



HAL
open science

Quelques exemples de biologie structurale. Mécanismes de reconnaissance et de signalisation de l'hôte face aux pathogènes et au soi-altéré; recherche de nouvelles molécules thérapeutiques

Jean-Baptiste Reiser

► To cite this version:

Jean-Baptiste Reiser. Quelques exemples de biologie structurale. Mécanismes de reconnaissance et de signalisation de l'hôte face aux pathogènes et au soi-altéré; recherche de nouvelles molécules thérapeutiques. Biologie structurale [q-bio.BM]. Université Joseph Fourier, 2012. tel-01302387

HAL Id: tel-01302387

<https://hal.univ-grenoble-alpes.fr/tel-01302387>

Submitted on 14 Apr 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.



Institut de Biologie Structurale



Mémoire

Pour obtenir le diplôme d'

HABILITATION A DIRIGER LES RECHERCHES

Spécialité : **Biologie**

Présenté par

Jean-Baptiste REISER

préparé à l'Institut de Biologie Structurale Jean-Pierre Ebel, Grenoble

Quelques exemples de biologie structurale

Mécanismes de reconnaissance et de signalisation de l'hôte face aux pathogènes et au soi-altéré ; recherche de nouvelles molécules thérapeutiques

Travaux présentés publiquement le **25 octobre 2012**,
devant le jury composé de :

Jacqueline CHERFILS

Directrice de Recherche CNRS, Gif sur Yvette (Rapporteur)

Lionel MOUREY

Directeur de Recherche CNRS, Toulouse (Rapporteur)

Patrice GOUET

Professeur, Université de Lyon (Rapporteur)

Marc JAMIN

Professeur, Université de Grenoble (Examineur)

Dominique HOUSSET

Ingénieur-chercheur, Grenoble (Membre)



Remerciements

Je tiens en tout premier lieu à exprimer toute ma gratitude à Dominique. Tu as un été tour à tour le pédagogue durant mes toutes premières années de thèse, le guide lors de mon escapade californienne, l'appui quand j'ai voulu rejoindre notre institut, le superviseur quand j'ai intégré ton équipe, l'éclaireur quand j'ai souhaité explorer de nouvelles voies et j'espère que tu seras le soutien quand il sera temps que j'apprenne à voler de mes propres ailes.

Mes remerciements vont aussi aux membres de notre modeste équipe qui nous suit depuis quelques années. J'aurai un mot tout particulier pour toi, Anne au caractère bien trempé. Notre relation est sûrement plus profonde que celle qui peut exister entre « chef » (comme tu aimes le dire) et « assistante ingénieure » (comme il est bienséant de dire). Elle a connu des heurts et des moments plus chaleureux, mais j'aimerais t'assurer qu'elle reste et restera une collaboration fructueuse. Je te remercie de ton enthousiasme parfois masqué sous une carapace de « ronchonades » (pour lesquelles je ne suis pas en reste). J'aurai une pensée aussi pour tous les assistants ingénieurs qui sont passés par notre laboratoire et sans qui le travail de paillasse prendrait un temps plus que certain : Madalène, Paul et Ombeline.

Je tiens à citer les étudiants qui m'ont déjà fait confiance : Renaud qui m'a impressionné par ses facultés, Marion qui a eu le courage de s'obstiner sur un projet qui n'a pas été facile et Marc qui part innocemment à la découverte de la biologie structurale. J'espère qu'ils n'auront pas de trop mauvais souvenirs à faire partager aux futurs étudiants qui auront l'audace de travailler avec moi.

Je tiens à remercier nos collaborateurs fidèles : les Nantais, Marc et Xavier. Par nos rencontres et discussions que j'ai eu parfois grande difficulté à suivre, tous les deux, vous avez su faire progresser mon intellect.

Mes remerciements chaleureux vont à mes collaborateurs proches physiquement et notamment à toi, Jean-Philippe pour ta volonté et tes envies de science mais surtout pour ton soutien au jour le jour depuis notre première rencontre en 2005. J'espère fortement continuer à travailler avec toi et ton équipe dans les meilleures conditions. Un mot aussi pour mes collègues du groupe IRPAS, Philippe, Nicole et leur équipe respective, avec qui une nouvelle aventure et une prochaine cohabitation commencent dans le nouvel IBS. Je n'oublie pas non plus Claude qui continue à me faire confiance sur les aspects structuraux et l'importance de l'expérience face à la modélisation. Un dernier mot de remerciements va à Jean-Michel et Anne-Emmanuelle pour une collaboration enrichissante et dynamique.

Mes dernières pensées, et pas des moindres, vont à ma famille de sang qui m'a laissé la liberté d'arriver là où je suis et à mes familles de cœur pour leur soutien, leur curiosité, leur amour et la fraternité qu'elles ont su me donner les yeux fermés.

Table des matières

Remerciements	3
Table des matières	5
Table des illustrations	9
Avant-propos	11
Chapitre I : Parcours professionnel	13
1. Curriculum Vitae	15
1.1. Détails personnels	15
1.2. Position actuelle	15
1.3. Parcours professionnel	15
1.4. Prix et distinction	16
1.5. Compétences scientifiques	16
1.6. Formation	16
2. Liste des publications	17
3. Liste des communications	19
4. Liste des personnes encadrées	21
4.1. Etudiants en thèse	21
4.2. Etudiant en Master 2	21
4.3. Etudiants BTS Assistants ingénieurs	21
4.4. Assistants ingénieurs	22
5. Activités de transmission du savoir	23
5.1. Enseignement	23
5.2. Vulgarisation	23
Chapitre II : Travaux postdoctoraux	25
1. Protéines de la réponse immunitaire chez la drosophile	27
1.1. Laboratoire d'accueil	27
1.2. Compétences et synergies	27
1.3. Contribution scientifique	27
1.4. Contexte scientifique	27
1.5. Résultats marquants	28
1.5.1. Les PGRP	28
1.5.2. Spätzle	28

2. Séjour sur la ligne de lumière BM14 de l'ESRF	31
2.1. Laboratoire d'accueil	31
2.2. Compétences et synergies	31
2.3. Contexte scientifique	31
2.3.1. Une ligne pour la diffusion anormale à haut-débit	31
2.3.2. Les TNF	32
2.4. Faits marquants	32
Chapitre III : Reconnaissance et réponse antigéniques spécifiques des lymphocytes T	33
1. Informations générales	35
1.1. Laboratoire et équipe d'accueil	35
1.2. Compétences, synergies et financements	35
1.3. Contribution scientifique	35
2. Contexte scientifique	37
2.1. Les lymphocytes T : acteurs de la réponse cellulaire et adaptative	37
2.2. L'antigène : le complexe pCMH	38
2.2.1. Le « peptide processing » et la présentation de l'antigène	38
2.2.2. La structure des molécules CMH	39
2.3. Le récepteur des cellules T : le TCR $\alpha\beta$	40
2.3.1. La structure génétique des TCR	40
2.3.2. La structure du TCR	41
2.3.3. La sélection thymique et la diversité du répertoire naïf	42
2.3.4. La dégénérescence : les réactivités croisées	42
2.3.5. La dégénérescence : la fréquence et la diversité des répertoires spécifiques	43
2.4. L'interaction TCR/pCMH	44
2.4.1. La structure du complexe TCR/pCMH	44
2.4.2. Les paramètres thermodynamiques de l'interaction TCR/pCMH	46
2.4.3. La synapse immunologique et la signalisation	46
3. Résultats marquants	49
3.1. L'immunogénicité différentielle de peptides	50
3.2. La fréquence et la diversité du répertoire T spécifique d'un antigène	51
3.2.1. Une relation entre les caractéristiques structurales des pCMH et la fréquence des clones T spécifiques ?	52
3.2.2. Un exemple structural d'une réponse T publique	54
4. Projet à venir	57

Chapitre IV : Une nouvelle voie de signalisation de la phagocytose	59
1. Informations générales	61
1.1. Laboratoire et équipe d'accueil	61
1.2. Compétences, synergies et financements	61
1.3. Contribution scientifique	61
2. Contexte scientifique	63
2.1. La phagocytose des corps apoptotiques : éliminer les cellules mortes	63
2.2. La reconnaissance et la signalisation : les acteurs moléculaires	64
2.2.1. La reconnaissance : entre signaux « mangez-moi » et récepteurs	64
2.2.2. La signalisation intracellulaire : une machinerie conservée	66
2.3. La famille DOCK : nouveaux facteurs d'échange de guanine	68
2.3.1. Les différents DOCK chez l'Homme	69
2.3.2. Les domaines fonctionnels	71
2.4. La famille ELMO : un régulateur ?	74
2.5. Une régulation complexe	77
2.5.1. Le recrutement et l'autorégulation d'ELMO	77
2.5.2. Le recrutement et l'auto-inhibition de DOCK	78
2.5.3. Un modèle d'activation et de régulation	80
2.6. Les pirates Nef et IpgB1	81
3. Résultats marquants	83
3.1. L'interaction entre ELMO et le domaine SH3 de DOCK	83
3.2. Caractérisation structurale d'ELMO1	85
3.3. Caractérisation préliminaire de l'interaction entre DOCK, ELMO et Nef	87
4. Projet à venir	89
Chapitre V : Recherche de nouveaux agents anticancéreux et antibactériens	91
1. Informations générales	93
1.1. Laboratoire et équipe d'accueil	93
1.2. Compétences, synergies et financements	93
1.2.1. Recherche de nouveaux agents anticancéreux	93
1.2.2. Caractérisation de nouvelles GTPases bactériennes	93
1.3. Contribution scientifique	94
2. Contexte scientifique : les protéines kinases	95
2.1. Les protéines kinases	95
2.1.1. Quelques généralités	95
2.1.2. Les protéines kinases : cibles thérapeutiques	96
2.1.3. La structure du domaine catalytique des protéines-kinases	97
2.1.4. Quelques modes d'action des inhibiteurs	99

2.2. La protéine kinase 2 : CK2	100
2.2.1. Une « Master kinase »	100
2.2.2. CK2 : une cible thérapeutique	102
2.2.3. La structure de CK2	102
2.2.4. Les bases structurales des inhibiteurs ATP-compétitifs	104
2.2.5. Les inhibiteurs non-ATP-compétitifs	106
3. Résultats marquants sur CK2	107
3.1. Les acides Difuranedicarboxyliques	108
3.2. Les dérivés d'Ellipticine	111
4. Projet à venir sur les inhibiteurs des protéines kinases	113
5. Contexte scientifique : les GTPases bactériennes	115
5.1. Les GTPases	115
5.1.1. Quelques généralités	115
5.1.2. La structure du domaine G	116
5.1.3. Le mécanisme d'hydrolyse	117
5.1.4. Les régulateurs GEF et GAP	119
5.1.5. Les GTPases : cibles antibactériennes ?	121
5.2. EngA	122
5.2.1. La structure de EngA	122
5.2.2. Les fonctions de EngA	124
5.2.3. La biochimie de EngA	124
6. Résultats marquants sur EngA	127
7. Projet à venir sur les inhibiteurs de EngA	131
Bibliographie	133

Table des illustrations

Figure 1 : Structure de PGRP-SA	28
Figure 2 : Structures des molécules CMH	40
Figure 3 : Représentation schématique des clusters génétiques du TCR	41
Figure 4 : Structure d'un TCR $\alpha\beta$	41
Figure 5 : Structure d'un complexe entre un TCR $\alpha\beta$ et son antigène peptide/molécule CMH	44
Figure 6 : Empreinte du TCR $\alpha\beta$ sur la surface pCMH	45
Figure 7 : Vues des peptides HA-1 (jaune ou orange) dans la gorge de fixation de la molécule CMH HLA-A2 (rose et vert).	51
Figure 8 : Relations entre fréquences des clones T, surface exposée et structures des antigènes peptidiques.	53
Figure 9 : Vues détaillés des interactions clés entre le TCR public RA14 et l'antigène NLV fixé à la molécule CMH HLA-A2	55
Figure 10 : Analyse par SPR des interactions entre le TCR RA14 et le peptide antigénique NLV ou ses mutants	55
Figure 11 : Superposition des sites antigéniques des TCR RA14 et RA15	56
Figure 12 : Représentation schématique des molécules connues pour être impliquées dans la reconnaissance des cellules apoptotiques par les cellules phagocytaires	65
Figure 13 : Représentation schématique des molécules connues pour être impliquées dans la signalisation de phagocytose des cellules apoptotiques	68
Figure 14 : Schéma de l'organisation en domaine des protéines de la superfamille DOCK	70
Figure 15 : Structure du domaine DHR1 (à droite)	72
Figure 16 : Structure du complexe entre un domaine DHR2 et une petite GTPase Rho (à gauche)	72
Figure 17 : Schéma de l'organisation en domaine des protéines de la famille ELMO	74
Figure 18 : Structure du complexe entre la partie N-terminale de DOCK et la partie C-terminale d'ELMO	76
Figure 19 : Schéma du modèle du réseau d'interactions connues entre DOCK180, ELMO1 et leurs partenaires	80

Figure 20 : Investigation par SPR de l'interaction ELMO1/domaine SH3 de DOCK180	84
Figure 21 : Modèle combiné de la structure d'un complexe entre ELMO1 et la partie N-terminale de DOCK180	86
Figure 22 : Première estimation par SPR de l'affinité de Nef pour les domaines SH3 de DOCK	87
Figure 23 : Le « kinome » humain	95
Figure 24 : Structure du domaine catalytique d'une protéine kinase en complexe avec l'ATP et un peptide substrat	98
Figure 25 : Structure de l'« holoenzyme » CK2	103
Figure 26 : Structure d'un analogue de l'ATP dans le site actif de la protéine kinase 2	104
Figure 27 : Structures des sites ATP des protéines kinases CK2 et PIM-1 complexées à des dérivés d'acides Difuranedicarboxyliques	110
Figure 28 : Structures du site ATP de la protéine kinase 2 complexée à des dérivés de l'Ellipticine	111
Figure 29 : Structure du domaine catalytique d'une GTPase	117
Figure 30 : Structure du site nucléotidique d'une GTPase en complexe avec un analogue de l'intermédiaire réactionnel et de sa GAP	118
Figure 31 : Structure de la GTPase EngA de B. subtilis	123
Figure 32 : Structure du site actif des domaines GD1 et GD2 de EngA	128

Avant-propos

Depuis le début de ma carrière de chercheur, j'ai essentiellement porté mes efforts d'activité de recherches à différents projets de biologie structurale en rapport avec la santé humaine. La diversité des questions de biologie que j'ai abordées depuis mes débuts dans la recherche académique reflète d'une part toute ma fascination pour les mécanismes que le monde du vivant est capable de mettre en œuvre et d'autre part mon attrait pour l'observation à l'échelle atomique de ces phénomènes. Je me suis ainsi rapidement tourné vers la cristallographie des protéines qui, bien que jalonnée de nombreuses difficultés, s'offre comme une des méthodes de choix pour étudier finement les interactions entre protéines ou entre protéines et substrats qui sont cœur des processus de la cellule et du métabolisme de l'organisme. Même si au premier abord, ce mémoire apparaît comme une juxtaposition de projets très divers, l'intérêt que je porte à des questions en rapport avec la santé humaine en est le fil conducteur. Les projets que j'ai pu aborder portent aussi bien sur les mécanismes immunologiques de reconnaissance, de signalisation et de réponse des organismes hôtes face aux pathogènes et au soi-altéré que sur la recherche de nouvelles molécules anticancéreuses ou antibiotiques.

