000. No evidence for a band of higher m.w. was obtained after incubation of C1r-LP for 90 min at 37°C in the presence of equimolar amounts of full-length C1 inhibitor or the C1inhΔ97 variant, indicating that a stable complex was not formed in either case.

We next used surface plasmon resonance spectroscopy to test the ability of C1inhΔ97 to interact with C1s, using a C1s mutant in which the active site Ser617 was mutated to Ala to prevent covalent reaction with the inhibitor. As shown in Fig. 3, C1inhΔ97 bound to immobilized C1s Ser617Ala to form a relatively stable complex. Interaction was clearly dose dependent, but the data could not be fitted properly, precluding accurate determination of the kinetic parameters.

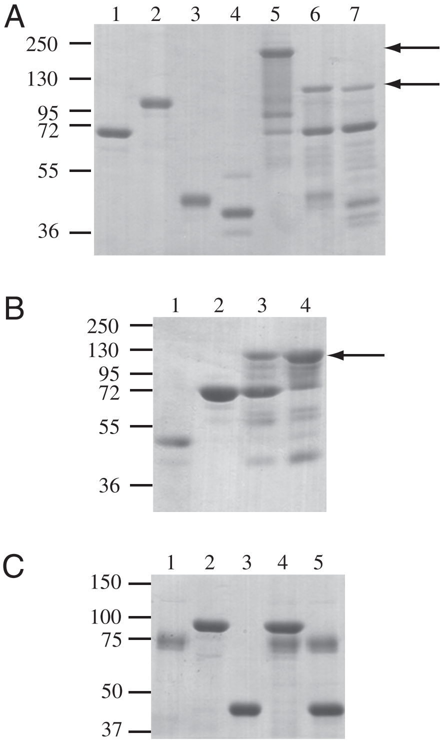


FIGURE 2. SDS-PAGE analysis of the rC1 inhibitor constructs and their ability to form covalent complexes with C1s and C1r-LP. *A*, *Lane 1*, activated C1s; *lane 2*, native C1 inhibitor; *lane 3*, C1inhΔ97; *lane 4*, C1inhΔ97DM; *lanes 5–7*, native C1 inhibitor, C1inhΔ97, and C1inhΔ97DM incubated for 90 min at 37°C with equimolar amounts of activated C1s. *B*, *Lane 1*, C1inhΔ97-ht; *lane 2*, purified C1s; *lanes 3*, *4*, C1inhΔ97-ht incubated for 90 min at 37°C with activated C1s in a 1:1 and 2:1 molar ratio, respectively. *C*, *Lane 1*, C1r-LP; *lane 2*, native C1 inhibitor; *lane 3*, C1inhΔ97; *lanes 4*, *5*, native C1 inhibitor and C1inhΔ97 incubated for 90 min at 37°C with equimolar amounts of C1r-LP. All samples were analyzed under nonreducing conditions. The molecular masses of standard proteins (expressed in kDa) are shown. Covalent complexes between C1s and the C1 inhibitor variants are indicated by arrows.

Inhibition of protease activity

The ability of C1 inhibitor variants to inhibit C1s enzymatic activity was next investigated using *N*-acetylglycine-L-lysine methyl ester as a synthetic substrate. As depicted in Fig. 4*A*, native C1 inhibitor and C1inhΔ97 each gradually titrated C1s esterolytic activity, and complete inhibition was reached in each instance at a inhibitor/C1s molar ratio of ~1:1. Similar inhibition curves were obtained in the case of C1inhΔ97DM and C1inhΔ97-ht (data not shown).

We next tested the ability of C1 inhibitor and its C1inhΔ97 variant to inhibit the proteolytic activity of C1r-LP toward proenzyme C1s. In line with previous findings (8), preincubation of C1r-LP with an equimolar amount of native C1 inhibitor nearly totally prevented its ability to activate proenzyme C1s (Fig. 4*B*, 4*C*). A similar result was obtained using C1inhΔ97 under the same conditions. Thus, despite the fact that C1 inhibitor or C1inhΔ97 did not form stable complexes with C1r-LP as revealed by SDS-PAGE analysis, both proteins readily inhibited C1r-LP proteolytic activity toward C1s. Whether C1inhΔ97DM also inhibited C1r-LP without forming a stable complex could not be determined because of the low amount of C1r-LP available.

Inhibition of C1 activation

To further functionally characterize the truncated C1 inhibitor variants, we checked their ability to prevent spontaneous C1 activation. In keeping with previous results (31), native C1 inhibitor inhibited C1 activation in a dose-dependent fashion, and maximal inhibitory effect was reached at a C1 inhibitor/C1 molar ratio of ~2:1 (Fig. 5). Similar inhibition curves were obtained using C1inhΔ97 or C1inhΔ97DM, demonstrating that the serpin domain of C1 inhibitor fully retained the ability to prevent C1 activation and that removal of the carbohydrates at positions 216 and 231 had no impact on this property.

Interaction with heparin

The ability of C1 inhibitor and its variants to interact with heparin was initially assessed by surface plasmon resonance spectroscopy, using the avidin-biotin system to immobilize 6-kDa heparin onto the biosensor surface through its reducing end. As illustrated in Fig. 6*A*, native C1 inhibitor readily bound to immobilized heparin, with a K_D of 16.7 × 10^{-8} M, a value in the same range as that determined previously by Caldwell et al. (35) under comparable experimental conditions. The truncated proteins C1inhΔ97 and

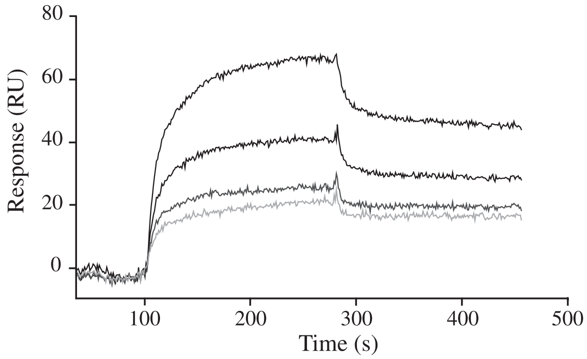


FIGURE 3. Surface plasmon resonance analysis of the interaction of C1inhΔ97 with the C1s Ser617Ala mutant. C1s Ser617Ala was activated by C1r, immobilized on a CM5 sensor chip, and allowed to interact with soluble C1inhΔ97. The interaction curves shown correspond to C1inhΔ97 concentrations of 0.4, 0.5, 0.75, and 1 μM (from bottom to top).

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

P16

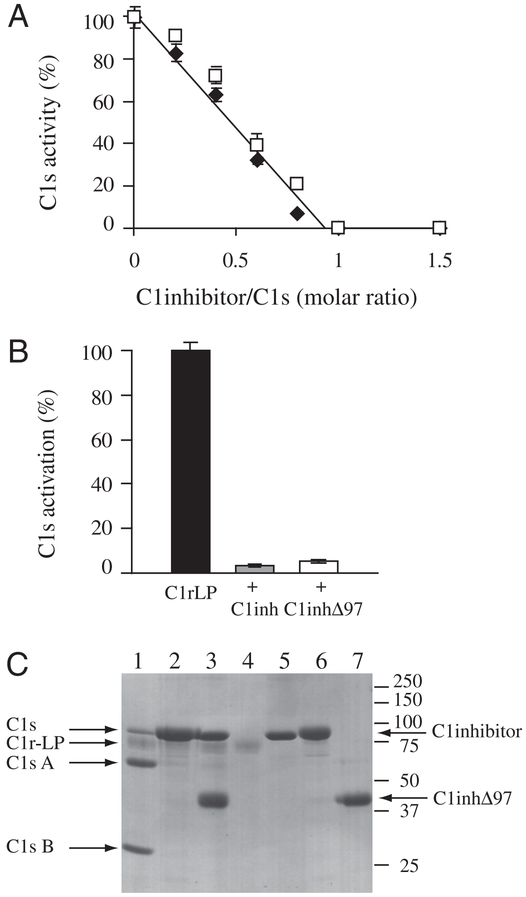


FIGURE 4. Inhibition of C1s and C1r-LP enzymatic activities. *A*, Inhibition of C1s esterolytic activity. Activated C1s was preincubated for 90 min at 37°C with increasing amounts of native C1 inhibitor (◆) or C1inhΔ97 (☐), and its residual esterolytic activity was measured using *N*-acetylglycine-L-lysine methyl ester. *B*, Inhibition of C1r-LP–mediated C1s activation. C1r-LP was preincubated for 30 min at 37°C with equimolar amounts of native C1 inhibitor or C1inhΔ97, and its residual ability to activate proenzyme C1s was measured by SDS-PAGE, as described in *Materials and Methods*. *C*, SDS-PAGE analysis corresponding to the experiment shown in *B*. *Lane 1*, C1r-LP + C1s; *lane 2*, C1r-LP + C1s + C1 inhibitor; *lane 3*, C1r-LP + C1s + C1inhΔ97; and *lane 4–7*, C1r-LP, proenzyme C1s, C1 inhibitor, and C1inhΔ97, respectively. All samples were analyzed under reducing conditions. The molecular masses of standard proteins (expressed in kDa) are shown.

C1inhΔ97DM each retained the ability to bind heparin, with even significantly lower K_D values resulting, for the most part, from decreased k_{off} values (Table I). Considering that proteins expressed in insect cells lack the terminal sialic acid characteristic of serum glycoproteins, this is expected to decrease their negative charge and may explain why C1inhΔ97 and C1inhΔ97DM each form heparin–inhibitor complexes with increased stability compared with purified C1 inhibitor. These latter data provided a clear indication that the serpin domain of C1 inhibitor contains the structural elements required for interaction with heparin. In comparative binding experiments, activated C1s also bound to immobilized heparin, although with a significantly higher K_D (108×10^{-8} M), resulting essentially from an increased k_{off} value, indicative of a much lower stability of the heparin–C1s complex.

To determine whether the size of heparin influenced its ability to interact with C1 inhibitor, native C1 inhibitor was allowed to bind to immobilized 6-kDa heparin in the presence of heparin oligomers of varying sizes. As illustrated in Fig. 6*B*, increasing the size of the heparin oligomers from 2 to 20 saccharide units progressively decreased binding of native C1 inhibitor, with 50% inhibition being achieved by oligomers containing eight saccharide units. A similar competition experiment performed using C1s yielded a comparable inhibition curve (Fig. 6*C*). However, in this case, half-maximal inhibition of the interaction could be achieved using heparin oligomers containing only six saccharide units.

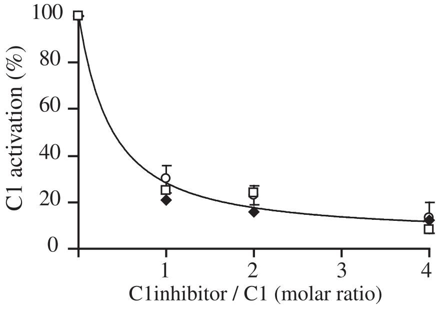


FIGURE 5. Inhibition of C1 activation. The C1 complex reconstituted from equimolar amounts of C1q and the proenzyme C1s-C1r-C1r-C1s tetramer was incubated for 90 min at 37°C in the presence of increasing amounts of native C1 inhibitor (◆), C1inhΔ97 (☐), or C1inhΔ97DM (○). The extent of C1s activation was measured by SDS-PAGE and Western blot analysis, as described in *Materials and Methods*.

We next investigated the effect of heparin size on the potentiation of the inhibition of C1s activity by C1 inhibitor. For this purpose, the association rate constant of the inhibition of C1s esterolytic activity by full-length C1 inhibitor and C1inhΔ97 was measured in the presence of the different heparin oligomers, as described in *Materials and Methods*. As illustrated for full-length C1 inhibitor, increasing the heparin oligomer size progressively increased the k_{ass} value and 50% potentiation was achieved using the oligomer containing 10 saccharide units (Fig. 6*D*). This value is slightly higher but comparable to those measured for the interaction of heparin with C1 inhibitor and C1s. Therefore, this result seems to be consistent with a "sandwich" mechanism rather than with a "bridging" mechanism (19); in the latter case, the oligomer size necessary to achieve half-maximal potentiation would be expected to be much higher. Similar potentiation curves were obtained using C1inhΔ97 instead of full-length C1 inhibitor (data not shown).

Thermal shift assays

A thermal shift assay based on the use of a fluorescent probe was used to measure the thermal stability of C1 inhibitor and its truncated form C1inhΔ97-ht. As shown in Fig. 7*A*, native C1 inhibitor displayed a melting temperature of 51.5°C, and this value remained unchanged when analysis was performed in the presence of a 2:1 molar excess of heparin. C1inhΔ97-ht exhibited a similar melting temperature at ~50.5°C (Fig. 7*B*). However, in contrast to native C1 inhibitor, this value was shifted to 47°C in the presence of heparin, indicating that interaction with the isolated serpin domain of C1 inhibitor significantly decreased its thermal stability. Interestingly, a minor negative peak was also observed at 66°C, possibly reflecting the presence of low amounts of a latent form of the serpin domain (19).

Analysis of active C1s by the same method yielded two minima at 39°C and 49°C, which are in good agreement with the melting temperatures of 37°C and 49°C previously determined by differential scanning calorimetry (36), corresponding to the CUB$_1$-EGF-CUB$_2$ segment and the SP domain, respectively (Fig. 7*C*). The third thermal transition at 60°C, corresponding to the CCP$_1$-CCP$_2$ segment (36), could not be detected by the method used. The transition at 49°C was shifted to 52°C in the presence of heparin, indicating

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

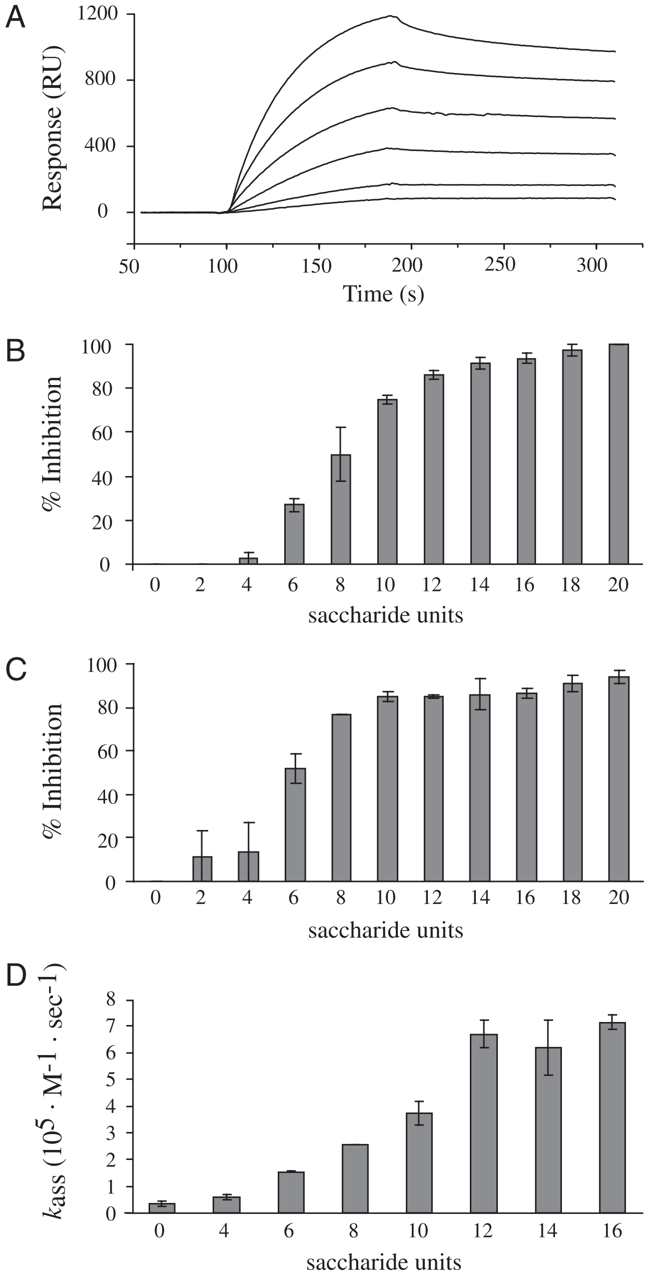


FIGURE 6. Interaction of C1 inhibitor and C1s with immobilized heparin. *A*, Surface plasmon resonance analysis of the interaction between native C1 inhibitor and immobilized heparin. Six-kDa biotinylated heparin was immobilized on the sensor chip, as described in *Materials and Methods*. Native C1 inhibitor was injected at six concentrations, ranging from 0.05–1 μM (from bottom to top). *B*, Inhibition of C1 inhibitor–heparin interaction by heparin oligomers. Native C1 inhibitor was allowed to bind to immobilized heparin in the presence of heparin oligomers of varying sizes, as indicated. The number of RUs determined in each case is expressed relative to the value measured in the absence of heparin. *C*, Inhibition of C1s–heparin interaction by heparin oligomers. Active C1s was allowed to bind to immobilized heparin in the presence of heparin oligomers of varying sizes, as indicated, and inhibition was determined as in *B*. *D*, Potentiation by heparin oligomers of the inhibition of C1s activity by native C1 inhibitor. The association rate constant of the inhibition of C1s esterolytic activity by full-length C1 inhibitor was measured in the presence of the different heparin oligomers, as described in *Materials and Methods*.

Table I. *Kinetic and dissociation constants for the interaction of C1 inhibitor, C1 inhibitor variants, and C1s with immobilized heparin*

Proteins	k_{on} ($M^{-1}.s^{-1}$)	k_{off} (s^{-1})	K_D (M)[a]
C1 inhibitor	1.94×10^4	3.32×10^{-3}	$16.7 \pm 2.7 \times 10^{-8}$
C1inhΔ97	3.75×10^4	0.87×10^{-3}	$2.31 \pm 0.3 \times 10^{-8}$
C1inhΔ97DM	1.83×10^4	0.66×10^{-3}	$3.61 \pm 0.4 \times 10^{-8}$
C1s	1.01×10^4	11.0×10^{-3}	$108 \pm 10 \times 10^{-8}$

[a] Mean values determined from three separate experiments.

that heparin bound to the SP domain and increased its thermal stability. Consistent with this observation, the CCP$_2$–SP segment of C1s only exhibited a transition at 47°C corresponding to the SP domain, and this value was shifted to 51°C in the presence of heparin (Fig. 7*D*). Using heparin oligomers, it was determined that full thermal stabilization of the SP domain could be achieved using oligomers containing 10 saccharide units, in agreement with the data obtained by surface plasmon resonance spectroscopy.

Discussion

Different variants of the C1inhΔ97 construct corresponding to the C-terminal serpin domain of human C1 inhibitor were expressed using a baculovirus/insect cells system and purified to homogeneity. The expression yields ranged from 3–10 mg/l and, therefore, were comparable to that reported previously for the production of full-length C1 inhibitor using the same expression system (16). In keeping with the report by Wolff et al. (16) and in contrast with the highly heterogeneous material produced in *P. pastoris* (12, 19), the C1 inhibitor serpin domain expressed in the current study was homogeneous in terms of size, as shown by SDS-PAGE and MALDI mass spectrometry analyses, indicating relative homogeneity of the *N*-linked oligosaccharides. In this regard, the mass spectrometry analyses performed on the recombinant proteins are consistent with the occurrence of short, oligomannose carbohydrates comprising two *N*-acetylglucosamine residues and three to four mannose residues (calculated mass = 893–1055), in keeping with previous analyses (16). For the first time, we have used site-directed mutagenesis to selectively prevent *N*-linked glycosylation at Asn residues at positions 216, 231, and 330. This strategy reveals that, although mutation of the former two residues has little effect on protein production, removal of the oligosaccharide linked to Asn^{330} results in a dramatic decrease in the expression yield, suggesting that this particular carbohydrate may play a role in the folding and/or the stability of the serpin domain. Previous attempts to prevent overall glycosylation of rC1 inhibitor or its serpin domain were made using expression in a bacterial system (18) or in the presence of tunicamycin (11, 15). Although these preparations were found to retain protease inhibitory activity, inhibition of glycosylation was reported to significantly decrease the expression yield (15) or to yield a major fraction of the recombinant material in insoluble form (18). In contrast, it is noteworthy that a latent form of the C1 inhibitor serpin domain was crystallized after enzymatic removal of *N*-linked carbohydrates by Endoglycosidase H (19). Taken together, these observations suggest that glycosylation at Asn^{330} may be crucial for proper folding of the serpin domain, but it is not required to stabilize the folded conformation.

Using active C1s as a target protease, it was found that formation of a covalent protease–inhibitor complex was comparatively less efficient for the truncated variants than for the full-length protein. Nevertheless, the truncated form C1inhΔ97 was shown to inhibit C1s esterolytic activity as efficiently as native C1 inhibitor and with the expected 1:1 stoichiometry. We also provide evidence that, as demonstrated previously for native C1 inhibitor (8), C1inhΔ97 efficiently inhibits the C1s-cleaving activity of C1r-LP. In contrast, neither of these molecules had the ability to form a stable complex with this protease, as judged

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

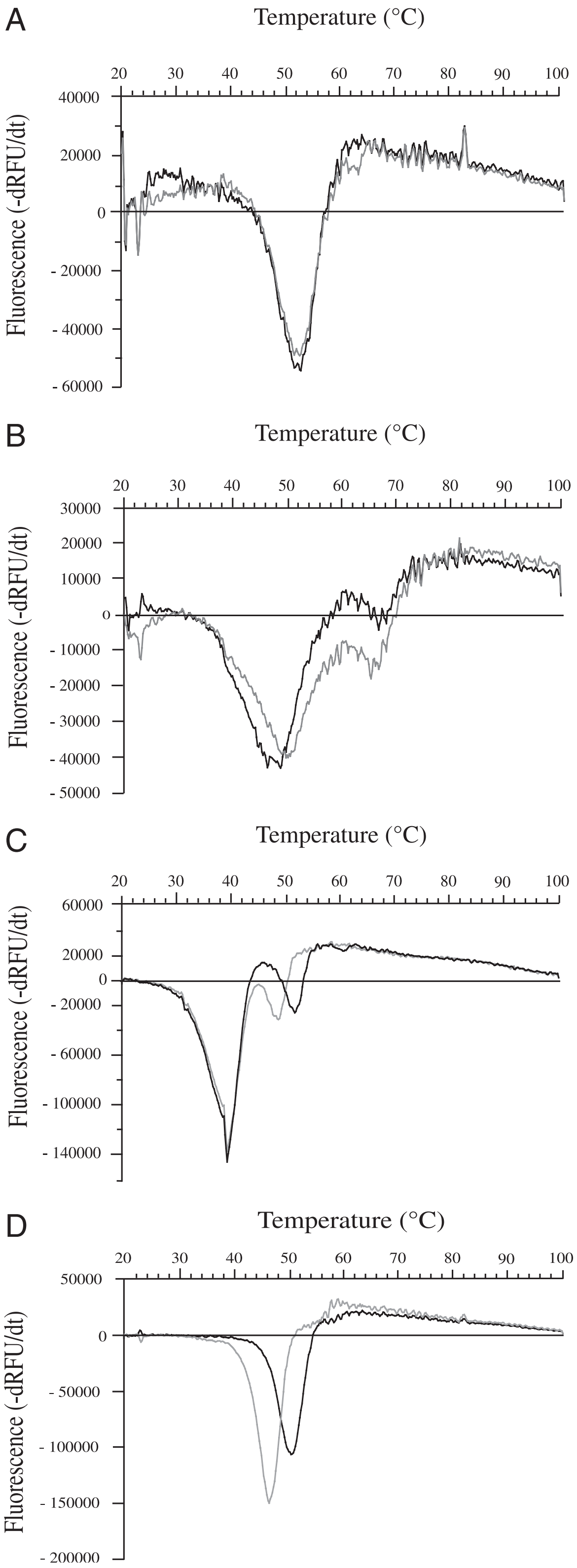


FIGURE 7. Thermal stability of C1 inhibitor and C1s and the effect of interaction with heparin. *A*, Native C1 inhibitor. *B*, C1inhΔ97-ht. *C*, Active C1s. *D*, C1s CCP$_2$-SP segment. Curves corresponding to analyses conducted in the absence of heparin and in the presence of a 2:1 molar excess of heparin are shown in gray and black, respectively.

from SDS-PAGE analysis. The example of C1r-LP shows that efficient inhibition of a protease activity by C1 inhibitor or its serpin domain does not necessarily require formation of a covalent protease–inhibitor linkage (i.e., cleavage of the reactive Arg-Thr bond), but that formation of a transient, noncovalent complex may, in certain cases, be sufficient to fulfill this function. The data obtained using active C1s as a target are more difficult to interpret, considering that small, but significant, amounts of the covalent complex were formed upon reaction with C1inhΔ97. In contrast, our finding that C1inhΔ97 forms a transient complex with the inactive C1s Ser617Ala mutant does not imply that the reaction does not proceed to a covalent complex when active C1s is the target.

In keeping with previous functional investigations (11, 18), the experiments carried out in the current study using C1s and C1r-LP as target proteases provide further evidence that the serpin domain of C1 inhibitor contains the structural determinants required for mediating protease inhibitory activity and that the N-terminal extension of C1 inhibitor is not implicated in this function.

Among the potential roles considered for the N-terminal extension of C1 inhibitor was the possibility that it participates in the control of C1 activation (11). In this respect, again, our data demonstrate that the C1inhΔ97 and C1inhΔ97DM variants share the ability of native C1 inhibitor to prevent C1 activation, demonstrating that the serpin domain alone is sufficient to fulfill this function. As observed in the case of C1r-LP, it appears likely that the ability of C1 inhibitor to prevent C1 activation involves formation of a reversible complex between its serpin domain and the proenzyme catalytic site of C1r, thereby inhibiting spontaneous C1r activation in the absence of activator or in the presence of mild activators (37).

In line with previous analyses using the same technique (35), our surface plasmon resonance spectroscopy data confirm that C1 inhibitor binds to heparin with high affinity and provide evidence that interaction is mediated by the serpin domain. C1s was also shown to bind heparin through its SP domain, but with a substantially lower affinity, reflecting the formation of a less stable complex. Consistent with this finding, full inhibition of the interaction by heparin oligomers was achieved more readily in the case of C1s than in the case of C1 inhibitor. An additional and intriguing difference between C1s and C1 inhibitor is that interaction with heparin seems to increase and decrease their thermal stability, respectively. Although these latter observations remain to be explained in molecular terms, our data provide further support to the hypothesis that heparin provides a link between C1 inhibitor and certain of its target proteases. With respect to the mechanism of this interaction, our observation that the heparin oligomer size required to achieve half-maximal potentiation of the inhibition of C1s activity by C1 inhibitor is on the same order as those necessary to inhibit binding of C1 inhibitor and C1s to 6-kDa heparin (Fig. 6*B*–*D*) seems to favor a "sandwich," rather than a "bridging," model. This conclusion is in line with the mechanism proposed by Beinrohr et al. (19) and seems to be fully consistent with the known characteristics of the potentiation of C1 inhibitor by polyanions, as reviewed by those investigators (19). Attempts to identify the C1 inhibitor residues involved in heparin interaction, by performing proteolytic digestion of C1 inhibitor–heparin cross-linked complexes followed by N-terminal sequence analysis (38), were unsuccessful, suggesting implication of residues located in various areas of the C1 inhibitor three-dimensional structure.

In summary, the current study provides further evidence that the serpin domain of C1 inhibitor is self-sufficient in terms of protease inhibitory activity and demonstrates that this also applies to the regulation of C1 activation and to heparin binding. The N-terminal extension of C1 inhibitor has no implication in these activities, and its precise role remains to be investigated further.

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

Acknowledgments

We thank Elodie Loizel for technical assistance with the thermal shift assay, Jean-Pierre Andrieu for performing N-terminal sequence analyses, and Izabel Bérard for mass spectrometry measurements.

Disclosures

The authors have no financial conflicts of interest.

References

1. Sim, R. B., A. Reboul, G. J. Arlaud, C. L. Villiers, and M. G. Colomb. 1979. Interaction of ^{125}I-labelled complement subcomponents C-1r and C-1s with protease inhibitors in plasma. *FEBS Lett.* 97: 111–115.
2. Rossi, V., S. Cseh, I. Bally, N. M. Thielens, J. C. Jensenius, and G. J. Arlaud. 2001. Substrate specificities of recombinant mannan-binding lectin-associated serine proteases-1 and -2. *J. Biol. Chem.* 276: 40880–40887.
3. Davis, A. E., 3rd, P. Mejia, and F. Lu. 2008. Biological activities of C1 inhibitor. *Mol. Immunol.* 45: 4057–4063.
4. Ambrus, G., P. Gál, M. Kojima, K. Szilágyi, J. Balczer, J. Antal, L. Gráf, A. Laich, B. E. Moffatt, W. Schwaeble, et al. 2003. Natural substrates and inhibitors of mannan-binding lectin-associated serine protease-1 and -2: a study on recombinant catalytic fragments. *J. Immunol.* 170: 1374–1382.
5. Schapira, M., C. F. Scott, and R. W. Colman. 1982. Contribution of plasma protease inhibitors to the inactivation of kallikrein in plasma. *J. Clin. Invest.* 69: 462–468.
6. Pixley, R. A., M. Schapira, and R. W. Colman. 1985. The regulation of human factor XIIa by plasma proteinase inhibitors. *J. Biol. Chem.* 260: 1723–1729.
7. Wuillemin, W. A., M. Minnema, J. C. Meijers, D. Roem, A. J. Eerenberg, J. H. Nuijens, H. ten Cate, and C. E. Hack. 1995. Inactivation of factor XIa in human plasma assessed by measuring factor XIa-protease inhibitor complexes: major role for C1-inhibitor. *Blood* 85: 1517–1526.
8. Ligoudistianou, C., Y. Xu, G. Garnier, A. Circolo, and J. E. Volanakis. 2005. A novel human complement-related protein, C1r-like protease (C1r-LP), specifically cleaves pro-C1s. *Biochem. J.* 387: 165–173.
9. Davis, A. E., 3rd, S. Cai, and D. Liu. 2007. C1 inhibitor: biologic activities that are independent of protease inhibition. *Immunobiology* 212: 313–323.
10. Bock, S. C., K. Skriver, E. Nielsen, H.-C. Thøgersen, B. Wiman, V. H. Donaldson, R. L. Eddy, J. Marrinan, E. Radziejewska, R. Huber, et al. 1986. Human C1 inhibitor: primary structure, cDNA cloning, and chromosomal localization. *Biochemistry* 25: 4292–4301.
11. Coutinho, M., K. S. Aulak, and A. E. Davis, 3rd. 1994. Functional analysis of the serpin domain of C1 inhibitor. *J. Immunol.* 153: 3648–3654.
12. Bos, I. G., Y. T. Lubbers, D. Roem, J. P. Abrahams, C. E. Hack, and E. Eldering. 2003. The functional integrity of the serpin domain of C1-inhibitor depends on the unique N-terminal domain, as revealed by a pathological mutant. *J. Biol. Chem.* 278: 29463–29470.
13. Salvesen, G. S., J. J. Catanese, L. F. Kress, and J. Travis. 1985. Primary structure of the reactive site of human C1-inhibitor. *J. Biol. Chem.* 260: 2432–2436.
14. Wuillemin, W. A., H. te Velthuis, Y. T. Lubbers, C. P. de Ruig, E. Eldering, and C. E. Hack. 1997. Potentiation of C1 inhibitor by glycosaminoglycans: dextran sulfate species are effective inhibitors of in vitro complement activation in plasma. *J. Immunol.* 159: 1953–1960.
15. Eldering, E., J. H. Nuijens, and C. E. Hack. 1988. Expression of functional human C1 inhibitor in COS cells. *J. Biol. Chem.* 263: 11776–11779.
16. Wolff, M. W., F. Zhang, J. J. Roberg, E. E. Caldwell, P. R. Kaul, J. N. Serrahn, D. W. Murhammer, R. J. Linhardt, and J. M. Weiler. 2001. Expression of C1 esterase inhibitor by the baculovirus expression vector system: preparation, purification, and characterization. *Protein Expr. Purif.* 22: 414–421.
17. Bos, I. G. A., E. C. de Bruin, Y. A. Karuntu, P. W. Modderman, E. Eldering, and C. E. Hack. 2003. Recombinant human C1-inhibitor produced in *Pichia pastoris* has the same inhibitory capacity as plasma C1-inhibitor. *Biochim. Biophys. Acta* 1648: 75–83.
18. Lamark, T., M. Ingebrigtsen, C. Bjørnstad, T. Melkko, T. E. Mollnes, and E. W. Nielsen. 2001. Expression of active human C1 inhibitor serpin domain in *Escherichia coli. Protein Expr. Purif.* 22: 349–358.
19. Beinrohr, L., V. Harmat, J. Dobó, Z. Lörincz, P. Gál, and P. Závodszky. 2007. C1 inhibitor serpin domain structure reveals the likely mechanism of heparin potentiation and conformational disease. *J. Biol. Chem.* 282: 21100–21109.
20. Arlaud, G. J., R. B. Sim, A.-M. Duplaa, and M. G. Colomb. 1979. Differential elution of Clq, Clr and Cls from human Cl bound to immune aggregates. Use in the rapid purification of Cl subcomponents. *Mol. Immunol.* 16: 445–450.
21. Arlaud, G. J., and N. M. Thielens. 1993. Human complement serine proteases C1r and C1s and their proenzymes. *Methods Enzymol.* 223: 61–82.
22. Bally, I., V. Rossi, N. M. Thielens, C. Gaboriaud, and G. J. Arlaud. 2005. Functional role of the linker between the complement control protein modules of complement protease C1s. *J. Immunol.* 175: 4536–4542.
23. Rossi, V., I. Bally, N. M. Thielens, A. F. Esser, and G. J. Arlaud. 1998. Baculovirus-mediated expression of truncated modular fragments from the catalytic region of human complement serine protease C1s. Evidence for the involvement of both complement control protein modules in the recognition of the C4 protein substrate. *J. Biol. Chem.* 273: 1232–1239.
24. Thielens, N. M., C. A. Aude, M. B. Lacroix, J. Gagnon, and G. J. Arlaud. 1990. Ca2+ binding properties and Ca2(+)-dependent interactions of the isolated NH2-terminal alpha fragments of human complement proteases C1-r and C1-s. *J. Biol. Chem.* 265: 14469–14475.
25. Pétillot, Y., P. Thibault, N. M. Thielens, V. Rossi, M. Lacroix, B. Coddeville, G. Spik, V. N. Schumaker, J. Gagnon, and G. J. Arlaud. 1995. Analysis of the *N*-linked oligosaccharides of human C1s using electrospray ionisation mass spectrometry. *FEBS Lett.* 358: 323–328.
26. Harpel, P. C. 1976. C1 inactivator. *Methods Enzymol.* 45: 751–750.
27. Gill, S. C., and P. H. von Hippel. 1989. Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* 182: 319–326.
28. Thielens, N. M., S. Cseh, S. Thiel, T. Vorup-Jensen, V. Rossi, J. C. Jensenius, and G. J. Arlaud. 2001. Interaction properties of human mannan-binding lectin (MBL)-associated serine proteases-1 and -2, MBL-associated protein 19, and MBL. *J. Immunol.* 166: 5068–5077.
29. King, L. A., and R. D. Possee. 1992. *The Baculovirus Expression System: A Laboratory Guide.* Chapman and Hall, Ltd, London.
30. Nagase, H., and G. Salvesen. 1989. Inhibition of proteolytic enzymes. In *Proteolytic Enzymes, a Practical Approach.* R. J. Beynon, and J. S. Bond, eds. IRL Press, Oxford, U.K., p. 144.
31. Tacnet-Delorme, P., S. Chevallier, and G. J. Arlaud. 2001. Beta-amyloid fibrils activate the C1 complex of complement under physiological conditions: evidence for a binding site for A beta on the C1q globular regions. *J. Immunol.* 167: 6374–6381.
32. Vivès, R. R., R. Sadir, A. Imberty, A. Rencurosi, and H. Lortat-Jacob. 2002. A kinetics and modeling study of RANTES(9-68) binding to heparin reveals a mechanism of cooperative oligomerization. *Biochemistry* 41: 14779–14789.
33. Attali, C., C. Frolet, C. Durmort, J. Offant, T. Vernet, and A.-M. Di Guilmi. 2008. *Streptococcus pneumoniae* choline-binding protein E interaction with plasminogen/plasmin stimulates migration across the extracellular matrix. *Infect. Immun.* 76: 466–476.
34. Rossi, V., C. Gaboriaud, M. Lacroix, J. Ulrich, J. C. Fontecilla-Camps, J. Gagnon, and G. J. Arlaud. 1995. Structure of the catalytic region of human complement protease C1s: study by chemical cross-linking and three-dimensional homology modeling. *Biochemistry* 34: 7311–7321.
35. Caldwell, E. E., A. M. Andreasen, M. A. Blietz, J. N. Serrahn, V. VanderNoot, Y. Park, G. Yu, R. J. Linhardt, and J. M. Weiler. 1999. Heparin binding and augmentation of C1 inhibitor activity. *Arch. Biochem. Biophys.* 361: 215–222.
36. Medved, L. V., T. F. Busby, and K. C. Ingham. 1989. Calorimetric investigation of the domain structure of human complement Cl-s: reversible unfolding of the short consensus repeat units. *Biochemistry* 28: 5408–5414.
37. Ziccardi, R. J. 1982. A new role for C1-inhibitor in homeostasis: control of activation of the first complement of human complement. *J. Immunol.* 128: 2505–2508.
38. Vivès, R. R., E. Crublet, J. P. Andrieu, J. Gagnon, P. Rousselle, and H. Lortat-Jacob. 2004. A novel strategy for defining critical amino acid residues involved in protein/glycosaminoglycan interactions. *J. Biol. Chem.* 279: 54327–54333.

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

OPEN ACCESS Freely available online

PLOS | ONE

The Serine Protease Domain of MASP-3: Enzymatic Properties and Crystal Structure in Complex with Ecotin

Christine Gaboriaud[1,2,3]*, Rajesh Kumar Gupta[1,2,3¤a], Lydie Martin[1,2,3], Monique Lacroix[1,2,3], Laurence Serre[3,1¤b], Florence Teillet[1,2,3], Gérard J. Arlaud[1,2,3], Véronique Rossi[1,2,3], Nicole M. Thielens[1,2,3]

1 Institut de Biologie Structurale (IBS), Direction des Sciences du Vivant, Commissariat à l'Energie Atomique et aux Energies Alternatives, Grenoble, France, **2** IBS, Centre National de la Recherche Scientifique, Grenoble, France, **3** IBS, Université Grenoble Alpes, Grenoble, France

Abstract

Mannan-binding lectin (MBL), ficolins and collectin-11 are known to associate with three homologous modular proteases, the MBL-Associated Serine Proteases (MASPs). The crystal structures of the catalytic domains of MASP-1 and MASP-2 have been solved, but the structure of the corresponding domain of MASP-3 remains unknown. A link between mutations in the *MASP1/3* gene and the rare autosomal recessive 3MC (Mingarelli, Malpuech, Michels and Carnevale,) syndrome, characterized by various developmental disorders, was discovered recently, revealing an unexpected important role of MASP-3 in early developmental processes. To gain a first insight into the enzymatic and structural properties of MASP-3, a recombinant form of its serine protease (SP) domain was produced and characterized. The amidolytic activity of this domain on fluorescent peptidyl-aminomethylcoumarin substrates was shown to be considerably lower than that of other members of the C1r/C1s/MASP family. The *E. coli* protease inhibitor ecotin bound to the SP domains of MASP-3 and MASP-2, whereas no significant interaction was detected with MASP-1, C1r and C1s. A tetrameric complex comprising an ecotin dimer and two MASP-3 SP domains was isolated and its crystal structure was solved and refined to 3.2 Å. Analysis of the ecotin/MASP-3 interfaces allows a better understanding of the differential reactivity of the C1r/C1s/MASP protease family members towards ecotin, and comparison of the MASP-3 SP domain structure with those of other trypsin-like proteases yields novel hypotheses accounting for its zymogen-like properties *in vitro*.

Citation: Gaboriaud C, Gupta RK, Martin L, Lacroix M, Serre L, et al. (2013) The Serine Protease Domain of MASP-3: Enzymatic Properties and Crystal Structure in Complex with Ecotin. PLoS ONE 8(7): e67962. doi:10.1371/journal.pone.0067962

Editor: Claudio M. Soares, Instituto de Tecnologica Química e Biológica, UNL, Portugal

Received March 27, 2013; **Accepted** May 23, 2013; **Published** July 4, 2013

Copyright: © 2013 Gaboriaud et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the Commissariat à l'Energie Atomique, the Centre National de la Recherche Scientifique, the Université Joseph Fourier (Grenoble, France) and partly by a grant from the French National Research Agency (ANR-09-PIRI-0021). RKG was supported by a post-doctoral scholarship of the Embassy of France in India. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: christine.gaboriaud@ibs.fr

¤a Current address: Institute for Glycomics, Griffith University, Queensland, Australia
¤b Current address: Institut des Neurosciences, Université Grenoble Alpes, Grenoble, France

Introduction

Tightly regulated cascades of proteolytic activations control the complement system, a key player of the host humoral defence, as well as essential physiological processes such as coagulation and fibrinolysis. Activation of the classical and lectin pathways of complement is mediated by homologous modular proteases of the C1r/C1s/MASP family (Fig. 1). This process involves large proteolytic complexes including a recognition molecule of the defence collagens family together with its cognate proteases. The recognition proteins of the lectin pathway identified so far encompass mannan-binding lectin (MBL) [1], collectin 11 (CL-11, CL-K1) [2,3], and ficolins M, L and H (also called ficolin-1, -2 and -3) [4]. Three homologous MBL-associated serine proteases (MASP)-1, -2 and -3, are found associated to these recognition proteins [5], as well as two non-enzymatic components called MAp19 (MBL-associated protein of 19 kDa) or sMAP (small MBL-associated protein) [6,7] and MAp44 (MBL-associated protein of 44 kDa) or MAP-1 (MBL-associated protein 1) [8,9]. MAp19 is an alternative splicing product of the *MASP2* gene whereas MASP-1, MASP-3, and MAp44 are all encoded by the *MASP1/3* gene. MASP-1 and MASP-3 only differ by their serine protease domains and the preceding 15 amino acid residues (Fig. 1). All MBL-associated proteins form homodimers able to interact individually with the lectin pathway recognition proteins through their N-terminal interaction domain.

Clear roles have been recently assigned to MASP-1 and -2 in the activation of the complement lectin pathway, paralleling the roles of C1r and C1s in the classical pathway (Fig. 1, [10,11]). The classical pathway C1 complex comprises a recognition protein C1q and a C1r$_2$C1s$_2$ tetrameric complex of serine proteases. Binding of C1q to suitable targets triggers self-activation of C1r, which in turn activates C1s, the protease responsible for cleavage of C4 and C2, leading to assembly of the C3 convertase C4b2a [12]. In a similar way, MASP-1 appears to be essential for the activation of MASP-2, the latter cleaving C4 and C2 [10,11]. MASP-1 can also activate proenzyme C2 in the C4bC2 complex [13].

A regulatory role has been suggested for MASP-3, MAp44 and MAp19, because of their potential ability to compete with MASP-1 and -2 for interaction with the recognition proteins [8,9,13,14,15]. Initial analysis of recombinant MASP-3 has revealed that it is produced in a proenzyme form, unable to

Publication du Chapitre IV

P14 *- Bally I. et al.,* ***(2009)*** *Identification of the C1q Binding Sites of Human C1r and C1s. A refined Three-dimensional Model of the C1 Complex of Complement.*

P14

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 284, NO. 29, pp. 19340–19348, July 17, 2009
© 2009 by The American Society for Biochemistry and Molecular Biology, Inc. Printed in the U.S.A.

Identification of the C1q-binding Sites of Human C1r and C1s

A REFINED THREE-DIMENSIONAL MODEL OF THE C1 COMPLEX OF COMPLEMENT[*]

Received for publication, February 23, 2009, and in revised form, April 7, 2009 Published, JBC Papers in Press, May 27, 2009, DOI 10.1074/jbc.M109.004473

Isabelle Bally[‡], Véronique Rossi[‡], Thomas Lunardi[‡], Nicole M. Thielens[‡], Christine Gaboriaud[§], and Gérard J. Arlaud[‡1]

From the [‡]Laboratoire d'Enzymologie Moléculaire and the [§]Laboratoire de Cristallographie et Cristallogénèse des Protéines, Institut de Biologie Structurale Jean-Pierre Ebel, CNRS-CEA-Université Joseph Fourier, UMR 5075, 41 rue Jules Horowitz, 38027 Grenoble Cedex 1, France

The C1 complex of complement is assembled from a recognition protein C1q and C1s-C1r-C1r-C1s, a Ca^{2+}-dependent tetramer of two modular proteases C1r and C1s. Resolution of the x-ray structure of the N-terminal CUB_1-epidermal growth factor (EGF) C1s segment has led to a model of the C1q/C1s-C1r-C1r-C1s interaction where the C1q collagen stem binds at the C1r/C1s interface through ionic bonds involving acidic residues contributed by the C1r EGF module (Gregory, L. A., Thielens, N. M., Arlaud, G. J., Fontecilla-Camps, J. C., and Gaboriaud, C. (2003) *J. Biol. Chem.* 278, 32157–32164). To identify the C1q-binding sites of C1s-C1r-C1r-C1s, a series of C1r and C1s mutants was expressed, and the C1q binding ability of the resulting tetramer variants was assessed by surface plasmon resonance. Mutations targeting the Glu^{137}-Glu-Asp^{139} stretch in the C1r EGF module had no effect on C1 assembly, ruling out our previous interaction model. Additional mutations targeting residues expected to participate in the Ca^{2+}-binding sites of the C1r and C1s CUB modules provided evidence for high affinity C1q-binding sites contributed by the C1r CUB_1 and CUB_2 modules and lower affinity sites contributed by C1s CUB_1. All of the sites implicate acidic residues also contributing Ca^{2+} ligands. C1s-C1r-C1r-C1s thus contributes six C1q-binding sites, one per C1q stem. Based on the location of these sites and available structural information, we propose a refined model of C1 assembly where the CUB_1-EGF-CUB_2 interaction domains of C1r and C1s are entirely clustered inside C1q and interact through six binding sites with reactive lysines of the C1q stems. This mechanism is similar to that demonstrated for mannan-binding lectin (MBL)-MBL-associated serine protease and ficolin-MBL-associated serine protease complexes.

The classical pathway of complement, a major component of innate immune defense against pathogens and altered self, is triggered by C1, a 790-kDa Ca^{2+}-dependent complex assembled from a recognition protein C1q and C1s-C1r-C1r-C1s, a tetramer of two modular proteases, C1r and C1s, that respectively mediate activation and proteolytic activity of the complex (1–3). C1q has the overall shape of a bunch of tulips and comprises six heterotrimeric collagen-like triple helices that assemble through their N-terminal moieties to form a "stalk" and then diverge to form individual "stems," each prolonged by a C-terminal globular recognition domain (4). C1r and C1s are homologous modular proteases each comprising, starting from the N-terminal end, a C1r/C1s, sea urchin EGF[2] (uEGF), bone morphogenetic protein (CUB) module (5), an EGF-like module (6), a second CUB module, two complement control protein modules (7), and a serine protease domain. This modular structure is shared by the mannan-binding lectin-associated serine proteases (MASPs), a group of enzymes that associate with mannan-binding lectin (MBL) and the ficolins and thereby trigger activation of the lectin pathway of complement (8).

Assembly of the C1s-C1r-C1r-C1s tetramer involves Ca^{2+}-dependent heterodimeric C1r-C1s interactions between the CUB_1-EGF segments of each protease (9–12). Similarly, MASP-1, MASP-2, MASP-3, and mannan-binding lectin-associated protein 19 (MAp19), an alternative splicing product of the *MASP-2* gene comprising the N-terminal CUB_1-EGF segment of MASP-2, all associate as homodimers through their N-terminal CUB_1-EGF moieties (13–15). The structures of human C1s CUB_1-EGF, human MAp19, human MASP-1/3 CUB_1-EGF-CUB_2, and rat MASP-2 CUB_1-EGF-CUB_2 have been solved by x-ray crystallography (16–19), revealing that these domains all associate as head-to-tail homodimers through a highly conserved interface involving interactions between the CUB_1 module of one monomer and the EGF module of its counterpart. In addition, all CUB modules contained in these structures were found to contain a hitherto unrecognized Ca^{2+}-binding site involving three conserved acidic residues (Glu^{45}, Asp^{53}, and Asp^{98} in C1s), defining a novel CUB module subset diverging from the type originally described in the spermadhesins (20).

Mutagenesis studies have recently established that assembly of the MBL- and ficolin-MASP complexes involves a major electrostatic interaction between two acidic Ca^{2+} ligands from the MASP CUB modules and a conserved lysine located in the collagen fibers of MBL and ficolins (16, 18, 21, 22). In the case of C1, a hypothetical model of the C1q/C1r/C1s interface, involving interaction between acidic residues mainly contributed by the C1r EGF module and unmodified lysine residues also located in the collagen-like stems of C1q, was derived from the

Downloaded from http://www.jbc.org/ by guest on March 9, 2015

---

[*] This work was supported by the Commissariat à l'Energie Atomique, the Centre National de la Recherche Scientifique, and the Université Joseph Fourier, Grenoble.

[1] To whom correspondence should be addressed. Tel.: 33-4-38-78-49-81; Fax: 33-4-38-78-54-94; E-mail: gerard.arlaud@ibs.fr.

[2] The abbreviations used are: EGF, epidermal growth factor; MAp19, mannan-binding lectin-associated protein 19; MASP, mannan-binding lectin-associated serine protease; MBL, mannan-binding lectin.

x-ray structure of the C1s CUB_1-EGF interaction domain (16, 23). The aim of this work was to use site-directed mutagenesis to delineate the sites of C1r and C1s involved in the interaction between C1s-C1r-C1r-C1s and C1q. Our data rule out our previous interaction model and provide evidence that C1 assembly involves the same basic Ca^{2+}-dependent mechanism as demonstrated in the case of MBL-MASP and ficolin-MASP complexes.

EXPERIMENTAL PROCEDURES

Reagents and Proteins—Diisopropyl phosphorofluoridate was purchased from Sigma-Aldrich. C1q was purified from human plasma as described previously (24). The Vent$_R$ DNA polymerase and restriction enzymes were from New England Biolabs (Beverly, MA). The pHC1r3 and pBS-C1s plasmids containing the full-length human C1r and C1s cDNAs (25, 26) were kindly provided by Dr. Agnès Journet (CEA, Grenoble, France) and. Dr. Mario Tosi (University of Rouen, Rouen, France), respectively. Oligonucleotides were purchased from MWG-BIOTECH (Courtaboeuf, France). Antibiotics and molecular biology reagents were obtained from Fermentas (Burlington, Canada). Cellfectin was from Invitrogen. The concentrations of purified proteins were determined using the following absorption coefficients ($A_{1\%,\ 1\ \text{cm}}$ at 280 nm) and molecular weights: C1q, 6.8 and 459,300; recombinant C1r (S637A mutant), 12.4 and 83,373; recombinant C1s (wild type), 14.5 and 77,465; and C1s-C1r-C1r-C1s tetramer assembled from recombinant C1r and C1s, 13.45 and 321,676. The molecular weights of recombinant C1r and C1s were determined by mass spectrometry analysis using the matrix-assisted laser desorption ionization technique as described previously (27).

Expression of C1r and C1s Variants—Recombinant C1r and C1s variants were expressed using a baculovirus/insect cells system. The protocol used for C1s has been previously described in detail (28). In the case of C1r, a DNA fragment encoding the signal peptide plus the full-length mature protein (amino acid residues 1–705) was amplified by PCR using the Vent$_R$ polymerase. The sequences of the sense (5′-CCGGAATTCATGTGGCTCTTGTAC-3′) and antisense (5′-CCCAAGCTTTCAGTCCTCCTCCTCCA-3′) primers introduced an EcoRI restriction site (underlined) at the 5′ end of the PCR product and a HindIII site (underlined) at the 3′ end. The amplified DNA was purified using the gel extraction kit QIAquick (Qiagen) and cloned into the pCR-Script Amp SK(+) intermediate vector (Stratagene) according to the manufacturer's instructions. The fragment was excised by digestion with EcoRI and HindIII and cloned into the corresponding sites of the pFastBac1 baculovirus transfer vector (Invitrogen). The resulting construct was characterized by restriction mapping and checked by double-stranded DNA sequencing (Cogenics, Meylan, France).

The expression plasmids coding for all C1r and C1s mutants were generated using the QuikChange^{TM} XL site-directed mutagenesis kit (Stratagene, La Jolla, CA). The mutagenic oligonucleotides were designed according to the manufacturer's recommendations, and a silent restriction site was introduced in each case for screening of positive clones. The pFastBac1/C1s expression plasmid coding for wild type C1s was used as a template for all C1s variants. The pFastBac1/C1r S637A plasmid coding for the S637A C1r mutant was used as a template for all C1r variants. The sequences of all variants were confirmed by double-stranded DNA sequencing.

The recombinant baculoviruses were generated using the Bac-to-Bac^{TM} system (Invitrogen) as described previously (13). The bacmid DNA was purified using the Qiagen midiprep purification system and used to transfect Sf21 insect cells with cellfectin in Sf900 II SFM medium (Invitrogen) as recommended by the manufacturer. Recombinant virus particles were collected 4 days later and amplified as described by King and Possee (29). High Five cells (1.75×10^7 cells/175-cm^2 tissue culture flask) were infected with the recombinant viruses at a multiplicity of infection of 2 in SF900 II SFM medium at 28 °C for 48 h (C1s variants) or 55 h (C1r variants). The culture supernatants containing recombinant C1r or C1s were collected by centrifugation, supplemented with 1 mM diisopropyl phosphorofluoridate, and stored frozen at −20 °C until use.

Reconstitution and Purification of the C1s-C1r-C1r-C1s Tetramer Variants—The relative C1r or C1s content of each culture supernatant was estimated by SDS-PAGE analysis (30) followed by Coomassie Blue staining and gel scanning. Based on this estimate, C1r- and C1s-containing supernatants were mixed in proportions appropriate to achieve a C1r:C1s ratio of ~1:1. The resulting mixture was dialyzed against 25 mM NaCl, 2 mM $CaCl_2$, 50 mM triethanolamine HCl, pH 7.4, and loaded onto a Q-Sepharose Fast Flow column (Pharmacia) (50 ml) equilibrated in the same buffer. Elution was carried out by applying a linear gradient from 25 to 250 mM NaCl in the same buffer. Fractions containing the C1s-C1r-C1r-C1s tetramer were identified by SDS-PAGE analysis, pooled, and concentrated by ultrafiltration to 0.2–0.4 mg/ml. Further purification was achieved by high pressure gel filtration on either a TSK G3000 SW column (7.5 × 60 cm) (Tosoh Bioscience, Tokyo, Japan) or a Superose 6 10/300 GL column (Amersham Biosciences), each equilibrated in 145 mM NaCl, 2 mM $CaCl_2$, 50 mM triethanolamine HCl, pH 7.4. The purified tetramer was concentrated by ultrafiltration to 0.2 mg/ml and stored at 4 °C until use.

Surface Plasmon Resonance Spectroscopy—Analyses were performed using a BIAcore X instrument (GE Healthcare). C1q was immobilized on the surface of a CM5 sensor chip (GE Healthcare) using the amine coupling chemistry, as described previously (31). Binding of the purified C1s-C1r-C1r-C1s tetramer variants was measured over 15,000 resonance units of immobilized C1q, at a flow rate of 20 μl/min in 145 mM NaCl, 2 mM $CaCl_2$, 50 mM triethanolamine HCl, pH 7.4, containing 0.005% surfactant P20 (GE Healthcare). Each sample was injected in parallel over a surface with immobilized bovine serum albumin for subtraction of the bulk refractive index background. Regeneration of the surfaces was achieved by injection of 10 μl of 145 mM NaCl, 5 mM EDTA, 50 mM triethanolamine HCl, pH 7.4. The data were analyzed by global fitting to a 1:1 Langmuir binding model of the association and dissociation phases simultaneously, using the BIAevaluation 3.1 software (GE Healthcare). The apparent equilibrium dissociation constants (K_D) were calculated from the ratio of the dissociation and association rate constants (k_{off}/k_{on}). Each C1s-C1r-

Downloaded from http://www.jbc.org/ by guest on March 9, 2015

P14

C1r-C1s variant was analyzed at six different concentrations, ranging from 1 to 10 nM.

Modeling of the C1r/C1s CUB$_1$-EGF-CUB$_2$ Heterodimer— Experimentally determined structures were used in the case of C1s CUB$_1$-EGF (16) and C1r EGF (32), whereas homology models were derived for the modules of unknown structure C1r CUB$_1$ and CUB$_2$ and C1s CUB$_2$. C1r CUB$_1$ was built using the scaffold common to its counterparts in C1s, MAp19, and MASP-1/3 (16–18). A few segments (residues 23–25, 62–67, 74–92, and 107–115) correspond to more variable areas, and their conformation is therefore somewhat arbitrary. The N-terminal residues 1–8 were not included because this stretch shows low structural and sequence homology in the available homologous structures. A Ca^{2+}-binding site, as seen in the C1s, MAp19, and MASP-1/3 structures, was introduced in the model.

Assembly of the head-to-tail C1r/C1s CUB$_1$-EGF heterodimer was carried out using the remarkably conserved intermonomer interface seen in the human C1s, human MAp19, rat MASP-2, and human MASP-1/3 homodimeric structures (16–19). Indeed, there are only a few conservative replacements in C1r compared with C1s, with Phe^{9}, Thr^{13}, Phe^{17}, Gln^{44}, Leu^{118}, Gln^{122}, and Val^{152} substituting for Tyr^{5}, Leu^{9}, Tyr^{13}, Thr^{40}, Ala^{109}, Val^{113}, and Ile^{136}, respectively (16). The relative positioning of the monomers observed in the C1s structure was used as a template, and assembly of the C1r/C1s heterodimer was achieved by superimposing the C1r CUB$_1$ model and the NMR-derived C1r EGF structure onto one of the C1s monomers. The root mean square deviations calculated for these superimpositions were 0.9 Å (CUB$_1$, 98 Cα pairs) and 1.04 Å (EGF, 33 Cα pairs). Homology models for C1r and C1s CUB$_2$ were derived from the scaffold common to their counterparts in rat MASP-2 and human MASP-1/3 (18, 19). The most arbitrary conformation corresponds to the two-residue insertion at residues 239 and 240 in C1s. Subtle variations may also occur in the C1r segments 242–248 and 254–257. Initial positioning of the C1r and C1s CUB$_2$ modules in the overall C1r-C1s CUB$_1$-EGF-CUB$_2$ heterodimeric structure was achieved using the plane configurations of human MASP-1/3 and rat MASP-2 as templates. The graphics program O (33) was used throughout the modeling procedure.

RESULTS

The objective of this study was to perform point mutations in various areas of the interaction domains of C1r and C1s and to measure the impact of these mutations on the assembly of the C1 complex with a view to delineate the sites of C1r and C1s responsible for the interaction between the C1s-C1r-C1r-C1s tetramer and C1q. For this purpose, C1r and C1s were expressed in a baculovirus/insect cells system and secreted in the culture supernatants, at concentrations of 1–3 and 5–10 mg/liter, respectively. The wild type C1s DNA sequence was used as a template for all of the mutations. In the case of C1r, in contrast, the template sequence used for subsequent mutations contained a Ser to Ala mutation at the active site residue Ser^{637}. The purpose of this modification was to stabilize all C1r variants in the proenzyme form and to prevent spontaneous activation during expression and upon subsequent interaction of the reconstituted C1s-C1r-C1r-C1s tetramer with C1q.