Ma curiosité à la fois pour la biologie structurale et pour les phénomènes immunologiques prend naissance dans mes travaux de thèse (1999-2002) sous la direction de Dominique Housset. Ces travaux que j'ai pu effectuer à l'Institut de Biologie Structurale Jean-Pierre Ebel de Grenoble ont eu pour objet la compréhension de mécanismes de reconnaissance des pathogènes qui sont au cœur de la réponse immunitaire cellulaire et adaptative chez l'Homme. Nous nous sommes tout particulièrement intéressés aux bases moléculaires et structurales de l'interaction entre le récepteur spécifique des lymphocytes T (TCR), l'antigène peptidique issu du pathogène ainsi que sa molécule présentatrice codée par les gènes du complexe majeur d'histocompatibilité (CMH). Ainsi, grâce à la résolution de structures par cristallographie des rayons X de plusieurs complexes binaires peptide/CMH et ternaires TCR/peptide/CMH (codes PDB : 1FO0, 1KJ2, 1KJ3, 1NAM et 1NAN), nous avons établi les bases structurales d'une part de la reconnaissance de l'antigène par les cellules T, et d'autre part de la réactivité croisée de ces cellules qui est à l'origine notamment des phénomènes des rejets de greffes et des maladies du greffon-contre-l'hôte. Ces travaux ont conduit à trois publications majeures dans les journaux *Nature Immunology* et *Immunity* et ils ont été récompensés par le prix Ernest Déchelle de l'académie des sciences en 2006.

Par la suite, les travaux postdoctoraux que j'ai pu réaliser dans l'équipe de Ian A. Wilson du Scripps Research Institute à La Jolla en Californie (2003-2004) m'ont permis d'une part de conforter mon expertise de cristallographe des protéines et de biologiste structural et d'autre part, d'un point de vue scientifique, d'explorer les mécanismes de reconnaissance et d'activation de la réponse immunitaire innée. En particulier, je me

suis intéressé chez la drosophile à la reconnaissance des motifs antigéniques bactériens qui est la première étape de la chaîne de réactions immunitaires. La résolution de la structure d'une protéine de reconnaissance des peptidoglycanes bactériens (code PDB : 1SXR) nous a permis d'identifier les caractéristiques structurales de cette famille de protéine et de proposer les premières bases structurales à l'origine de l'activation de la réponse immune antibactérienne.

Depuis mon intégration à l'Institut de Biologie Structurale de Grenoble en tant que chargé de recherche du CNRS (2006-présent), mes activités de recherche sont toujours marquées par l'intérêt que je porte aux aspects structuraux et fonctionnels des protéines de l'immunité. En rejoignant le groupe « Réponse Immunitaire aux pathogènes et au soi altéré », je continue à orienter l'ensemble de mes activités autour de différents projets de biologie structurale s'inscrivant pleinement dans un des axes majeurs de recherche de l'institut : l'axe « Immunité et Interactions Hôtes-Pathogènes ».

Outre la cristallographie des protéines qui reste une de mes méthodes de prédilection pour caractériser les macromolécules et leurs interactions à l'échelle atomique, j'ai su développer et intégrer dans mon travail diverses méthodes de biologie moléculaire et de biochimie pour la production d'échantillons biologiques, ainsi que des méthodes biophysiques pour la caractérisation des interactions protéine-protéine comme la résonance plasmonique de surface (SPR). De plus, avec l'apparition et le développement sur le bassin grenoblois de plateaux techniques comme l'IBS et le PSB regroupant des méthodes à haut débit de production d'échantillon et de caractérisation structurale, il nous est possible aujourd'hui de sonder à toutes les échelles, allant de la cellule à l'atome, les mécanismes mis en jeu dans les processus biologiques grâce à l'accès à de multiples méthodes.

Ce mémoire n'a pas la prétention d'être une revue exhaustive de chacun des sujets que j'ai pu aborder mais il se veut être un aperçu rapide de mon travail de biologiste structural. Je m'attache à y apporter les clés nécessaires à la compréhension des questions diverses que nous nous sommes posées. Ce document comprend un premier chapitre contenant mon curriculum vitae pour donner les détails de mon parcours de chercheur. Ce volet rapide est suivi d'un second chapitre qui reviendra rapidement sur les projets de ma période de « post-doc » au Scripps Research Institute et sur la ligne de lumière BM14 de l'ESRF. Enfin, les chapitres suivant s'attardent plus longuement sur les projets de recherche que je développe depuis mon intégration à l'IBS en 2006 dans l'équipe de Dominique Housset. Je consacre ainsi mon temps entre trois grands projets de biologie structurale concernant :

- *la reconnaissance et la réponse antivirales et anti-tumorales des lymphocytes T,*
- *de nouvelles voies de signalisation de la phagocytose,*
- *des protéines liant les nucléotides et la recherche de leurs inhibiteurs.*

Chapitre I

Parcours professionnel

Ce chapitre retrace mon parcours professionnel en une liste exhaustive de mon *curriculum vitae*, depuis ma formation initiale en biologie structurale, en passant par mes contributions scientifiques et jusqu'à une liste des personnes dont j'ai assuré l'encadrement du travail technique et scientifique. Ce chapitre n'a pas de prétention littéraire et se limite à une vue détaillée de la chronologie de mon parcours professionnel pour donner les éléments clés de ma carrière scientifique.

1. Curriculum Vitae

1.1. Détails personnels

Né le 16 janvier 1977 à Les Andelys (France, Eure)

De nationalité française

Vie maritale sans enfant

1.2. Position actuelle

**Chargé de Recherche de classe 1 du CNRS
détaché à l'Institut de Biologie Structurale Jean-Pierre Ebel de Grenoble**

1.3. Parcours professionnel

Chargé de Recherche CNRS **2006-présent**

Laboratoire : Institut de Biologie Structurale (IBS), CEA-CNRS-UJF, Grenoble, France

Directeur : Eva Pebay-Peyroula

Projets : Etudes structurales et fonctionnelles (i) de l'interaction hôte-pathogène chez le lymphocyte T (ii) de la signalisation chez les cellules phagocytaires (iii) de la protéine kinase 2 chez l'Homme et de ses inhibiteurs (iv) de nouvelles GTPases bactériennes

Chargé de cours de cristallographie des protéines **2006-2009**

Masters 2 – Physique des sciences de la vie (2006-2007) et Méthodes Physiques et Physico-Chimiques (2007-2009) – Université Joseph Fourier Grenoble I

Associé de Recherche en cristallographie **2005**

Laboratoire : Ligne de lumière CRG BM14 de l'European Synchrotron Radiation Facility (ESRF) – Medical Research Council, Grenoble, France

Superviseurs : Martin Walsh – David Stuart

Projets : Etudes de génomique structurale de molécules d'intérêt immunologique / Installation et utilisation des outils de cristallographie des protéines à haut débit appliqués à une ligne de lumière

Associé de Recherche en immunologie structurale **2003-2004**

Laboratoire : The Scripps Research Institute (TSRI), Department of molecular biology, La Jolla, California, USA

Principal Investigator : Ian A. Wilson

Projet : Etudes structurales de protéines d'intérêt immunologique par cristallographie des rayons X

Thèse en cristallographie des protéines **1999-2002**

Laboratoire : Laboratoire de Cristallographie et Cristallogenèse des Protéines (LCCP), Institut de Biologie Structurale (IBS), CEA-CNRS-UJF, Grenoble, France

Directeur de thèse : Dominique Housset

Sujet : La spécificité du récepteur du lymphocyte T pour l'antigène et la molécule présentatrice : une approche structurale par cristallographie des rayons X

1.4. Prix et distinction

Prix Ernest Déchelle (Biologie intégrative) de l'Académie des Sciences 2006

1.5. Compétences scientifiques

Biologie structurale

Biologie moléculaire et biochimie des protéines

Cristallisation des protéines

Cristallographie aux rayons X

Résonance des Plasmons de Surface (Biacore)

1.6. Formation

Doctorat de 3^e cycle en Physique, cristallographie des protéines 2002

Université J. Fourier, Grenoble, France

Mention Très Honorable avec félicitations du jury

Higher European Research Course for Users of Large Experimental Systems (HERCULES) 2000

Université J. Fourier, Grenoble, France

Diplôme d'Etudes Appliquées de Cristallographie et RMN biologique 1999

Université J. Fourier, Grenoble, France

Mention Bien

Magistère de Physique générale 1999

Université J. Fourier, Grenoble, France

Mention Bien

Maîtrise de Physique générale 1998

Université J. Fourier, Grenoble, France

Mention Bien

Licence de Physique générale 1997

Université J. Fourier, Grenoble, France

Mention Assez Bien

Diplôme d'Enseignement Universitaire Général A 1996

Université J. Fourier, Grenoble, France

Mention Assez Bien

Baccalauréat série C 1994

Lycée Marie Curie, Echirolles, Isère, France

Mention Assez Bien

2. Liste des publications

Nombre de publications : 13 dont 7 en premier auteur et 2 en auteur correspondant

Nombre de citations : 531

Facteur H : 9

P13. Foucher A.-M., **Reiser J.-B.**, Ebel C., Housset D., Jault J.-M.

Potassium acts as a GTPase-acting element on each nucleotide-binding domain of the essential Bacillus subtilis EngA

Article soumis à Plos One, 2012

P12. Sevajol M., **Reiser J.-B.**[#], Ayala I., Gans P., Perard J., Kleman J.P., Housset D.

The C-terminal polyproline motif of ELMO1 contributes to increase the life-time of the ELMO-DOCK complex.

Biochimie, 2012. **94**, p. 823-8.

P11. Prudent R., Moucadel V., Nguyen C., Barette C., Schmidt F., Florent J., Lafanechère L., Sautel C., Duchemin-Pelletier E., Spreux E., Filhol O., **Reiser J.-B.**^{*}, Cochet C.

Antitumor activity of pyridocarbazole and benzopyridoindole derivatives that inhibit protein kinase CK2.

Cancer Reas., 2010. **23**: p. 9865-74.

P10. López-Ramos M., Prudent R., Moucadel V., Sautel C.F., Barette C., Lafanechère L., Mouawad L., Grierson D., Schmidt F., Florent J.C., Filippakopoulos P., Bullock A.N., Knapp S., **Reiser J.B.**^{*}, Cochet C.

New potent dual inhibitors of CK2 and Pim kinases: discovery and structural insights.

FASEB J., 2010. **24**: p. 3171-85.

P9. **Reiser J.-B.**, Legoux F., Machillot P., Debeaupuis E., Le Moullac-Vaydie, Chouquet A., Saulquin X., Bonneville M., Housset D.

Crystallization and preliminary X-ray crystallographic characterization of a public CMV specific TCR in complex with its cognate antigen.

Acta Cryst. F, 2009. **65**: p. 1157-61.

P8. Gras S., Saulquin X., **Reiser J.-B.**, Debeaupuis E., Echasserieau K., Kissenpfennig A., Legoux F., Chouquet A., Le Gorrec M., Machillot P., Neveu B., Thielens N., Malissen B., Bonneville M., Housset D.

Structural bases for the Affinity-driven Selection of a Public T cell Receptor against a dominant human cytomegalovirus epitope.

J. Immunol., 2009. **183**: p. 430-7.

P7. **Reiser J.-B.**[#], Spierings E., Gras S., Mommaas B., Almekinders M., Kester M.G.D., Chouquet A., Le Gorrec M., Drijfhout J.W., Ossendorp F., Housset D., Goulmy E.

Steric Hindrance and Fast Dissociation Explain the Lack of Immunogenicity of the Minor Histocompatibility HA-1Arg Null Allele

J. Immunol., 2009. **182**: p. 4809-16.

- P6.** Holland Z, Prudent R, **Reiser J.-B.**, Cochet C, Doerig C.
Functional Analysis of Protein Kinase CK2 of the Human Malaria Parasite Plasmodium falciparum.
Eukaryot Cell. 2009. **8**: p. 388-97.
- P5.** Garlatti V, Martin L, Gout E, **Reiser J.-B.**, Fujita T, Arlaud GJ, Thielens N, Gaboriaud C.
Structural basis for innate immune sensing by M-ficolin and its control by a pH-dependent conformational switch.
J Biol Chem., 2007. **282**: p. 35814-20.
- P4.** **Reiser J.-B.**, Teyton L., Wilson A.I.
Crystal structure of the Drosophila peptidoglycan recognition protein (PGRP)-SA at 1.56 Angström resolution.
J. Mol. Biol., 2004. **340**: p. 909-917.
- P3.** **Reiser J.-B.**, Darnault C., Grégoire C., Mosser T., Mazza G., Kearney A., van der Merwe P.A., Fontecilla-Camps J.C., Housset D., Malissen B.
CDR3 loop flexibility contributes to degeneracy of TCR recognition.
Nature Immunol., 2003. **4**: p. 241-247.
- P2.** **Reiser J.-B.**, Grégoire C., Darnault C., Mosser T., Guimezanes A., Schmitt-Verhults A.-M., Fontecilla-Camps J.C., Mazza G., Malissen B., Housset D.
A T cell receptor CDR3beta loop undergoes conformational changes of unprecedented magnitude upon binding to a peptide/CMH class I complex.
Immunity, 2002. **16**: p. 345-354.
- P1.** **Reiser J.-B.**, Darnault C., Guimezanes A., Grégoire C., Mosser T., Schmitt-Verhults A.-M., Fontecilla-Camps J.C., Malissen B., Housset D. and Mazza G.
The crystal structure of a T cell receptor bound to an allogeneic CMH molecule
Nature Immunol., 2000. **1**: p. 291-297.