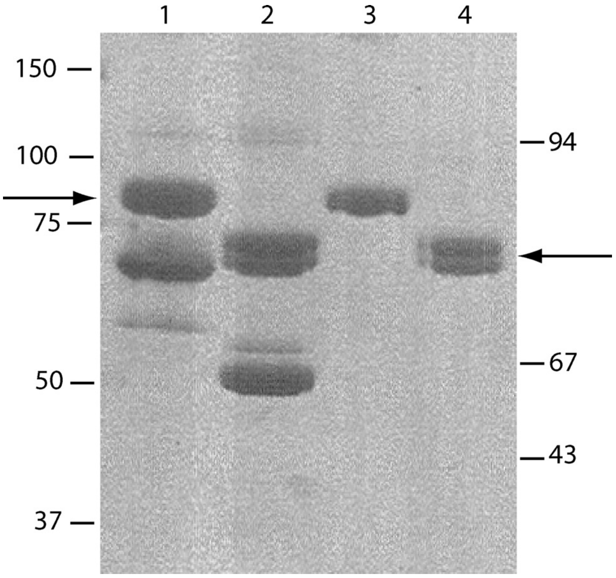
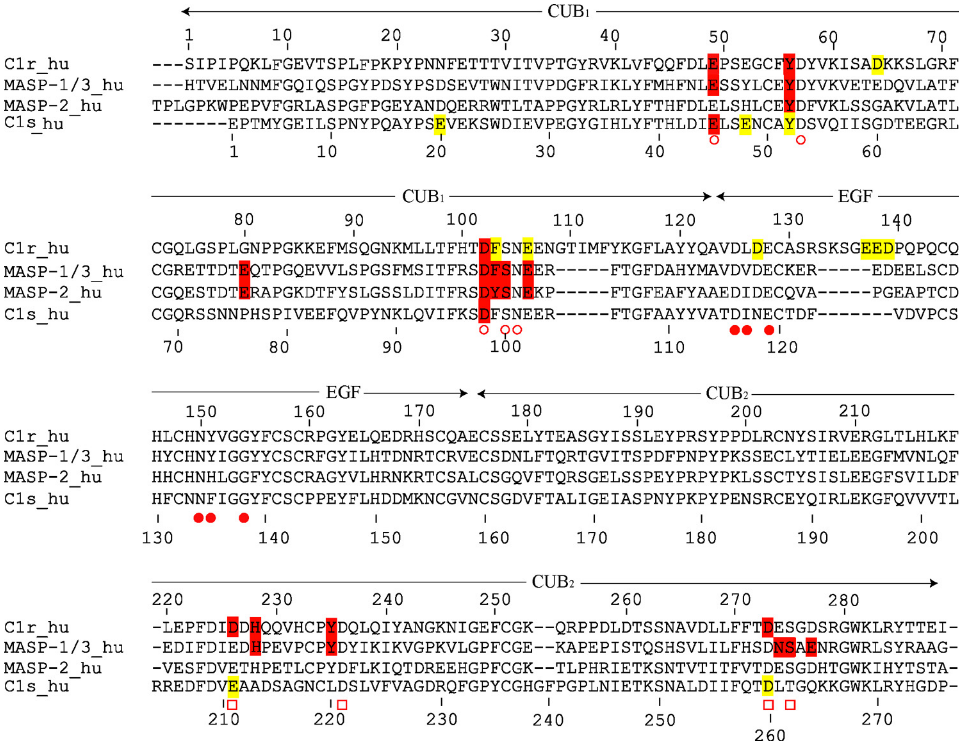


FIGURE 1. **Purification of the recombinant C1s-C1r-C1r-C1s tetramer.** The tetramer was reconstituted from wild type C1s and the S637A C1r mutant. SDS-PAGE analysis was performed after ion exchange chromatography (*lanes 1* and *2*) and after high pressure gel permeation (*lanes 3* and *4*), under reducing (*lanes 1* and *3*), and nonreducing conditions (*lanes 2* and *4*). The *arrows* indicate the position of recombinant C1r and C1s under both conditions. The positions of reduced and unreduced standard proteins are indicated on the *left* and *right* sides of the figure, respectively.

We initially attempted to express and purify recombinant C1r and C1s separately, with a view to subsequently reconstitute the C1s-C1r-C1r-C1s tetramer variants. A major drawback of this protocol arose from the relatively poor expression yield of recombinant C1r, combined with a considerable loss of protein during ion exchange purification. This did not allow us to produce purified C1r variants in sufficient amounts to perform subsequent analyses by surface plasmon resonance spectroscopy. To circumvent this problem, an alternative strategy was used. The culture supernatants containing the C1r and C1s variants were first combined in the presence of Ca^{2+} ions under appropriate proportions to achieve a C1r:C1s molar ratio of ~1:1 and thereby reconstitute the C1s-C1r-C1r-C1s tetramer. The resulting mixture was then submitted to ion exchange chromatography on a Q-Sepharose Fast Flow column as described under "Experimental Procedures." SDS-PAGE analysis of the elution profile provided evidence that C1r and C1s eluted as a single peak containing equimolar amounts of each protein (data not shown), confirming formation of the C1s-C1r-C1r-C1s tetramer. Excess C1s was observed occasionally and eluted in a peak immediately following the tetramer. All of the C1r/C1s combinations tested yielded similar elution profiles, providing a first indication that none of the mutations performed in C1r or C1s had a significant impact on the assembly of the tetramer. The tetramer pool from the ion exchange chromatography column was always contaminated by 40–50% bovine serum albumin, originating from the fetal calf serum contained in the culture medium and partially overlapping with the tetramer (Fig. 1, *lanes 1* and *2*). Further purification was achieved by high pressure gel permeation,

Downloaded from http://www.jbc.org/ by guest on March 9, 2015

P14



```
                         CUB1
             1        10        20        30        40        50        60        70
C1r_hu       ---SIPIPQKLFGEVTSPLFPKPYPNNFETTTVITVPTGYRVKLVFQQFDLEPSEGCFYDYVKISADKKSLGRF
MASP-1/3_hu  ---HTVELNNMFGQIQSPGYPDSYPSDSEVTWNITVPDGFRIKLYFMHFNLESSYLCEYDYVKVETEDQVLATF
MASP-2_hu    TPLGPKWPEPVFGRLASPGFPGEYANDQERRWTLTAPPGYRLRLYFTHFDLELSHLCEYDFVKLSSGAKVLATL
C1s_hu       -------EPTMYGEILSPNYPQAYPSEVEKSWDIEVPEGYGIHLYFTHLDIELSENCAYDSVQIISGDTEEGRL
                    1        10        20        30        40        50        60

                         CUB1                                       EGF
                  80        90       100       110       120       130       140
C1r_hu       CGQLGSPLGNPPGKKEFMSQGNKMLLTFHTDFSNEENGTIMFYKGFLAYYQAVDLDECASRSKSGEEDPQPQCQ
MASP-1/3_hu  CGRETTDTEQTPGQEVVLSPGSFMSITFRSDFSNEER-----FTGFDAHYMAVDVDECKER-----EDEELSCD
MASP-2_hu    CGQESTDTERAPGKDTFYSLGSSLDITFRSDYSNEKP-----FTGFEAFYAAEDIDECQVA-----PGEAPTCD
C1s_hu       CGQRSSNNPHSPIVEEFQVPYNKLQVIFKSDFSNEER-----FTGFAAYYVATDINECTDF-------VDVPCS
              70        80        90       100            110        120

                  EGF                          CUB2
              150       160       170       180       190       200       210
C1r_hu       HLCHNYVGGYFCSCRPGYELQEDRHSCQAECSSELYTEASGYISSLEYPRSYPPDLRCNYSIRVERGLTLHLKF
MASP-1/3_hu  HYCHNYIGGYYCSCRFGYILHTDNRTCRVECSDNLFTQRTGVITSPDFPNPYPKSSECLYTIELEEGFMVNLQF
MASP-2_hu    HHCHNHLGGFYCSCRAGYVLHRNKRTCSALCSGQVFTQRSGELSSPEYPRPYPKLSSCTYSISLEEGFSVILDF
C1s_hu       HFCNNFIGGYFCSCPPEYFLHDDMKNCGVNCSGDVFTALIGEIASPNYPKPYPENSRCEYQIRLEKGFQVVVTL
             130       140       150       160       170       180       190       200

                                       CUB2
             220       230       240       250          260       270       280
C1r_hu       -LEPFDIDDHQQVHCPYDQLQIYANGKNIGEFCGK--QRPPDLDTSSNAVDLLFFTDESGDSRGWKLRYTTEI-
MASP-1/3_hu  -EDIFDIEDHPEVPCPYDYIKIKVGPKVLGPFCGE--KAPEPISTQSHSVLILFHSDNSAENRGWRLSYRAAG-
MASP-2_hu    -VESFDVETHPETLCPYDFLKIQTDREEHGPFCGK--TLPHRIETKSNTVTITFVTDESGDHTGWKIHYTSTA-
C1s_hu       RREDFDVEAADSAGNCLDSLVFVAGDRQFGPYCGHGFPGPLNIETKSNALDIIFQTDLTGQKKGWKLRYHGDP-
                  210       220        230       240       250        270
                                                                 260
```

FIGURE 2. **Sequence alignment of selected CUB$_1$-EGF-CUB$_2$ segments.** The amino acid numberings of both C1r (*top*) and C1s (*bottom*) are shown, and the approximate domain boundaries are indicated. Residues known to contribute Ca^{2+} ligands in C1s, MAp19, and MASP-1/3 are marked with *open circles* (CUB$_1$ site), *closed circles* (EGF site), and *open squares* (CUB$_2$ site). C1r and C1s residues interacting with C1q, and MASP residues interacting with MBL and ficolins are colored *red*. Mutations in C1r and C1s having little or no effect on C1 assembly are colored *yellow*. The implication of C1s Tyr^{52} in C1q binding cannot be excluded.

TABLE 1

Kinetic and dissociation constants for the interaction between C1s-C1r-C1r-C1s variants and immobilized C1q

Shown are the effects of mutations in the C1r EGF module.

Location of mutation	Tetramer variant	k_{on}	k_{off}	K_D (nM)	K_D/K_D,wt	χ^2
		$M^{-1}s^{-1}$	s^{-1}	nM		
	Wild type[a]	0.82×10^6	1.21×10^{-3}	1.47 ± 0.6[b]	1.0	2.8
EGF	D127N	0.89×10^6	1.29×10^{-3}	1.45	1.4	1.9
	E137A	0.73×10^6	1.47×10^{-3}	2.03	1.3	5.1
	E138A	1.09×10^6	2.14×10^{-3}	1.97	1.2	4.8
	D139A	0.68×10^6	1.17×10^{-3}	1.72	0.9	1.3
	Δ137–139	1.05×10^6	1.42×10^{-3}	1.35		4.4
CUB$_1$	D64N	0.83×10^6	1.27×10^{-3}	1.53	1.0	1.7

[a] Tetramer assembled from wild type C1s and the S637A C1r mutant.

[b] Mean value determined from three separate experiments.

allowing complete elimination of bovine serum albumin and providing further evidence that, in the presence of Ca^{2+} ions, all of the C1r/C1s combinations tested yielded a peak co-eluting with the "native" tetramer reconstituted from wild type C1s and the S637A C1r mutant (data not shown). As expected, gel filtration analysis in the presence of EDTA resulted in disruption of the tetramer, yielding two peaks corresponding to monomeric C1s and the C1r-C1r dimer. SDS-PAGE analysis of the purified tetramer preparation indicated that C1r and C1s each behaved as single-chain proteins under nonreducing and reducing conditions, indicative of their proenzyme state. Both C1r and C1s had similar apparent molecular masses of 85–90 kDa under reducing conditions (Fig. 1, *lanes 3* and *4*).

Mutations in the C1r EGF Module Have No Effect on C1 Assembly—A series of C1s-C1r-C1r-C1s variants were produced, and their ability to associate with C1q was analyzed by surface plasmon resonance spectroscopy, using the tetramer variants as soluble ligands and immobilized C1q. To test the validity of the interaction model derived from the x-ray structure of the C1s interaction domain (16), we initially targeted acidic residues located in the C1r EGF module. Residues Glu^{137}, Glu^{138}, and Asp^{139}, which form an acidic cluster in a large insertion loop of the C1r EGF module (Fig. 2), were each individually mutated to alanine, and the whole Glu^{137}–Asp^{139} sequence stretch was deleted. As listed in Table 1, none of these modifications significantly decreased the ability of the resulting C1s-C1r-C1r-C1s tetramer variants to associate with C1q, with K_D ratios relative to the native tetramer ranging from 0.9 to 1.4. Two other acidic residues of C1r also considered as possible C1q ligands (16), Asp^{127} (in the EGF module) and Asp^{64} (in the CUB$_1$ module), were also individually mutated to asparagine. Again, these mutations did not alter the C1q binding ability of C1s-C1r-C1r-C1s. These results were therefore clearly not compatible with a major contribution of the C1r EGF module in the tetramer/C1q interaction, ruling out our previous model of C1 assembly essentially based on this assumption (16).

Mutations in the Ca^{2+}-binding Sites of the C1r and C1s CUB Modules Reveal Their Implication in C1q Binding—Recent studies on human MAp19 (17) and MASP-3 (18) have provided experimental evidence that both proteins associate with MBL and the ficolins through residues involved in the Ca^{2+}-binding sites contained in their CUB modules. Given the strong structural homology between C1r, C1s, and the MASPs, this prompted us to target residues expected to participate in the Ca^{2+}-binding sites of the C1r and C1s CUB$_1$ and CUB$_2$ modules. We particularly focused our attention on the residues homologous to two of the three conserved acidic residues known to coordinate Ca^{2+} in C1s CUB$_1$, Glu^{45} and Asp^{98} (Fig. 2). The reason for this choice is that, unlike the third Ca^{2+} ligand (homologous to Asp^{53} in C1s CUB$_1$), these residues are always positioned on the outer part of the Ca^{2+}-binding site, as shown by the x-ray crystallography analyses performed on C1s (Fig. 3), MAp19 and MASP-1/3 (16–18), and are therefore available for contributing ionic interactions with a protein ligand.

Mutations to alanine of the corresponding residues Glu^{49} and Asp^{102} in the C1r CUB$_1$ module each virtually abolished binding of the C1s-C1r-C1r-C1s tetramer to C1q (Fig. 4 and

Downloaded from http://www.jbc.org/ by guest on March 9, 2015

P14

Table 2), providing a clear indication of their implication in the interaction. Mutation of the neighboring residue Tyr^{56} also had a marked inhibitory effect, in keeping with the observation that replacement of the homologous residues Tyr^{59} and Tyr^{56} of human MAp19 and MASP-1/3, respectively, inhibit their interaction with MBL and the ficolins (17, 18). In contrast, mutation of residues Phe^{103} and Glu^{106} only had a slight or no significant inhibitory effect, at variance with previous data obtained upon replacement of the corresponding residues of MAp19 and MASP-1/3. We next targeted two of the acidic residues thought to coordinate Ca^{2+} in the C1r CUB$_2$ module, namely Asp^{226} and Asp^{273}. Again, mutation of each residue to alanine strongly inhibited the ability of the resulting C1s-C1r-C1r-C1s variants to bind C1q (Fig. 4 and Table 2). Comparable inhibitory effects were observed upon mutation of the neighboring residues His^{228} and Tyr^{235}, as observed for the homologous residues His^{218} and Tyr^{225} of MASP-1/3 (18).

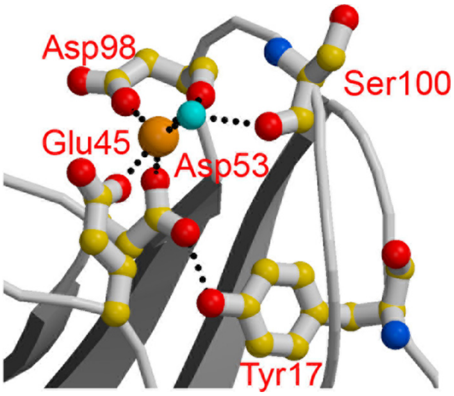


FIGURE 3. **Structure of the Ca^{2+}-binding site of the C1s CUB$_1$ module.** Oxygen atoms are shown in *red*, and nitrogen atoms are in *blue*. The *light blue* sphere represents a water molecule. Ionic bonds and hydrogen bonds are represented by *dotted lines* (adapted from Gregory *et al.* (16)).

A similar strategy was applied to C1s. In the CUB$_1$ module, mutations of the Ca^{2+}-binding residues Glu^{45} and Asp^{98} each reproducibly yielded slight inhibitory effects (Fig. 4 and Table 2), suggesting a minor contribution to the tetramer/C1q interaction. Further support for this hypothesis came from the observation that double C1s-C1r-C1r-C1s mutants with an D273A mutation in C1r CUB$_1$ and either a E45A or an D98Ala mutation in C1s CUB$_1$ completely lost their C1q binding ability. In contrast, the single D273A C1r variant retained low, yet measurable binding activity (Table 2). Mutations at Glu^{20} and Glu^{48}, two acidic residues exposed in the C1s CUB$_1$ structure (16), had no significant inhibitory effect. Mutation at Tyr^{52} yielded similar results (Table 2), although a contribution of this residue cannot be entirely excluded given the overall low impact of the mutations performed in the C1s CUB$_1$ module. In contrast, the mutations to alanine of the Ca^{2+}-binding residues Glu^{211} and Asp^{260} in C1s CUB$_2$ clearly had no effect on the binding of C1s-C1r-C1r-C1s to C1q (Table 2), ruling out an implication of this module in the interaction.

As illustrated on the three-dimensional model of the C1r/C1s CUB$_1$-EGF-CUB$_2$ heterodimer shown in Fig. 5, the above results provided strong experimental evidence for the occurrence of high affinity C1q-binding sites contributed by the CUB$_1$ and CUB$_2$ modules of C1r, and lower affinity sites contributed by the C1s CUB$_1$ module, each of these sites implicating Ca^{2+}-binding residues and residues located in the vicinity of the Ca^{2+}-binding site. In contrast, the mutagenesis data clearly demonstrated that the corresponding site in the C1s CUB$_2$ module, as well as the C1r EGF module, have no significant role in the interaction with C1q.

DISCUSSION

The aim of this study was to delineate in C1r and C1s the sites mediating interaction between the C1s-C1r-C1r-C1s tetramer

Downloaded from http://www.jbc.org/ by guest on March 9, 2015

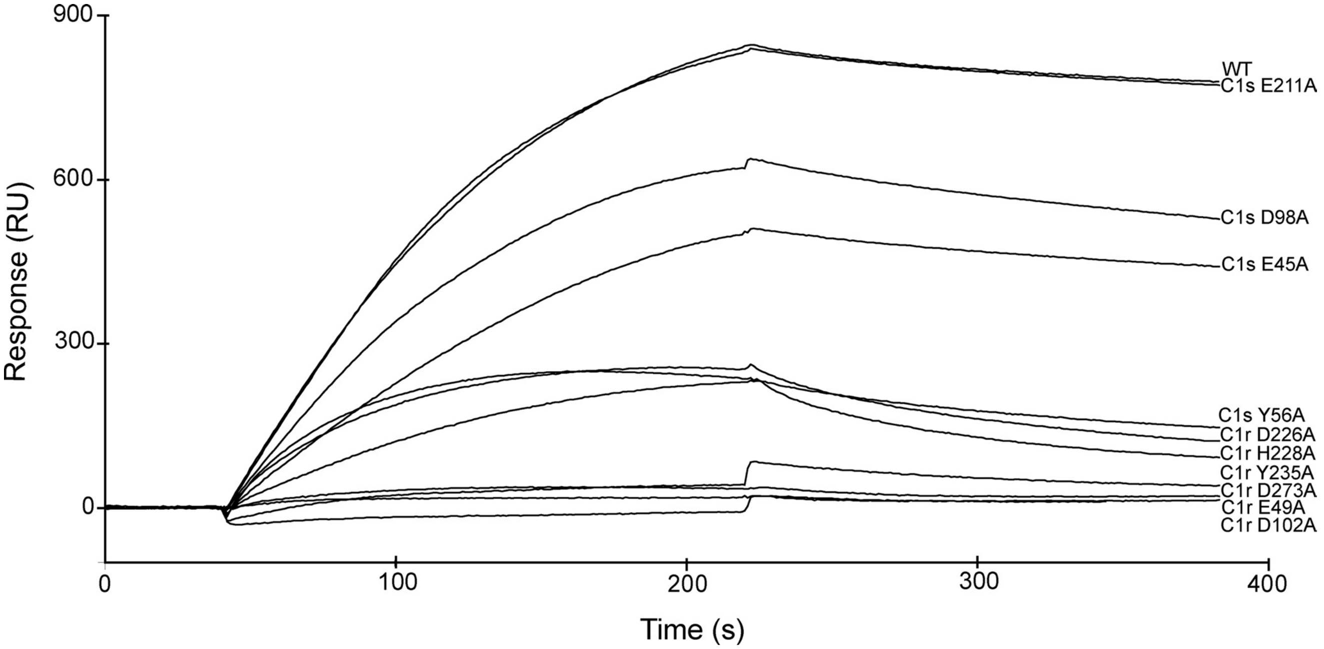


FIGURE 4. **Analysis by surface plasmon resonance spectroscopy of the interaction between selected C1s-C1r-C1r-C1s variants and immobilized C1q.** C1q was immobilized on the sensor chip as described under "Experimental Procedures." The native C1s-C1r-C1r-C1s tetramer reconstituted from wild type C1s and the S637A C1r mutant, and selected variants with mutations in either C1r or C1s were injected at a concentration of 10 nM.

TABLE 2

Kinetic and dissociation constants for the interaction between C1s-C1r-C1r-C1s variants and immobilized C1q

Shown are the effects of mutations in the Ca2+-binding sites of the C1r and C1s CUB modules.

Location of mutation(s)	Tetramer variant	k_{on}	k_{off}	K_D	K_D/ K_D,wt	χ^2
		$M^{-1} \cdot s^{-1}$	s^{-1}	nM		
	Wild type[a]	0.98×10^6	0.64×10^{-3}	0.68 ± 0.4[b]		3.2
C1r CUB_1	E49A	0.12×10^6	9.55×10^{-3}	82.4	121	3.9
	Y56A	0.61×10^6	3.44×10^{-3}	5.64	8.3	1.1
	D102A	ND[c]	ND	ND	ND	
	F103A	0.75×10^6	0.95×10^{-3}	1.26	1.8	0.9
	E106A	1.14×10^6	1.18×10^{-3}	1.03	1.5	3.9
C1r CUB_2	D226A	0.56×10^6	4.16×10^{-3}	7.45	10.9	1.2
	H228A	0.97×10^6	4.86×10^{-3}	5.02	7.4	1.8
	Y235A	0.85×10^6	24.3×10^{-3}	28.6	42	2.9
	D273A	0.15×10^6	6.12×10^{-3}	40.8	60	4.3
C1s CUB_1	E20A	0.61×10^6	0.71×10^{-3}	1.15	1.7	3.1
	E45A	0.35×10^6	1.34×10^{-3}	3.85	5.7	3.4
	E48A	0.60×10^6	0.80×10^{-3}	1.33	1.9	3.7
	Y52A	0.90×10^6	1.40×10^{-3}	1.55	2.3	4.2
	D98A	1.00×10^6	1.70×10^{-3}	1.67	2.5	4.2
C1s CUB_2	E211A	0.77×10^6	0.47×10^{-3}	0.61	0.9	2.2
	D260A	1.06×10^6	0.56×10^{-3}	0.53	0.8	3.9
C1r CUB_1/	D273A/D98A	ND	ND	ND	ND	
C1s CUB_2	D273A/E45A	ND	ND	ND	ND	

[a] Tetramer assembled from wild type C1s and the S637A C1r mutant.

[b] Mean value determined from three separate experiments.

[c] ND, value not measurable due to the weakness of the binding.

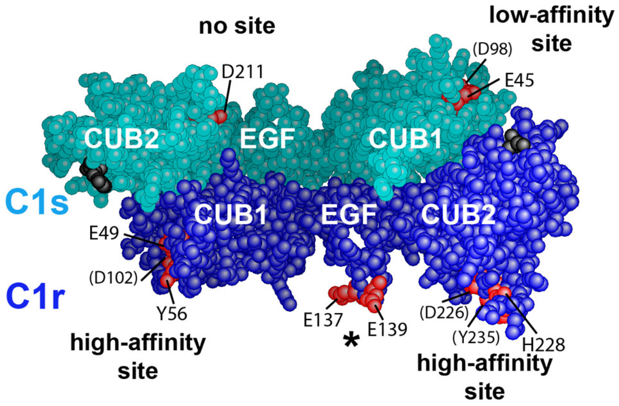


FIGURE 5. **Three-dimensional space-filling representation of the C1r/C1s CUB_1-EGF-CUB_2 heterodimer.** The head-to-tail C1r/C1s CUB_1-EGF-CUB_2 heterodimeric structure was built as described under "Experimental Procedures." C1r and C1s domains are colored *dark blue* and *turquoise*, respectively. The Ca^{2+}-binding sites of the C1r and C1s CUB modules are marked by residues Glu^{49}, Tyr^{56}, and Asp^{102} (C1r CUB_1); Asp^{226}, His^{228}, and Tyr^{235} (C1r CUB_2); Glu^{45} and Asp^{98} (C1s CUB_1); and Glu^{211} (C1s CUB_2), shown in *red*. Residues barely visible are indicated in *parentheses*. The acidic cluster Glu^{137}-Glu-Asp^{139} in the C1r EGF module (indicated by an *asterisk*) is also shown in *red*. The C-terminal ends of the C1r and C1s CUB_2 modules are shown in *black*.

and C1q. For this purpose, C1r and C1s were each expressed in a baculovirus/insect cells system, and a series of mutations were carried out in different areas of their interaction domains with a view to measure the ability of the resulting tetramer variants to associate with C1q. In agreement with earlier observations (34), expression of C1r in insect cells resulted in poor yields compared with C1s (28), precluding purification of sufficient amounts of protein for subsequent reconstitution of the tetramer and analysis of its interaction properties. This problem could be overcome in a simple way, by first combining the C1r- and C1s-containing culture supernatants and then purifying the reconstituted C1s-C1r-C1r-C1s tetramer variants by ion exchange and gel permeation chromatography.

Twenty-two mutations were performed in either C1r or C1s, targeting residues located in their CUB_1-EGF-CUB_2 interaction domains while avoiding those expected to mediate heterodimeric C1r-C1s interaction in the C1s-C1r-C1r-C1s tetramer, as predicted from the homologous homodimeric x-ray structures of human C1s CUB_1-EGF (16), rat MASP-2 CUB_1-EGF-CUB_2 (19), human MAp19 (17), and human MASP-1/3 CUB_1-EGF-CUB_2 (18). In agreement with these structures, none of the mutations performed in this study had a significant impact on the assembly of the C1s-C1r-C1r-C1s tetramer, as shown by both ion exchange chromatography and gel filtration analyses. Thus, the inhibitions of C1 assembly caused by mutations in C1r or C1s, as observed in this study, solely reflect direct effects on the interactions between C1s-C1r-C1r-C1s and C1q.

A first lesson from this work is that none of the mutations targeting the sequence stretch Glu^{137}-Glu-Asp^{139} in the C1r EGF module had any detectable effect on C1 assembly. This demonstrates that, contrary to our earlier hypothesis, this acidic cluster has no significant contribution to the interaction with C1q. These data are therefore not compatible with our previous model of C1 assembly (16, 23), where the collagen triple helix of C1q was proposed to bind at the interface of the C1r/C1s CUB_1-EGF heterodimer, through ionic interactions with acidic residues mainly contributed by the C1r EGF module. Consistent with these findings, the second round of mutations targeting residues expected to participate in the Ca^{2+}-binding sites of the C1r and C1s CUB modules provided clear experimental evidence for the occurrence of C1q-binding sites contributed by C1r CUB_1 and CUB_2 and C1s CUB_1; each of these sites involving residues also engaged in Ca^{2+} coordination or located in close vicinity of the Ca^{2+}-binding site. These results are strikingly similar to those obtained previously in the case of MAp19 and MASP-3 (17, 18) and clearly indicate that C1s-C1r-C1r-C1s and the MASPs make use of the same basic mechanism to associate with their partner proteins, C1q, MBL, and the ficolins, respectively. These observations provide indirect evidence that, as already demonstrated by x-ray crystallography in the case of C1s CUB_1 (16), both C1r CUB modules contain a Ca^{2+}-binding site. If this hypothesis is correct, then this would mean that substitution of Asp for Glu at position 226 in C1r CUB_2 (Fig. 2) is tolerated by the Ca^{2+}-binding site. C1s CUB_2 possesses the canonical Glu-Asp-Asp consensus sequence (Fig. 2), and therefore the fact that it is not involved in the interaction with C1q is likely connected to structural and/or functional constraints at the level of the whole C1 complex rather than to a lack of Ca^{2+}-binding site. As proposed for the MASPs (18), a plausible hypothesis is that the two outer acidic Ca^{2+} ligands in C1r CUB_1 and CUB_2 and in C1s CUB_1 coordinate Ca^{2+} on one side and mediate interaction with C1q on the other side. However, there are other possibilities, including displacement of the Ca^{2+} ion upon interaction with C1q. The precise mechanism involved in this interaction therefore remains to be fully elucidated.

In addition, as observed in the MASPs, other interactions are contributed by residues located in the vicinity of the Ca^{2+}-

Downloaded from http://www.jbc.org/ by guest on March 9, 2015

binding sites, such as Tyr^{56} in C1r CUB$_1$ and His^{228} and Tyr^{235} in C1r CUB$_2$ (Table 2). There are, however, subtle differences between the two types of complexes, as exemplified by the fact that, unlike their counterparts in the MASPs (17, 18), Phe^{103} and Glu^{106} in C1r CUB$_1$ do not seem to play a major role in the interaction with C1q (Table 2). Notwithstanding these differences, the striking homology, in terms of assembly, between C1 and the MBL/ficolin-MASP complexes lends further credit to the hypothesis that, as demonstrated by site-directed mutagenesis for MBL and L- and H-ficolins (21, 22), interaction in C1 also involves a conserved Lys residue from C1q. As discussed previously (16, 23), the unmodified Lys residues at positions A59, B61, and C58, located about half-way along the C1q collagen stems, are likely candidates for this function.

From a general standpoint, our finding that C1s-C1r-C1r-C1s associates with C1q through sites contributed by both C1r and C1s is fully consistent with a series of earlier data (reviewed in Ref. 35). In the same way, the fact that, unlike C1s CUB$_1$, C1s CUB$_2$ does not contribute a C1q-binding site, is in complete agreement with previous reports showing that the N-terminal CUB$_1$-EGF segment of C1s is sufficient to promote interaction of C1r with C1q (11, 36). Actually, the lack of interaction between C1s CUB$_2$ and C1q is expected to yield more flexibility at the CUB$_2$-CCP$_1$ interface and hence more freedom to the CCP$_1$-CCP$_2$-SP C1s catalytic domain, consistent with its role in the proteolytic activity of C1 (23).

Our mutagenesis data provide evidence that each C1r/C1s CUB$_1$-EGF-CUB$_2$ heterodimer contains two high affinity C1q-binding sites centered on residues Glu^{49}/Asp^{102} (C1r CUB$_1$) and Asp^{226}/Asp^{273} (C1r CUB$_2$), and one lower affinity site involving residues Glu^{45}/Asp^{98} of C1s CUB$_1$. This heterodimer occurs twice in C1s-C1r-C1r-C1s, and therefore the whole tetramer contributes six C1q-binding sites, which is one per each C1q collagen-like stem. Based on the location of these sites within the C1r/C1s CUB$_1$-EGF-CUB$_2$ assembly, we have explored different ways of positioning the two heterodimers with respect to C1q to identify the configuration(s) fulfilling the following constraints: (i) Both heterodimers should interact with C1q in the same way, implying that they be positioned symmetrically about the C1q axis. (ii) Each site contributed by C1r and C1s should make a contact with the corresponding site on a C1q stem, but interaction should be stronger where mutations have the strongest inhibitory effects. Conversely, areas where mutations are ineffective should not be close to a reactive lysine on a C1q stem. Additional constraints arise from the fact that the CUB$_2$ modules of C1r and C1s have to connect with their respective catalytic domains. This question is particularly crucial in the case of C1r, because its catalytic domains are expected to lie inside the C1q cone, underneath the C1r/C1s interaction domains (23). In this regard, if one considers the "plane" configuration described for the homologous MASP-1/3 CUB$_1$-EGF-CUB$_2$ homodimer (18), the most favorable positioning is clearly the one where the C-terminal ends of the CUB$_2$ modules are oriented toward the C1q globular domains.