3. Liste des communications

- C14.** Réunion du groupe thématique de biologie de l'Association française de cristallographie, Paris, France **Avril 2012**
Etudes structurales d'ELMO1 : une protéine impliquée dans la phagocytose des corps apoptotiques (poster)
Rôle des deux domaines de fixation nucléotidiques dans l'activité GTPase stimulée par le potassium de la protéine EngA de Bacillus subtilis (poster)
- C13.** Réunion du groupe thématique de biologie de l'Association française de cristallographie, Paris, France **Décembre 2009**
Etudes structurales d'ELMO1 : une protéine impliquée dans la phagocytose des corps apoptotiques (poster)
Steric hindrance and fast dissociation rate explain the lack of immunogenicity of the minor histocompatibility HA-1Arg null allele (poster)
- C12.** 2e Congrès européen d'immunologie, Berlin, Allemagne **Septembre 2009**
Structural bases for the affinity-driven selection of a public T cell receptor against the pp65₄₉₅₋₅₀₃ human cytomegalovirus epitope (poster)
Steric hindrance and fast dissociation rate explain the lack of immunogenicity of the minor histocompatibility HA-1Arg null allele (poster)
- C11.** Réunion internationale de cristallographie, Osaka, Japon **Août 2008**
Structural insights into an affinity-based selection of virus-specific public T cell receptors (poster)
- C10.** Réunion de l'Association française de cristallographie, Rennes, France **Juillet 2008**
Bases structurales de la sélection d'un répertoire public de récepteurs de cellule T (TCR) lors d'une infection virale (présentation orale)
- C9.** 12^e Congrès international d'immunologie, Rio de Janeiro, Brésil **Août 2007**
Pursuing structural bases of peptide-CMH immunogenicity (poster)
- C8.** 12^e Congrès international d'immunologie, Montreal, Canada **Août 2004**
Structure of a peptidoglycan recognition protein (PGRP) at 1.56Å resolution (poster)
- C7.** Réunion du groupe thématique de biologie de l'Association française de cristallographie, Marseille, France **Octobre 2002**
Two new TCR structures highlight the allogeneic CMH molecule recognition (présentation orale)
- C6.** Réunion internationale de cristallographie, Genève, Suisse **Août 2002**
Two TCR structures highlight the allogeneic CMH recognition (poster)
- C5.** 20^e réunion européenne de cristallographie, Cracovie, Pologne **Août 2001**
Structures of murine TCRs bound to an allogeneic CMH molecule (présentation orale)
- C4.** 11^e Congrès international d'immunologie, Stockholm, Suède **Juillet 2001**
Structures of murine TCRs bound to an allogeneic CMH molecule (poster)

C3. Réunion de l'Association française de cristallographie, Orsay, France) **Juillet 2001**
Structure of murine TCRs bound to an allogeneic CMH molecule (présentation orale)

C2. 19^e réunion européenne de cristallographie, Nancy, France **Août 2000**
The crystal structure of a murine TCR bound to an allogeneic CMH molecule (poster)

C1. 48 heures de biologie structurale, 3^e édition, Autrans, France **Octobre 1999**
Specificity of T cell receptor for the antigenic peptide and the presenting molecule (poster)

4. Liste des personnes encadrées

Depuis mon arrivée dans notre groupe de recherche de l'Institut, j'encadre et j'ai encadré, le travail de plusieurs personnes : (i) du personnel technique qualifié ou en formation dans notre laboratoire et (ii) d'étudiants en Master 2 ou en thèse dont les travaux de biochimie et de cristallographie ont nécessité mon expertise.

4.1. Etudiants en thèse

Marion Sévajol

2008-présent

Titre des travaux : Etudes structurales et fonctionnelles de la protéine ELMO impliquée dans la signalisation de la phagocytose et de la motilité cellulaire

Directeur de thèse : Dominique Housset

% encadrement : 50%

Publications et communications : **P12.**, **C13.** et **C14.**

Renaud Prudent

2006-2009

Titre des travaux : Identification et caractérisation d'inhibiteurs de la protéine-kinase CK2

Directeur de thèse : Claude Cochet

% encadrement : 15%

Publications : **P6.**, **P10.** et **P11.**

4.2. Etudiant en Master 2

Marc Heitz

2012-présent

Titre des travaux : Etudes structurales du mécanisme catalytique de la GTPase bactérienne EngA de *Bacillus subtilis* possédant 2 domaines G.

% encadrement : 100%

4.3. Etudiants BTS Assistants ingénieurs

Mélanie Vo Van

2010-2011

(BTS Analyses Agricoles Biologiques et Biotechnologies, MFR de Moirans)

Titre du projet de stage : Mise au point du protocole d'expression et de purification d'un fragment de la protéine DOCK180 impliquée dans la signalisation de la phagocytose des corps apoptotiques

% encadrement : 100%

Romain Ruggierri

2010

(BTS bioanalyses et contrôles, Lycée Louise Michel, Grenoble)

Titre du projet de stage : Mise au point du protocole de purification de la protéine Nef de HIV-1 et de sa cible cellulaire

% encadrement : 100%

4.4. Assistants ingénieurs

Ombeline Pessey

CDD – 2012-présent

Projet : Reconnaissance et réponse antivirales et anti-tumorales des Lymphocytes T

Anne Chouquet

Permanent – CEA

Projets : De nouvelles voies de signalisation de la phagocytose / Reconnaissance et réponse antivirales et anti-tumorales des Lymphocytes T

Publications et communications : **P7., P8., P9., C9., C10., C11., C12. et C13.**

Paul Machillot

CDD – 2007-2008

Projet : Reconnaissance et réponse antivirales et anti-tumorales des Lymphocytes T

Publications : **P7., P8., P9., C9., C10. et C11.**

Madalène Le Gorrec

CDD – 2006-2007

Projets : De nouvelles voies de signalisation de la phagocytose / Reconnaissance et réponse antivirales et anti-tumorales des Lymphocytes T

Publications : **P7., P8., P9., C9., C10. et C11.**

5. Activités de transmission du savoir

En annexe de mon travail de recherche, j'ai participé à plusieurs activités de transmission du savoir. J'ai d'abord été en charge d'un cours de détermination des structures de macromolécules par cristallographie au niveau Master 2 « Physique des sciences de la vie (PSV) » puis « Méthodes physiques et Physico-chimiques (MPPC) » de l'Université Joseph Fourier de Grenoble. Je participe aussi régulièrement à l'animation d'activités scientifiques pour les lycéens ou pour le grand public lors de journées portes ouvertes de nos laboratoires lors de la fête de la science ou lors de l'accueil de stagiaires de collèges ou de lycées.

5.1. Enseignement

Cours de détermination des structures de macromolécules par Cristallographie des rayons X **2006-2009**

Contenu du cours : théorie de la cristallographie par rayons X ; principes des grands instruments ; cristallogénèse des protéines ; méthodes d'enregistrement des données, de résolution et d'affinement des structures ; représentations et analyses des structures

Volume horaire : 18h cours/TD – 8h TP

Masters 2 : Physique des sciences de la vie (2006-2007), Méthodes physiques et Physico-chimiques – parcours biologie structurale (2007-2009)

5.2. Vulgarisation

Membre de l'équipe d'organisation de la fête de la science, responsable de l'atelier cristallographie (2007-2009)

Accueil de stagiaires de collège (3^e) et lycée (Terminale)

Chapitre II

Travaux postdoctoraux

Suite à l'obtention du diplôme de doctorat de 3^e cycle de l'Université Joseph Fourier de Grenoble en 2002, j'ai eu la chance de séjourner pendant deux ans (2003-2004) dans le laboratoire de Ian A. Wilson du Scripps Research Institute (La Jolla, Californie). Renom de l'immunologie structurale, Ian A. Wilson s'est tout d'abord illustré dans les années 80 par la résolution de la structure de l'hémagglutinine du virus de la grippe lors de son séjour dans le laboratoire de Don Wiley. Arrivé à La Jolla, il s'intéresse au monde des protéines à domaine immunoglobuline et notamment aux anticorps et ensuite aux récepteurs des lymphocytes T. Mon séjour dans le laboratoire de Ian A. Wilson a été pour moi l'occasion d'aborder un nouvel aspect de l'immunologie structurale et d'enrichir mes compétences en biologie de laboratoire.

L'ère post-génomique et de l'automatisation à haut débit qui s'applique à la cristallographie des protéines à la fin des années 90 et au début des années 2000 me voit explorer ces nouvelles approches techniques et scientifiques par un séjour sur une ligne de lumière du synchrotron de Grenoble (BM14 - 2005) et par une collaboration avec un large consortium de protéomique structurale. Cette courte expérience a profité à mon expertise en cristallographie par l'utilisation et le « beta-testing » des outils de robotisation qui font aujourd'hui notre quotidien de cristallographe.

Ce premier chapitre aborde rapidement les différents travaux de recherche que j'ai pu alors effectuer pendant mes deux séjours « postdoctoraux ».

1. Protéines de la réponse immunitaire chez la drosophile

1.1. Laboratoire d'accueil

The Scripps Research Institute, Department of Molecular Biology, La Jolla California USA, Laboratory of Ian A. Wilson, D.Phil.

1.2. Compétences et synergies

Ce projet de recherche est le résultat d'une collaboration entre le laboratoire de Ian A. Wilson pour la purification, la cristallisation et la détermination des structures, et l'équipe de Luc Teyton (The Scripps Research Institute, Department of Immunology, La Jolla, Californie) pour le clonage et l'expression des différentes protéines d'intérêt.

1.3. Contribution scientifique

Publications : P4.

Communications : C8.

Structures déposées (code PDB) : 1SXR

1.4. Contexte scientifique

La réponse immunitaire innée est la plus ancienne et la plus universelle voie de défense contre les microorganismes capables d'envahir l'organisme hôte. Elle forme la première ligne de réponse de l'hôte contre les infections. Son action quasi-immédiate fait intervenir, dans l'étape de reconnaissance du pathogène, un nombre limité de molécules réceptrices qui sont identiques pour toutes les cellules d'une même lignée cellulaire et qui ont évolué avec le temps pour ne reconnaître que des motifs conservés chez les microbes invasifs mais absents de l'organisme hôte [1]. Ces dernières années, nos connaissances des voies cellulaires d'activation de la réponse innée chez les insectes et les vertébrés ont largement avancé grâce à la caractérisation de certaines molécules (Famille Toll, CD14, facteurs de transcriptions). Le travail remarquable de Jules Hoffman, qui a reçu le prix Nobel de médecine en 2011 pour ses travaux sur l'immunité innée chez la drosophile, est d'ailleurs à citer ici largement [2, 3].

Dans le but de comprendre les mécanismes de reconnaissance du pathogène et d'activation de la réponse innée contre les bactéries, nos efforts de recherche, au sein du laboratoire de Ian A. Wilson, se sont focalisés sur la détermination des structures de différentes protéines impliquées, soit directement dans la reconnaissance des éléments pathogènes, soit dans les mécanismes d'activation de la réponse immunitaire contre ces pathogènes. Pour ma part, je me suis concentré sur deux familles de protéines impliquées dans l'activation de la voie Toll qui est une des voies de la réponse innée chez la drosophile : les protéines de reconnaissance des peptidoglycanes (PGRP), élément primordial pour la reconnaissance et l'activation de

la réponse immune et la protéine Spätzle, un élément clef dans l'activation du récepteur Toll et la cascade de signalisation qui suit la reconnaissance bactérienne.

1.5. Résultats marquants

1.5.1. Les PGRP

Pendant mon séjour postdoctoral au Scripps Research Institute, j'ai essentiellement porté mes efforts à la détermination de la structure de différentes protéines de reconnaissance des peptidoglycanes (ou PGRP pour Peptidoglycan Recognition Proteins). Ces protéines, présentes dans tous les organismes et conservées de l'insecte à l'Homme, sont capables d'interagir avec ou de dégrader les peptidoglycanes qui sont des éléments essentiels de la paroi cellulaire des bactéries. Elles sont ensuite capables d'activer une réponse immunitaire ciblée contre les bactéries [4]. En particulier, mes travaux ont porté sur plusieurs PGRPs de drosophile et de souris qui venaient d'être nouvellement identifiées.

Mes travaux ont conduit en outre à la résolution de la première structure d'une PGRP de reconnaissance : PGRP-SA de *Drosophila melanogaster* (voir Figure 1) [5] par opposition aux PGRP catalytiques capables de dégrader leurs ligands. La comparaison avec PGRP-LB, une protéine catalytique de drosophile dont la structure était la seule disponible à l'époque [6], a révélé la conservation de la structure globale avec la présence d'une gorge hydrophile en forme de « L ». Cette dernière a été pressentie pour être le site de fixation des carbohydrates et des acides aminés communs formant le cœur conservé des peptidoglycanes. Par la suite, de nombreuses études structurales et fonctionnelles de cette famille de protéines issues de l'insecte ou de mammifères, seules ou en complexe avec leur ligand, ont apporté de nombreux éléments sur les bases structurales du mode de reconnaissance des motifs peptidoglycanes et de leur discrimination, mais aussi sur l'activité amidase dépendante du zinc des PGRP [7-15].

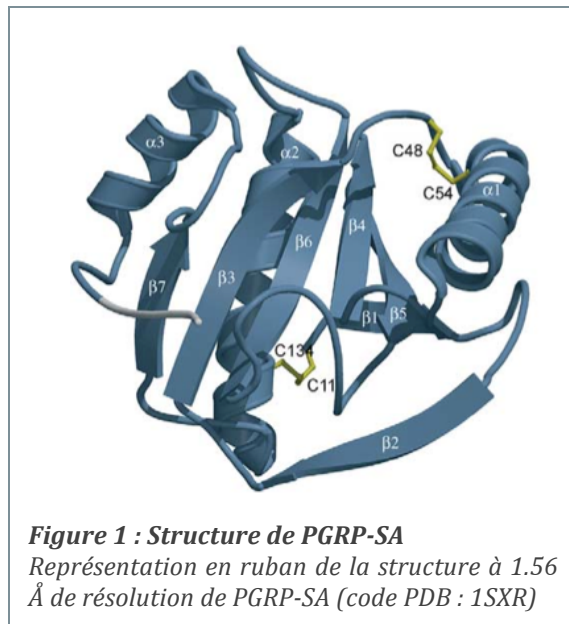


Figure 1 : Structure de PGRP-SA
Représentation en ruban de la structure à 1.56 Å de résolution de PGRP-SA (code PDB : 1SXR)

1.5.2. Spätzle

Le second objectif de mon travail en tant que « post-doc » a été la caractérisation et la résolution de la structure des différents états de la protéine Spätzle. Cette protéine est impliquée dans les mécanismes d'activation du récepteur de la voie Toll chez la Drosophile : PGRP-SA après complexation aux peptidoglycanes est capable d'induire une cascade protéolytique, dont les composants moléculaires

sont encore peu connus et qui se conclue par le clivage de Spätzle. La partie C-terminale de cette dernière se fixe au récepteur membranaire Toll pour activer la réponse immunitaire innée chez *Drosophila melanogaster* contre les bactéries gram-positives et les champignons [16].

La protéine Spätzle a pu être clonée et produite par l'équipe de Luc Teyton sous différentes formes proactives. Lors de mon séjour dans le laboratoire de Ian A. Wilson, je me suis consacré à la purification et la cristallisation des différentes constructions. Les premiers cristaux ont pu être obtenus pour un mutant de Spätzle mais les premiers essais de diffraction ont révélé la mauvaise qualité de ces cristaux. Par la suite, le projet n'a pas trouvé de suite dans le laboratoire et une seule structure d'un fragment de la protéine Spätzle a été déposée dans la PDB en 2008 [17].

2. Séjour sur la ligne de lumière BM14 de l'ESRF Les Facteurs de Nécrose Tumorale

2.1. Laboratoire d'accueil

European Synchrotron Radiation Facility (ESRF) – Beamline BM14 – Medical Research Council (MRC), Grenoble, France.

2.2. Compétences et synergies

Pour son développement, la ligne de lumière BM14 a principalement été soutenue par différents conseils de recherche de Grande-Bretagne : le Medical Research Council (MRC), le Biotechnology and Biological Sciences Research Council (BBSRC) et l'Engineering and Physical Sciences Research Council (EPSRC), mais aussi par l'European Molecular Biology Laboratory (EMBL) de Grenoble.