We tested different configurations and found that none of those involving location of the C1r/C1s CUB$_1$-EGF-CUB$_2$ heterodimers on the outside part of the C1q stems, as proposed in our previous model (16, 23), fulfilled the interaction constraints defined above. We next tested the hypothesis that both heterodimers are located inside the cone defined by the C1q stems. In this case, there are only two possible ways of assembling the heterodimers in a head-to-tail manner with respect to the inner molecules, with either CUB$_1$-CUB$_2$ contacts (not shown) or CUB$_1$-EGF contacts (Fig. 6*A*). In the former configuration, the C1q-binding sites cannot be distributed in a radial fashion because the assembly is too elongated, and the distance between the C-terminal ends of the outer molecules is too large to fit inside the C1q cone. In the other alternative, the configuration is more favorable because the assembly is more compact, and provided that

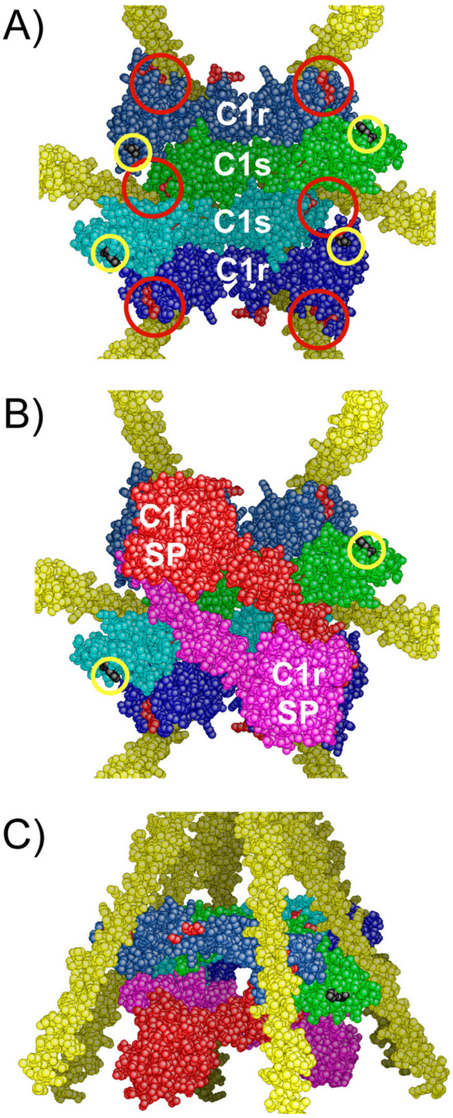


FIGURE 6. **A refined three-dimensional model of C1 assembly.** *A*, bottom view of the assembly between the C1r/C1s CUB$_1$-EGF-CUB$_2$ interaction domains and the C1q collagen stems. C1r and C1s domains are colored *dark blue* and *light blue*, respectively. The approximate positioning of the six C1q-binding sites contributed by C1r and C1s is indicated by *red circles*. The C-terminal ends of the C1r and C1s CUB$_2$ modules are marked by *yellow circles*. *B*, bottom view of the assembly featuring the CCP$_1$-CCP$_2$-SP catalytic domains of C1r, shown in *red* and *pink*. The C1s catalytic domains, emerging from the C-terminal end of both C1s CUB$_2$ modules (*yellow circles*), are not shown for clarity. *SP*, serine protease domain. *C*, side view of the assembly shown in *B*.

Downloaded from http://www.jbc.org/ by guest on March 9, 2015

C1s is the inner molecule, then the six binding sites are distributed radially at positions roughly appropriate to make individual contacts with the target lysine residues on the C1q collagen-like stems (Fig. 6*A*). Consistent with this configuration, C1r is unlikely to be the inner molecule because this would conceal the C1q-binding site contributed by its CUB_2 module.

In this configuration, however, if the C1r-C1s CUB1-EGF-CUB2 heterodimers have the planar orientation seen in the corresponding MASP-1/3 x-ray structure (18), then some of the ineffective mutations are not far from the putative binding sites on the C1q stems. In addition, there would be steric clashes between the C1s CUB_2 modules and the C1q stems, and the C termini of the outer molecules (C1r) are still too far away from each other. Nevertheless, the above problems can be overcome by orienting the CUB_2 modules toward the C1q globular domains. In other words, the model becomes plausible if the C1r-C1s CUB_1-EGF-CUB_2 heterodimers deviate from the planar configuration observed in the MASP-1/3 structure (18). Such changes are indeed feasible because (i) only three of the four potential CUB binding sites are functional, and therefore all four sites do not need to lie in the same plane as observed in MASP-1/3, and (ii) the interaction strength varies from one binding site to the other. Thus, a more compact model can be readily obtained by twisting the EGF-CUB_2 interfaces and tilting the two CUB_1-EGF-CUB_2 heterodimers about the C1q axis to bring them closer to each other in their central part. These modifications also modulate the relative positioning of the areas where mutations are ineffective.

In the resulting nonplanar model (Fig. 6*A*), the interaction domains of C1r and C1s are entirely located inside the C1q cone, in sharp contrast with our previous hypothesis (23). The C1s CUB_1-EGF-CUB_2 moieties occupy the inner part of the assembly, the interface between them appearing to be sufficiently flat to allow the sliding movements expected to take place upon activation (23). Compared with our previous C1 model (23), the remainder of the complex is essentially unchanged; the C1r catalytic domains occupy the lower part of the C1q cone (Fig. 6, *B* and *C*), beneath the interaction domains, the C1s serine protease domains being likely partly located inside the C1q cone, at least transiently during the activation process. Thus, in this refined version of the C1 complex, most of the C1r and C1s domains are located inside C1q, in full agreement with earlier analyses of the C1 complex by electron microscopy (37) and neutron scattering (38). From a functional standpoint, a prominent feature of this model is that each of the six C1q stems is engaged in the interaction with the tetramer. Thus, distortion of any of the C1q stems upon binding of C1 to a target surface is expected to generate part of the mechanical stress thought to trigger C1r activation (39).

From a general standpoint, it should be emphasized that the C1s-C1r-C1r-C1s moiety of C1 undergoes large changes upon C1 assembly and activation. The model arising from this study must therefore be viewed as a ground state representation of a sophisticated machinery undergoing multiple conformational changes.

Acknowledgments—We are grateful to Guy Schoehn and Wai-Li Ling for performing electron microscopy analyses and for helpful discussions.

REFERENCES

1. Cooper, N. R. (1985) *Adv. Immunol.* **37,** 151–216
2. Arlaud, G. J., Colomb, M. G., and Gagnon, J. (1987) *Immunol. Today* **8,** 106–111
3. Arlaud, G. J., Gaboriaud, C., Thielens, N. M., Rossi, V., Bersch, B., Hernandez, J. F., and Fontecilla-Camps, J. C. (2001) *Immunol. Rev.* **180,** 136–145
4. Kishore, U., and Reid, K. B. M. (2000) *Immunopharmacology* **49,** 159–170
5. Bork, P., and Beckmann, G. (1993) *J. Mol. Biol.* **231,** 539–545
6. Campbell, I. D., and Bork, P. (1993) *Curr. Opin. Struct. Biol.* **3,** 385–392
7. Reid, K. B. M., Bentley, D. R., Campbell, R. D., Chung, L. P., Sim, R. B., Kristensen, T., and Tack, B. F. (1986) *Immunol. Today* **7,** 230–234
8. Fujita, T. (2002) *Nat. Rev. Immunol.* **2,** 346–353
9. Thielens, N. M., Aude, C. A., Lacroix, M. B., Gagnon, J., and Arlaud, G. J. (1990) *J. Biol. Chem.* **265,** 14469–14475
10. Busby, T. F., and Ingham, K. C. (1990) *Biochemistry* **29,** 4613–4618
11. Tsai, S. W., Poon, P. H., and Schumaker, V. N. (1997) *Mol. Immunol.* **34,** 1273–1280
12. Thielens, N. M., Enrie, K., Lacroix, M., Jaquinod, M., Hernandez, J. F., Esser, A. F., and Arlaud, G. J. (1999) *J. Biol. Chem.* **274,** 9149–9159
13. Thielens, N. M., Cseh, S., Thiel, S., Vorup-Jensen, T., Rossi, V., Jensenius, J. C., and Arlaud, G. J. (2001) *J. Immunol.* **166,** 5068–5077
14. Zundel, S., Cseh, S., Lacroix, M., Dahl, M. R., Matsushita, M., Andrieu, J. P., Schwaeble, W. J., Jensenius, J. C., Fujita, T., Arlaud, G. J., and Thielens, N. M. (2004) *J. Immunol.* **172,** 4342–4350
15. Chen, C. B., and Wallis, R. (2001) *J. Biol. Chem.* **276,** 25894–25902
16. Gregory, L. A., Thielens, N. M., Arlaud, G. J., Fontecilla-Camps, J. C., and Gaboriaud, C. (2003) *J. Biol. Chem.* **278,** 32157–32164
17. Gregory, L. A., Thielens, N. M., Matsushita, M., Sorensen, R., Arlaud, G. J., Fontecilla-Camps, J. C., and Gaboriaud, C. (2004) *J. Biol. Chem.* **279,** 29391–29397
18. Teillet, F., Gaboriaud, C., Lacroix, M., Martin, L., Arlaud, G. J., and Thielens, N. M. (2008) *J. Biol. Chem.* **283,** 25715–25724
19. Feinberg, H., Uitdehaag, J. C., Davies, J. M., Wallis, R., Drickamer, K., and Weis, W. I. (2003) *EMBO J.* **22,** 2348–2359
20. Romero, A., Romão, M. J., Varela, P. F., Kölln, I., Dias, J. M., Carvalho, A. L., Sanz, L., Töpfer-Petersen, E., and Calvete, J. J. (1997) *Nat. Struct. Biol.* **4,** 783–788
21. Teillet, F., Lacroix, M., Thiel, S., Weiguny, D., Agger, T., Arlaud, G. J., and Thielens, N. M. (2007) *J. Immunol.* **178,** 5710–5716
22. Lacroix, M., Dumestre-Pérard, C., Schoehn, G., Houen, G., Cesbron, J. Y., Arlaud, G. J., and Thielens, N. M. (2009) *J. Immunol.* **182,** 456–465
23. Gaboriaud, C., Thielens, N. M., Gregory, L. A., Rossi, V., Fontecilla-Camps, J. C., and Arlaud, G. J. (2004) *Trends Immunol.* **25,** 368–373
24. Arlaud, G. J., Sim, R. B., Duplaa, A. M., and Colomb, M. G. (1979) *Mol. Immunol.* **16,** 445–450
25. Journet, A., and Tosi, M. (1986) *Biochem. J.* **240,** 783–787
26. Luo, C., Thielens, N. M., Gagnon, J., Gal, P., Sarvari, M., Tseng, Y., Tosi, M., Zavodszky, P., Arlaud, G. J., and Schumaker, V. N. (1992) *Biochemistry* **31,** 4254–4262
27. Lacroix, M., Rossi, V., Gaboriaud, C., Chevallier, S., Jaquinod, M., Thielens, N. M., Gagnon, J., and Arlaud, G. J. (1997) *Biochemistry* **36,** 6270–6282
28. Bally, I., Rossi, V., Thielens, N. M., Gaboriaud, C., and Arlaud, G. J. (2005) *J. Immunol.* **175,** 4536–4542
29. King, L. A., and Possee, R. D. (1992) in *The Baculovirus Expression System: A Laboratory Guide*, pp. 111–114, Chapman and Hall, Ltd., London
30. Laemmli, U. K. (1970) *Nature* **227,** 680–685
31. Cseh, S., Vera, L., Matsushita, M., Fujita, T., Arlaud, G. J., and Thielens, N. M. (2002) *J. Immunol.* **169,** 5735–5743
32. Bersch, B., Hernandez, J. F., Marion, D., and Arlaud, G. J. (1998) *Biochemistry* **37,** 1204–1214

Downloaded from http://www.jbc.org/ by guest on March 9, 2015

33. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A* **47,** 110–119
34. Gál, P., Sárvári, M., Szilágyi, K., Závodszky, P., and Schumaker, V. N. (1989) *Complement Inflamm.* **6,** 433–441
35. Arlaud, G. J., Gaboriaud, C., Thielens, N. M., Budayova-Spano, M., Rossi, V., Fontecilla-Camps, J. C. (2002) *Mol. Immunol.* **39,** 383–394
36. Thielens, N. M., Illy, C., Bally, I. M., and Arlaud, G. J. (1994) *Biochem. J.* **301,** 378–384
37. Strang, C. J., Siegel, R. C., Phillips, M. L., Poon, P. H., and Schumaker, V. N. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79,** 586–590
38. Perkins, S. J., Villiers, C. L., Arlaud, G. J., Boyd, J., Burton, D. R., Colomb, M. G., and Dwek, R. A. (1984) J. *Mol. Biol.* **179,** 547–557
39. Budayova-Spano, M., Lacroix, M., Thielens, N. M., Arlaud, G. J., Fontecilla-Camps, J. C., and Gaboriaud, C. (2002) *EMBO J.* **21,** 231–239

Downloaded from http://www.jbc.org/ by guest on March 9, 2015

Publication du Chapitre V

P15 *-* ***Rossi V.*** *et al.,* ***(2010)*** *Topology of the membrane-bound form of complement protein C9 probed by glycosylation mapping, anti-peptide antibody binding, and disulfide modification*

Molecular Immunology 47 (2010) 1553–1560

Contents lists available at ScienceDirect

Molecular Immunology

journal homepage: www.elsevier.com/locate/molimm

Topology of the membrane-bound form of complement protein C9 probed by glycosylation mapping, anti-peptide antibody binding, and disulfide modification

Véronique Rossi[a,b], Yunxia Wang[a,1], Alfred F. Esser[a,*]

[a] Division of Cell Biology and Biophysics, School of Biological Sciences, University of Missouri-Kansas City, 5100 Rockhill Road, Kansas City, MO 64110, USA
[b] Laboratoire d'Enzymologie Moléculaire, Institut de Biologie Structurale Jean-Pierre Ebel, 38027 Grenoble Cedex 1, France

ARTICLE INFO

Article history:
Received 20 October 2009
Received in revised form 11 January 2010
Accepted 17 January 2010
Available online 12 February 2010

Keywords:
Complement
Complement 9
MACPF
(Poly)C9
Monotopic membrane protein
Membrane protein anchoring

ABSTRACT

The two N-linked oligosaccharides in native human C9 were deleted by site-specific mutagenesis. This aglycosyl-C9 did not differ from its native form in hemolytic and bactericidal activity. A new N-glycosylation site (K311N/E313T) was introduced into the turn of a helix–turn–helix [HTH] fold that had been postulated to form a transmembrane hairpin in membrane-bound C9. This glycosylated form of human C9 was as active as the native protein suggesting that the glycan chain remains on the external side of the membrane and that translocation of this hairpin is not required for membrane anchoring. Furthermore, flow cytometry provided evidence for the recognition of membrane-bound C9 on complement-lysed ghosts by an antibody specific for the HTH fold. A new N-glycosylation site (P26N) was also introduced close to the N-terminus of C9 to test whether this region was involved in C9 polymerization, which is thought to be required for cytolytic activity of C9. Again, this glycosylated C9 was as active as native C9 and could be induced to polymerize by heating or incubation with metal ions. The two C-terminal cystines within the MACPF domain could be eliminated partially or completely without affecting the hemolytic activity. Free sulfhydryl groups of unpaired cysteines in such C9 mutants are blocked since they could not be modified with SH-specific reagents. These results are discussed with respect to a recently proposed model that, on the basis of the MACPF structure in C8α, envisions membrane insertion of C9 to resemble the mechanism by which cholesterol-dependent cytolysins enter a membrane.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Human complement protein C9 is a hydrophilic serum glycoprotein responsible for the efficient expression of cytolytic and bacteriolytic functions of complement. It assembles on the surface of a target cell together with C5b, C6, C7, and C8 to form the C5b-9 or membrane attack complex (MAC1) and in the process refolds to become an integral membrane protein (Esser, 1994; Müller-Eberhard, 1986; Plumb and Sodetz, 1998). Based on amino acid sequence alignments it is evident that C6, C7, C8α, C8β and C9 share a homologous central region called the MACPF domain (Plumb and Sodetz, 1998) and a MACPF fragment of C8α expressed in bacteria could substitute for the complete C8α molecule in hemolytic and bacteriolytic assays (Slade et al., 2006). Although C5b-8 precursor complexes at high concentrations are weakly hemolytic, killing of nucleated cells and Gram-negative bacteria is dependent on C9. With respect to bacterial killing it has been useful conceptually to consider the C5b-8 complex as a receptor to which C9 binds and then translocates across the outer membrane to reach the periplasm (Dankert and Esser, 1987; Taylor, 1992). Bacterial cell death can be produced solely by C9 (Dankert and Esser, 1987). Strong evidence has been published showing that C5b-8 by itself is not a transmembrane structure but is situated in the membrane as a monotopic complex (McCloskey et al., 1989). Many investigators favor the view that polymerization of C9 leads to formation of transmembrane protein channels that are responsible for complement-mediated cytotoxicity (Podack and Tschopp, 1982; Tschopp et al., 1982, 1985). Early models designed to explain how C9 anchors the MAC in a membrane postulated that amphipathic α-helices generate transmembrane structures in membrane-bound C9 (Peitsch et al., 1990). The formation of disulfide bonds between adjacent C9 molecules in the assembled MAC has also received much attention (Hatanaka et al., 1994; Ware and Kolb, 1981; Yamamoto and Migita, 1983) and glycosylation of C9 was thought to be important for function (Kontermann and Rauterberg, 1989).

In order to test these models we decided to seek further experimental evidence for the topography of membrane-bound C9. After

---

Abbreviations: HTH, helix–turn–helix; MAC, membrane attack complex; PAGE, polyacrylamide gel electrophoresis; TBS, Tris buffered saline; SNA, *Sambucus nigra* agglutinin; GNA, *Galanthus nivalis* agglutinin; KLH, keyhole limpet hemocyanin; TMH, transmembrane β-hairpin; CDC, cholesterol-dependent cytolysin; wtC9, wild type C9; rC9, recombinant wtC9.

[*] Corresponding author. Tel.: +1 816 235 5247.
E-mail address: essera@umkc.edu (A.F. Esser).

[1] Current address: PEL-FREEZ Clinical Systems, LLC, Brown Deer, WI 53223, USA.

0161-5890/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.molimm.2010.01.013

removing the two consensus N-glycosylation sites in C9 we used glycosylation mapping to probe the surface topology of C9. Thus, new N-linked glycosylation sites were created by site-directed mutagenesis and glycosylation of these new sites was verified by lectin blotting. We also generated several mutants with completely deleted disulfide bonds and tested their importance for hemolytic activity and those having an unpaired cysteine were tested for their potential use in cysteine-scanning experiments. Recently two groups reported the crystal structure of the C8α MACPF domain and noticed that it shares a common folding pattern with cholesterol-dependent cytolysins (Hadders et al., 2007; Slade et al., 2008). On the basis of this common folding pattern it was proposed (Hadders et al., 2007; Rosado et al., 2008) that C9 polymerizes and enters a target membrane in a fashion similar to CDC proteins. According to this model the membrane-inserting regions are two β-strands that are connected by disulfide bonds. Whereas our results are not in conflict with this newly proposed membrane insertion mechanism some details, however, are difficult to reconcile with earlier results from our laboratory and others.

2. Materials and methods

2.1. Complement proteins and assays

Hemolytic activity of complement proteins was assayed according to standard procedures and bactericidal activity using the *Escherichia coli* strain C600 was assayed as described (Tomlinson et al., 1993). Cell survival is presented as the ratio of CFU/mL incubated with serum (sample) to CFU/mL of cells incubated in buffer (control).

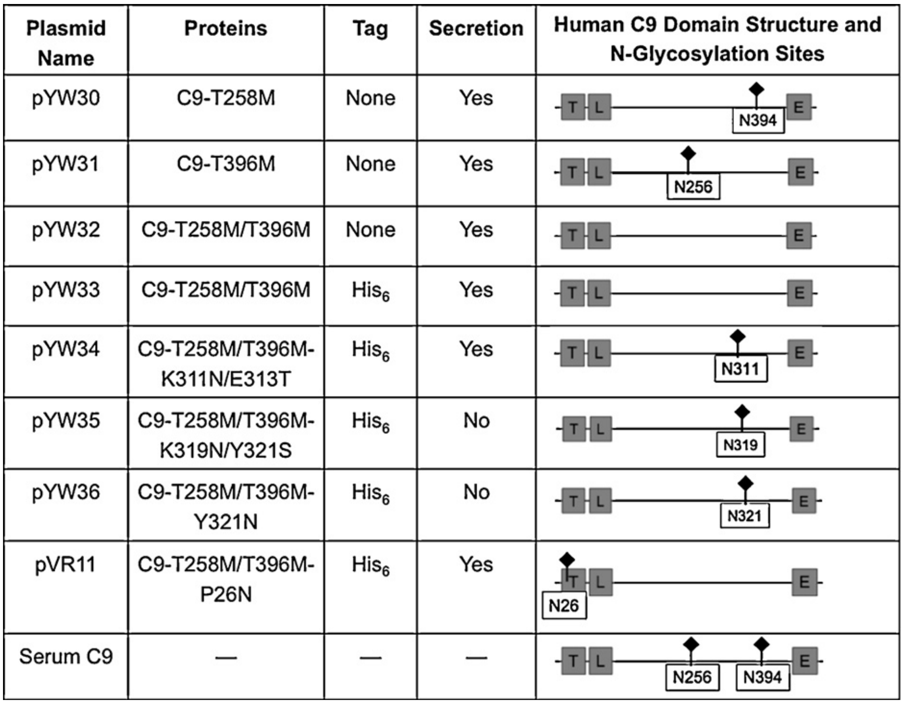
2.2. Construction of prokaryotic C9 expression plasmids

All DNA manipulations were carried out using standard techniques (Sambrook et al., 1989). The previously characterized C9 vectors pSL301HuC9 (Tomlinson et al., 1993) and pYW29 (Wang et al., 2000), which contains two mutations (L532S/K538R) and a His_6 C-terminal tag, were used as the starting plasmids for the expression vectors listed in Fig. 1. A KpnI–SacI fragment from pSL301HuC9 was cloned into pSelect-1 to eliminate the two existing N-glycosylation sites individually using the “Altered Sites Mutagenesis System” (Promega, Madison, WI). The mutated genes were then transferred into pSVL (Pharmacia) as XbaI–SacI fragments to generate pYW30 and pYW31, respectively. Plasmid pYW32 was prepared by exchanging the XbaI–NruI fragment in pYW31 with the corresponding one from pYW30. The hexahistidine tag was added by exchanging the XhoI–Tth111I fragment in pYW29 with the respective XhoI–Tth111I fragment of pYW32 to generate pYW33.

In order to introduce new glycosylation sites into aglycosyl-C9 a KpnI–SacI fragment of pYW33 was transferred into pSelect-1 and mutated at the desired positions. The mutated genes were transferred back into pYW33 by exchanging the respective XbaI–SacI fragments to yield pYW34, pYW35, and pYW36. To generate the P26N mutation an XbaI–SacI fragment was moved from pYW33 to pUC19 and mutagenesis was performed with the “Transformer Site-directed Mutagenesis System” from Clontech (Madison, WI) according to the manufacturers protocol. The mutated gene was then transferred back into pSVL as an XbaI–SacI fragment to create pVR11. All C9 mutants are listed in Fig. 1 together with a sketch showing the location of the various glycosylation sites within the C9 domain structure.

2.3. Construction of recombinant baculovirus

A baculovirus shuttle vector, or bacmid system (Luckow et al., 1993) kindly provided by Dr. Verne Luckow (Monsanto Corporation, Chesterfield, MO) was used for the generation of recombinant viruses. In brief, XbaI–SacI fragments were excised from pYW33, pYW34, and pVR11 and inserted individually between the NheI and SacI sites of the donor plasmid pMON27045. Transposition of the various C9 genes was carried out by transforming the recombinant donor plasmids into *E. coli* DH10B harboring the transposition helper plasmid pMon7124 and the bacmid bMON14272. Selec-



Plasmid Name	Proteins	Tag	Secretion	Human C9 Domain Structure and N-Glycosylation Sites
pYW30	C9-T258M	None	Yes	T L N394 E
pYW31	C9-T396M	None	Yes	T L N256 E
pYW32	C9-T258M/T396M	None	Yes	T L E
pYW33	C9-T258M/T396M	His_6	Yes	T L E
pYW34	C9-T258M/T396M-K311N/E313T	His_6	Yes	T L N311 E
pYW35	C9-T258M/T396M-K319N/Y321S	His_6	No	T L N319 E
pYW36	C9-T258M/T396M-Y321N	His_6	No	T L N321 E
pVR11	C9-T258M/T396M-P26N	His_6	Yes	N26 T L E
Serum C9	—	—	—	T L N256 N394 E

Fig. 1. C9 glycan mutants and location of the respective glycosylation sites. Tag refers to the addition of a C-terminal hexahistidine sequence that was not removed from the individual proteins. Secretion indicates whether the proteins were secreted by COS-7 and also by Sf9 cells.

tion of the correct composite bacmids was performed as described (Tomlinson et al., 1993) and purified bacmid DNA was used for the transfection of insect cell lines.

2.4. Expression of C9 mutants in COS-7 and insect cells

COS-7 cells were transfected with the pSVL-derived expressions vectors by electroporation and the different proteins were isolated as described previously (Tomlinson et al., 1993). C9 mutants were expressed at 27 °C by baculovirus-infected Sf9 cells growing in suspension in serum-free medium contained in 250 mL shaker flasks rotating at 135 rpm as described (Rossi et al., 1998). Protein purifications were performed as published (Tomlinson et al., 1993) except those that carried a C-terminal hexahistidine tag were purified by metal-affinity chromatography using the manufacturer's (Qiagen, Valencia, CA) protocol.

2.5. Quantification of secreted proteins

Proteins were separated by 7.5% acrylamide SDS-PAGE and transferred to nitrocellulose using standard techniques. Transferred C9 was immunostained by means of mAb216 and anti-mouse IgG-alkaline phosphatase conjugate according to the protocol supplied by BioRad. Concentrations of recombinant wild type and mutant forms of C9 were determined by a dot blot procedure using the "ELIFA™" apparatus (Pierce, Rockford, IL). Samples were adsorbed to nitrocellulose, blocked with 5% "Blotto" solution and incubated overnight with mAb216 anti-C9 (20 μg/mL), washed with 0.3% Tween-20 TBS buffer, and then incubated for 2 h with goat anti-mouse 125I-IgG (10 μg/mL) as detection antibody. The amount of radioiodinated IgG bound to serial dilutions of respective test samples was then quantified in a PhosphorImager™ (Molecular Dynamics, Sunnyvale, CA). Purified human C9 was used to generate a standard curve from which the amounts of secreted C9 were determined. The hemolytic activity of the different recombinant C9 mutants was determined as published before (Tomlinson et al., 1993).

2.6. Analysis of glycans on secreted C9 proteins

The presence of carbohydrates on native and secreted C9 mutants was detected by specific binding of lectins using the DIG Glycan Differentiation Kit purchased from Boehringer Mannheim (Indianapolis, IN) following the procedures supplied with the kit. Among the digoxigenin-labeled lectins only the *Sambucus nigra* agglutinin (SNA) recognized C9 purified from human serum demonstrating that it contains complex sialylated N-glycan chains.

2.7. Epitope mapping by sequence-specific anti-peptide antibodies

Antisera specific for two overlapping peptides, C-L_{305}PTTYEKGEYFAFLE_{319} and C-K_{311}GEYFAFLETYGTH_{324}, coupled via their N-terminal cysteines to keyhole limpet hemocyanin (KLH) were raised in rabbits as described (Laine and Esser, 1989a). Monospecific IgG was isolated individually from each antiserum by immunoaffinity chromatography using as an antigen the peptide C-TTYEKGEYFA coupled to ReactiGel (Pierce). Recognition of C9 on the surface of complement-lysed erythrocyte ghosts by these two antibodies was assayed by flow cytometry as published (Laine and Esser, 1989a) using a Becton-Dickson FacScan analyzer.

2.8. Modification of disulfides in C9

Human C9 contains 12 disulfide bonds, three in each one of the four recognized domains (www.UniProt.org), that is, the TSP type-1 and the LDL-receptor class A at the N-terminus, the MACPF in the center, and the EGF-like domain at the C-terminus. The C9 gene constructs were first cloned into the PUC19 plasmid and the desired cystine changes within the MACPF domain were accomplished using the Clontech (Palo Alto, CA) Transformer Site-directed Mutagenesis kit. The new constructs were then transferred either into a pSVL vector for expression in COS-7 cells or into pMON27045 for expression in insect cells.

2.9. Labeling of free sulfhydryl groups

Fluorescence labeling of free sulfhydryls in C9 and mutants was attempted using reagents with either a methanthiosulfonate (MTS) reactive group, such as MTS-4 fluoresceine, and MTS-pyrene (Toronto Research Chemicals, Inc.), or iodoacetamide [BODIPY-IA] and maleimide leaving groups (7-diethylamino-3-(4′-maleimidyl-phenyl)-4-methylcoumarin [CPM]), and also with monobromobimane (all purchased from Molecular Probes, Eugene, OR). Labeling reactions were done in 100 mM Hepes buffer at pH 7–7.5 and in the absence of any denaturing reagents.

3. Results

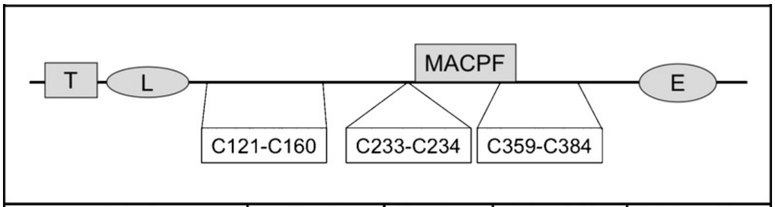
3.1. Cystine modification in C9

Recently two groups reported the crystal structure of the C8α MACPF domain (residues 103–462) and noticed that it shares a common folding pattern with CDC proteins (Hadders et al., 2007; Slade et al., 2008). On the basis of this common folding pattern it was proposed (Hadders et al., 2007; Rosado et al., 2008) that C9 polymerizes and enters a target membrane in a fashion similar to CDC proteins, that is, residues 197–270 (TMH1) and 343–400 (TMH2) in C9 refold into amphipathic transmembrane β-sheets that provide both the anchoring and the walls of the central aqueous pore in the poly(C9) ring structure. The MACPF domain in C9 contains three disulfide bonds and two of them are located within these putative transmembrane hairpin structures (Fig. 2). Thus it was of interest to test whether any one of these cystines is required for the functional activity of C9. Using site-specific mutagenesis the six cysteines were changed individually or together to serines. C9 mutants in which the disulfide bond between C121 and C160 was altered or deleted were not secreted from COS-7 cells or SF9 cells while alteration of the other two cystines (at position 233/234 and 359/384 in the TMH1 and 2, respectively) was not

```
TMH1
C8α(201-256)    DNANDLLSKVKKDKSDSFGVTIGIGPAGSPLL------------------VGVGVSHSQDTSFLNELNKYNEKK
C9 (197-270)    EEQIEAFKSIIQEKTSNFNAAISLKFTPTETNKAEQCCEETASSISLHGKGSFRFSYSKNETYQLFLSYSSKKE
                                                    ::

TMH2
C8α(329-385)    KAKMESLGITSRDITTCFGGSLGIQYE-DKINVGGGLSGDHCKKFGGG-KT-ERARKAMA
C9 (343-400)    KASMKRKGVELKDIKRCLGYHLDVSLAFSEISVGAEFNKDDCVKRGEG-RA-VNITSENL
                                :........................:
```

Fig. 2. Putative TMH sequences in human C8 and C9 (Hadders et al., 2007). Dotted lines indicate disulfide bonds.



Mutant	Cell Line	Secretion	Hemolytic Activity	SH-Labeling
C9-C121S	COS-7	No	—	—
C9-121S/160S	COS-7	No	—	—
C9-121S/160S/233T/234T/359S/384S	COS-7	No	—	—
C9-C233T	COS-7	Yes	Positive	Low level
C9-C359S	COS-7 & Sf9	Yes	Positive	Low level
C9-C384S	Sf9	Yes	Positive	Negative
C9-233T/234T	COS-7 & Sf9	Yes	Positive	Low level
C9-359S/384S	COS-7 & Sf9	Yes	Positive	Negative
C9-234T/359S/384S	COS-7 & Sf9	Yes	Positive	Negative

Fig. 3. Hemolytic activity and sulfhydryl labeling of C9-cysteine mutants. C9 mutant proteins with modified cystine residues were expressed in COS-7 and Sf9 cells and tested for secretion into the culture fluid, hemolytic activity, and labeling by SH-specific fluorescent probes. All secreted proteins have a C-terminal His6 tag allowing isolation of the secreted proteins by affinity chromatography. Protein concentrations were determined by immunoblotting and equivalent amounts of proteins were used to determine hemolytic activities which were not significantly different from recombinant wild type protein. SH-specific labeling of C9 proteins was attempted using a variety of different reagents (listed in the Section 2) with bovine serum albumin (BSA) serving as a control. Low level labeling of C9 proteins was about 20% of that obtained with BSA.

detrimental and all secreted mutants retained hemolytic activity (Fig. 3).

Since it was possible to create functional C9 mutant proteins with unpaired cysteines we attempted to label such cysteines with fluorescent reporter groups. After performing several control experiments it became evident, however, that even native C9 despite having only disulfide-paired cysteines could be labeled slightly with SH-specific labels in the native state but not under denaturing conditions, e.g., in the presence of 6 M GuHCl, in agreement with observations published earlier (Yamamoto and Migita, 1983). The authors attributed the low level labeling to disulfide exchange within the native protein. The labile disulfide bond responsible for the exchange in the native state was shown to be the bond between C359 and C384 by Hatanaka et al. (1994). Thus, we created a C9 molecule having a single cysteine at position 234 and a complete deletion of the C359 to C384 disulfide bond. Surprisingly this mutant protein could not be labeled with a variety of SH-specific reagents even under denaturing conditions nor was it possible to label C359 when it was the single cysteine in C9-C384S. This is in contrast to bovine serum albumin, which contains a free cysteine and could be labeled successfully under the same conditions (data not shown). We analyzed the secreted proteins by SDS-PAGE under non-reducing conditions and also by gel chromatography and could not find any evidence for a disulfide-linked dimer or multimers (data not shown). We conclude that the sulfhydryl groups in these secreted mutant proteins were perhaps chemically blocked by glutathione or cysteine present in culture media or provided by the cells during secretion. Others encountered a similar situation with an engineered thrombin mutant in which an engineered free sulfhydryl group was blocked (Chen et al., 1996). Attempts to remove the blocking substance by limited reduction under mild conditions were not successful. Conditions that would remove it also inactivated the proteins.

3.2. Hemolytic and bacteriocidal activity of aglycosyl-C9

The two N-glycosylation sites in human C9 at N256 and N394 were mutated (T258M and T396M) individually and together to provide a non-glycosylated protein. This protein was expressed in COS-7 cells and secreted into the culture medium. The absence of the carbohydrate moiety had little effect on the hemolytic (Fig. 4A) and the bactericidal activity of aglycosyl-C9 (Fig. 4B). It is known that glycosylation can have an effect on the stability of some proteins and protect them against degradation. Aglycosyl-C9 was as stable as native C9 when incubated in buffer for several days at room temperature (data not shown).

3.3. Expression and activity of C9 with different glycosylation sites

The fact that aglycosyl-C9 retained its hemolytic and bactericidal activity made it possible to introduce new consensus glycosylation sites into human C9 and test for their effect on folding and function of the protein. Therefore, four new consensus sites, P26N, K311N/E313T, Y321N and E319N/Y321S, were introduced into human aglycosyl-C9. COS-7 cells did not secrete the last two mutants but the other two were and at rates similar to the wild type recombinant protein. To verify that the consensus N-glycosylation acceptor sites were indeed recognized by the oligosaccharyl transferase we used lectin blotting to demonstrate the presence of carbohydrate moieties on the C9 constructs. Using the DIG glycan differentiation kit we could confirm that serum C9 contains complex, sialylated carbohydrates (DiScipio et al., 1984) since it was recognized by *S. nigra* agglutinin (SNA) (Fig. 5, top panel, lane 1) as was rC9 secreted from COS-7 cells (data not shown). Proteins secreted from SF9 cells were not recognized by SNA (data not shown) consistent with many observations that indicate that

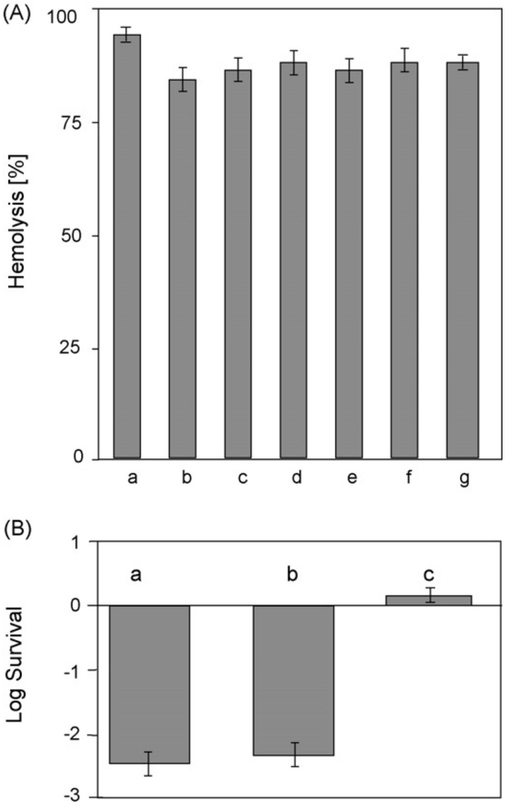


Fig. 4. Functional activity of C9 glycan mutants. Panel A: hemolytic activity of C9 mutants expressed in COS-7 and Sf9 cells. (a) Serum C9, (b) rC9 expressed in COS-7 cells, (c) C9-T258M (pYW30) expressed in COS-7 cells, (d) C9-T396M (pYW31) expressed in COS-7 cells, (e) C9-T258M/T396M (pYW33) expressed in Sf9 cells, (f) C9-258M/396M–K311N/E313T (pYW34) expressed in Sf9 cells, (g) C9-258M/396M–P26N (pVR11) expressed in Sf9 cells. Standard assays of C9 hemolytic activity (Esser and Sodetz, 1988) consisted of 5 × 108 sheep EAC1-8 and 50 μL of sample dilution in a total volume of 0.1 mL isotonic buffer. Incubations were conducted for 45 min at 37 °C followed by the addition of ice-cold buffer and centrifugation. 100% lysis was achieved by addition of 1 mL H2O in place of buffer after incubation. The amount of recombinant C9 used in all assays was approximately 5 ng/mL as measured by ELISA (Tomlinson et al., 1993). Panel B: killing of *E. coli* strain C600 by (a) C9-deficient serum reconstituted with 8 μg of rC9 or (b) with 8 μg aglycosyl-C9 (C9-258M/396M) per mL. (c) C9-deficient human serum. Error bars indicate means ± standard deviations (from 3 to 5 experiments).

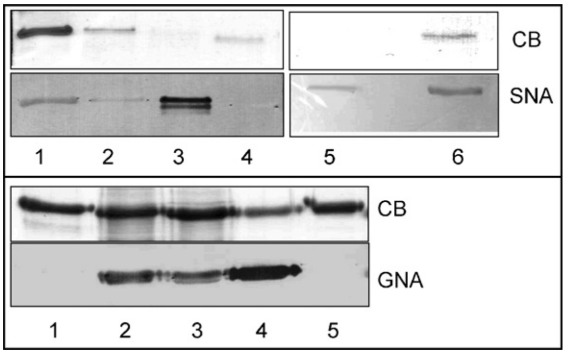


Fig. 5. Lectin blotting of C9 and glycan mutant proteins after SDS-PAGE under non-reducing conditions. Top panel: detection by Coomassie blue (CB) staining and *Sambucus nigra* agglutinin (SNA) blotting. Lane 1: 5 μg serum C9, Lane 2: 0.5 μg serum C9, Lanes 3 and 5: fetuin (positive control), Lane 4: carboxpetidase Y (negative control), Lane 6: C9-258M/396M–K311N/E313T expressed in COS-7 cells. Bottom panel: detection by Coomassie blue (CB) staining and *Galanthus nivalis* agglutinin (GNA) blotting: Lane 1: 5 μg serum C9, Lane 2: rC9 expressed in Sf9 cells, Lane 3: C9-258M/396M-P26N expressed in Sf9 cells, Lane 4: Carboxpetidase Y (positive control), Lane 5: rC9 expressed in COS-7 cells.

insect cells produce glycoproteins without terminal sialic acid and with mostly terminal oligo-mannose structures. Indeed, a lectin specific for such structures, *Galanthus nivalis* agglutinin (GNA), recognizes C9 secreted from SF9 cells but not serum C9 (Fig. 5, bottom panel, lanes 1 and 2) or rC9 when expressed in COS-7 cells (Fig. 5, bottom panel, lane 5). However, both lectins detect the C9 mutants with newly created N-glycosylation sites demonstrating that these mutant proteins are indeed glycosylated. SNA binds to C9-258M/396M–K311N/E313T when expressed in COS-7 (Fig. 5, top panel, lane 6) and GNA binds to C9-258M/396M-P26N when produced in SF9 cells (Fig. 5, bottom panel, lane 3) and the reverse is correct also, that is, the K311N/E313T mutant is recognized by GNA when expressed in Sf9 cells (data not shown). Importantly, hemolytic assays indicated that the aglycosyl-C9 mutants and those with new N-glycosylation sites are as active as the recombinant wild type C9 (Fig. 4).

Earlier reports (Laine and Esser, 1989a; Scibek et al., 2002; Taylor et al., 1994) suggested that the N-terminal portion of C9 and specifically the few amino acids preceding the TSP-1 domain are important for the polymerization properties of C9. Our results with the P26N mutant indicate that addition of a glycan chain at position 26 does not affect the function of the protein (Fig. 4) and the ability of the protein to aggregate when incubated with Zn^{2+} (Amiguet et al., 1985) to promote polyC9 formation (data not shown). This result is somewhat surprising since an antibody specific for amino acids 19–28 inhibited polymerization very strongly (Laine and Esser, 1989a).

3.4. Surface location of the EKGE peptide in membrane-bound C9

The observation that a carbohydrate structure at residue 311 in C9 did not alter its hemolytic activity suggested that the region around this carbohydrate attachment point remained on the surface of a target membrane during hemolysis. To test this possibility further we used binding of anti-peptide antibodies coupled with flow cytometry. We had used this methodology previously to locate C9 peptide regions on the surface of complement-lysed erythrocyte ghosts (Laine and Esser, 1989a). Two overlapping peptides with C-terminal cysteines covering the amino acid sequence between L305 and H324 were synthesized, coupled to KLH, and used as immunogens to raise antisera in two rabbits. Peptide-specific antibodies from both sera were purified separately on an immobilized antigen containing the overlapping region. As shown in Fig. 6, both antibody preparations bound to complement-lysed ghosts indicating that this C9 region is located on the membrane surface.

4. Discussion

Although several different models, inspired by electron microscopy data, primary sequence analyses, or biochemical studies, have been proposed to explain the anchoring of complement protein C9 in a target membrane the details remain elusive. The current favorite is a model (Podack, 1984) in which C9, after binding to a C5b-8 complex, changes conformation and polymerizes into a ring structure called poly(C9) that not only anchors the MAC to a membrane but also forms the wall of a transmembrane pore structure that is responsible for its function. The latest entry is a model based on X-ray crystallographic studies on the core portion of the C8α chain that support a role for β-sheets in the membrane anchoring of C9 within the C5b-9 complex.

The participation of TMH1 – also called the C9 hinge region in earlier publications – in these processes is indeed very attractive since it could provide answers to some intriguing unsolved puzzles. For example, when the primary structure of horse C9 was established it became immediately apparent that it was very similar to

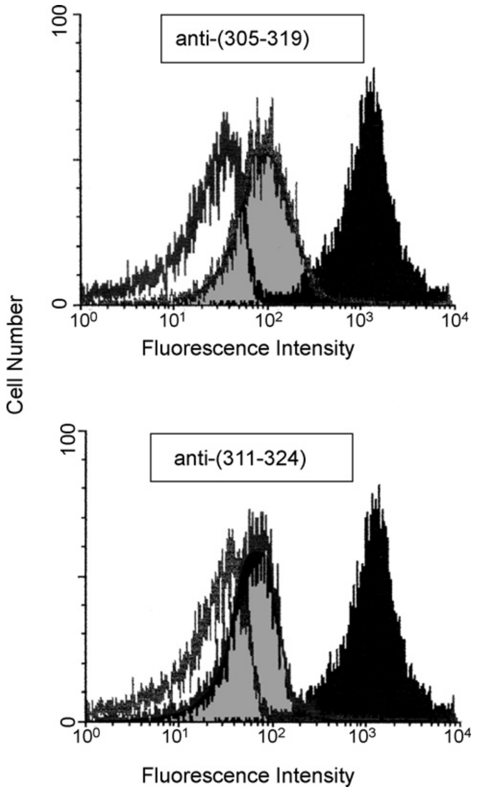


Fig. 6. Recognition of membrane-bound C9 by anti-peptide antibody binding as measured by flow cytometry. The left trace (white peak in both panels) shows the autofluorescence of complement-lysed ghosts incubated with non-immune IgG. Incubation with anti-(305–319) IgG (top panel) or anti-(311–324) (bottom panel) shifts the fluorescence (middle trace, gray peak in both panels) to the right and maximum fluorescence is obtained after incubation with polyclonal anti-C9 (black peak in both panels).

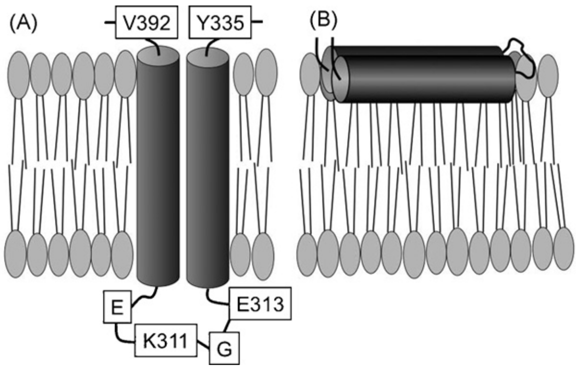

(A)
(B)
V392
Y335
E
E313
K311
G

Fig. 8. Membrane anchoring of C9. Panel A: transmembrane orientation of C9 as proposed by Peitsch et al. (1990). Panel B: monotopic anchoring of C9. The C9 sequence shown is the same in both panels and is homologous with helices D and E in the MACPF C8α–γ structure of Slade et al. (2008).

the human structure (Fig. 7) although horse C9 lacked the ability to lyse erythrocytes (Esser et al., 1996). The only significant differences noted were the two TMH sequences and indeed replacement of the hinge region in horse C9 with the human hinge rendered horse C9 hemolytic (Tomlinson et al., 1995). Furthermore, a peptide comprising residues 247–261 of human C9 was shown to be weakly hemolytic (Chang et al., 1994) indicating that TMH1 does have an affinity for membranes The absence of hemolytic activity in horse C9 could be the result of an additional oligosaccharide structure at position N256 that prevents membrane insertion of TMH1 and formation of a lytic pore. However, in order for this explanation to hold true it would be necessary to verify that horse C9 is indeed glycosylated at this site and that other C9 hemolytic molecules from species such as rabbit, mouse, rat and trout that have similar potential glycosylation sites in this loop region, are secreted without such oligosaccharides (Fig. 7). The alternating hydrophilic–hydrophobic pattern of the C-terminal part of horse TMH1 is also less pronounced compared to other C9 molecules but whether this could explain the lack of hemolytic power is not clear.

While it is extremely likely that the homologous MACPF core region in C9 is folded like the one in C8α it does not necessarily follow that C9 uses the same structural elements to enter a membrane as similarly folded CDC proteins. First, cutting the peptide chain between H244 and G245 in TMH1 of human C9 has no effect on its lytic activity (Dankert and Esser, 1985). It is difficult to imagine how such a severed β-strand could insert and find its partner to form a functional transmembrane β-sheet. Second, previous photolabeling data by Ishida et al. (1982) and by Tschopp and collaborators (Amiguet et al., 1985; Peitsch et al., 1990) are in severe conflict with the proposed membrane association of TMH1 and TMH2. In these studies no evidence was obtained for any labeling of TMH2 and the N-terminal β-strand in TMH1 by membrane-restricted radioactive photolabels. The radioactivity was associated only with the stretch from G245 to S345 leaving out most of TMH1 and all of TMH2. Perhaps these two regions are more important for the C5b-8/C9 oligomerization process than the actual membrane insertion event, a possibility that is reinforced by the fact that cutting the protein in

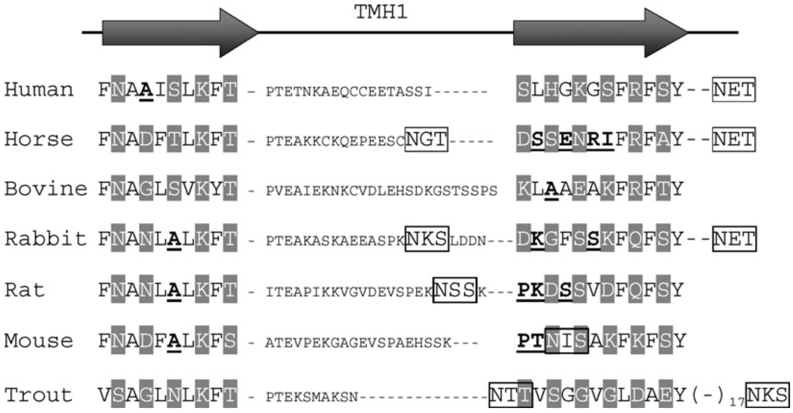

TMH1
Human FNAAISLKFT - PTETNKAEQCCEETASSI------ SLHGKGSFRFSY--NET
Horse FNADFTLKFT - PTEAKKCKQEPEESCNGT----- DSSENRIFRFAY--NET
Bovine FNAGLSVKYT - PVEAIEKNKCVDLEHSDKGSTSSPS KLAAEAKFRFTY
Rabbit FNANLALKFT - PTEAKASKAEEASPKNKSLDDN---DKGFSSKFQFSY--NET
Rat FNANLALKFT - ITEAPIKKVGVDEVSPEKNSSK---PKDSSVDFQFSY
Mouse FNADFALKFS - ATEVPEKGAGEVSPAEHSSK--- PTNISAKFKFSY
Trout VSAGLNLKFT - PTEKSMAKSN-------------NTTVSGGVGLDAEY(-)17NKS

Fig. 7. Comparison of the first putative transmembrane hairpin (TMH1) sequence in C9 from different species. The cartoon shows locations of two β-strands within the putative transmembrane hairpin regions. Amino acids are coded according to character: hydrophilic (white letters in gray box), hydrophobic (black letters) and those that do not fit the alternating hydrophilic/hydrophobic amphipathic pattern are shown in underlined bold black letters. Potential glycosylation sites [NXS/T] are placed within rectangles.

half at H244 prevents the formation of circular poly(C9) (Dankert and Esser, 1985). Since there is no direct experimental evidence that TMH1 and TMH2 are indeed located in a lysed target membrane it is therefore important to consider again the other modes of membrane attachment that were postulated earlier.

In the topological model for membrane-bound C9 of Stanley and coworkers (Stanley et al., 1986) two amphipathic α-helixes and three β-sheets transverse the membrane and two short and one longer loop of about 25–30 amino acids in the middle of the protein (between the major chymotrypsin and the single thrombin cleavage sites) protrude on the cytoplasmic side of the membrane. Peitsch combined secondary structure prediction, helix packing constraints, and energy minimization to construct a model for the membrane-spanning domain in C9 (Peitsch et al., 1990). In this model, residues 292–308 and 313–334 form an amphipathic α-helical hairpin, or HTH motif, that spans the membrane with the connecting short turn (E-K-G-E) protruding on the trans (cytoplasmic) side of the membrane (Fig. 8A). Removing the two recognition sites N-X-T for N-glyosylation in human C9 by site-specific mutation without affecting the specific hemolytic and bactericidal activity of the secreted proteins allowed us to use glycosylation mapping to probe structural aspects of C9. This technique of placing novel carbohydrate structures has already shown its usefulness in defining the folding patterns of polytopic membrane proteins. It is based on the observation that N-linked carbohydrates are always located on the exoplasmic site of a membrane in animal cells. It has not been used to study the topography of soluble, membrane-inserting proteins but since it is energetically unfavorable to move a large hydrophilic mass across a lipid bilayer the same technique should also allow the identification of surface exposed peptide loops of the membrane-bound form of such proteins. Because the hemolytic activity of aglycosyl-C9 was unaffected by the addition of a new carbohydrate moiety at position 311 we assume that it is unlikely that an HTH structure as envisioned by Peitsch and coworkers spans the membrane in a membrane-bound C5b-9 complex. Additional support for a surface location of the K-E-G-E stretch was provided by the recognition of the sequence K311–A316 by two different sequence-specific anti-peptide antibodies. Since it is unlikely that the immobilized, short peptide that was used to purify the antibodies had much secondary structure we assume that these antibodies recognize non-helical segments and most likely a turn or random structures. Such a turn was indicated by the modeling experiments of Peitsch and indeed the homologous region in C8α MACPF was shown recently to fold as such a helix–turn–helix motif (denoted as helix D and helix E in the topographical map of Slade et al.). Finally our earlier results (Laine and Esser, 1989b; Laine et al., 1988) identified residues 292–296 to be part of a discontinuous epitope that continues passed the EGF domain into the C-terminal end of C9 suggesting that the C-terminal end folds up against this putative hairpin.

However, we wish to emphasize that our results do not preclude the possibility that this pair of helices could lie more or less parallel to the membrane surface in the interfacial region analogous to the anchoring of monotopic membrane proteins (Picot and Garavito, 1994). Such topography is certainly correct for the membrane anchoring of the C5b-8 complex (McCloskey et al., 1989) and may be also correct for the initial C9 binding mode. The importance of this hairpin structure for C9 folding and function is evident from the mutagenesis studies of Dupuis et al. (1993) who attempted to change its amphipathic character. All non-conservative amino acid replacements introduced on either side of the α-helices – specifically residues 305, 318, 319 and 325 – resulted in non-secreted C9. Thus, the lack of secretion of the mutants we constructed with a consensus glycosylation site at position 319 (pYW35 and pYW36 in Fig. 1) could be the result of replacing a tyrosine at position 321 with either a serine or an asparagine and not necessarily from an interfering glycan. Changing E319 to N319 does not affect secretion and has no influence on hemolytic activity (Dupuis et al., 1993).

Considering the combined evidence currently available it is clear that in membrane-bound C9 hydrophobic or amphipathic transmembrane α-helical hairpins are not possible. Thus, it is necessary to invoke either a monotopic arrangement, as shown in Fig. 8B, or amphipathic transmembrane β-structures like those in the first Stanley model or those proposed for the membrane-inserted CDC proteins. However, it is also important to consider that in CDC proteins the transmembrane β-hairpins are α-helical or unordered in the native proteins and assume their new conformation already during pre-pore formation and before membrane insertion. Thus, it is tempting to speculate that in a C5b-9 complex containing only a few (less than four) C9 monomers these conserved α-helices D and E in C6, C7, C8, and in C9 anchor the complex in a monotopic fashion as shown in Fig. 8 whereas in a fully formed poly(C9) structure other structural elements may refold and become the staves of a transmembrane barrel. Since we used normal human serum, or C9-deficient serum reconstituted with mutant C9 proteins, in which a C9:C8 ratio of 2:1 was never exceeded the experiments reported here apply to C5b-9 complexes containing very few if any poly(C9) structures. Therefore, our results may not be applicable when the majority of MACs contain poly(C9) rings.

Much of the success in understanding the membrane insertion of the CDC proteins is based on the ingenious use of cysteine-scanning mutagenesis combined with fluorescence labeling of these newly created sulfhydryls in CDC proteins by Tweten and Johnson and their collaborators (Heuck et al., 2001). Unfortunately, as our attempts to create such mutant proteins indicate, such a strategy may not be possible with C9 unless one can find culture conditions and/or cell lines that will not block free sulfhydryl groups during secretion. Alternatively, glycosylation scanning on a larger scale than presented here should allow identification of those amino acids that must enter a target membrane to achieve hemolysis and a more detailed analysis of C9 molecules with consensus glycosylation sites within the TMH1 and TMH2 regions should provide much useful information.

Acknowledgements

We would like to thank Lilla Khalifah for excellent technical assistance. This work was supported by National Institutes of Health Grant GM53748 and by a Marion Merrell Dow Professorship Endowment (to A.F.E.).

References

Amiguet, P., Brunner, J., Tschopp, J., 1985. The membrane attack complex of complement: lipid insertion of tubular and nontubular polymerized C9. Biochemistry 24, 7328–7334.

Chang, C.P., Husler, T., Zhao, J., Wiedmer, T., Sims, P.J., 1994. Identity of a peptide domain of human C9 that is bound by the cell-surface complement inhibitor CD59. J. Biol. Chem. 269, 26424–26430.

Chen, Q., Lord, S.T., Lentz, B.R., 1996. Construction, properties and specific fluorescent labeling of a bovine prothrombin mutant engineered with a free C-terminal cysteine. Protein Eng. 9, 545–553.

Dankert, J.R., Esser, A.F., 1985. Proteolytic modification of human complement protein C9: loss of poly(C9) and circular lesion formation without impairment of function. Proc. Natl. Acad. Sci. U.S.A. 82, 2128–2132.

Dankert, J.R., Esser, A.F., 1987. Bacterial killing by complement C9-mediated killing in the absence of C5b-8. Biochem. J. 244, 393–399.

DiScipio, R.G., Gehring, M.R., Podack, E.R., Kan, C.C., Hugli, T.E., Fey, G.H., 1984. Nucleotide sequence of cDNA and derived amino acid sequence of human complement component C9. Proc. Natl. Acad. Sci. U.S.A. 81, 7298–7302.

Dupuis, M., Peitsch, M.C., Hamann, U., Stanley, K.K., Tschopp, J., 1993. Mutations in the putative lipid-interaction domain of complement C9 result in defective secretion of the functional protein. Mol. Immunol. 30, 95–100.

Esser, A.F., Sodetz, J.M., 1988. Membrane attack complex proteins C5b-6, C7, C8, and C9 of human complement. Methods Enzymol. 162, 551–578.

Esser, A.F., 1994. The membrane attack complex of complement, assembly, structure and cytotoxic activity. Toxicology 87, 229–247.

Esser, A.F., Tarnuzzer, R.W., Tomlinson, S., Tatar, L.D., Stanley, K.K., 1996. Horse complement protein C9: primary structure and cytotoxic activity. Mol. Immunol. 33, 725–733.

Hadders, M.A., Beringer, D.X., Gros, P., 2007. Structure of C8alpha-MACPF reveals mechanism of membrane attack in complement immune defense. Science 317, 1552–1554.

Hatanaka, M., Seya, T., Inai, S., Shimizu, A., 1994. The functions of the ninth component of human complement are sustained by disulfide bonds with different susceptibilities to reduction. Biochim. Biophys. Acta 1209, 117–122.

Heuck, A.P., Tweten, R.K., Johnson, A.E., 2001. Beta-barrel pore-forming toxins: intriguing dimorphic proteins. Biochemistry 40, 9065–9073.

Ishida, B., Wisnieski, B.J., Lavine, C.H., Esser, A.F., 1982. Photolabeling of a hydrophobic domain of the ninth component of human complement. J. Biol. Chem. 257, 10551–10553.

Kontermann, R., Rauterberg, E.W., 1989. N-deglycosylation of human complement component C9 reduces its hemolytic activity. Mol. Immunol. 26, 1125–1132.

Laine, R.O., Esser, A.F., 1989a. Detection of refolding conformers of complement protein C9 during insertion into membranes. Nature 341, 63–65.

Laine, R.O., Esser, A.F., 1989b. Identification of the discontinuous epitope in human complement protein C9 recognized by anti-melittin antibodies. J. Immunol. 143, 553–557.

Laine, R.O., Morgan, B.P., Esser, A.F., 1988. Comparison between complement and melittin hemolysis: anti-melittin antibodies inhibit complement lysis. Biochemistry 27, 5308–5314.

Luckow, V.A., Lee, S.C., Barry, G.F., Olins, P.O., 1993. Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in Escherichia coli. J. Virol. 67, 4566–4579.

McCloskey, M.A., Dankert, J.R., Esser, A.F., 1989. Assembly of complement components C5b-8 and C5b-9 on lipid bilayer membranes: visualization by freeze-etch electron microscopy. Biochemistry 28, 534–540.

Müller-Eberhard, H.J., 1986. The membrane attack complex of complement. Annu. Rev. Immunol. 4, 503–528.

Peitsch, M.C., Amiguet, P., Guy, R., Brunner, J., Maizel Jr., J.V., Tschopp, J., 1990. Localization and molecular modelling of the membrane-inserted domain of the ninth component of human complement and perforin. Mol. Immunol. 27, 589–602.