En ce qui concerne les projets de détermination de structure, l'équipe de BM14 entretenait des collaborations étroites avec de nombreuses équipes et en particulier avec le consortium européen pour la protéomique structurale (SPINE, Structural Proteomics in Europe consortium) qui regroupait près de 20 laboratoires répartis en Europe et qui concentrait les efforts pour le développement des technologies dédiées à la cristallographie à haut débit sur des cibles d'importance immunologique et biomédicale. Au sein de ce large regroupement de compétences, une collaboration plus étroite prédominait avec les groupes de David Stuart et Yvonne E. Jones de la division de biologie structurale du Wellcome Trust Centre for Human Genetics (Headington, Université d'Oxford, United Kingdom).

2.3. Contexte scientifique

2.3.1. Une ligne pour la diffusion anormale à haut-débit

Un des obstacles majeurs en cristallographie des protéines vient de notre seule capacité à mesurer l'intensité des ondes diffractées par un cristal et notre incapacité à mesurer directement l'information de phase de ces mêmes ondes, pourtant essentielle pour la détermination de la structure tridimensionnelle. Différentes approches ont été élaborées pour résoudre ce « problème des phases », et notamment des méthodes de phasage en utilisant la diffusion anormale de certains atomes déjà présents ou incorporés au sein du cristal (voir *Methods in Enzymology*, vol. 276, 1997). BM14 est une ligne de lumière à l'European Synchrotron Radiation Facility (ESRF, Grenoble) dédiée à la cristallographie des macromolécules. A l'ère de la robotisation et des méthodes à haut débit, BM14 a été une des premières lignes de lumières développées pour les expériences de diffraction et les méthodes de phasage utilisant la diffusion anormale à haut débit par l'automatisation.

2.3.2. Les TNF

En plus des aspects généraux et techniques de la ligne de lumière, mes activités de recherche ont été focalisées sur la détermination de la structure de différents facteurs de nécroses tumorales (TNF) et de leurs récepteurs (TNF-R). Ces protéines sont à l'origine de nombreux mécanismes physiologique et pathologiques d'une importance majeure en induisant un large spectre de réponses cellulaires, allant de l'activation de la transcription des gènes à l'activation de la mort cellulaire, notamment pour réguler un grand nombre de fonctions des réponses immunitaires (par exemple : induction de la mort cellulaire ou régulation de l'inflammation) [18-21]. L'intérêt clinique de cette étude réside dans la manipulation de ce type de molécules, notamment pour le développement de nouvelles thérapies contre le cancer. Par exemple, il a été montré que la cytokine TRAIL (TNR Related Apoptosis-Inducing Ligand) est capable d'induire la mort de diverses cellules tumorales, sans toutefois montrer d'activité contre les cellules saines de l'organisme [22]. Au moment de mon intégration à l'équipe de BM14, six structures de TNF étaient connues, ainsi que la structure d'un récepteur et seulement trois structures de complexes [23].

2.4. Faits marquants

Le développement de la ligne de lumière BM14 a eu pour objectif de procurer un équipement de haute technologie pour la résolution à haut débit de structures de protéines par cristallographie des rayons X. Il s'est inscrit dans une dynamique générale du développement des techniques de cristallographie biologique qui a permis la robotisation de toutes les étapes menant à la structure, depuis la cristallisation jusqu'à l'expérience de diffraction. A mon arrivée, de nombreux équipements d'automatisation développés entre l'EMBL et l'ESRF (micro-diffractomètre, robot de chargement des échantillon, systèmes automatiques de contrôle de la ligne, système de suivi des échantillons) avaient déjà été implantés à la ligne d'expérience. Mon rôle s'est essentiellement résumé au « beta-testing » de l'automatisation et à l'encadrement des équipes d'utilisateurs de la ligne (aide à l'enregistrement, traitement des données et résolution de structures). Aujourd'hui, le MRC s'est totalement désengagé du financement de la ligne (au bénéfice du synchrotron britannique Diamond), mais BM14 reste une ligne de lumière de premier plan grâce à une collaboration entre l'EMBL et le ministère des sciences et technologies indien.

Grâce aux installations de l'OPPF, une quinzaine de TNF, ainsi que différents récepteurs, ont pu être clonés et exprimés en système bactérien ou de cellules d'insecte ; et des essais de cristallisation à grande échelle ont permis d'obtenir plusieurs cristaux de TNF différents. Plusieurs jeux de données de diffraction ont pu être obtenus grâce aux équipements de BM14 mais à des résolutions et des qualités insuffisantes pour résoudre les structures tridimensionnelles.

Chapitre III

Reconnaissance et réponse antigéniques spécifiques des lymphocytes T

Les lymphocytes T sont des acteurs centraux du système immunitaire en protégeant l'organisme contre la prolifération des agents infectieux qui s'introduisent dans les cellules de l'hôte mais aussi contre la prolifération des cellules tumorales. Pour arriver à ses fins, les cellules T disposent à leur surface d'un récepteur spécifique, le TCR $\alpha\beta$. Le système immunitaire a évolué pour générer un très grand nombre de lymphocytes T, chacun exprimant un récepteur unique. Ce dernier permet la discrimination entre des fragments peptidiques qui sont issus de molécules du soi ou issus de molécules étrangères et qui sont présentés à la surface des cellules par les molécules du Complexe Majeur d'Histocompatibilité (CMH). De façon très élégante, très efficace et très régulée, la spécificité du répertoire T repose sur des processus de sélection qui permettent de générer un répertoire naïf capable de réagir contre des antigènes du « non-soi » qu'il n'a jamais rencontré, tout en étant tolérant vis-à-vis d'antigènes du « soi ».

Comprendre la reconnaissance des complexes CMH/peptide par le TCR depuis l'échelle atomique jusqu'au niveau systémique reste un réel défi. En effet, l'intensité de la réponse contre un antigène spécifique dépend d'un grand nombre de paramètres allant de l'aptitude de cet antigène à être efficacement présenté jusqu'à la capacité de réactivité du répertoire de cellules T. Notre compréhension de l'interaction TCR/peptide/CMH a grandement avancé grâce aux diverses structures atomiques de complexes. Elles ont fourni des données de grande importance sur des caractéristiques clefs comme la plasticité du TCR, la restriction CMH, la cross-réactivité, l'alloréactivité, l'immuno-dominance ou encore la reconnaissance auto-immune. Néanmoins et en partie à cause de la grande versatilité du TCR pour se lier aux complexes CMH/peptide et l'extrême sensibilité de l'événement de reconnaissance, aucune règle générale gouvernant cette interaction n'a pu être dévoilée.

Ayant déjà abordé ces problématiques pendant ma thèse de doctorat sous la direction de Dominique Housset, il a été naturel lors de mon retour dans son équipe d'aborder une nouvelle fois ces questions sous sa direction scientifique. Nous essayons ainsi d'apporter de nouveaux éléments pour clarifier à différents niveaux le lien entre la présentation de certains antigènes et la réponse spécifique du répertoire de cellules T chez l'Homme.

1. Informations générales

1.1. Laboratoire et équipe d'accueil

Institut de Biologie Structurale Jean-Pierre Ebel, Grenoble

Laboratoire de cristallogénèse et cristallographie des protéines (2006-2011)

Groupe Réponse immunitaire aux pathogènes et soi altéré (2012-présent)

Equipe Housset

Partnership for Structural Biology, Grenoble

1.2. Compétences, synergies et financements

Les avancées de ce projet reposent sur la collaboration de notre équipe avec trois partenaires pour mettre en commun nos différentes expertises en immunologie, biologie cellulaire, biologie fonctionnelle et biologie structurale. Les approches d'immunologie cellulaire et fonctionnelle sont essentiellement menées par les équipes de Marc Bonneville et Xavier Saulquin (Centre de Recherche en Cancérologie Nantes-Angers, Institut de Biologie, INSERM U892, Nantes), de Bernard Malissen (Centre d'Immunologie de Marseille-Luminy, Marseille) et d'Els Goulmy (Leiden University Medical Center, Leiden, Netherlands). Les approches structurales sont entreprises par notre équipe sous la direction scientifique de Dominique Housset.

Ce projet a été et est soutenu par différents financements ANR (Acronymes « human TCRs » 2005-2009 et « pCMHantigenicity » 2012-2015) et ARC (2010-2011), tous portés par Dominique Housset comme responsable du projet.

1.3. Contribution scientifique

Publications : P7., P8., P9.

Communications : C9., C10., C11., C12., C13.

Structures déposées (code PDB) : 3FT2, 3FT3, 3FT4, 3GSN, 3GSO, 3GSQ, 3GSR, 3GSU, 3GSV, 3GSW, 3GSX, 3MR9, 3MRB, 3MRC, 3MRD, 3MRE, 3MRF, 3MRG, 3MRH, 3MRI, 3MRJ, 3MRK, 3MRL, 3MRM, 3MRN, 3MRO, 3MRP, 3MRQ, 3MRR.

2. Contexte scientifique

2.1. Les lymphocytes T : acteurs de la réponse cellulaire et adaptative

Les lymphocytes T sont les acteurs centraux de la réponse cellulaire et adaptative du système immunitaire des vertébrés en protégeant l'organisme contre la prolifération des agents infectieux comme les virus, les bactéries ou les champignons, mais aussi contre la prolifération de certaines cellules tumorales. Les cellules T sont issues d'un progéniteur lymphoïde commun aux lymphocytes B (responsable de la sécrétion des anticorps spécifiques) qui prend naissance dans la moelle osseuse et mature dans le thymus.

Relâchés dans les organes et tissus périphériques, les cellules T vont jouer un rôle de surveillance et d'élimination des pathogènes intracellulaires, c'est-à-dire les pathogènes capables de s'introduire à l'intérieur des cellules de l'organisme (virus, bactéries), mais aussi des cellules au caractère tumoral et qui échappent à la vigilance des cellules phagocytaires de l'immunité innée et des lymphocytes B. De plus, les lymphocytes T participent à la régulation des nombreuses lignées cellulaires de l'immunité innée ou adaptative pour permettre les phénomènes de mémoire immunologique et d'homéostasie. Bien qu'élégante, très efficace et très régulée, l'activité aberrante de ce système peut conduire à des conséquences graves de maladies auto-immunes, d'allergie ou encore aux rejets de greffe et de maladies du greffon-contre-l'hôte.

Les lymphocytes T ont la particularité de ne pas reconnaître le pathogène dans son intégralité, mais seulement des fragments peptidiques présentés à la surface des cellules infectées ou des cellules spécialisées dans la présentation des antigènes (APC, macrophages, cellules dendritiques...). Il existe plusieurs formes de cellules T aux rôles distincts :

- Les lymphocytes T cytotoxiques : caractérisés par la présence du corécepteur CD8, ils sont activés par des peptides issus de pathogènes se développant dans le cytosol. Après activation par les APC et prolifération, ces cellules sont capables de reconnaître et de tuer directement les cellules infectées.

- Les lymphocytes T auxiliaires : caractérisés par la présence du corécepteur CD4, ils sont activés par des peptides issus de pathogènes ou de leurs fragments localisés dans les vésicules des macrophages, des cellules dendritiques ou des cellules B. Ils jouent plus un rôle d'intermédiaire en étant capable d'activer différents acteurs de la réponse immune pour une action plus directe contre les pathogènes comme la sécrétion d'anticorps spécifiques par les lymphocytes B.

- Les lymphocytes T régulateurs : comme leur nom l'indique, ils jouent un rôle de régulation de l'action immune en réprimant l'activité des cellules de l'immunité d'une part pour le maintien de la tolérance au soi et la prévention de la prolifération de cellules auto-immunes et d'autre part pour la régulation négative de la réponse immune lorsque l'infection a été maîtrisée.

2.2. L'antigène : le complexe pCMH

Pour permettre la surveillance par les lymphocytes T, la grande majorité des cellules de l'organisme présentent à leur surface des fragments peptidiques issus de la dégradation des protéines endogènes ou pathogènes par les machineries moléculaires et enzymatiques de la cellule, comme le protéasome ou les peptidases endosomales. Les peptides, de 8 à 20 acides aminés ainsi synthétisés, sont alors associés aux molécules codées par les gènes du Complexe Majeur d'Histocompatibilité (CMH ou HLA pour Human Leucocyte Antigen) pour être présentés à la surface des cellules sous la forme d'un complexe protéique binaire peptide/molécule CMH ou pCMH.

Les gènes CMH (ou aussi appelée HLA chez l'Homme) ont la particularité d'être à la fois polygéniques et polymorphiques. Le caractère polygénique se traduit par la présence chez le même individu de plusieurs gènes (ou loci) codant pour ces molécules (14 chez l'Homme). Le caractère polymorphique consiste en l'existence de nombreux allèles différents de chaque locus (plusieurs centaines parfois). Ainsi, la chance de trouver la même combinaison de loci (ou haplotype) chez deux individus différents est extrêmement mince.

Il existe deux classes de molécules CMH, dénommées tout simplement classe I et classe II. Les molécules CMH de classe I présentent des peptides de 8 à 13 acides aminés issus de la dégradation des protéines cytosoliques et qui sont reconnus par les lymphocytes T cytotoxiques. Les molécules CMH de classe II présentent des peptides plus longs (jusqu'à 20 acides aminés) provenant des systèmes vésiculaires des cellules phagocytaires et qui sont reconnus par les lymphocytes T auxiliaires.

2.2.1. Le « peptide processing » et la présentation de l'antigène

Deux voies distinctes sont utilisées par les cellules pour produire les complexes pCMH et conduire à la présentation des antigènes peptidiques aux lymphocytes T (pour revue : [24, 25]).

La voie de présentation des antigènes associés aux molécules CMH de classe I est active dans presque toutes les cellules et elle permet la présentation de peptides issus de protéines qui ont été synthétisées dans le cytosol à tout moment de la vie cellulaire. Les peptides sont générés par la dégradation des protéines grâce à l'action du protéasome et de différentes aminopeptidases, puis transportés jusqu'au réticulum endoplasmique par le transporteur TAP (Transportor associated with Antigen Processing). Le complexe binaire pCMH s'assemble alors par une série d'évènements impliquant notamment le complexe de chargement du peptide (PLC pour Peptide-Loading Complex) qui comprend des protéines comme la Tapasine, la Calreticuline ou la Thiol-oxydoréductase ERp57. Après le chargement, les complexes pCMH sont transportés à travers le système de Golgi vers la surface cellulaire.

Les molécules CMH de classe II sont quant à elles initialement assemblées en complexe avec une chaîne invariante « li » au niveau du réticulum endoplasmique des cellules présentatrices d'antigènes et capables d'endocytose et de phagocytose des

pathogènes (macrophages, cellules dendritiques, cellules B). Les complexes CMH/li migrent ensuite vers les compartiments endosomaux. La protéine li est alors relâchée après une série de digestions par des protéases ne laissant qu'un peptide résiduel (CLIP) occupant la gorge peptidique de la molécule CMH. Une seconde molécule CMH chaperonne (HLA-DM) catalyse le relargage de CLIP et l'échange avec le peptide antigène. Ces derniers sont essentiellement produits par les différentes voies de dégradation endosomale des protéines exogènes qui ont été internalisées.