Picot, D., Garavito, R.M., 1994. Prostaglandin H synthase: implications for membrane structure. FEBS Lett. 346, 21–25.

Plumb, M.E., Sodetz, J.M., 1998. Proteins of the membrane attack complex. In: Volanakis, J.E., Frank, M.M. (Eds.), The Human Complement System in Health and Disease. Marcel Dekker, New York, NY, pp. 119–148.

Podack, E.R., 1984. Molecular composition of the tubular structure of the membrane attack complex of complement. J. Biol. Chem. 259, 8641–8647.

Podack, E.R., Tschopp, J., 1982. Polymerization of the ninth component of complement (C9): formation of poly(C9) with a tubular ultrastructure resembling the membrane attack complex of complement. Proc. Natl. Acad. Sci. U.S.A. 79, 574–578.

Rosado, C.J., Kondos, S., Bull, T.E., Kuiper, M.J., Law, R.H., Buckle, A.M., Voskoboinik, I., Bird, P.I., Trapani, J.A., Whisstock, J.C., Dunstone, M.A., 2008. The MACPF/CDC family of pore-forming toxins. Cell Microbiol. 10, 1765–1774.

Rossi, V., Bally, I., Thielens, N.M., Esser, A.F., Arlaud, G.J., 1998. Baculovirus-mediated expression of truncated modular fragments from the catalytic region of human complement serine protease C1s. Evidence for the involvement of both complement control protein modules in the recognition of the C4 protein substrate. J. Biol. Chem. 273, 1232–1239.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual. Cold Spring HarborLaboratory, Cold Spring Harbor, NY.

Scibek, J.J., Plumb, M.E., Sodetz, J.M., 2002. Binding of human complement C8 to C9: role of the N-terminal modules in the C8 alpha subunit. Biochemistry 41, 14546–14551.

Slade, D.J., Chiswell, B., Sodetz, J.M., 2006. Functional studies of the MACPF domain of human complement protein C8alpha reveal sites for simultaneous binding of C8beta, C8gamma, and C9. Biochemistry 45, 5290–5296.

Slade, D.J., Lovelace, L.L., Chruszcz, M., Minor, W., Lebioda, L., Sodetz, J.M., 2008. Crystal structure of the MACPF domain of human complement protein C8 alpha in complex with the C8 gamma subunit. J. Mol. Biol. 379, 331–342.

Stanley, K.K., Page, M., Campbell, A.K., Luzio, J.P., 1986. A mechanism for the insertion of complement component C9 into target membranes. Mol. Immunol. 23, 451–458.

Taylor, K.M., Morgan, B.P., Campbell, A.K., 1994. Altered glycosylation and selected mutation in recombinant human complement component C9: effects on haemolytic activity. Immunology 83, 501–506.

Taylor, P.W., 1992. Complement-mediated killing of susceptible gram-negative bacteria: an elusive mechanism. Exp. Clin. Immunogenet. 9, 48–56.

Tomlinson, S., Ueda, E., Maruniak, J.E., Garcia-Canedo, A., Bjes, E.S., Esser, A.F., 1993. The expression of hemolytically active human complement protein C9 in mammalian, insect, and yeast cells. Protein Exp. Purif. 4, 141–148.

Tomlinson, S., Wang, Y., Ueda, E., Esser, A.F., 1995. Chimeric horse/human recombinant C9 proteins identify the amino acid sequence in horse C9 responsible for restriction of hemolysis. J. Immunol. 155, 436–444.

Tschopp, J., Podack, E.R., Müller-Eberhard, H.J., 1982. Ultrastructure of the membrane attack complex of complement: detection of the tetramolecular C9-polymerizing complex C5b-8. Proc. Natl. Acad. Sci. U.S.A. 79, 7474–7478.

Tschopp, J., Podack, E.R., Müller-Eberhard, H.J., 1985. The membrane attack complex of complement: C5b-8 complex as accelerator of C9 polymerization. J. Immunol. 134, 495–499.

Wang, Y., Bjes, E.S., Esser, A.F., 2000. Molecular aspects of complement-mediated bacterial killing. Periplasmic conversion of C9 from a protoxin to a toxin. J. Biol. Chem. 275, 4687–4692.

Ware, C.F., Kolb, W.P., 1981. Assembly of the functional membrane attack complex of human complement: formation of disulfide-linked C9 dimers. Proc. Natl. Acad. Sci. U.S.A. 78, 6426–6430.

Yamamoto, K., Migita, S., 1983. Mechanisms for the spontaneous formation of covalently linked polymers of the terminal membranolytic complement protein (C9). J. Biol. Chem. 258, 7887–7889.

Publication du Chapitre VI

P17 – *Jacquet M. et al.,* ***(2013)****, Deciphering complement receptor type 1 (CR1) interactions with recognition proteins of the lectin complement pathway****.***

P17

Deciphering Complement Receptor Type 1 Interactions with Recognition Proteins of the Lectin Complement Pathway

Mickaël Jacquet, Monique Lacroix, Sarah Ancelet, Evelyne Gout, Christine Gaboriaud, Nicole M. Thielens, and Véronique Rossi

Complement receptor type 1 (CR1) is a membrane receptor expressed on a wide range of cells. It is involved in immune complex clearance, phagocytosis, and complement regulation. Its ectodomain is composed of 30 complement control protein (CCP) modules, organized into four long homologous repeats (A–D). In addition to its main ligands C3b and C4b, CR1 was reported to interact with C1q and mannan-binding lectin (MBL) likely through its C-terminal region (CCP22–30). To decipher the interaction of human CR1 with the recognition proteins of the lectin complement pathway, a recombinant fragment encompassing CCP22–30 was expressed in eukaryotic cells, and its interaction with human MBL and ficolins was investigated using surface plasmon resonance spectroscopy. MBL and L-ficolin were shown to interact with immobilized soluble CR1 and CR1 CCP22–30 with apparent dissociation constants in the nanomolar range, indicative of high affinity. The binding site for CR1 was located at or near the MBL-associated serine protease (MASP) binding site in the collagen stalks of MBL and L-ficolin, as shown by competition experiments with MASP-3. Accordingly, the mutation of an MBL conserved lysine residue essential for MASP binding (K55) abolished binding to soluble CR1 and CCP22–30. The CR1 binding site for MBL/ficolins was mapped to CCP24–25 of long homologous repeat D using deletion mutants. In conclusion, we show that ficolins are new CR1 ligands and propose that MBL/L-ficolin binding involves major ionic interactions between conserved lysine residues of their collagen stalks and surface exposed acidic residues located in CR1 CCP24 and/or CCP25. ***The Journal of Immunology*****, 2013, 190: 3721–3731.**

The lectin pathway of complement is an element of the innate immune system that plays an essential role in host resistance to pathogens and protection against damaged self-structures. It is triggered by soluble oligomeric proteins of the collectins and ficolins family able to recognize pathogen- and damage-associated molecular patterns and to trigger complement activation through associated proteases called the mannan-binding lectin (MBL)–associated serine proteases (MASPs) (1, 2). Five recognition proteins of the lectin pathway have been described in humans: MBL (3), the recently identified collectin 11 (CL-11, CL-K1) (4), and the three ficolins M, L and H (also termed ficolin-1, -2, and -3) (5). These proteins are multimers of homotrimeric subunits comprising a collagen-like triple helix and a C-terminal globular trimeric recognition domain of the C-lectin or fibrinogen-like type. They are structurally related to C1q, the recognition protein of the classical complement pathway, whereas the MASPs are modular proteins homologous to C1q-associated proteases C1r and C1s (6). Three MASPs have been identified, MASP-1, MASP-2, and MASP-3, among which only MASP-1 and MASP-2 are essential for complement activation (7), whereas the role of MASP-3 still requires further elucidation. The three MASPs and the C1s-C1r-C1r-C1s tetramer interact in a similar way with the collagen-like stalks of the recognition proteins, which allows in vitro cross-interaction between subcomponents of the classical and lectin pathways (8, 9). Following recognition of a target, the activated C1 and MBL/ficolin–MASP proteolytic complexes cleave complement components C4 and C2 to generate the C3 convertase responsible for cleavage of C3 and subsequent opsonization of the target by C3 fragments such as C3b. In addition to C3b and C4b, C1q, MBL, and ficolins may also act as opsonins, thereby facilitating phagocytosis of their targets through direct interaction with cell surface receptors.

The major complement receptor for C3b C4b, complement receptor type 1 (CR1; CD35), is a polymorphic membrane glycoprotein found on a wide range of cells including erythrocytes, monocytes and lymphocytes (10). CR1 on erythrocytes has a major role in the clearance of C3b/C4b-opsonized immune complexes from the circulation. It also acts as a regulator of complement by accelerating the decay of C3 and C5 convertases and by serving as a cofactor in factor I–mediated cleavage of C3b/C4b to iC3b/iC4b and subsequently to C3d/C4d, which are ligands for other complement receptors types of the Ig superfamily, CR2, CR3, and CR4 (11). Additional CR1 ligands include C1q (12), MBL (13), and two *Plasmodium falciparum* proteins, the erythrocyte membrane protein 1 (PfEMP1) and the reticulocyte-binding protein homolog 4 (PfRh4) involved in rosette formation between infected and noninfected erythrocytes and sialic acid–independent cell invasion, respectively (14, 15).

Commissariat à l'Energie Atomique, Institut de Biologie Structurale Jean-Pierre Ebel, 38027 Grenoble Cedex 1, France; Centre National de la Recherche Scientifique, Institut de Biologie Structurale Jean-Pierre Ebel, 38027 Grenoble Cedex 1, France; and Université Joseph Fourier, Grenoble 1, Institut de Biologie Structurale Jean-Pierre Ebel, 38027 Grenoble Cedex 1, France

Received for publication August 31, 2012. Accepted for publication February 2, 2013.

This work was supported by the Commissariat à l'Energie Atomique, the Centre National de la Recherche Scientifique, the Université Joseph Fourier (Grenoble, France), and by a grant from the French National Research Agency (ANR-09-PIRI-0021).

Address correspondence and reprint requests to Dr. Nicole M. Thielens, Institut de Biologie Structurale Jean-Pierre Ebel, 41 Rue Jules Horowitz, 38027 Grenoble Cedex 1, France. E-mail address: nicole.thielens@ibs.fr

The online version of this article contains supplemental material.

Abbreviations used in this article: CCP, complement control protein; CD, circular dichroism; CR1, complement receptor type 1; LHR, long homologous repeat; MASP, mannan-binding lectin–associated serine protease; MBL, mannan-binding lectin; PfEMP1, *Plasmodium falciparum* erythrocyte membrane protein 1; PfRh4, *Plasmodium falciparum* reticulocyte binding protein homolog 4; RU, resonance unit; sCR1, soluble CR1; SPR, surface plasmon resonance.

Copyright © 2013 by The American Association of Immunologists, Inc. 0022-1767/13/$16.00

www.jimmunol.org/cgi/doi/10.4049/jimmunol.1202451

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

CR1 is a multimodular membrane glycoprotein composed of an extracellular stretch of 30 complement control protein (CCP) modules, with the 28 N-terminal CCPs organized into four long homologous repeats (LHR-A, -B, -C, and -D) of 7 CCPs each, a transmembrane domain and a short cytoplasmic tail (16–18). LHR-A contains the major C4b binding site located in CCPs 1–3, also involved in PfRh4 binding, whereas LHR-B and LHR-C contain homologous C4b/C3b binding sites in CCPs 8–10 and 15–17, which are also candidates for interaction with PfEMP1 (19-21). C1q and MBL have been proposed to bind LHR-D, but the localization of their interaction sites to particular CCP modules has not been described so far (12, 13). Interestingly, LHR-D carries all known Knops blood groups Ags (resulting from single nucleotide polymorphisms), the most common Knops ($Kn^{a/b}$), McCoy ($McC^{a/b}$), and Swain–Langley (Sl1/2) being located in CCPs 24–25 (22).

The fact that MBL was characterized previously as a ligand of CR1 prompted us to investigate the possibility that ficolins, which are structurally and functionally related to MBL and share the capacity to activate the lectin complement pathway and to interact with cell surface proteins such as calreticulin (23) and CD91 (24), could also interact with CR1. For this purpose, a recombinant fragment encompassing CCP modules 22–30 of CR1 was produced in eukaryotic cells, and its interaction with MBL and ficolins was investigated using surface plasmon resonance (SPR) spectroscopy. The CR1 binding site for MBL/ficolins was mapped using deletion mutants and shown to involve the CCPs 24 and 25 of LHRD. Moreover, we show that MBL and L-ficolin share a common binding site for CR1, located in the collagen stalks, most probably at or in close proximity to the MASP interaction site.

Materials and Methods

Proteins and reagents

Oligonucleotides were purchased from Eurogentec (Liège, Belgium). Restriction enzymes were from New England Biolabs (Beverly, MA). Recombinant soluble human CR1 was purchased from R&D Systems Europe (Lille, France). Acetylated BSA and diisopropylphoshorofluoridate were from Sigma-Aldrich (St. Quentin Fallavier, France).

Recombinant human L- and H-ficolins were produced in Chinese hamster ovary cells (25) and purified using a one-step affinity chromatography on *N*-acetylcysteine–Sepharose for L-ficolin (25) and on acetylated BSA–Sepharose for H-ficolin. Elution from the acetylated BSA–Sepharose column (prepared by coupling 125 mg acetylated BSA to 15 ml CNBr-activated Sepharose 4B [GE Healthcare, Vélizy, France]) was carried out using 1 M sodium acetate and 5 mM EDTA (pH 7.4). Recombinant MBL and its K55A and K55E variants, produced in Freestyle 293-F cells (Invitrogen Life Technologies, St. Aubin, France) and purified as previously described (26) were provided by NatImmune (Copenhagen, Denmark). The molar concentrations of dimeric MASP-3, tetrameric MBL, tetrameric L-ficolin, and tetrameric H-ficolin were estimated using M_r values of 175,200, 305,400, 406,300, and 396,000 and absorbance coefficients at 280 nm ($A_{1\%,\ 1\ cm}$) of 13.5, 7.8, 17.6, and 19.4, respectively (25, 27).

Construction of the CR1 CCP22–30 and CR1 CCP22–30 size variants expression plasmids

The expression plasmids coding for CR1 CCP22–30 (amino acid residues 1395–1969 of CR1 F allotype) and its size variants were generated using the QuickChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies, Massy, France), according to the manufacturer's protocol for simple mutations or a modified protocol for the large insertions/deletions (28). The template containing the DNA coding sequence for the soluble part of CR1 was a generous gift from Prof. J. Cohen (Université de Reims Champagne-Ardennes, Reims, France) (29). First, BglII and KpnI restriction sites were introduced, respectively, at the 5′ and 3′ ends of the CCP 22–30 coding sequence. The BglII/KpnI restriction fragment was purified and cloned into the BamHI–KpnI sites of the pNT-Bac vector (30), in-frame with the melittin signal peptide. Finally, a fusion 6× His-tag was introduced at the C-terminal end of CR1 CCP22–30 by site-directed mutagenesis. The resulting vector pNT-Bac_CR1 CCP22–30_His was used as a template for the deletion of CCP modules to obtain the CR1 CCP22–30_His size variants. A DNA fragment corresponding to CR1 CCP22–30 was also cloned into the pcDNA3.1 mammalian expression vector, in-frame with the native signal peptide of CR1 (that was inserted by site-directed mutagenesis). The sequences of all constructs were verified by dsDNA sequencing (GATC Biotech, Mulhouse, France).

Production and purification of CR1 CCP22–30_His and its deletion variants in insect cells

The insect cells *Spodoptera frugiperda* (Sf21) and *Trichoplusia ni* (High Five) were routinely grown and maintained as described previously (31), and recombinant baculoviruses were generated using the Bac-to-Bac system (Invitrogen). The bacmid DNA was purified using the Qiagen midiprep purification system (Qiagen, Courtaboeuf, France) and used to transfect Sf21 insect cells with cellfectin in Sf900 II serum-free medium, as recommended by the manufacturer (Invitrogen). Recombinant virus particles were collected 4 d later and amplified as described previously (31). High Five cells (1.75×10^7 cells/175-cm^2 tissue culture flask) were infected with the recombinant viruses at a multiplicity of infection of 2 in Sf900 II serum-free medium at 27°C for 72 h. The culture supernatants containing the CR1 CCP22–30 fragment or its size variants were collected by centrifugation, supplemented with 1 mM diisopropylphosphorofluoridate, and purified immediately or stored at −20°C until use.

The same procedure was used for the purification of CR1 CCP22–30_His and its size variants. Insect cell culture supernatants (250 or 500 ml) were dialyzed against 50 mM sodium phosphate, 300 mM NaCl, and 10 mM imidazole buffer (pH 7.4), centrifuged, and loaded at a flow-rate of 2.5 ml/min onto a 20-ml HIS-Select Nickel Affinity (Sigma-Aldrich) home-packed column (2.5 × 10 cm) equilibrated in the same buffer. After washing, proteins were eluted with a 150 mM imidazole buffer. Fractions containing the recombinant protein were identified by SDS-PAGE analysis, pooled, and dialyzed against PBS (pH 7.2). The purified proteins were concentrated to 0.7–2.6 mg/ml by ultrafiltration (Vivaspin Ultra4; Millipore, Molsheim, France) and stored at −20°C. The concentrations of the purified CR1 CCP22–30_His variants were estimated using the following absorption coefficients $A_{1\%,\ 1\ cm}$ at 280 nm calculated in silico using the PROTPARAM program on Expasy Server (http://web.expasy.org/protparam/), and experimental molecular weights were determined by mass spectrometry: CR1 CCP22–30, 12.0 and 71156; Δ22–23, 12.4 and 56132; Δ22–25, 13.9 and 38894; Δ27–30, 10.1 and 40107; Δ25–26, 12.5 and 57083; Δ26–30, 9.8 and 32708; Δ29–30, 9.7 and 56567; and CCP25–26, 9.4 and 15434.

Production and purification of untagged CR1 CCP22–30

FreeStyle 293-F cells grown in FreeStyle 293 Expression Medium (Invitrogen) were transiently transfected using 293fectin (Invitrogen) as recommended by the manufacturer. The medium was harvested 72 h posttransfection. The following procedure was used to purify untagged CR1 CCP22–30 from the 293-F or High Five cells supernatants. The supernatant was dialyzed against 40 mM NaCl and 25 mM sodium phosphate (pH 7.5), and loaded at 1.5 ml/min onto a Q-Sepharose Fast Flow column (GE Healthcare) (2.8 × 10 cm) equilibrated in the same buffer. Elution was carried out by applying a 1-l linear gradient to 500 mM NaCl in the same buffer. The recombinant protein, which was identified in the flow through fractions by SDS-PAGE analysis, was concentrated and applied onto a Superose 12 10/30GL column (GE Healthcare) at a flow rate of 0.5 ml/min in 150 mM NaCl and 50 mM sodium phosphate (pH 7.5). The purified protein was concentrated by ultrafiltration to 0–5–1.3 mg/ml.

Expression and purification of recombinant MASP-3

A DNA segment encoding the human MASP-3 signal peptide plus the 709-aa residues of the mature protein was amplified by PCR using the pFastBac-MASP-3 plasmid (32) as a template, according to established procedures. The amplified MASP-3 DNA was cloned into the NheI and EcoRI restriction sites of the pcDNA3.1 plasmid. Transient transfection of 293-F cells with the pcDNA3.1 vector encoding MASP-3 was performed as described above for untagged CR1 CCP22–30. The medium was harvested 72 h posttransfection, and diisopropylphosphorofluoridate was added to a final concentration of 1 mM. The recombinant protein was purified using a one-step affinity chromatography on a C1q–Sepharose column (prepared by coupling 30 mg purified serum-derived C1q (33) to 15 ml CNBr-activated Sepharose 4B according to the manufacturer's protocol). Briefly, the supernatant was loaded onto the C1q–Sepharose column equilibrated in 50 mM triethanolamine-HCl, 145 mM NaCl, and 2 mM $CaCl_2$ (pH 7.4), and MASP-3 was eluted using the same buffer containing 5 mM EDTA instead of $CaCl_2$. The purified protein was dialyzed against

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

50 mM triethanolamine-HCl, 145 mM NaCl, and 2 mM $CaCl_2$ (pH 7.4) and concentrated to ~1 mg/ml by ultrafiltration.

SPR analyses and data evaluation

All experiments were performed on a BIAcore 3000 instrument (GE Healthcare). Recombinant soluble CR1 (sCR1), CR1 CCP22–30, and its size variants were covalently immobilized on CM5 sensor chips in 10 mM HEPES, 145 mM NaCl, and 0.005% surfactant P20 (pH 7.4) (HBS-P) using the amine coupling chemistry according to the manufacturer's instructions (GE Healthcare). The protein ligands were diluted in 10 mM sodium acetate at the following concentration and pH values prior to immobilization: sCR1, CR1 CCP22–30 and its deletion variants: 25, 20, and 5 μg/ml (pH 4.2); MBL and BSA: 25 μg/ml (pH 4); L-ficolin: 21 μg/ml (pH 5). Binding was measured at a flow rate of 20 μl/min in 150 mM NaCl, 10 mM HEPES (pH 7.4), and 0.005% surfactant P20 (HBS-P), containing 2 mM EDTA (HBS-EP) or 2 mM $CaCl_2$ and 10 mM mannose, to prevent unwanted interaction between the high-mannose–type N-linked carbohydrates of the recombinant fragment produced in insect cells and the C-type lectin domains of MBL. Sixty microliters of each soluble analyte at desired concentrations were injected over 13,700–14,900 resonance units (RU) of immobilized sCR1 and 2,500–6,000 RU of CR1 CCP22–30. A flow cell submitted to the coupling steps without immobilized protein was used as blank, and the specific binding signal was obtained by subtracting the background signal over the blank surface. The immobilization levels for the CR1 CCP22–30 size variants varied between 900 RU (CR1 CCP25–26) and 3700 RU (CR1 CCP22–30Δ25–26), according to the relative m.w. of each size variant. For the determination of the binding capacity of the size variants, several sensorchips were prepared, each with immobilized CR1 CCP22–30 (as a reference) and one or two more variants, given that a sensorchip has only four flow cells, with one reserved for the blank. For competition assays, MBL or L-ficolin was incubated for 20 min at room temperature with recombinant MASP-3 in HBS-P containing 2 mM $CaCl_2$ and 10 mM mannose before injection. Binding of sCR1 and CR1 CCP22–30 to immobilized MBL (13,000 RU) and L-ficolin (17,000 RU) was measured in HBS-EP at a flow rate of 20 μl/min, using immobilized BSA (10,200 RU) as reference. Regeneration of the surfaces was achieved by 10-μl injections of 1 M NaCl, 10 mM EDTA.

Data were analyzed by global fitting to a 1:1 Langmuir binding model of both the association and dissociation phases for at least five analyte concentrations simultaneously, using the BIAevaluation 3.2 software (GE Healthcare). Buffer blanks were subtracted from the data sets used for kinetic analysis (double referencing). The apparent equilibrium dissociation constants (K_D) were calculated from the ratio of the dissociation and association rate constants (k_d/k_a). χ^2 values were below 7 in all cases.

SDS-PAGE, N-terminal sequence, and MALDI-TOF mass spectrometry analyses

Recombinant CR1 CCP22–30 and its size variants were analyzed by SDS-PAGE under nonreducing or reducing conditions using Tris-HCl gels containing 10% polyacrylamide. N-terminal sequence determination was performed using an Applied Biosystems gas-phase sequencer model 492 coupled online with an Applied Biosystems Model 140C HPLC system. MALDI-TOF mass spectrometry analyses were performed using a Brucker Daltonics Autoflex mass spectrometer.

Circular dichroism spectroscopy

Circular dichroism (CD) spectra were recorded on a thermostated Jasco J-810 spectropolarimeter using a 1-mm pathlength cuvette (Hellma France, Paris, France) at a temperature of 20°C. The CR1 fragments were dialyzed against 20 mM sodium phosphate and 130 mM NaF (pH 7.3) and used at concentrations of 10–20 μM. Five spectra were acquired in the far-UV region (200–250 nm) at a scan rate of 50 nm/min.

Electron microscopy analyses

CR1 CCP22–30 samples were diluted to 20–140 nM in 50 mM Tris-HCl and 150 mM NaCl (pH 7.4), applied to carbon-coated mica, and negatively stained with 2% sodium silicotungstate (pH 7). A grid was placed on top of the carbon film, which was subsequently dried in air prior to transmission electron microscopy observation. Images were taken under low-dose conditions with a CM12 FEI electron microscope operating at 120 kV.

Results

Production and biochemical characterization of the recombinant CR1 CCP22–30 fragment

A fragment of CR1, encompassing the nine C-terminal CCP22–30 modules of the ectodomain (LHRD + CCP29–30) was produced in a baculovirus/insect cells system and purified from the cell culture supernatant by ion-exchange and size-exclusion chromatography, as described in *Materials and Methods*. The same fragment with a C-terminal 6× His-tag was also produced and purified using a one-step nickel-affinity chromatography. The mean yields of purified proteins from 1-l insect cells culture supernatant were 0.76 and 3.3 mg for the untagged and His-tagged proteins, respectively. SDS-PAGE analysis of the purified fragments yielded in both cases a single band migrating with an apparent mass of 80 kDa under reducing conditions, as illustrated for the His-tagged protein (Fig. 1A). The protein migrated much faster under nonreducing conditions, which accounts for the presence of 18 disulfide bridges (two in each CCP module). N-terminal sequence analysis of the fragments yielded the single sequence *Asp-Leu*-His-(Cys)-Lys-Thr-Pro-Glu.., corresponding to the segment His^{1395}-Glu^{1400} of human CR1 preceded by two Asp-Leu residues expected to be added at the N-terminus because of in-frame cloning with the melittin signal peptide (30). Mass spectrometry analysis yielded heterogeneous peaks centered on mass values of 70,498 ± 70 Da and 71,156 ± 70 Da for the untagged and tagged fragments, respectively, accounting for the polypeptide chains (predicted masses, 63,679 and 64,502 Da) plus extra masses (6,819 ± 70 and 6,654 ± 70 Da) likely corresponding to N-linked carbohydrates. High-mannose N-linked saccharide chains found on recombinant proteins produced by High Five cells usually comprise 2 *N*-acetylglucosamine and four to eight mannose residues (calculated masses, 1055–1704 Da). The deduced values could thus account for the presence of four to six N-linked glycans. The CCP22–30 fragment contains eight predicted *N*-glycosylation sites (34), but it is likely that all positions are not occupied by carbohydrates, in accordance with previous carbohydrate analyses on full-length CR1 suggesting *N*-glycosylation of 14 among 25 predicted sites (34).

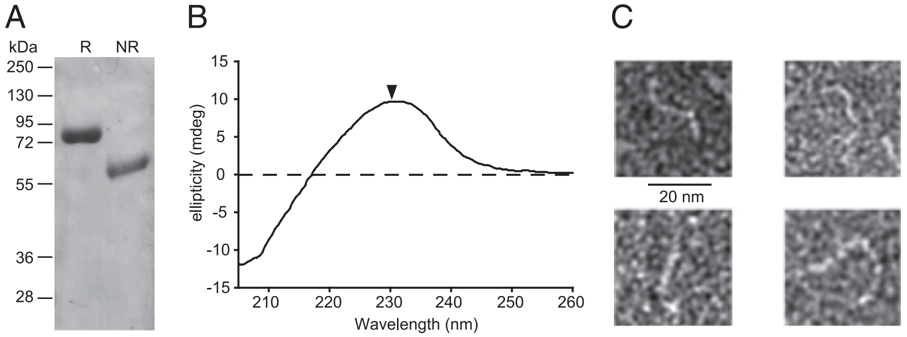


FIGURE 1. Characterization of the recombinant CR1 CCP 22–30_His fragment. (**A**) SDS-PAGE analysis of 4 μg CR1 CCP22-30 under reducing (R) and nonreducing (NR) conditions. The positions of the m.w. markers are indicated. (**B**) CD spectroscopy of CR1 CCP22–30. A spectrum was recorded in the far-UV (200–260 nm) and collected six times. The mean values for each wavelength were calculated. The maximal ellipticity is indicated by an arrowhead. (**C**) Electron microscopy analysis of CR1 CCP22–30 (×75000) after negative staining with 2% sodium silicotungstate.

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

CD spectroscopy was used to assess the correct folding of the fragments. As illustrated in Fig. 1B for the His-tagged protein, the spectrum was characterized by a maximal molar ellipticity value at ~230 nm, in accordance with CD spectra described for other complement proteins containing CCP modules, such as C4BP α-chain, complement factor H, CR2, and a CR1 fragment (CCP15–17) (35–37). The presence of the peak at 230 nm provides evidence for the correct folding of the CR1 CCP22–30 fragment. Indeed, this peak is generated by an invariant tryptophan residue located between the third and fourth Cys residues of the CCP modules (35) and is lost upon disruption of disulfide bonds (38).

Size-exclusion chromatography showed that the CCP22–30 fragments eluted much earlier than expected from their mass (at a position corresponding to a globular protein of 189 kDa; data not shown), a behavior previously observed with proteins assembled from tandem CCP modules such as factor H (39) and CR2 (37). The elongated shape of the fragment was confirmed by electron microscopy analysis of negatively stained samples showing flexible strings of beads with a length estimated to 25–30 nm, as shown in Fig. 1C for the His-tagged fragment. Given that the mean size of a single CCP module is ~3.6 nm (34), our electron microscopy data are compatible with the expected size of a molecule containing nine CCP modules.

The chemical and biophysical data presented above show that both untagged and 6× His-tagged recombinant CCP22–30 fragments secreted by insect cells are correctly folded glycoproteins with an elongated shape. Comparable results were obtained for the recombinant untagged fragment produced in 293-F mammalian cells (data not shown).

MBL and ficolins interact with immobilized sCR1 and CR1 CCP 22–30

The interaction properties of the purified recombinant CR1 CCP22–30 fragment produced in insect cells with MBL and ficolins were investigated using SPR spectroscopy and compared with those of sCR1. MBL, L-ficolin, and H-ficolin interacted with immobilized sCR1 and CCP22–30, although H-ficolin interaction was weaker, with binding levels at a concentration of 30 nM ~4- and 2-fold lower than those obtained with 10 nM L-ficolin, respectively (Fig. 2A, 2B). Comparable binding curves were obtained when using a recombinant CR1 CCP22–30 fragment produced by the human 293-F cell line (Fig. 2C), thus indicating that the interaction was independent of the nature of the N-linked glycans of the recombinant CR1 fragment (sialylated complex carbohydrates in mammalian cells versus high-mannose saccharidic chains in insect cells). In the case of L-ficolin, the interaction was fully maintained in the presence of 10 mM *N*-acetylglucosamine, a known ligand of the fibrinogen-like recognition domains, which indicates no significant contribution of these domains to the interaction with CR1 and its CCP22–30 fragment (data not shown). Moreover, both MBL and L-ficolin bound in a similar way to the untagged and His-tagged CR1 fragments produced in insect cells (Fig. 2C, 2D), thus indicating that addition of the C-terminal tag had no influence on the interaction properties of the CR1 fragment. Because the purification of the tagged fragment was faster and yielded higher amounts of purified protein, CR1 CCP22–30 with a His-tag was used for further characterization of the interaction with MBL and L-ficolin.

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

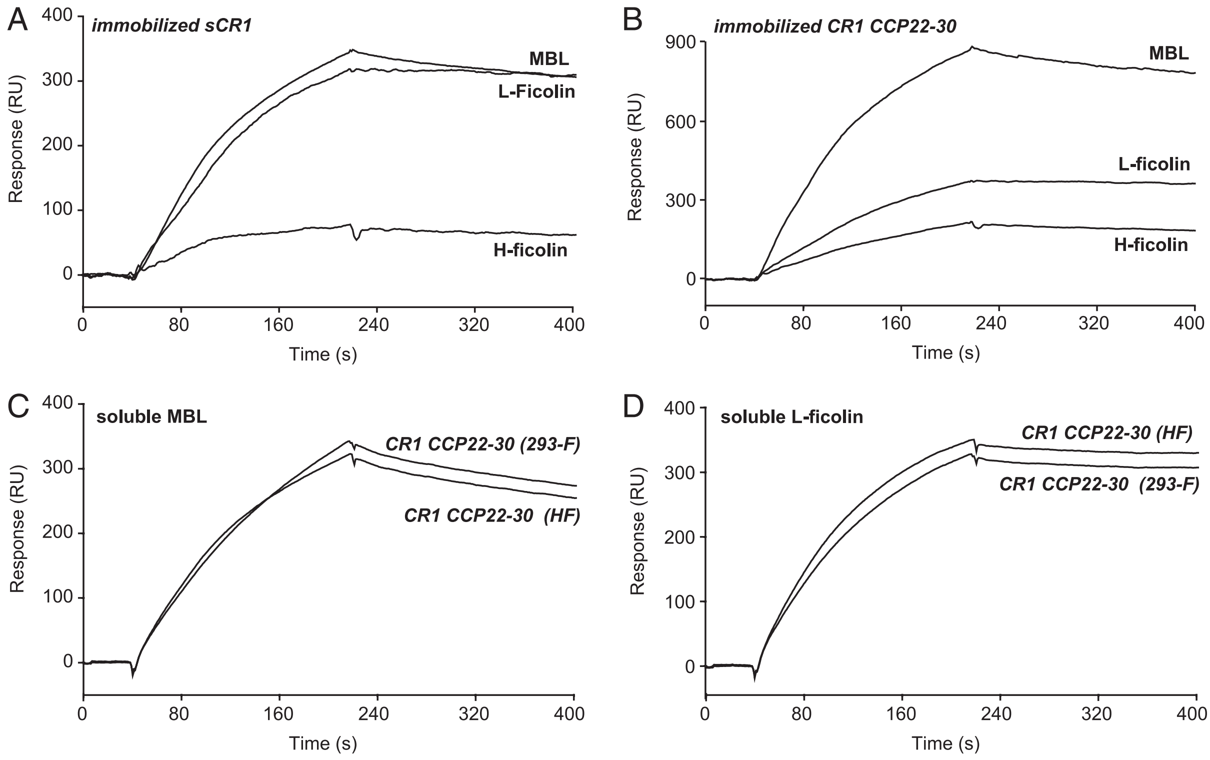


FIGURE 2. SPR analysis of the interaction of MBL, L-ficolin, and H-ficolin with immobilized sCR1 and CR1 CCP22–30. Sixty microliters of 20 nM MBL, 10 nM L-ficolin, and 30 nM H-ficolin were injected over 13,700 RU immobilized sCR1 (**A**) and 5,700 RU immobilized CR1 CCP22–30 (**B**) in 150 mM NaCl, 2 mM EDTA, 20 mM HEPES (pH 7.4), and 0.005% surfactant P20 (HBS-EP) at a flow rate of 20 μl/min. MBL (10 nM) (**C**) and L-ficolin (10 nM) (**D**) were injected over immobilized His-tagged CR1 CCP22–30 produced in High Five (HF) insect cells (3353 RU) and CR1 CCP22–30 produced in 293-F mammalian cells (4040 RU) in the above conditions. The specific binding signals shown were obtained by subtracting the background signal over a reference surface with no protein immobilized.

It has been shown previously that the MBL–sCR1 interaction was sensitive to ionic strength (13). In accordance with these observations, only 13% of MBL binding to CR1 CCP22–30 was maintained when the experiment was performed in a running buffer containing 250 mM NaCl, and the interaction was abolished in the presence of 400 mM NaCl. The interaction with L-ficolin was also salt sensitive, although to a lesser extent, because 73 and 23% of the binding were observed in the presence of 250 and 400 mM NaCl, respectively. Complete elution of bound L-ficolin could be achieved by a pulse injection of 1 M NaCl.

The kinetic parameters of the interactions were determined by recording binding of varying concentrations of MBL and L-ficolin to immobilized sCR1 and its CCP22–30 fragment (Fig. 3). Data were evaluated by global fitting using a 1:1 Langmuir interaction model as described under *Materials and Methods* and are reported in Table I. MBL and L-ficolin bound to sCR1 with comparable association and dissociation rate constants, yielding resulting apparent K_D values in the same range (0.76–0.92 nM). L-ficolin bound to sCR1 and CR1 CCP22–30 with comparable kinetic and dissociation constants, whereas MBL interacted with the fragment slightly faster than with full-length CR1. In all cases, the apparent equilibrium dissociation constants were in the nanomolar range, indicative of tight binding. It has been reported previously that MBL interacted with sCR1 with a K_D of 3.0–5.2 nM (13). Although this value is slightly higher than that determined in this study, it is compatible with our SPR results, given that the experimental settings (radioiodinated MBL and sCR1-coated plastic wells) and the MBL origin (plasma-derived) were different.

CR1 interacts with MBL and L-ficolin at or close to the MASP binding site

Previous studies have suggested that the CR1 binding site of C1q is located in its collagen-like stalks and that the C1s-C1r-C1r-C1s complex interferes with C1q–CR1 interaction (13, 40). To investigate whether CR1 binds to the homologous region within MBL and L-ficolin, we tested the ability of MASP-3 to compete with MBL and L-ficolin for binding to sCR1 and its CCP22–30 fragment. MBL and L-ficolin were preincubated in the presence of varying amounts of recombinant MASP-3, and the mixture was injected over immobilized sCR1 and CR1 CCP22–30. Because the interaction of the MASP-3 dimer with MBL and ficolins is Ca^{2+} dependent (32), the experiments were performed in the presence of 2 mM $CaCl_2$ and 10 mM mannose to prevent interaction of CR1 CCP22–30 with the lectin domain of MBL. Recombinant MASP-3 was produced in a mammalian expression system (see *Materials and Methods*) for the same reason. As shown in Fig. 4A, MASP-3 clearly inhibited MBL binding to sCR1 and its CCP22–30 fragment in a dose-dependent manner, reaching almost maximal inhibition on both surfaces at a MBL:MASP-3 dimer molar ratio of 1. Comparable results were obtained when performing the competition experiments using sCR1 and the recombinant CCP22–30 fragment produced in mammalian cells in the presence of Ca^{2+} and in the absence of mannose (data not shown), thus confirming that the MBL lectin domain was not involved in the interaction. Similar results were obtained with L-ficolin (Fig. 4B), except that the maximal inhibition was ~70 and 80% for sCR1 and its fragment, respectively. MASP-3 only partially interfered with H-ficolin binding to immobilized sCR1 and CCP22–30, with maximal inhibition of 44% (sCR1) and 38% (CCP22–30) at a H-ficolin:MASP-3 dimer molar ratio of 2 (data not shown).

We have previously generated MBL mutants devoid of MASP-binding capacity after mutation of the Lys55 residue of the collagenous region (25). The ability of the MBL K55A and K55E mutants to bind immobilized sCR1 and CR1 CCP22–30 was next compared with that of wild-type MBL. As shown in Fig. 5, no detectable interaction was obtained with these mutants, providing further support to the hypothesis that CR1 binds at or close to the MASP binding site of MBL. Similar results were obtained when

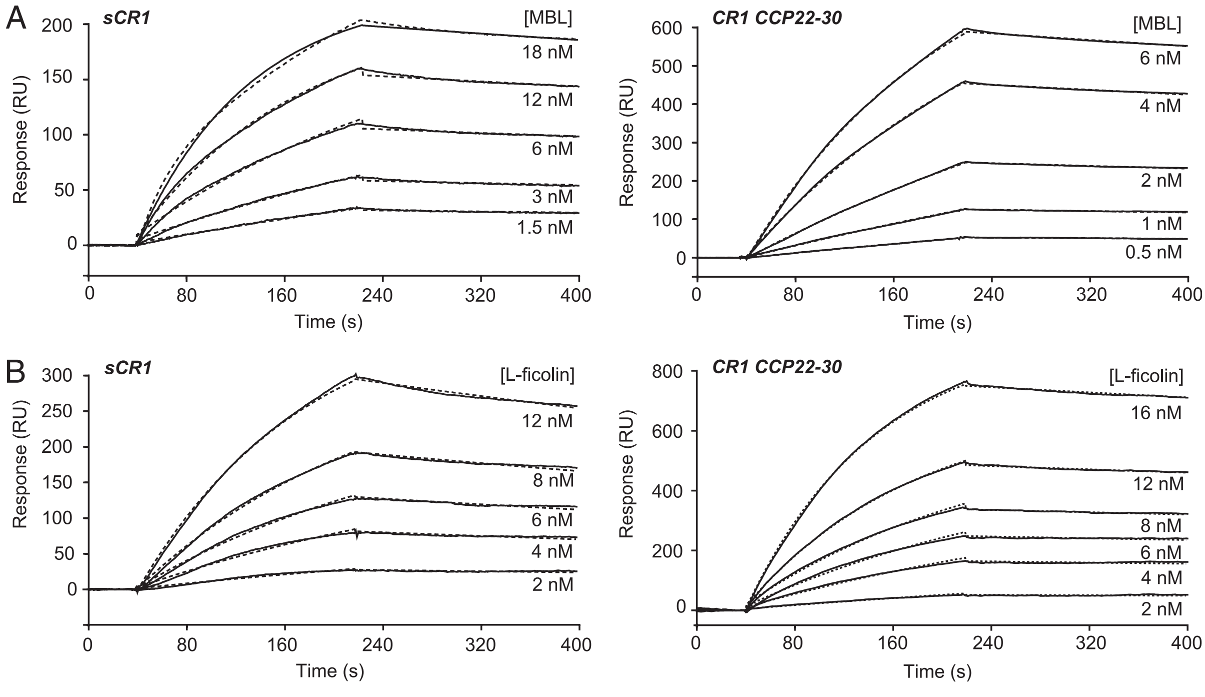


FIGURE 3. Kinetic analysis of the interaction between MBL or L-ficolin and immobilized sCR1 or CR1 CCP22-30. (**A**) MBL was injected at the indicated concentrations over immobilized sCR1 (14,700 RU) or CR1 CCP22–30 (5,200 RU) in HBS-EP. (**B**) L-ficolin was injected at the indicated concentrations over immobilized sCR1 (14,700 RU) and CR1 CCP22–30 (2,500 RU) in HBS-EP and HBS-P, respectively. Fits are shown as dotted lines and were obtained by global fitting of the data using a 1:1 Langmuir binding model. The results shown are representative of three to eight experiments (see number of experiments in Table I).

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

Table I. Kinetic and dissociation constants for binding of MBL and L-ficolin to immobilized sCR1 and CR1 CCP22–30

	Immobilized sCR1				Immobilized CR1 CCP22–30			
	k_a (M^{-1} s^{-1})	k_d (s^{-1})	K_D (nM)	n^a	k_a (M^{-1} s^{-1})	k_d (s^{-1})	K_D (nM)	n^a
MBL	$7.15 \pm 0.91 \times 10^5$	$5.19 \pm 0.54 \times 10^{-4}$	0.76 ± 0.07	8	$1.35 \pm 0.37 \times 10^6$	$4.93 \pm 1.14 \times 10^{-4}$	0.37 ± 0.01	3
L-ficolin	$7.72 \pm 0.60 \times 10^5$	$6.65 \pm 0.99 \times 10^{-4}$	0.92 ± 0.17	6	$7.30 \pm 1.71 \times 10^5$	$3.80 \pm 0.57 \times 10^{-4}$	0.63 ± 0.20	4

Values are expressed as mean ± SE.

[a] Number of separate experiments on different sensorchips

the MBL K55E mutant was injected over sCR1 in the presence of Ca^{2+}, thus providing further evidence that the carbohydrate recognition domains of MBL do not contribute to the interaction.

Location of the binding site in CR1 CCP22–30

With a view to determine which CCP modules interact with MBL and ficolins, we generated a set of seven deletion mutants by removing CCP modules at both extremities and in the central part of the CR1 CCP22–30 segment (Fig. 6A). The mutated proteins were produced using a baculovirus–insect cells system and purified in the same way as CR1 CCP22–30, using Ni^{2+}-affinity chromatography. SDS-PAGE and mass spectrometry analyses indicated that all fragments were glycosylated, given the diffuse character of the bands (Fig. 6B) and their measured molecular masses (Table II). Although the accuracy of the mass measurements and the glycosylation heterogeneity preclude the exact determination of the number of occupied N-linked glycosylation sites and their location within CCP modules, some information can be inferred from the deduced masses of the glycosylated moieties of the fragments (Table II). For example, the CCP22–23 and 25–26 module pairs seem to contain only one N-linked glycan, which means that only one of the two potential sites (Fig. 6A) is occupied. Comparison of the various fragments also shows unambiguously that the CCP23, CCP24, and CCP27 modules are glycosylated, but it is difficult to draw conclusions about the presence of one or two N-linked glycans in CCP module 24. CD analysis of the fragments yielded in all cases spectra with ellipticity maxima centered on 223–230 nm, typical of CCP modules (41), thus indicating that the deletions did not induce significant changes in the secondary structure of the fragments (data not shown).

The ability of the deletion mutants to interact with MBL and L-ficolin was investigated using SPR spectroscopy, as described for the parent CR1 CCP22–30 molecule. Equimolar amounts of the different constructs were immobilized and the binding signals of MBL and L-ficolin, recorded as described under *Materials and Methods*, are presented in Supplemental Fig. 1. A summary of the data, expressed as percentage of the reference signal obtained with immobilized CCP22–30, is presented in Fig. 6C. Deletion of the CCP modules located at the N- or C-terminal ends (ΔCCP22–23 and ΔCCP27–30) had no or little impact on the MBL and L-ficolin binding capacities. In contrast, removal of the central CCP module pairs 24–25 (ΔCCP22–25) and 25–26 (ΔCCP25–26) reduced the interaction in a significant manner. Interestingly, the CCP22–25 fragment (ΔCCP26–30) had a higher binding capacity than the CCP22–26 (ΔCCP27–30) fragment, suggesting that deletion of

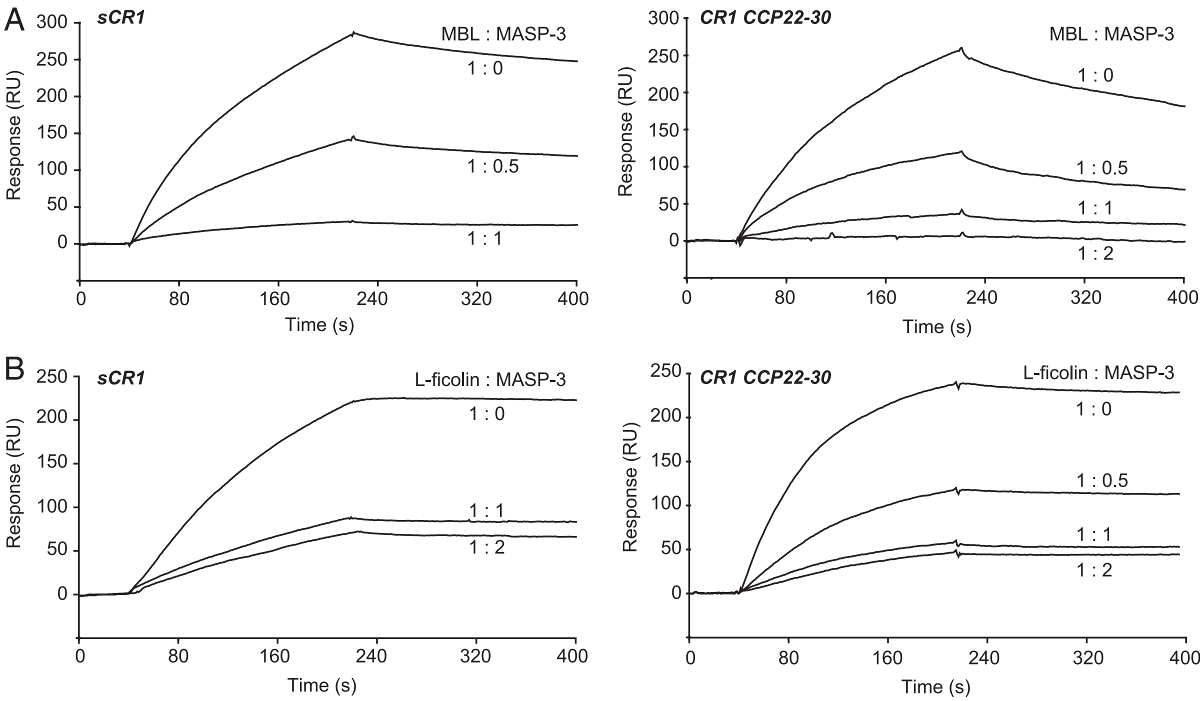


FIGURE 4. MASP-3 prevents the interaction between MBL or L-ficolin and immobilized sCR1 and CR1 CCP22–30. MBL (**A**) or L-ficolin (**B**) was incubated 20 min at room temperature in the absence or presence of recombinant MASP-3 at the indicated molar ratios. The mixture was then injected in HBS-P containing 2 mM $CaCl_2$ (HBS-CaP) over immobilized sCR1 (14,900 RU) and in HBS-CaP containing 10 mM mannose over immobilized CR1 CCP22–30 (3,400 RU for MBL and 4,300 RU for L-ficolin). The concentrations of soluble analytes were as follows: MBL, 20 nM (sCR1) and 10 nM (CCP22–30); L-ficolin, 5 nM (sCR1) and 10 nM (CCP22–30). MBL and L-ficolin are considered as tetramers and MASP-3 as a dimer.

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

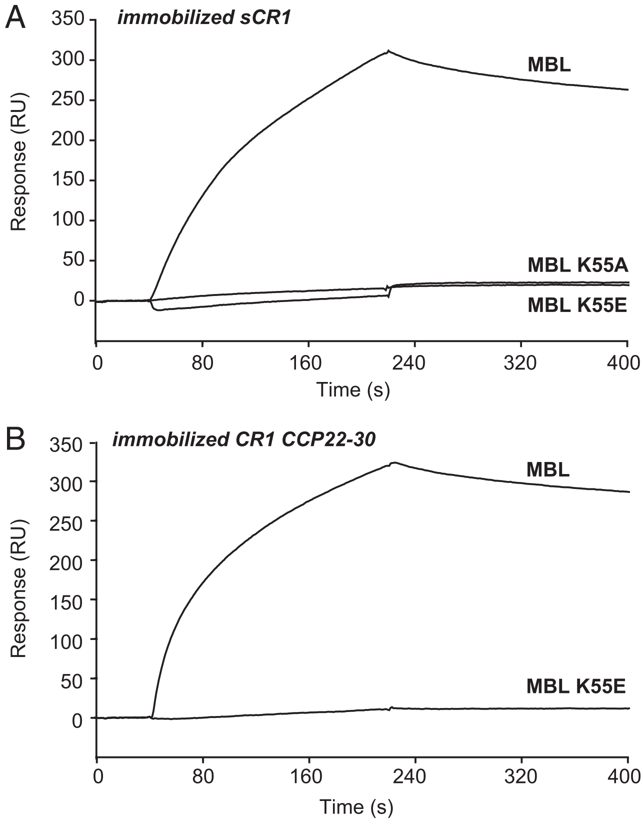


FIGURE 5. Lys^{55} of MBL is essential for the interaction with sCR1 and CR1 CCP22–30. The binding of the K55E or K55A MBL mutants was analyzed by SPR using wild-type MBL as a reference. Each analyte was injected over immobilized sCR1 (14,900 RU) (**A**) and immobilized CR1 CCP22–30 (6,000 RU) (**B**) in 150 mM NaCl, 50 mM triethanolamine-HCl (pH 7.4), and 0.005% surfactant P20 at the following concentrations: 20 nM (sCR1) and 10 nM (CCP22–30).

the CCP26 module enhanced the interaction with both MBL and L-ficolin. These results suggest that the binding site for MBL and L-ficolin involves both CCP 24 and 25 of CR1. The fact that the CCP25–26 fragment has a low binding capacity suggests that the CCP24 module plays an essential role in the interaction. However, it should be kept in mind that covalent immobilization of the short CCP25–26 fragment on the sensor chip might restrict access to the binding site for MBL and L-ficolin. Surprisingly, deletion of CCP29–30 resulted in an important decrease in the binding signal. It seems unlikely that these modules are directly involved in the interaction because other fragments devoid of these modules have a binding capacity comparable to or even higher than that of the CCP22–30 fragment. A hypothesis might be that the conformation of the immobilized CCP22–28 fragment is not favorable to the interaction with MBL and L-ficolin. It should also be mentioned that, although MBL and L-ficolin yielded essentially comparable interaction profiles, small differences were observed, the most noteworthy regarding ΔCCP27–30. Indeed, removal of CCP27–30 yielded reduced L-ficolin binding (61% of the reference value) but had no significant effect on the interaction with MBL (Fig. 6C, Supplemental Fig. 1), which suggests slight differences in the binding sites for MBL and L-ficolin.

Discussion

We show in this study for the first time, to our knowledge, that, in addition to MBL, ficolins are ligands of CR1. Both L- and H-ficolins interacted with immobilized sCR1, although the binding level of H-ficolin was much weaker (Fig. 2A, 2B). This observation may be related to the shorter collagen-like segment of H-ficolin (11 GXY triplets compared with 15 for L-ficolin and 19 for MBL). Indeed, longer collagen-like regions are expected to bring more flexibility, which could facilitate interactions involving the collagen stalks. Although no structural data are available regarding soluble ficolins, recent studies using solution X-ray scattering of MBL oligomers revealed highly flexible molecules with near planar fan-like structures in some instances (42, 43). Moreover, MBL bound to surface ligands was shown to adopt a stretched conformation, as shown by atomic force microscopy (44). MBL and H-ficolin bound more to the CR1 CCP22–30 fragment than to full-length CR1, which might result from increased accessibility of their binding site in the absence of the rest of the CR1 molecule. However, the fact that comparable binding levels were obtained in the case of L-ficolin likely reflects slightly different requirements for CR1 binding among the three recognition proteins.

L-ficolin and MBL bound to sCR1 with comparable nanomolar affinities, in a range similar to that previously reported for MBL (13). Previous studies have provided evidence for a single CR1 binding site for C1q and MBL, likely located in a segment encompassing the nine C-terminal CCP modules of the ectodomain (LHR-D + CCP29–30), thus differing from the binding sites for C3b/C4b located in LHR-A, -B, and -C (12, 13). However, these experiments were performed by capturing the C-terminal recombinant CR1 fragment from lysates of transfected Chinese hamster ovary cells using an appropriate Ab coated on an ELISA plate. To investigate the MBL/L-ficolin–CR1 interaction at the molecular level, the CCP22–30 fragment of CR1 was expressed in a recombinant form in eukaryotic cells, which allowed production of a glycosylated fragment and subsequent use of the purified protein in SPR interaction studies. MBL and L-ficolin were found to bind to immobilized CR1 CCP22–30 with affinities in the same range as those obtained with sCR1. In accordance with previous studies (13), the interaction with MBL was observed in the presence of EDTA and was thus not contingent upon the Ca^{2+}-dependent lectin activity of MBL. It should be mentioned that we were not able to observe the CR1–MBL/L-ficolin interaction in the reverse configuration of the SPR experiments (i.e., using immobilized MBL or L-ficolin and soluble CR1 or its CCP22–30 fragment) up to concentrations of 100 nM and 1 μM, respectively (data not shown) (see also *Materials and Methods*). This might arise from the process of covalent immobilization of the ligand onto the sensor chip, possibly restricting access to the soluble interactant or putative conformational changes associated with the binding process, as previously described in the case of calreticulin, the receptor for the collagen fragments of collectins (25). Oriented capture of MBL might also favor the interaction, because solid-phase binding of sCR1 to MBL immobilized through its CRD domain was observed previously, although the affinity was ~10 times lower than that obtained in the reverse configuration (13). It should also be mentioned that the SPR configuration used allows multivalent interactions between oligomeric MBL/ficolin molecules and immobilized CR1 or its CCP22–30 fragment, which are likely strengthened through an avidity phenomenon. This might have physiological consequences because avidity is known to play a dominant role in immune recognition and ligand–receptor interactions, as underlined in a recent review (45).

With a view to explore the CR1–MBL interaction in a cellular context, we also transfected 293-F cells with a plasmid containing a CR1 fragment encompassing CCP modules 22–30, the transmembrane segment and the intracytoplasmic tail. The CR1 fragment could be detected at the cell surface by flow cytometry and

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

P17

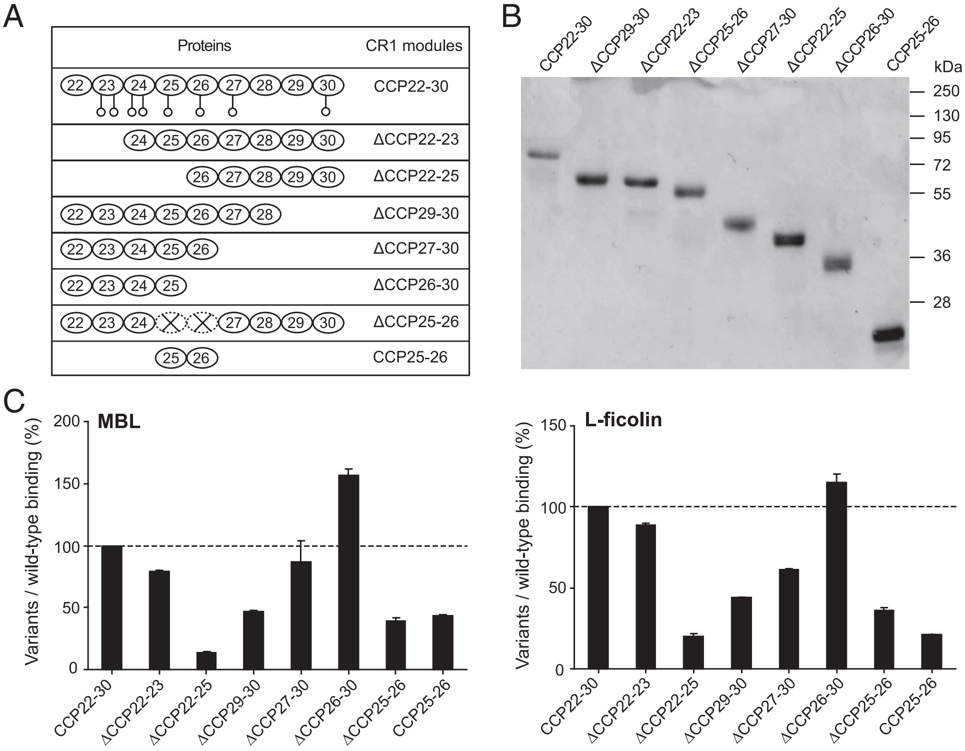


FIGURE 6. Location of the MBL/L-ficolin binding site in CR1 CCP22–30 using deletion mutants. (**A**) Schematic view of the size variants of CR1 CCP22–30. Potential N-glycosylation sites are indicated by ○. (**B**) SDS-PAGE analysis of the reduced form of CR1 CCP22–30 and its size variants (2 μg). The positions of the m.w. markers are indicated. (**C**) Binding of MBL (*left panel*) and L-ficolin (*right panel*) to the CCP22–30 size variants was analyzed by SPR. The CCP22–30 variants were immobilized on CM5 sensor chips and 10 nM MBL and L-ficolin were injected over the surfaces as described in *Materials and Methods*. Results are expressed in percentage (%) of CR1 CCP22-30 binding (mean ± SE of three injections). The binding data (in resonance units) used to calculate the percentages are presented in Supplemental Fig. 1.

immunostaining, but we obtained no evidence for specific MBL binding to the transfected cells (data not shown). However, ligand binding to cell surface CR1 is known to induce clustering of the receptor (46), which may not happen at the surface of transiently transfected 293-F cells.

It has been shown previously that the C1s-C1r-C1r-C1s tetramer interferes with C1q binding to CR1, suggesting a steric hindrance effect or direct competition with the receptor for a CCP binding site on the collagen stalks of C1q (40). It was therefore of interest to investigate whether the MASP dimers had a similar effect on the binding of MBL/ficolins to CR1 and its CCP22–30 fragment. Preincubation with MASP-3 was shown to inhibit binding in all cases, although to a lesser extent for H-ficolin, with maximal inhibition obtained at an equimolar ratio of MBL or L-ficolin (considered as tetramers) and MASP-3 dimers. These results are in accordance with our previous findings that isolated trimeric and tetrameric forms of serum-derived MBL are able to accommodate only one MASP-3 dimer (47). Moreover, the affinities of recombinant MASP-3 produced in mammalian cells for MBL and L-ficolin are in the namolar range (1.0 and 1.1 nM, respectively, for the MASP-3 monomer and thus 0.5–0.55 nM for the dimer) (N.M. Thielens, unpublished observations), which indicates that most available MBL and L-ficolin are complexed with MASP-3 at the concentrations used in our experiments (10–20 nM MBL and 5–10 nM L-ficolin). These data thus strongly suggest that the CR1 binding site is located at or near the sites occupied by the MASPs on the collagen stalks of MBL and L-ficolin. This hypothesis is further supported by the fact that mutations of MBL lysine 55, which impair binding to the MASPs (25), also abolish the interaction of MBL with CR1 and its CCP22–30 fragment.