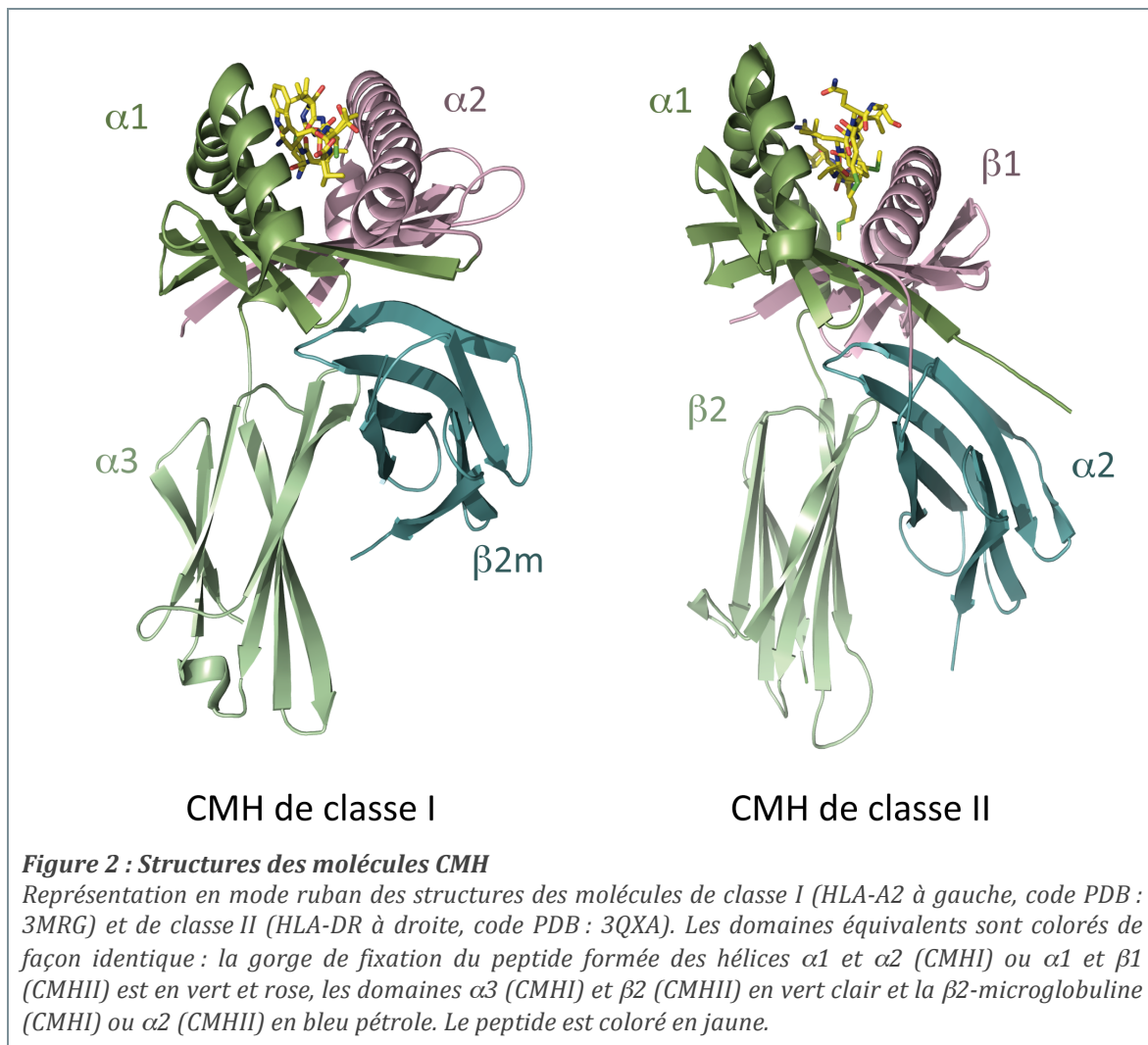
L'ensemble de ces mécanismes produit un système robuste d'exportation et de présentation de pCMH à la surface des cellules. En effet, bien que le protéasome ou les peptidases endosomales produisent un très grand nombre de peptides différents, tous les peptides ne sont pas mis à disposition et seuls les peptides correctement processés par toutes les étapes et produisant des complexes pCMH thermodynamiquement stables seront à même d'être présentés aux cellules T. On comprend alors aisément que la séquence et la structure des peptides présentés ainsi que le temps de demi-vie des complexes pCMH pourront avoir des conséquences directes sur la capacité des cellules T à reconnaître efficacement le pathogène et notamment sur les processus de sélection thymique et de reconnaissance dans les tissus périphériques [26, 27].

2.2.2. La structure des molécules CMH

Les structures des molécules CMH de classe I et II sont très similaires, mais diffèrent par la connectivité des différents domaines [28-30]. Les molécules CMH de classe I sont formées de deux chaînes polypeptidiques. La chaîne α , dite aussi chaîne lourde forme trois domaines : les domaines $\alpha 1$ et $\alpha 2$ composant le site de fixation du peptide et le domaine $\alpha 3$ à la topologie immunoglobuline qui est ancré à la membrane cellulaire. La seconde chaîne est la $\beta 2$ -microglobuline ($\beta 2$ -m), molécule soluble et non codée par les gènes du CMH, qui interagit de façon non-covalente avec la chaîne α . Les molécules CMH de classe II sont formées de deux chaînes, α et β , toutes deux codées par les gènes du CMH et ancrées à la membrane. Elles forment les domaines $\alpha 1$ et $\alpha 2$, équivalents aux domaines $\alpha 1$ et $\beta 2$ -m des molécules de classe I, et les domaines $\beta 1$ et $\beta 2$, équivalents aux domaines $\alpha 2$ et $\alpha 3$ (voir Figure 2).

Le site de fixation du peptide est formé par l'entrelacement de deux domaines : $\alpha 1$ - $\alpha 2$ pour les molécules CMH de classe I ou $\alpha 1$ - $\beta 1$ pour les molécules CMH de classe II, et peut être comparé à une « gorge » formée par deux longues hélices α et un large feuillet β . Chaque molécule CMH est aussi capable de fixer de façon extrêmement stable un vaste répertoire de peptides différents, dont le nombre peut atteindre plusieurs millions. Mais les caractères polymorphiques et polygéniques du Complexe Majeur d'Histocompatibilité s'expriment essentiellement au niveau du domaine de fixation du peptide par la variation de résidus situés dans la gorge. Ils confèrent ainsi des propriétés structurales et physico-chimiques uniques à chaque molécule CMH. Chaque molécule CMH présente ainsi un motif spécifique de fixation caractérisé par quelques résidus du peptide (dits « ancrés ») enfouis dans la gorge. Au nombre de 2, 3 ou 4, la nature des ancrés est entièrement conditionnée par

la séquence de la molécule CMH et par la stéréochimie de la gorge. La variabilité, bien plus grande, des autres résidus exposés permet ainsi à chaque molécule CMH de présenter plusieurs dizaines de millions de peptides différents.



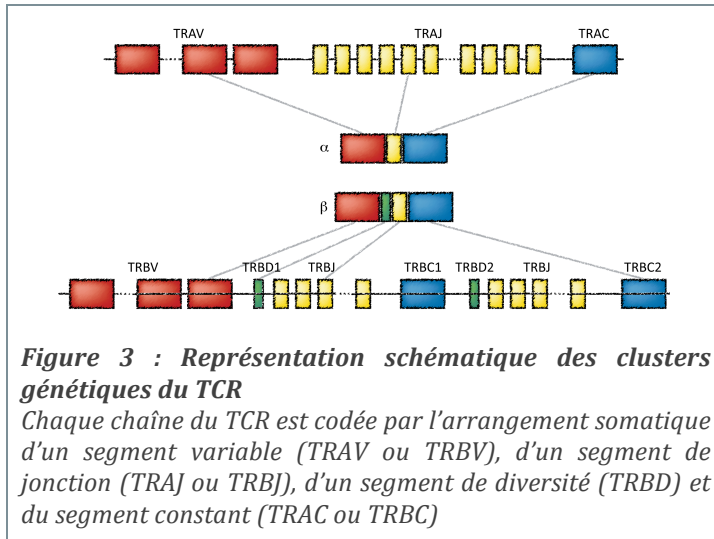
2.3. Le récepteur des cellules T : le TCR $\alpha\beta$

Afin d'assurer la reconnaissance du plus grand nombre d'antigènes possible, l'organisme a développé une stratégie clonale des lymphocytes T similaire à celle des lymphocytes B et des anticorps. La reconnaissance des complexes pCMH présentés et l'induction d'une réponse spécifique reposent sur un récepteur spécifique hétérodimérique, le TCR $\alpha\beta$, qui est capable de discriminer entre des fragments peptidiques issus de molécules du soi ou issus de molécules étrangères.

2.3.1. La structure génétique des TCR

La structure des gènes et les mécanismes de réarrangement des différents segments permettent de générer un très grand nombre de TCR $\alpha\beta$, chaque TCR caractérisant un clone T. En effet, chaque chaîne du TCR est la transcription d'un gène codant qui est le résultat de la combinaison d'un segment variable (TRAV ou TRBV),

d'un segment jonctionnel (TRAJ ou TRBJ) et d'un segment constant (TRAC ou TRBC). Ils ont pour origine des clusters génétiques contenant plusieurs exemplaires différents de chaque segment. Les clusters variables contiennent quarante à cinquante segments, les clusters jonctionnels une cinquantaine de segments et les clusters constants un ou deux segments. La chaîne β possède deux segments



supplémentaires (TRBD) qui augmentent la diversité de ces chaînes en s'intercalant entre les segments TRBV et TRBJ. La diversité est encore accrue au niveau du segment J qui peut subir de façon aléatoire l'insertion ou la délétion de bases nucléotidiques lors du réarrangement génétique. La combinatoire permet ainsi de construire de façon théorique 10^{16} TCR $\alpha\beta$ différents (voir Figure 3).

2.3.2. La structure du TCR

Comme le révèle la séquence et leur structure génétique, la structure tridimensionnelle des TCR $\alpha\beta$ est similaire aux fragments Fab des anticorps. Les deux chaînes forment un hétérodimère composé de quatre domaines immunoglobulines, associés deux à deux pour former un fragment variable portant le site d'interaction avec l'antigène et un fragment constant ancré à la membrane de la cellule T. Le site d'interaction avec l'antigène est un site combiné de six boucles variables, les boucles CDR (pour Complementarity Determining Regions). Les CDR1 et CDR2 ont une variabilité qui est limitée au nombre de segments TRAV et TRBV car ils sont exclusivement codés par ces segments, alors que le répertoire de CDR3 est beaucoup plus vaste car la séquence de ces boucles est codée à la jonction entre les segments TRAV et TRAJ ou entre TRBJ, TRBD et TRBJ (voir Figure 4).

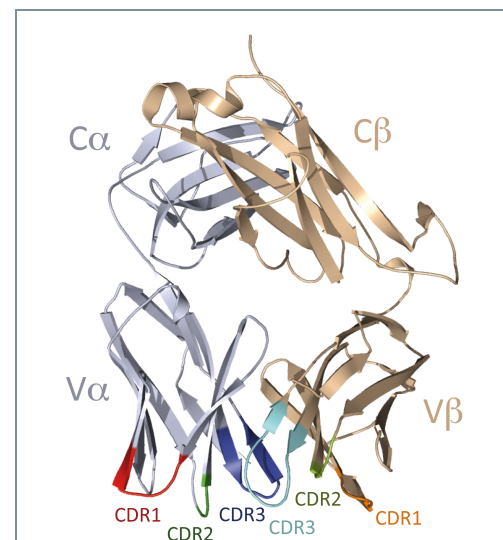


Figure 4 : Structure d'un TCR $\alpha\beta$
Représentation en mode ruban d'un TCR $\alpha\beta$ (RA14, code PDB : 3GSN). La structure est similaire à celle des fragments Fab des immunoglobulines. Le site de reconnaissance de l'antigène est composé de 6 boucles variables CDR représentées en couleur.

2.3.3. La sélection thymique et la diversité du répertoire naïf

Comme nous l'avons rapidement évoqué précédemment, la structure génétique en segments polygéniques des chaînes du TCR $\alpha\beta$ et les mécanismes de recombinaison pour générer une chaîne unique permettent de créer de façon théorique une très grande diversité estimée à 10^{16} récepteurs différents (voir section 2.3.1).

Lors de la maturation des cellules T dans le thymus et avant leur libération dans les systèmes périphériques où elles vont répondre à la présentation des antigènes, le répertoire T est façonné par des mécanismes de double sélection pour être capable de reconnaître les antigènes issus de pathogènes tout en étant tolérant aux peptides du soi. Basé seulement sur l'interaction avec les molécules CMH du soi présentant des peptides du soi, le processus de sélection positive permet de générer un répertoire de cellules T exprimant des TCR capables de reconnaître les molécules CMH du soi [31]. Dans le même temps, la sélection négative élimine les clones T présentant une auto-réactivité contre des peptides du soi présentés lors du développement [32]. Le répertoire T naïf ainsi créé présente un large spectre de TCR pour la reconnaissance d'antigènes étrangers tout en limitant les possibilités d'auto-immunité. Cette double sélection intra-thymique réduit remarquablement la diversité du répertoire généré entre trois et cent fois, conduisant le nombre théorique de clones T différents en circulation dans les systèmes périphérique à 10^{13} [33-36].

Toutefois, il est estimé que l'organisme ne peut contenir que 10^{12} cellules T à la fois [37] et il est évident qu'un même TCR peut être exprimé par plusieurs cellules T à la fois grâce aux expansions intra- et post-thymiques de la population naïve et à l'expansion des cellules mémoires. Des études récentes estiment que la diversité des TCR $\alpha\beta$ dans la périphérie est de l'ordre de $2,5 \cdot 10^7$ pour le répertoire naïf et 100 fois moins pour le répertoire mémoire [38] avec entre 1000 et 4000 cellules T présentant le même récepteur [37].

Ainsi et bien que le répertoire T naïf exprimé soit de plusieurs ordres de magnitude inférieur au répertoire théoriquement possible, il conserve une très grande diversité pour assurer la reconnaissance d'antigènes étrangers sans activer une réponse contre des antigènes du soi.

2.3.4. La dégénérescence : les réactivités croisées

Bien que le répertoire naïf présente un degré de spécificité pour le complexe pCMH qu'il reconnaît et contre qui il agit, de nombreuses observations révèlent une incroyable dégénérescence de la reconnaissance de l'antigène par le TCR. En effet, non seulement les clones T sont individuellement capables de reconnaître et de réagir à de très nombreux peptides différents (cross-réactivité) ou à différents allèles CMH (cas de allo-réactivité), mais de nombreux clones T peuvent aussi répondre à un même antigène pCMH.

Il est essentiel que le répertoire T naïf, pris dans son ensemble, soit capable de répondre à un large panel de peptides différents présentés par l'ensemble des molécules CMH de l'organisme pour s'assurer qu'aucun pathogène n'échappe à la

reconnaissance. Le nombre de peptides présentables par les molécules CMH est toutefois très largement supérieur à la diversité de clones T en circulation. Il est donc impératif qu'un clone T pris individuellement soit capable de cross-réactivité et donc être capable de reconnaître un large panel de peptides dans le contexte de restriction du CMH [39]. De nombreuses études ont permis de quantifier cette dégénérescence de la spécificité peptidique qui peut être basée sur des peptides aux séquences très homologues ou sur des peptides aux séquences éloignées (revue complète [40]).

L'allo-réactivité est un type de cross-réactivité des cellules T qui se manifeste par une très forte réponse T lors des rejets de greffe de tissus ou d'organe entre individus. Elle est basée sur la capacité du répertoire T naïf à interagir avec les différents variants alléliques des molécules CMH qui sont issus du greffon et qui sont reconnus comme les molécules de l'hôte. Cette réactivité aberrante est ainsi le résultat de la reconnaissance des peptides, normalement du soi, mais considérés comme peptides étrangers car ils sont présentés dans un contexte CMH pour lequel le répertoire T naïf n'a pas été sélectionné négativement. Cette forte réaction immune peut avoir différents niveaux suivant le degré de similitude des molécules CMH et des répertoires de peptides présentés (pour revue [41]).

2.3.5. La dégénérescence : la fréquence et la diversité des répertoires spécifiques

A l'inverse, un même peptide antigénique peut être reconnu par plusieurs clones T aux caractéristiques qui peuvent être différentes. La fréquence du répertoire T spécifique d'un antigène, c'est à dire le nombre de cellules T spécifiques de cet antigène parmi la population T naïve, doit être suffisamment importante pour assurer une réponse rapide et efficace contre le pathogène. Jusqu'à présent et du fait des fréquences relativement faibles (de l'ordre d'un clone T spécifique pour un à dix millions de cellules T) [42], l'estimation précise de répertoires T spécifiques pour des antigènes viraux ou tumoraux a été difficile. Mais l'étude dans un contexte « sauvage » a été rendue possible par de nouvelles techniques de biologie cellulaire [43-47]. Ces études ont permis de précisément mesurer ces fréquences pour certains antigènes. Il apparaît ainsi que le nombre de cellules T spécifiques peut varier fortement d'un antigène à l'autre et de plusieurs ordres de magnitudes. De plus, il apparaît que la taille du répertoire T spécifique d'un pCMH est directement corrélée à l'intensité de la réponse immune contre cet antigène.

La réponse T spécifique contre un antigène viral ou tumoral particulier n'est pas seulement caractérisée par la fréquence des clones T qu'il est capable d'activer, mais aussi par la diversité clonale du répertoire, c'est-à-dire du nombre de clones T ou TCR différents. Les réponses T contre des complexes pCMH spécifiques sont généralement diverses (un grand nombre de clones T différents sont activés) et privées (les clones T activés sont différents d'un individu à l'autre), mais elles peuvent dans certains cas être dominées et restreintes à des clones T présentant des récepteurs TCR fortement apparentés ou identiques chez des individus différents [48]. Ce type d'immuno-dominance, aussi décrite comme réponse publique, a été longtemps considérée comme un phénomène inhabituel du fait de la faible

probabilité de trouver un même TCR spécifique répondant au même épitope chez des individus différents. Toutefois, l'importance d'une réponse publique a été montrée dans de nombreux cas de réponses antivirales aiguës mais aussi de réponses auto-immunes ou alloréactives (pour revue [49])

2.4. L'interaction TCR/pCMH

Au cœur de la réponse immunitaire adaptative induite par les cellules T repose l'interaction moléculaire entre le récepteur TCR et l'antigène composite entre le peptide et la molécule CMH. Depuis la fin des années 90 et la détermination des premières structures de complexe, notre compréhension de l'interaction TCR/peptide/CMH a grandement avancé grâce aux diverses structures atomiques de près de 70 complexes ternaires composés de TCR humains ou murins, de molécules CMH de classes I ou II syngénique, allogénique ou xenogénique, ou encore de peptides du soi ou antigéniques [50].

2.4.1. La structure du complexe TCR/pCMH

L'ensemble de ces structures révèle un mode de fixation au complexe pCMH commun à tous les TCR et relativement conservé : le TCR adopte une orientation diagonale de son site de reconnaissance antigénique (la surface formée par les CDR) par rapport à l'axe de la gorge du CMH dans laquelle est fixé le peptide. Le domaine variable α contacte l'hélice $\alpha 2$ des molécules CMH de classe I ou $\beta 1$ des molécules CMH de classe II ainsi que la partie N-terminale du peptide ; et le domaine variable β contacte l'hélice $\alpha 1$ des molécules CMH de classe I ou II ainsi que la partie C-terminale du peptide (voir Figure 5). De manière générale, les CDR2 contactent principalement les régions hélicales des molécules CMH, les CDR1 la surface CMH et le peptide et les CDR3 le peptide antigénique.

Les données structurales comparées pour chaque partenaire sous forme libre et complexée ont mis en évidence une certaine rigidité des complexes binaires peptide/molécule CMH, bien que quelques exemples de changements de conformation du peptide ou de la surface CMH soient documentés [51]. Inversement, les structures des TCR non-complexés et complexés à leur antigène ont révélé la grande flexibilité du

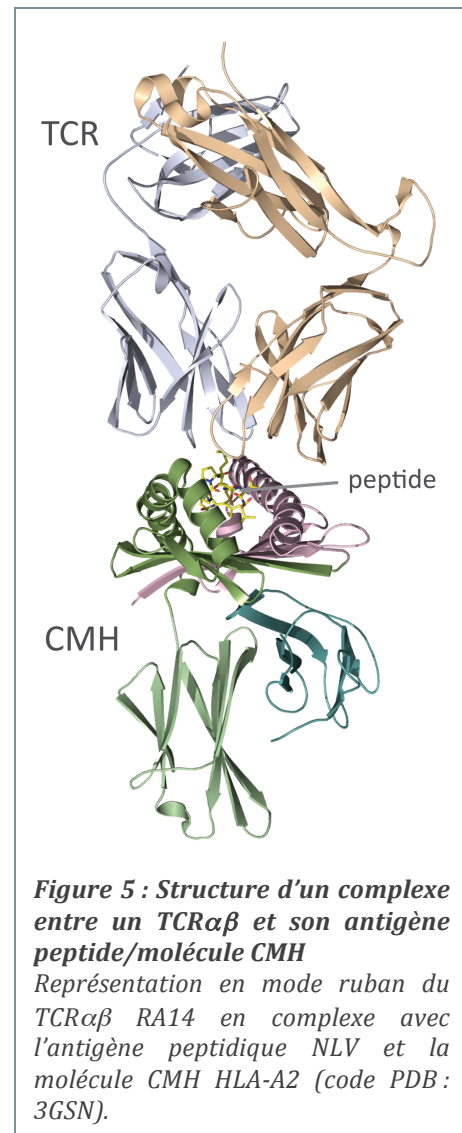


Figure 5 : Structure d'un complexe entre un TCR $\alpha\beta$ et son antigène peptide/molécule CMH

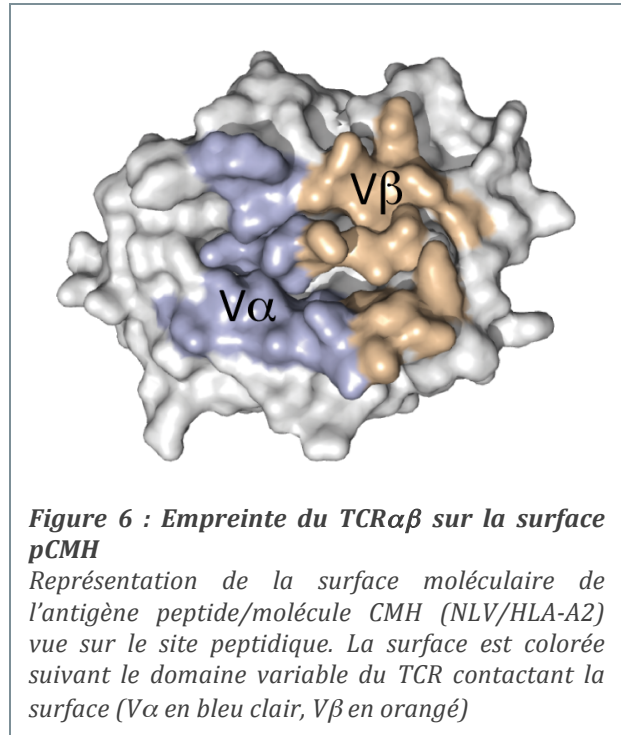
Représentation en mode ruban du TCR $\alpha\beta$ RA14 en complexe avec l'antigène peptidique NLV et la molécule CMH HLA-A2 (code PDB : 3GSN).

site antigénique du TCR. En effet, de grands changements et/ou la stabilisation de conformations des CDR sont observés lors de la complexation au pCMH [45]. Cette plasticité du TCR est sans doute un déterminant majeur de la spécificité pour l'antigène ainsi que des mécanismes de dégénérescence de reconnaissance des peptides antigènes et de différentes molécules CMH [40].

Structuralement parlant, l'interaction TCR/pCMH est complexe certainement due à une interface large, conformationnellement flexible et d'une grande variabilité. Les surfaces enfouies par chaque partenaire sont relativement importantes et comprises entre 600 et 1000 Å², mais de façon surprenante avec une très faible complémentarité ; ce qui peut aisément expliquer les faibles affinités observées du TCR pour son antigène (voir section 2.4.2). L'interaction du TCR n'est pas symétrique et la contribution de chaque domaine variable, ainsi que la contribution individuelle de chaque CDR, ne sont pas équivalentes et

peuvent varier d'un TCR à l'autre (voir Figure 6). Enfin, l'interaction TCR/pCMH est essentiellement dominée par des interactions hydrophobes et quelques interactions électrostatiques sont observées dans certains complexes. Toutefois, bien que les zones de la surface CMH contactées soient similaires, aucun ensemble d'interactions communes à tous les complexes dont la structure est connue ne permet d'expliquer de façon générale les mécanismes de spécificité du TCR pour les complexes peptide/molécule CMH. Comme une illustration, seulement 3 résidus des molécules CMH de classe I ont été identifiés comme très souvent contactés par les TCR [52] sans pourtant être des points d'ancrage essentiels [53]. Leur contribution varie d'un complexe à l'autre et ils constituent plus une triade d'acides aminés que le TCR ne peut éviter.

L'ensemble de ces structures de complexes ternaires a pu ainsi fournir des données structurales de grande importance sur les caractéristiques clés de la reconnaissance TCR/peptide/CMH comme la restriction CMH [54-56], la plasticité du TCR et la cross-réactivité [57-61], l'alloréactivité, [62, 63] l'immuno-dominance [61, 64] ou encore la reconnaissance auto-immune [65]. Néanmoins et en partie à cause de la grande versatilité du TCR pour se lier aux complexes pCMH et à cause de l'extrême sensibilité de l'événement de reconnaissance qui repose sur un grands nombres de paramètres différents, toutes ces données structurales n'ont pas permis à



l'heure actuelle d'établir de règle générale gouvernant cette interaction clef de reconnaissance et d'activation de la réponse cellulaire.

2.4.2. Les paramètres thermodynamiques de l'interaction TCR/pCMH

Bien que les bases moléculaires de l'interaction TCR/pCMH qui gouverne la spécificité des lymphocytes T restent encore élusives, il apparaît tout de même que le TCR est finement modelé pour reconnaître des peptides étrangers avec des caractéristiques suffisantes pour induire une activation tout en restant largement tolérant et inactif face aux tissus et cellules de l'organisme. De nombreuses études biophysiques viennent compléter les données structurales et apporter des éléments sur les paramètres thermodynamiques et cinétiques de l'interaction TCR/pCMH essentiels à l'activation de la cellule [66].

Ces mesures obtenues par des techniques de résonance des plasmons de surface (SPR) ou de titration calorimétrique (ITC) et faites sur une grande variété de systèmes ont mis en évidence que les clones T peuvent répondre aux antigènes avec une affinité du TCR pour le pCMH de faible ampleur ($K_D < 200 \mu\text{M}$) caractérisée par une association lente (k_{on} de l'ordre de 10^3 - $10^5 \text{ M}^{-1}\cdot\text{s}^{-1}$) et une dissociation rapide (k_{off} de l'ordre de 10^{-2} - 10^{-1} s^{-1}). L'accumulation des données biophysiques nous montre une hiérarchie des affinités : les TCR se lient aux antigènes issus du soi avec une plus faible affinité (K_D moyen de $100 \mu\text{M}$) qu'aux antigènes agonistes du non-soi (K_D moyen de $8 \mu\text{M}$). Cette corrélation entre affinité TCR/pCMH et activation de la réponse T avait déjà été pressentie par la biologie cellulaire et des mesures de compétition [67]. Plus que l'affinité, de nombreuses études apportent des indices expérimentaux sur l'importance du temps de demi-vie ou de la vitesse de dissociation du complexe ternaire dans la sensibilité du lymphocyte T. En effet, ce paramètre cinétique est le plus souvent corrélé avec l'effet biologique dans la majorité des études combinant mesures de paramètres biophysiques de l'interaction et études fonctionnelles. Cette corrélation est en accord avec le modèle d'activation de la réponse T qui est le plus communément admis. Ce modèle est basé sur l'idée que le TCR engage un pCMH agoniste suffisamment longtemps pour produire un signal d'activation, tout en se désengageant assez rapidement pour permettre au pCMH d'engager d'autres TCR dans la zone de contact entre les deux cellules [68].

2.4.3. La synapse immunologique et la signalisation

L'engagement des complexes ternaires entre TCR, peptide et molécule CMH à l'interface entre les cellules T et les cellules qui présentent l'antigène reste l'événement clé de reconnaissance. Néanmoins, il est associé à de nombreux éléments à l'interface cellulaire pour permettre la transduction du signal de reconnaissance à la cellule T. Les nouvelles techniques d'imagerie et de biophysique ont permis ces dernières années d'apporter de nombreuses informations sur les éléments clés jusqu'à présent inaccessibles au niveau de l'interface cellulaire.

L'évènement de reconnaissance et la signalisation qui s'en suit interviennent ainsi au sein d'un assemblage supramoléculaire : la synapse immunologique. Elle est

assemblée en clusters moléculaires associant les TCR mais aussi un ensemble de corécepteurs, de molécules d'adhésion et de signalisation. Les données récentes montrent que la synapse immunologique peut être vue comme un arrangement de trois couches de clusters ou d'îlots protéiques spécialisés pour l'engagement du récepteur T, la transduction du signal et le transport cytosquelettique (pour une revue complète : [69, 70]).

Ainsi, la couche du récepteur qui entre directement en contact avec la cellule cible contient des îlots de complexes entre les TCR et les corécepteurs CD3, CD4 ou CD8, des molécules de co-stimulation ou d'inhibition comme CD28 et CTLA4. Ces dernières vont moduler la signalisation et ainsi le devenir de la cellule T [71]. La présence de molécules de type intégrines comme LFA1 (Leukocyte function-associated antigen 1) qui sont responsables de l'adhésion intercellulaire est aussi observée [72].