Interestingly, similar observations have been made previously regarding the interaction of MBL and ficolins with two other cell surface MBL/ficolin-binding proteins, calreticulin (25, 48) and the low-density lipoprotein receptor-related protein CD91 (24). Whereas the MBL/L-ficolin binding sites in CD91 and CRT remain to be identified, site-directed mutagenesis and structural studies demonstrated that the MBL–MASP interface involves major ionic interactions between acidic Ca^{2+} ligands of the MASP C1r/C1s, sea urchin epidermal growth factor, bone morphogenetic protein domains, and a conserved lysine residue of the collagen-like region of MBL (25–27, 49, 50). Considering also the fact that the interactions are sensitive to high ionic strength, these results

Table II. Summary of the theoretical and experimental masses of the recombinant fragments derived from CR1 CCP22–30

CR1 Fragment	Theoretical Mass	Experimental Mass	Δ[a]	No. of Predicted Glycosylation Sites
CCP22–30	64,502	71,156 ± 70	6,654	8
ΔCCP22–23	50,978	56,132 ± 56	5,154	6
ΔCCP22–25	36,493	38,894 ± 39	2,401	3
ΔCCP29–30	50,470	56,567 ± 56	6,097	7
ΔCCP27–30	35,835	40,107 ± 40	4,272	6
ΔCCP25–30	29,234	32,708 ± 33	3,474	5
ΔCCP25–26	51,163	57,083 ± 57	5,920	6
CCP25–26	14,408	15,434 ± 15	1,026	2

[a] Difference between the experimental mass determined by mass spectrometry and the calculated mass of the polypeptide chain.

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

altogether strongly suggest a contribution of ionic bonds in the MBL/ficolin–CR1 interaction, likely involving the conserved lysine residue of the MASP binding site in the collagen stalks and surface-exposed acidic residues of CR1.

With a view to delineate the MBL/L-ficolin binding site within the CR1 CCP22–30 segment, we generated a series of deletion mutants and investigated their MBL/L-ficolin binding properties. The fragments devoid of modules 24 and/or 25 had a reduced binding capacity compared with the full-length segment, whereas the deletion of the other modules had less impact. On the basis of these data, we propose that CCPs 24 and 25 contain the interaction site of CR1 CCP22–30 with the collagen stalks of MBL and L-ficolin. We aligned the sequences of the 30 CCP modules of CR1 (Supplemental Fig. 2) to search for specific acidic residues of CCP24 and 25 potentially involved in MBL/ficolin binding. We have identified Glu^{1555}, a CCP24 residue located in a loop, which is shared by a limited number of CCP modules and contains a glutamic acid in CCP24 exclusively. In addition, this residue may be part of a patch of negative charges contributed by two additional acidic residues located nearby, Asp^{1553} and Glu^{1559} (Supplemental Fig. 2). Regarding CCP25, we have identified two close glutamate residues (Glu^{1595} and Glu^{1597}) (Supplemental Fig. 2). Although the presence of these negatively charged residues, taken individually, is not specific to CCP25, such proximity is unique to CCP25.

The CCP modules of CR1 have been classified into four different types, based on homology criteria (17). Interestingly, the CCPs 22–25 are unique and divergent from the other CCPs of CR1 (17, 51). This divergence suggests a specific role for these modules, which could be in accordance with our experimental hypothesis regarding the involvement of CCPs 24 and 25 in the interaction with MBL and L-ficolin. It is also noteworthy that CCPs 24 and 25 contain all of the known Knops blood group polymorphisms (22). The ethnic-specific distribution of two of these blood group variants, Sl1/2 ($\mathrm{Lys/Glu}^{1590}$ at the CCP24–25 junction) and $\mathrm{McC}^{a/b}$ ($\mathrm{Arg/Gly}^{1601}$ in CCP25) has been proposed to result from selective pressures in malarial endemic regions of Africa. However, contradictory results have been obtained regarding their association with malarial severity in different African populations (52–55). CR1 interacts with two *P. falciparum* proteins, the malarial adhesin PfEMP1 (14) and the PfRh4 invasion ligand (15), but the binding sites have been located to the C3b and C4b binding sites of CR1, respectively (20, 21). Consistent with these locations, Tetteh-Quarcoo et al. (56) observed no difference between the Knops blood group variants $\mathrm{McC}^{a/b}$ and Sl1/2 in the ability of the CR1 fragment CCP15–25 to interact with PfRh4 or to disrupt PfEMP1-dependent *P. falciparum* rosetting (56). It should be mentioned that the same study assigned the C1q binding site to a CR1 region encompassing CCPs 15–25, with a suggested major contribution of CCPs 15–17, using recombinant CR1 fragments produced in yeasts and treated with endoglycosidase H (56). However, binding of MBL or L-ficolin was not reported in that study. Interestingly, MBL has been shown previously to interact with *P. falciparum* through its carbohydrate recognition domain (57, 58), raising the possibility of MBL acting as an opsonin for uptake of the malaria parasite by uninfected CR1-bearing erythrocytes.

CR1 has also been reported to facilitate the entry into host monocytes/macrophages of various infectious agents, including *Leishmania major* (59, 60) and *Mycobacterium tuberculosis* (61). Interestingly, MBL has been shown to bind to both pathogens (60, 62), whereas L-ficolin binding was demonstrated in the case of *M. tuberculosis* (62). The molecular basis of the interaction of CR1 with these pathogens has not been established yet, but these observations raise the possibility of a bridging role of the defense collagens between the pathogens and CR1-bearing host cells. However, it cannot be excluded that the interaction of MBL/L-ficolin with the pathogen might trigger activation of the lectin pathway, leading to C3 cleavage and subsequent interaction of the C3b-opsonized pathogen with CR1.

Independently of its role in infection, CR1 has been identified in recent genome-wide association studies as a susceptibility locus in the late-onset form of Alzheimer's disease (63). It has been proposed that CR1 may participate in Aβ peptides clearance, which is supported by the observation that circulating Aβ42 adheres to erythrocytes, a process that is impaired in Alzheimer's disease cases (64). In a recent study, Keenan et al. (65) reported that a CR1 polymorphism resulting in substitution of a threonine for a serine residue at position 1610 of CCP25 of CR1 was associated with episodic memory decline in Alzheimer's disease. The authors suggested that this mutation could reduce the affinity of CR1 for C1q and/or MBL, thus impairing the clearance of opsonized Aβ, an assumption that is supported by the fact that both C1q and MBL interact with amyloid β protein (66, 67). We have located the MBL binding site in the CCP24–25 segment of CR1 in the current study, which could support the above assumption. However, it appears unlikely at first sight that the conservative serine/threonine substitution could destabilize the structure of CCP25 or influence the CR1–MBL interaction, which appears to be driven by electrostatic forces. Nevertheless, a new conformational epitope might arise from this change as suggested by the production of specific Ags (68). Site-directed mutagenesis studies will be clearly needed to solve this issue.

In summary, our results reveal that ficolins are novel CR1 ligands and that MBL and L-ficolin share homologous, although not identical CR1 binding sites at or near the known binding site for their associated proteases and for previously described collectin receptors such as calreticulin and CD91. CR1 binding should thus take place in the absence of the MASPs, which are circulating in serum in association with MBL and ficolins. Whether the MASPs could be released from these complexes by C1 inhibitor following complement activation, in the same way as the C1s-C1r-C1r-C1s tetramer from the C1 complex, is presently not known. Alternatively, CR1 might interact with locally synthesized MBL and ficolins. MBL/L-ficolin binding likely involves major ionic interactions between conserved lysine residues of their collagen stalks and surface-exposed acidic residues located in CR1 CCP24 and/or CCP25, which remain to be identified. Polymorphisms resulting in substitution of charged residues like some of the Knops blood groups might result in changes in the surface electrostatic properties of CCP modules and impair electrostatic interactions involving these residues or their close vicinity (69).

Acknowledgments

We thank Prof. Jacques Cohen for providing the vector containing the CR1 coding sequence. We thank Isabelle Bally, Julien Pérard, and Françoise Lacroix and Jean-Philippe Kleman from the Institut de Biologie Structurale/Unit of Virus Host Cell Interactions platforms of the Partnership for Structural Biology in Grenoble for assistance and access to the SPR, Biophysics (CD), and flow cytometry facilities, respectively. We also thank Jean-Pierre Andrieu, Luca Signor, and Drs. Chistine Moriscot and Guy Schoehn from the Institut de Biologie Structurale/Unit of Virus Host Cell Interactions platforms of the Partnership for Structural Biology in Grenoble for the protein sequencing, mass spectrometry, and electron microscopy analyses.

Disclosures

The authors have no financial conflicts of interest.

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

P17

References

1. Fujita, T., M. Matsushita, and Y. Endo. 2004. The lectin-complement pathway—its role in innate immunity and evolution. *Immunol. Rev.* 198: 185–202.
2. Thiel, S., and M. Gadjeva. 2009. Humoral pattern recognition molecules: mannan-binding lectin and ficolins. *Adv. Exp. Med. Biol.* 653: 58–73.
3. Dommett, R. M., N. Klein, and M. W. Turner. 2006. Mannose-binding lectin in innate immunity: past, present and future. *Tissue Antigens* 68: 193–209.
4. Hansen, S., L. Selman, N. Palaniyar, K. Ziegler, J. Brandt, A. Kliem, M. Jonasson, M. O. Skjoedt, O. Nielsen, K. Hartshorn, et al. 2010. Collectin 11 (CL-11, CL-K1) is a MASP-1/3-associated plasma collectin with microbial-binding activity. *J. Immunol.* 185: 6096–6104.
5. Garred, P., C. Honoré, Y. J. Ma, S. Rørvig, J. Cowland, N. Borregaard, and T. Hummelshøj. 2010. The genetics of ficolins. *J. Innate Immun.* 2: 3–16.
6. Gaboriaud, C., F. Teillet, L. A. Gregory, N. M. Thielens, and G. J. Arlaud. 2007. Assembly of C1 and the MBL- and ficolin-MASP complexes: structural insights. *Immunobiology* 212: 279–288.
7. Héja, D., A. Kocsis, J. Dobó, K. Szilágyi, R. Szász, P. Závodszky, G. Pál, and P. Gál. 2012. Revised mechanism of complement lectin-pathway activation revealing the role of serine protease MASP-1 as the exclusive activator of MASP-2. *Proc. Natl. Acad. Sci. USA* 109: 10498–10503.
8. Bally, I., V. Rossi, T. Lunardi, N. M. Thielens, C. Gaboriaud, and G. J. Arlaud. 2009. Identification of the C1q-binding Sites of human C1r and C1s: a refined three-dimensional model of the C1 complex of complement. *J. Biol. Chem.* 284: 19340–19348.
9. Phillips, A. E., J. Toth, A. W. Dodds, U. V. Girija, C. M. Furze, E. Pala, R. B. Sim, K. B. Reid, W. J. Schwaeble, R. Schmid, et al. 2009. Analogous interactions in initiating complexes of the classical and lectin pathways of complement. *J. Immunol.* 182: 7708–7717.
10. Fearon, D. T. 1980. Identification of the membrane glycoprotein that is the C3b receptor of the human erythrocyte, polymorphonuclear leukocyte, B lymphocyte, and monocyte. *J. Exp. Med.* 152: 20–30.
11. Krych-Goldberg, M., and J. P. Atkinson. 2001. Structure-function relationships of complement receptor type 1. *Immunol. Rev.* 180: 112–122.
12. Klickstein, L. B., S. F. Barbashov, T. Liu, R. M. Jack, and A. Nicholson-Weller. 1997. Complement receptor type 1 (CR1, CD35) is a receptor for C1q. *Immunity* 7: 345–355.
13. Ghiran, I., S. F. Barbashov, L. B. Klickstein, S. W. Tas, J. C. Jensenius, and A. Nicholson-Weller. 2000. Complement receptor 1/CD35 is a receptor for mannan-binding lectin. *J. Exp. Med.* 192: 1797–1808.
14. Rowe, J. A., J. M. Moulds, C. I. Newbold, and L. H. Miller. 1997. *P. falciparum* rosetting mediated by a parasite-variant erythrocyte membrane protein and complement-receptor 1. *Nature* 388: 292–295.
15. Tham, W. H., D. W. Wilson, S. Lopaticki, C. Q. Schmidt, P. B. Tetteh-Quarcoo, P. N. Barlow, D. Richard, J. E. Corbin, J. G. Beeson, and A. F. Cowman. 2010. Complement receptor 1 is the host erythrocyte receptor for *Plasmodium falciparum* PfRh4 invasion ligand. *Proc. Natl. Acad. Sci. USA* 107: 17327–17332.
16. Hourcade, D., D. R. Miesner, J. P. Atkinson, and V. M. Holers. 1988. Identification of an alternative polyadenylation site in the human C3b/C4b receptor (complement receptor type 1) transcriptional unit and prediction of a secreted form of complement receptor type 1. *J. Exp. Med.* 168: 1255–1270.
17. Klickstein, L. B., T. J. Bartow, V. Miletic, L. D. Rabson, J. A. Smith, and D. T. Fearon. 1988. Identification of distinct C3b and C4b recognition sites in the human C3b/C4b receptor (CR1, CD35) by deletion mutagenesis. *J. Exp. Med.* 168: 1699–1717.
18. Klickstein, L. B., W. W. Wong, J. A. Smith, J. H. Weis, J. G. Wilson, and D. T. Fearon. 1987. Human C3b/C4b receptor (CR1): demonstration of long homologous repeating domains that are composed of the short consensus repeats characteristics of C3/C4 binding proteins. *J. Exp. Med.* 165: 1095–1112.
19. Krych, M., D. Hourcade, and J. P. Atkinson. 1991. Sites within the complement C3b/C4b receptor important for the specificity of ligand binding. *Proc. Natl. Acad. Sci. USA* 88: 4353–4357.
20. Rowe, J. A., S. J. Rogerson, A. Raza, J. M. Moulds, M. D. Kazatchkine, K. Marsh, C. I. Newbold, J. P. Atkinson, and L. H. Miller. 2000. Mapping of the region of complement receptor (CR) 1 required for *Plasmodium falciparum* rosetting and demonstration of the importance of CR1 in rosetting in field isolates. *J. Immunol.* 165: 6341–6346.
21. Tham, W. H., C. Q. Schmidt, R. E. Hauhart, M. Guariento, P. B. Tetteh-Quarcoo, S. Lopaticki, J. P. Atkinson, P. N. Barlow, and A. F. Cowman. 2011. *Plasmodium falciparum* uses a key functional site in complement receptor type-1 for invasion of human erythrocytes. *Blood* 118: 1923–1933.
22. Moulds, J. M. 2010. The Knops blood-group system: a review. *Immunohematol.* 26: 2–7.
23. Stuart, G. R., N. J. Lynch, A. J. Day, W. J. Schwaeble, and R. B. Sim. 1997. The C1q and collectin binding site within C1q receptor (cell surface calreticulin). *Immunopharmacology* 38: 73–80.
24. Duus, K., N. M. Thielens, M. Lacroix, P. Tacnet, P. Frachet, U. Holmskov, and G. Houen. 2010. CD91 interacts with mannan-binding lectin (MBL) through the MBL-associated serine protease-binding site. *FEBS J.* 277: 4956–4964.
25. Lacroix, M., C. Dumestre-Pérard, G. Schoehn, G. Houen, J. Y. Cesbron, G. J. Arlaud, and N. M. Thielens. 2009. Residue Lys^{57} in the collagen-like region of human L-ficolin and its counterpart Lys^{47} in H-ficolin play a key role in the interaction with the mannan-binding lectin-associated serine proteases and the collectin receptor calreticulin. *J. Immunol.* 182: 456–465.
26. Teillet, F., M. Lacroix, S. Thiel, D. Weilguny, T. Agger, G. J. Arlaud, and N. M. Thielens. 2007. Identification of the site of human mannan-binding lectin involved in the interaction with its partner serine proteases: the essential role of Lys55. *J. Immunol.* 178: 5710–5716.
27. Teillet, F., C. Gaboriaud, M. Lacroix, L. Martin, G. J. Arlaud, and N. M. Thielens. 2008. Crystal structure of the CUB1-EGF-CUB2 domain of human MASP-1/3 and identification of its interaction sites with mannan-binding lectin and ficolins. *J. Biol. Chem.* 283: 25715–25724.
28. Wang, W., and B. A. Malcolm. 1999. Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange Site-Directed Mutagenesis. *Biotechniques* 26: 680–682.
29. Oudin, S., M. T. Libyh, D. Goossens, X. Dervillez, F. Philbert, B. Réveil, F. Bougy, T. Tabary, P. Rouger, D. Klatzmann, and J. H. Cohen. 2000. A soluble recombinant multimeric anti-Rh(D) single-chain Fv/CR1 molecule restores the immune complex binding ability of CR1-deficient erythrocytes. *J. Immunol.* 164: 1505–1513.
30. Rossi, V., I. Bally, N. M. Thielens, A. F. Esser, and G. J. Arlaud. 1998. Baculovirus-mediated expression of truncated modular fragments from the catalytic region of human complement serine protease C1s: evidence for the involvement of both complement control protein modules in the recognition of the C4 protein substrate. *J. Biol. Chem.* 273: 1232–1239.
31. Thielens, N. M., S. Cseh, S. Thiel, T. Vorup-Jensen, V. Rossi, J. C. Jensenius, and G. J. Arlaud. 2001. Interaction properties of human mannan-binding lectin (MBL)-associated serine proteases-1 and -2, MBL-associated protein 19, and MBL. *J. Immunol.* 166: 5068–5077.
32. Zundel, S., S. Cseh, M. Lacroix, M. R. Dahl, M. Matsushita, J. P. Andrieu, W. J. Schwaeble, J. C. Jensenius, T. Fujita, G. J. Arlaud, and N. M. Thielens. 2004. Characterization of recombinant mannan-binding lectin-associated serine protease (MASP)-3 suggests an activation mechanism different from that of MASP-1 and MASP-2. *J. Immunol.* 172: 4342–4350.
33. Arlaud, G. J., R. B. Sim, A. M. Duplaa, and M. G. Colomb. 1979. Differential elution of Clq, Clr and Cls from human Cl bound to immune aggregates: use in the rapid purification of Cl subcomponents. *Mol. Immunol.* 16: 445–450.
34. Furtado, P. B., C. Y. Huang, D. Ihyembe, R. A. Hammond, H. C. Marsh, and S. J. Perkins. 2008. The partly folded back solution structure arrangement of the 30 SCR domains in human complement receptor type 1 (CR1) permits access to its C3b and C4b ligands. *J. Mol. Biol.* 375: 102–118.
35. Kask, L., B. O. Villoutreix, M. Steen, B. Ramesh, B. Dahlbäck, and A. M. Blom. 2004. Structural stability and heat-induced conformational change of two complement inhibitors: C4b-binding protein and factor H. *Protein Sci.* 13: 1356–1364.
36. Kirkitadze, M. D., M. Krych, D. Uhrin, D. T. Dryden, B. O. Smith, A. Cooper, X. Wang, R. Hauhart, J. P. Atkinson, and P. N. Barlow. 1999. Independently melting modules and highly structured intermodular junctions within complement receptor type 1. *Biochemistry* 38: 7019–7031.
37. Moore, M. D., R. G. DiScipio, N. R. Cooper, and G. R. Nemerow. 1989. Hydrodynamic, electron microscopic, and ligand-binding analysis of the Epstein-Barr virus/C3dg receptor (CR2). *J. Biol. Chem.* 264: 20576–20582.
38. Hawrot, E., Y. Xiao, Q. L. Shi, D. Norman, M. Kirkitadze, and P. N. Barlow. 1998. Demonstration of a tandem pair of complement protein modules in GABA (B) receptor 1a. *FEBS Lett.* 432: 103–108.
39. Pangburn, M. K., N. Rawal, C. Cortes, M. N. Alam, V. P. Ferreira, and M. A. Atkinson. 2009. Polyanion-induced self-association of complement factor H. *J. Immunol.* 182: 1061–1068.
40. Tas, S. W., L. B. Klickstein, S. F. Barbashov, and A. Nicholson-Weller. 1999. C1q and C4b bind simultaneously to CR1 and additively support erythrocyte adhesion. *J. Immunol.* 163: 5056–5063.
41. Blein, S., R. Ginham, D. Uhrin, B. O. Smith, D. C. Soares, S. Veltel, R. A. McIlhinney, J. H. White, and P. N. Barlow. 2004. Structural analysis of the complement control protein (CCP) modules of GABA(B) receptor 1a: only one of the two CCP modules is compactly folded. *J. Biol. Chem.* 279: 48292–48306.
42. Gjelstrup, L. C., J. D. Kaspersen, M. A. Behrens, J. S. Pedersen, S. Thiel, P. Kingshott, C. L. Oliveira, N. M. Thielens, and T. Vorup-Jensen. 2012. The role of nanometer-scaled ligand patterns in polyvalent binding by large mannan-binding lectin oligomers. *J. Immunol.* 188: 1292–1306.
43. Miller, A., A. Phillips, J. Gor, R. Wallis, and S. J. Perkins. 2012. Near-planar solution structures of mannose-binding lectin oligomers provide insight on activation of lectin pathway of complement. *J. Biol. Chem.* 287: 3930–3945.
44. Dong, M., S. Xu, C. L. Oliveira, J. S. Pedersen, S. Thiel, F. Besenbacher, and T. Vorup-Jensen. 2007. Conformational changes in mannan-binding lectin bound to ligand surfaces. *J. Immunol.* 178: 3016–3022.
45. Vorup-Jensen, T. 2012. On the roles of polyvalent binding in immune recognition: perspectives in the nanoscience of immunology and the immune response to nanomedicines. *Adv. Drug Deliv. Rev.* 64: 1759–1781.
46. Ghiran, I., A. M. Glodek, G. Weaver, L. B. Klickstein, and A. Nicholson-Weller. 2008. Ligation of erythrocyte CR1 induces its clustering in complex with scaffolding protein FAP-1. *Blood* 112: 3465–3473.
47. Teillet, F., B. Dublet, J. P. Andrieu, C. Gaboriaud, G. J. Arlaud, and N. M. Thielens. 2005. The two major oligomeric forms of human mannan-binding lectin: chemical characterization, carbohydrate-binding properties, and interaction with MBL-associated serine proteases. *J. Immunol.* 174: 2870–2877.
48. Pagh, R., K. Duus, I. Laursen, P. R. Hansen, J. Mangor, N. Thielens, G. J. Arlaud, L. Kongerslev, P. Højrup, and G. Houen. 2008. The chaperone and potential mannan-binding lectin (MBL) co-receptor calreticulin interacts with MBL through the binding site for MBL-associated serine proteases. *FEBS J.* 275: 515–526.
49. Wallis, R., J. M. Shaw, J. Uitdehaag, C. B. Chen, D. Torgersen, and K. Drickamer. 2004. Localization of the serine protease-binding sites in the collagen-like domain of mannose-binding protein: indirect effects of naturally occurring mutations on protease binding and activation. *J. Biol. Chem.* 279: 14065–14073.

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

50. Gingras, A. R., U. V. Girija, A. H. Keeble, R. Panchal, D. A. Mitchell, P. C. Moody, and R. Wallis. 2011. Structural basis of mannan-binding lectin recognition by its associated serine protease MASP-1: implications for complement activation. *Structure* 19: 1635–1643.
51. Barlow, P. N., and D. C. Soares. 2005. Complement control protein modules in the regulators of complement activation. In *Structural Biology of the Complement System.* D. Morikis, and J. D. Lambris, eds. CRC Press, Baton Rouge, FL, p. 19–62.
52. Jallow, M., Y. Y. Teo, K. S. Small, K. A. Rockett, P. Deloukas, T. G. Clark, K. Kivinen, K. A. Bojang, D. J. Conway, M. Pinder, et al. 2009. Genome-wide and fine-resolution association analysis of malaria in West Africa. *Nat. Genet.* 41: 657–665.
53. Moulds, J. M., L. Kassambara, J. J. Middleton, M. Baby, I. Sagara, A. Guindo, S. Coulibaly, D. Yalcouye, D. A. Diallo, L. Miller, and O. Doumbo. 2000. Identification of complement receptor one (CR1) polymorphisms in west Africa. *Genes Immun.* 1: 325–329.
54. Thathy, V., J. M. Moulds, B. Guyah, W. Otieno, and J. A. Stoute. 2005. Complement receptor 1 polymorphisms associated with resistance to severe malaria in Kenya. *Malar. J.* 4: 54.
55. Zimmerman, P. A., J. Fitness, J. M. Moulds, D. T. McNamara, L. J. Kasehagen, J. A. Rowe, and A. V. Hill. 2003. CR1 Knops blood group alleles are not associated with severe malaria in the Gambia. *Genes Immun.* 4: 368–373.
56. Tetteh-Quarcoo, P. B., C. Q. Schmidt, W. H. Tham, R. Hauhart, H. D. Mertens, A. Rowe, J. P. Atkinson, A. F. Cowman, J. A. Rowe, and P. N. Barlow. 2012. Lack of evidence from studies of soluble protein fragments that Knops blood group polymorphisms in complement receptor-type 1 are driven by malaria. *PLoS ONE* 7: e34820.
57. Garred, P., M. A. Nielsen, J. A. Kurtzhals, R. Malhotra, H. O. Madsen, B. Q. Goka, B. D. Akanmori, R. B. Sim, and L. Hviid. 2003. Mannose-binding lectin is a disease modifier in clinical malaria and may function as opsonin for *Plasmodium falciparum*-infected erythrocytes. *Infect. Immun.* 71: 5245–5253.
58. Klabunde, J., A. C. Uhlemann, A. E. Tebo, J. Kimmel, R. T. Schwarz, P. G. Kremsner, and J. F. Kun. 2002. Recognition of *Plasmodium falciparum* proteins by mannan-binding lectin, a component of the human innate immune system. *Parasitol. Res.* 88: 113–117.
59. Da Silva, R. P., B. F. Hall, K. A. Joiner, and D. L. Sacks. 1989. CR1, the C3b receptor, mediates binding of infective *Leishmania major* metacyclic promastigotes to human macrophages. *J. Immunol.* 143: 617–622.
60. Green, P. J., T. Feizi, M. S. Stoll, S. Thiel, A. Prescott, and M. J. McConville. 1994. Recognition of the major cell surface glycoconjugates of *Leishmania* parasites by the human serum mannan-binding protein. *Mol. Biochem. Parasitol.* 66: 319–328.
61. Schlesinger, L. S., C. G. Bellinger-Kawahara, N. R. Payne, and M. A. Horwitz. 1990. Phagocytosis of *Mycobacterium tuberculosis* is mediated by human monocyte complement receptors and complement component C3. *J. Immunol.* 144: 2771–2780.
62. Carroll, M. V., N. Lack, E. Sim, A. Krarup, and R. B. Sim. 2009. Multiple routes of complement activation by *Mycobacterium bovis* BCG. *Mol. Immunol.* 46: 3367–3378.
63. Lambert, J. C., S. Heath, G. Even, D. Campion, K. Sleegers, M. Hiltunen, O. Combarros, D. Zelenika, M. J. Bullido, B. Tavernier, et al. 2009. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nat. Genet.* 41: 1094–1099.
64. Rogers, J., R. Li, D. Mastroeni, A. Grover, B. Leonard, G. Ahern, P. Cao, H. Kolody, L. Vedders, W. P. Kolb, and M. Sabbagh. 2006. Peripheral clearance of amyloid beta peptide by complement C3-dependent adherence to erythrocytes. *Neurobiol. Aging* 27: 1733–1739.
65. Keenan, B. T., J. M. Shulman, L. B. Chibnik, T. Raj, D. Tran, M. R. Sabuncu, A. N. Allen, J. J. Corneveaux, J. A. Hardy, M. J. Huentelman, et al. 2012. A coding variant in CR1 interacts with APOE-ε4 to influence cognitive decline. *Hum. Mol. Genet.* 21: 2377–2388.
66. Larvie, M., T. Shoup, W. C. Chang, L. Chigweshe, K. Hartshorn, M. R. White, G. L. Stahl, D. R. Elmaleh, and K. Takahashi. 2012. Mannose-binding lectin binds to amyloid β protein and modulates inflammation. *J. Biomed. Biotechnol.* 2012: 929803.
67. Tacnet-Delorme, P., S. Chevallier, and G. J. Arlaud. 2001. β-Amyloid fibrils activate the C1 complex of complement under physiological conditions: evidence for a binding site for Aβ on the C1q globular regions. *J. Immunol.* 167: 6374–6381.
68. Moulds, J. M., P. A. Zimmerman, O. K. Doumbo, D. A. Diallo, J. P. Atkinson, M. Krych-Goldberg, D. E. Hourcade, and J. J. Moulds. 2002. Expansion of the Knops blood group system and subdivision of Sl(a). *Transfusion* 42: 251–256.
69. Soares, D. C., D. L. Gerloff, N. R. Syme, A. F. Coulson, J. Parkinson, and P. N. Barlow. 2005. Large-scale modelling as a route to multiple surface comparisons of the CCP module family. *Protein Eng. Des. Sel.* 18: 379–388.

Downloaded from http://www.jimmunol.org/ by guest on March 9, 2015

Soutenance

Institut de Biologie Structurale

HDR

Vendredi 18 Septembre 2015 à 14h

Salle des séminaires

Institut de biologie structurale - 71 avenue des Martyrs CS 10090 38044 Grenoble Cedex 9 - T.+33 (0)4 57 42 85 00

www.ibs.fr

par Véronique Rossi

Institut de Biologie Structurale

Groupe Réponse immunitaire aux pathogènes et au soi altéré

Les protéines modulaires : un socle architectural pour la multifonctionnalité

Cas des protéines du complément

Diplôme d'Habilitation à Diriger des Recherches de l'Université de Grenoble

La plupart des fonctions multiples des protéines du complément humain est assurée par l'association d'un ensemble relativement restreint de modules. Dans le cadre de mes activités de recherche j'ai étudié le rôle de ces structures modulaires dans la spécificité fonctionnelle de plusieurs protéines du complément humain. Ce travail, essentiellement basé sur une stratégie de dissection moléculaire couplée à des études structurales de cristallographie aux rayons X, a mené à la cartographie fonctionnelle des protéases à sérine initiatrices du complément humain (C1s, C1r et MASPs 1 et 2 (MBL-Associated Serine Proteases)). Il a permis de définir le rôle des modules dans leur spécificité enzymatique et de proposer un modèle d'association aux complexes initiateurs des voies classique et lectine. Un volet de mes travaux a également consisté à étudier le mode d'interaction d'un régulateur de C1s, une serpine (SERine Protease INhibitor) nommée C1-inhibiteur et de proposer un modèle inhabituel de type « sandwich » de coopération de l'héparine dans la stimulation de l'inhibition. Depuis quelques années, mon activité de recherche est consacrée à l'étude des relations structure-fonction de récepteurs multimodulaires tels que CR1 (Complement Receptor 1) et LRP1 (LDL receptor-Related Protein 1) impliqués dans l'élimination par le complément des éléments étrangers ou du soi-modifié.

L'accès au campus EPN nécessite un avis de rendez-vous

Contactez Odile Cavoret - T. 04 57 42 87 04 - email : ibs.seminaires@ibs.fr