La couche de signalisation voit l'agrégation de protéines kinases comme LCK, ZAP70, ITK (IL-2-inducible T cell kinase), PKC θ ou de phospholipases PLC γ qui vont relayer par des phosphorylations en cascade l'événement de reconnaissance jusqu'à l'activation de facteurs de transcription NF- κ B. Le TCR n'a aucune activité catalytique intrinsèque mais en formant un complexe avec les chaînes CD3 qui contiennent des motifs de phosphorylation (ITAMs pour Immunoreceptor tyrosine-based activation motifs), il permet de relayer le signal de reconnaissance [73]. Les protéines kinases de la famille SRC comme LCK peuvent ainsi, sous certaines conditions, rapidement phosphoryler ces motifs ITAM [74] qui interagissent alors avec ZAP70 [75]. Les mécanismes précis du recrutement des différents éléments restent à ce jour encore très élusifs mais il semble que le niveau d'agrégation au niveau de la synapse des différentes kinases serait suffisant pour amplifier le signal de reconnaissance. Les cibles suivantes dans la cascade de signalisation sont deux protéines adaptatrices SLP76 et LAT [76], une protéine transmembranaire qui contient de nombreux sites de phosphorylation. Les mécanismes d'intervention de LAT restent à ce jour très mal compris et interviennent à différents niveaux de la synapse, soit en îlots à la surface cellulaire soit en étant inclus dans des vésicules sous-membranaires.

Enfin, la couche cytosquelettique est assemblée autour des filaments d'actine, de Myosine et de Tubuline grâce notamment à deux molécules, la Taline [77] et la Vinculine [78], qui pontent les molécules de types intégrines à la F-actine. Le TCR est lié aussi à l'actine par l'intermédiaire de protéines adaptatrices et de l'Actin Régulateur Wiskott-Aldrich syndrom protéine (WASP) [79]. Un concept récent suggère que le TCR serait un transducteur mécanique qui permet la conversion de l'énergie mécanique en signaux chimiques, un processus qui est observé lorsqu'une protéine change de conformation pour révéler de nouveaux sites de fixation à l'actine [80]. Toutefois, les résultats expérimentaux allant dans ce sens restent encore à être explorés.

Les nouvelles méthodes d'imagerie et de caractérisation de l'événement de reconnaissance et de signalisation à l'interface entre la cellule T et la cellule présentatrice d'antigène ont apporté ainsi des nombreuses données qui restent

encore à ce jour à une étape préliminaire d'interprétation pour comprendre pleinement la complexité des mécanismes. Il semble finalement raisonnable de penser que la synapse immunologique est un système complexe qui permet à la cellule d'utiliser la dynamique à la membrane et au niveau de l'actine pour amplifier de faibles différences cinétiques observées au niveau de l'interaction TCR/pCMH.

3. Résultats marquants

La réponse immune cellulaire induite par les lymphocytes T repose ainsi sur la présence d'une population de cellules T qui vont répondre aux peptides dérivés de pathogènes tout en restant tolérants aux peptides du soi. Toutefois et comme l'organisme ne peut prédire avec précision à quel antigène il est susceptible d'être confronté, la réponse T repose sur la génération et la maintenance d'un répertoire TCR très divers. Ceci implique que la taille et la diversité du répertoire T pour un antigène spécifique sont cruciales pour donner une réponse immune adaptée. Comprendre les mécanismes à la base de cette reconnaissance des complexes pCMH par le TCR et de l'activation de la réponse T qui s'en suit depuis l'échelle atomique jusqu'au niveau systémique reste un réel défi. L'intensité de la réponse cellulaire dépend en effet d'un grand nombre de paramètres allant de l'aptitude d'un antigène à être efficacement présenté par les molécules CMH (processing par le protéasome, transport par TAP, association et stabilité du complexe peptide/CMH ...) jusqu'à la capacité de réactivité du répertoire de cellules T spécifiques (biais lors de la recombinaison des gènes TCR, de la sélection intra-thymique et du polymorphisme des gènes du CMH, processus homéostatiques périphériques ou de la pression antigénique, avidité et affinité des cellules T réactives, réactivités croisées ...) [49].

Depuis plusieurs années l'équipe de Dominique Housset en collaboration avec les équipes de Marc Bonneville (Nantes), Bernard Malissen (Marseille) ou Els Goulmy (Leiden) essaie d'apporter de nouveaux éléments pour clarifier les relations entre d'un côté la présentation de certains antigènes, leur reconnaissance et la réponse spécifique du répertoire de cellule T chez l'Homme et de l'autre, les caractéristiques structurales de ces antigènes complexés à leurs molécules présentatrices et des complexes TCR/pCMH. Ces dernières années, nous avons abordé ainsi plusieurs aspects : (i) *l'étude des mécanismes de présentation impliqués dans l'immunogénicité différentielle de peptides* (ii) *l'étude des liens possibles entre la fréquence et la diversité du répertoire T spécifique d'un antigène et les caractéristiques structurales des complexes peptide/CMH, ainsi que celles de l'interaction entre TCR et pCMH.*

Différentes équipes de par le monde ont développé différentes stratégies pour cloner et produire les différents partenaires sous forme soluble et en quantité suffisante pour les études structurales. Au laboratoire, nous avons mis en place le clonage, l'expression bactérienne, la purification et la cristallisation de la molécule CMH HLA-A2 chargée avec les peptides de notre étude ainsi que différents TCR spécifiques. La principale difficulté ici réside dans l'obtention d'échantillons de protéines recombinantes solubles, homogènes et en quantité suffisante pour les études structurales (les rendements sont faibles) mais aussi dans l'obtention de cristaux de complexes ternaires entre TCR, peptide et CMH dont les affinités sont de l'ordre du millimolaire. En effet, ces étapes restent difficiles dans la détermination par cristallographie de structures de complexes TCR/pCMH et expliquent en partie le faible nombre de ces structures déposées à ce jour dans la PDB.

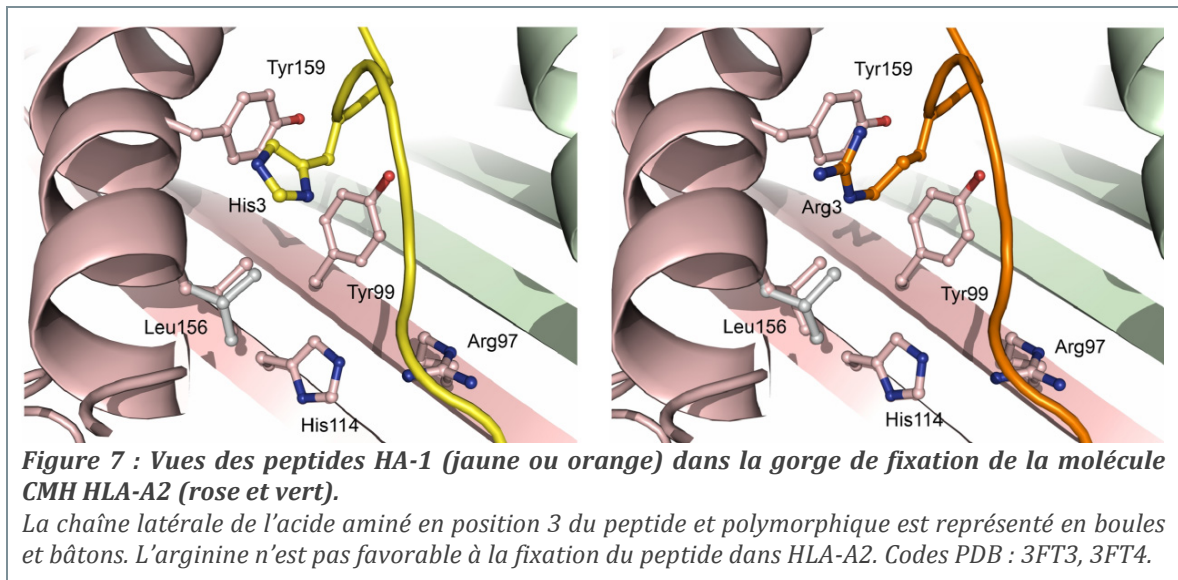
3.1. L'immunogénicité différentielle de peptides

Comme nous l'avons vu brièvement, les mécanismes de présentation d'un peptide aux cellules T par les molécules CMH sont des étapes primordiales dans la mise en place d'une réponse immune spécifique (voir section 2.2.1). La génération des peptides par le protéasome ou l'immuno-protéasome, le transport par chaperonne, l'association et la dissociation des complexes peptide/CMH ou encore le temps de demi-vie de ces complexes à la surface des cellules présentatrices sont autant de paramètres qui déterminent le devenir d'une réponse appropriée des cellules T. Dans ce contexte et avec l'équipe d'Els Goulmy, nous nous sommes intéressés aux antigènes mineurs du Complexe Mineur d'Histocompatibilité. Ces peptides jouent un rôle important dans les maladies du greffon-contre-l'hôte, notamment après transplantation de tissus entre individus HLA compatibles [81]. Ces antigènes sont issus de protéines cellulaires codées par des gènes polymorphiques et sont présentés à la surface des cellules par les molécules CMH. Ils sont alors capables d'induire une forte réponse immunitaire cellulaire chez les individus antigènes mineurs incompatibles conduisant ainsi au rejet de greffes ou aux maladies de greffon-contre-l'hôte. Ces réponses deviennent toutefois très intéressantes pour les thérapies anticancéreuses en exploitant les réponses immunitaires de type greffon-contre-tumeur après transplantation de cellules HLA compatibles mais antigènes mineurs incompatibles [82].

En particulier, nous nous sommes focalisés avec l'équipe d'Els Goulmy sur l'antigène mineur HA-1. Cet antigène, qui est spécifique des cellules hématopoïétiques, présente une surexpression dans les cellules épithéliales cancéreuses, mais est absent dans les cellules saines [83]. Ainsi, il représente un outil thérapeutique pour la transplantation de cellules souches comme l'immunothérapie cellulaire avec des cellules T HA-1 spécifiques ou l'activation de réponses greffon-contre-tumeur avec des peptides HA-1. Des peptides HA-1 sont d'ailleurs à l'heure actuelle en phase I et II d'essais cliniques de thérapies anti-leucémiques et anti-tumorales. Dans le cadre restrictif de HLA-A2, l'épitope HA-1 est un nona-peptide issu d'une protéine codée par un locus polymorphique contenant deux allèles qui diffèrent d'un seul acide aminé (Histidine ou Arginine). L'antigène HA-1^{His} est capable d'induire une forte réponse cellulaire chez les individus HA-1^{His} négatifs (autrement dit HA-1^{Arg} positifs), alors que l'inverse n'apparaît pas. Les précédentes études faite par l'équipe d'Els Goulmy ont mis en évidence une absence de complexes HLA-A2/HA-1^{Arg} à la surface des cellules bien que le peptide soit capable d'être fixé sur la molécule CMH.

Courant 2007, l'équipe hollandaise a pu mettre en évidence qu'aucun mécanisme connu dans la présentation des antigènes n'était affecté par le polymorphisme de HA-1 : ni les mécanismes de digestion des peptides antigènes par le protéasome, ni la translocation des peptides HA-1 par TAP vers le réticulum endoplasmique. Par contre, les mesures de demi-vie des complexes peptide/HLA-A2 montrent que l'interaction HLA-A2 avec le peptide HA-1^{Arg} est moins stable que

l'interaction avec le peptide HA-1^{His}. Grâce aux structures de HLA-A2 en complexe avec HA-1^{His} et HA-1^{Arg} que nous avons résolues (voir Figure 7), nous avons pu déterminer que la substitution Histidine en Arginine provoquait l'instabilité du peptide HA-1, essentiellement due à l'encombrement stérique et à la charge positive portée par l'arginine en condition physiologique. Nos résultats structuraux montrent ainsi clairement que l'absence de présentation d'un allèle HA-1 par HLA-A2 est le résultat d'une instabilité accrue de cet allèle dans la gorge de HLA-A2. Ce travail a conduit à un article publié dans The Journal of Immunology (publication P7.).



3.2. La fréquence et la diversité du répertoire T spécifique d'un antigène

Comprendre les principes gouvernant la génération de la fréquence et de la diversité d'un répertoire T est essentiel pour comprendre la spécificité de ce répertoire pour l'antigène donné et l'efficacité de la réponse immunitaire cellulaire et adaptative contre les pathogènes et tumeurs associés. Avec cet objectif, nous nous sommes focalisés sur différents épitopes immuno-dominants présentés par la molécule HLA-A2 et issus de protéines virales (HCV, EBV, HCMV) et tumorales (Melan-A) et dont les fréquences et diversités des répertoires T spécifiques diffèrent de façon significative.

Ainsi, avec nos collaborateurs, nous menons en parallèle deux approches : (i) une analyse précise et comparative des répertoires des cellules T naïves (c'est-à-dire n'ayant jamais rencontré l'antigène) spécifiques de différents complexes HLA-A2 chargés avec les peptides viraux et tumoraux naturels ainsi que différents mutants (ii) une analyse cristallographique détaillée de tous ces complexes peptide/HLA-A2, ainsi que différents complexes ternaires entre peptide/HLA-A2 et des TCR spécifiques.

3.2.1. Une relation entre les caractéristiques structurales des pCMH et la fréquence des clones T spécifiques ?

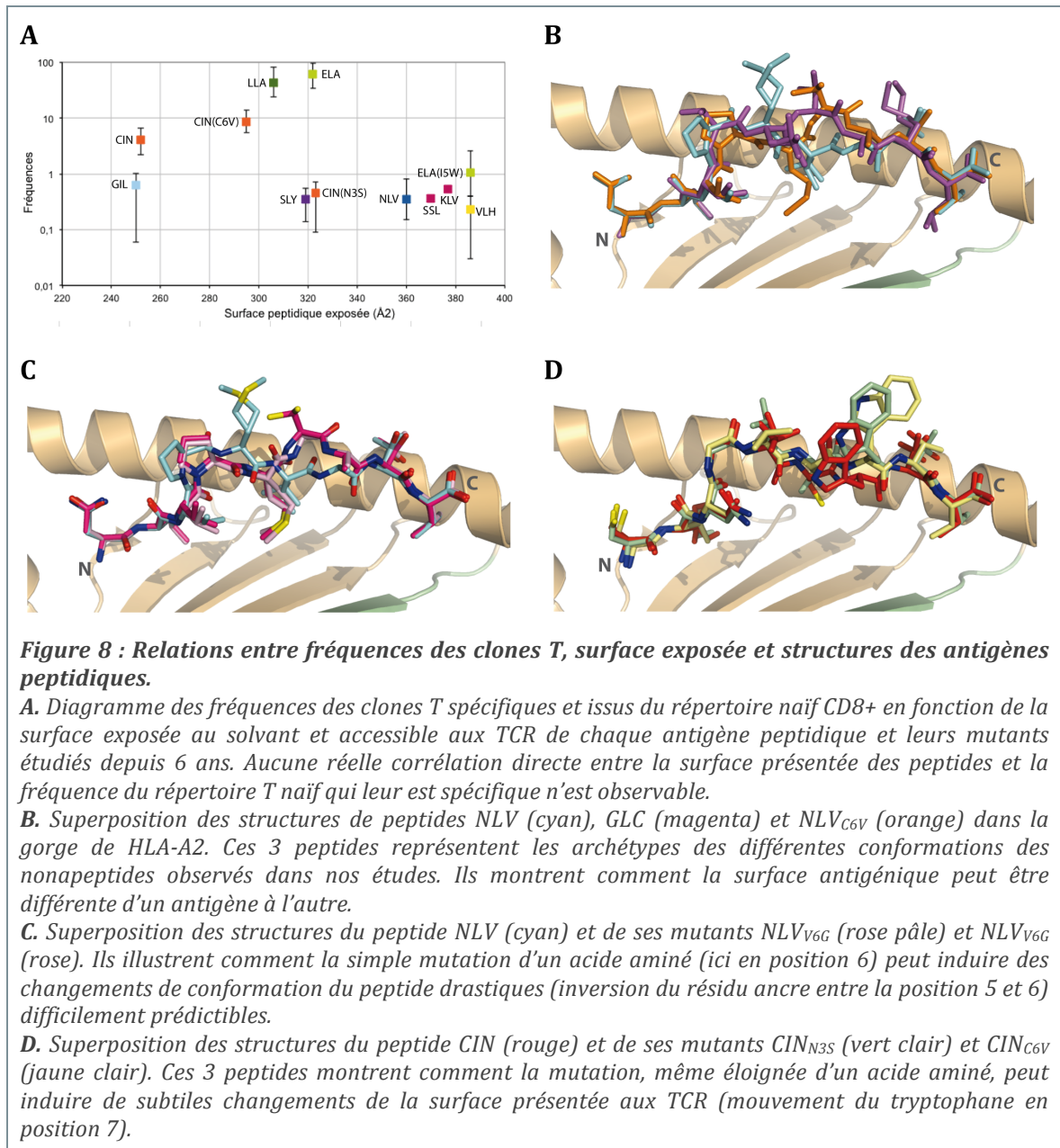
Dans le but d'apporter des éléments sur les règles qui gouvernent la reconnaissance TCR/pCMH, nous nous sommes focalisés sur les possibles corrélations entre les caractéristiques structurales des complexes binaires pCMH et les caractéristiques de fréquence des répertoires T spécifiques des différents épitopes que nous étudions. Cette approche comparative doit nous permettre d'obtenir des informations essentielles sur l'impact de la présentation de l'antigène sur la réponse T. En effet, comme nous avons pu le voir précédemment, la capacité d'un antigène à être efficacement présenté, la fréquence des cellules T réactives ou l'avidité des lymphocytes T sont potentiellement autant de paramètres corrélés avec les propriétés structurales et thermodynamiques des complexes pCMH.

A l'heure actuelle, nous avons pu résoudre la structure de plus d'une vingtaine de structures de complexes peptides/HLA-A2. Ainsi, nous avons maintenant à notre disposition les structures de HLA-A2 portant différents peptides viraux issus de HCV (Hépatite C Virus), d'EBV (Epstein-Barr Virus), de HCMV (Human Cytomegalovirus) et d'un peptide tumoral Melan-A présenté par les cellules du mélanome, ainsi que différents mutants de ces peptides. En parallèle, l'équipe de Nantes a effectué des mesures d'immunogénicité de ces différents peptides : grâce à des techniques d'isolation et d'enrichissement de clones T spécifiques pour un complexe peptide/HLA-A2 donné, basées notamment sur le tri cellulaire par le biais de billes magnétiques recouvertes par des complexes peptide/HLA-A2, nos collaborateurs de Nantes ont pu étudier en détails l'étendue des différents répertoires de cellules T capables d'être activés par les différents peptides naturels et mutants chez différents donneurs [44].

Précédant notre étude, Doherty et ses collaborateurs ont suggéré une corrélation entre la diversité/fréquence du répertoire T pour un antigène et l'accessibilité au solvant de l'antigène présenté par la molécule CMH. En effet, ils ont pu observer qu'un peptide issu du virus de la grippe (NP₃₆₆), qui est plutôt enfoui dans la gorge de la molécule CMH (H-2D^b), active un répertoire moins divers qu'un autre épitope plus exposé (PA₂₂₄). De plus, la mutation de ce dernier (Arginine en Alanine) permet de réduire à la fois la surface exposée du peptide et la diversité du répertoire T spécifique [84]. Aujourd'hui, nos résultats nous permettent de suggérer que les mécanismes gouvernant la fréquence de cellules T spécifiques d'un antigène donné sont beaucoup plus complexes que précédemment prédits.

Notre hypothèse de départ de corrélation entre la surface exposée du peptide hors de la gorge de fixation dans HLA-A2 et la capacité de ce peptide à activer un répertoire de cellules T ne semble pas être le seul paramètre structural mis en jeu. En effet, grâce à l'ensemble des données biologiques et structurales pour les peptides antigéniques dont nous disposons, nous n'observons pas de lien évident entre les variations de la surface antigénique et le nombre de clones T spécifiques (voir Figure 8). Par exemple, nous observons de très fortes fréquences pour des peptides présentant une surface se situant dans la moyenne (ELA, LLA) et de très faibles

fréquences à la fois pour des peptides enfouis (*i.e.* GIL) et le peptide le plus exposé (VLH). De plus, certaines mutations augmentant la surface peptidique présentée conduit à une diminution drastique de la fréquence des clones T spécifiques (*i.e.* ELA et ELA_{15W}).



Un autre paramètre qui pourrait influencer le caractère antigénique d'un peptide est sa mobilité dans la gorge du CMH. Dans certains cas, des changements de conformations du peptide et la stabilisation de chaînes latérales sont observés après complexation avec le TCR. Les coûts énergétiques de cette stabilisation pourraient jouer un rôle défavorable dans l'affinité globale du complexe TCR/pCMH : nous pouvons imaginer qu'un peptide moins mobile peut favoriser une meilleure interaction avec le TCR. Nos analyses structurales montrent une certaine variation dans l'agitation des peptides en complexe avec HLA-A2 qui pourrait être aussi un

paramètre à prendre en compte. Toutefois, là encore, il est difficile d'établir des règles directes entre la mobilité de l'antigène et les caractéristiques du répertoire T spécifique.

Enfin, des changements structuraux importants de l'antigène ou encore de fines variations structurales de chaînes latérales à la surface de la molécule HLA et imposées par les mutations du peptide peuvent apparaître sans avoir pu être prédits auparavant. Par exemple, bien que la plupart de peptides mutants que nous avons analysés adoptent une conformation identique à celle du peptide sauvage, nous observons une inversion des positions ancres pour certains antigènes mutés (*i.e.* peptide NLV, voir Figure 8) ; ou encore des mutations à une position du peptide peuvent avoir des répercussions structurales indirectes sur la conformation d'autres chaînes latérale du peptide (*i.e.* peptide CIN, voir Figure 8). Ces observations mettent en avant l'importance de la caractérisation structurale des antigènes pCMH pour évaluer l'impact des changements en acides aminés du peptide sur son immunogénicité et la réponse T qu'il est capable d'induire. En effet, il apparaît finalement qu'une seule modification d'un acide aminé peut changer de façon importante le registre d'atomes présentés au TCR et fortement impacter une large proportion de l'épitope.

En conclusion, les paramètres structuraux de la surface peptide/CHM et gouvernant la réponse T nous apparaissent aujourd'hui extrêmement complexes, multifactoriels et probablement caractéristiques à chaque antigène peptidique. Les données de fréquence des répertoires naïfs et spécifiques des antigènes que nous étudions nous laissent à penser que des mécanismes liés à la cross-réactivité inhérente des clones T et à la sélection thymique sont aussi mis en jeu et restent aujourd'hui à explorer. Un manuscrit décrivant ces travaux est en cours de préparation.

3.2.2. Un exemple structural d'une réponse T publique

En parallèle, la structure d'un complexe ternaire entre un TCR public (RA14) et son antigène viral NLV/HLA-A2 issu de HCMV a été résolue au laboratoire (S. Gras, étudiante en thèse 2003-2006). RA14 est un TCR issu d'un clone T commun à différents individus immunodéprimés et chez qui l'infection latente par HCMV est contrôlée par une réponse immunitaire présentant des clones T présentant des TCR fortement apparentés. Cette structure de complexe ternaire nous a permis de décrire un mécanisme de sélection de TCR public qui contraste avec d'autres exemples d'immuno-dominance de clones T lors d'infections virales comme la grippe ou HCV. Lors de ces infections aux caractères aigus, la sélection du TCR public se fait sur la base d'un seul « hot spot » peptidique reconnu de façon exclusive par un nombre limité de TCR [61, 64]. Dans notre cas, RA14 interagit avec l'ensemble du peptide et forme des interactions fortes avec trois résidus du peptide et un résidu de HLA-A2 (voir Figure 9). Les mesures d'affinités de RA14 par SPR que j'ai pu mener avec des mutants de NLV et de HLA-A2, ont confirmé le rôle majeur de ces quatre hot spots (voir Figure 10).

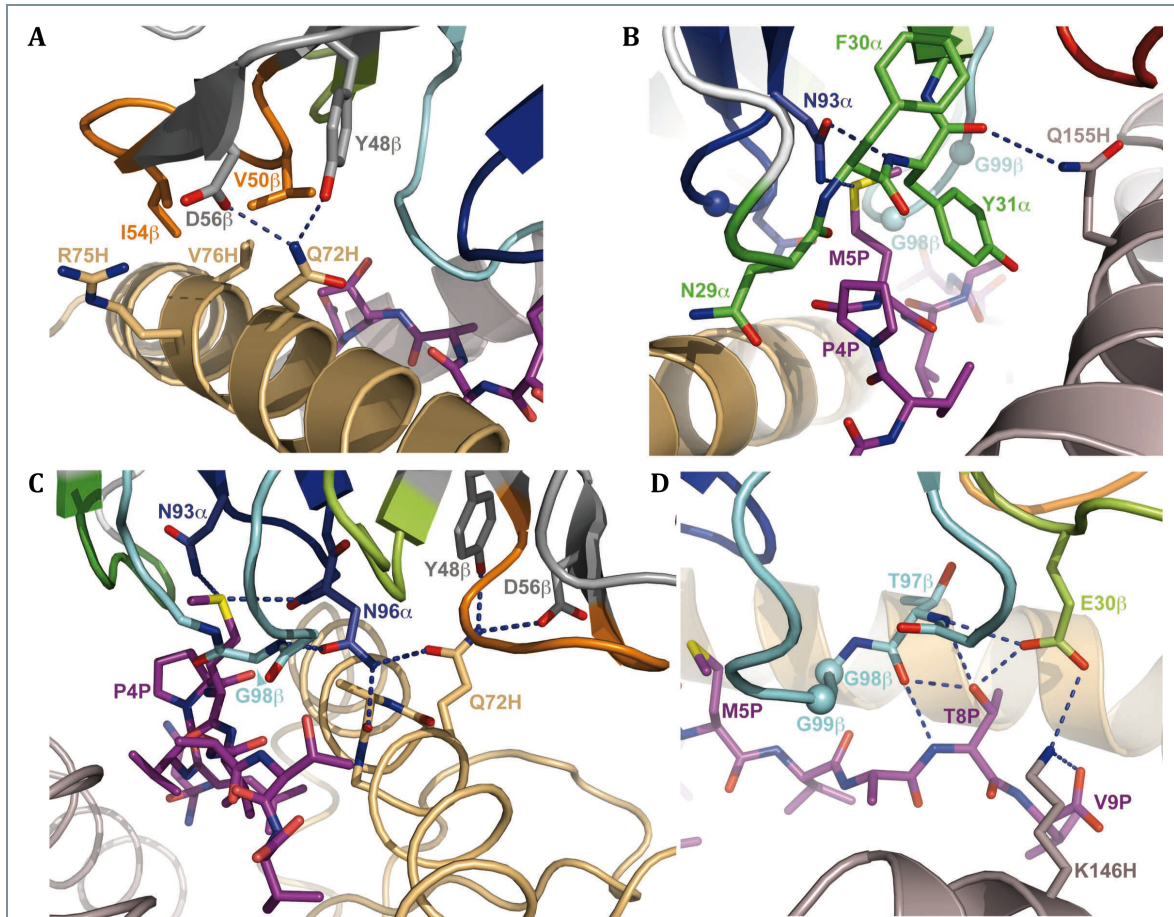


Figure 9 : Vues détaillées des interactions clés entre le TCR public RA14 et l'antigène NLV fixé à la molécule CMH HLA-A2
 au niveau **A.** du résidu Q72 de HLA-A2 (beige) et le CDR2β de RA14 (orange et gris), **B.** du résidu M5 du peptide NLV (magenta) et le CDR1α (vert), **C.** des résidus P4 et M5 (magenta) et les CDR3 (bleu clair et foncé), **D.** du résidu T8 (magenta) et les CDR1β (vert) et CDR3β (bleu clair).

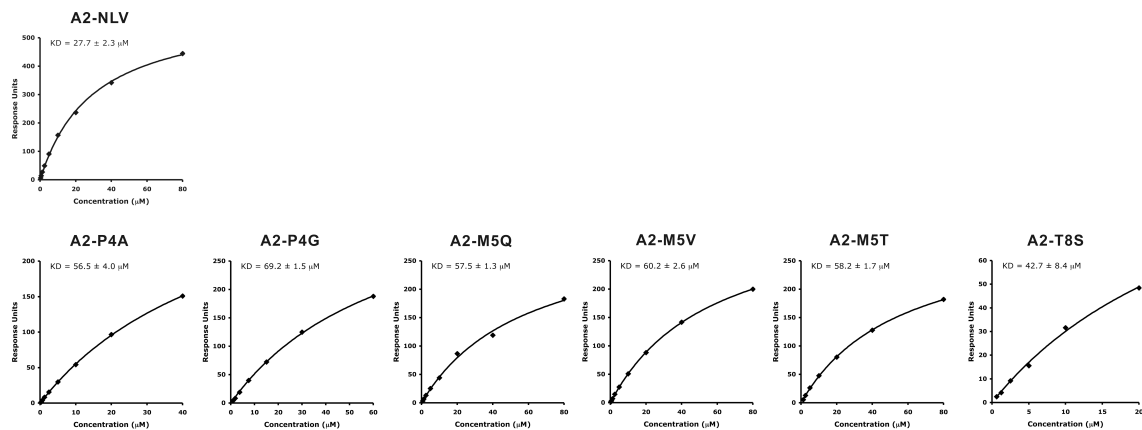


Figure 10 : Analyse par SPR des interactions entre le TCR RA14 et le peptide antigénique NLV ou ses mutants

L'affinité du TCR public RA14 pour différents complexes binaires HLA-A2/NLV et HLA-A2/peptide mutant a été mesurée par la technique de Résonance des Plasmons de Surface (SPR). En mutant les acides aminés qui forment des interactions clés avec le TCR, cette analyse nous permet de confirmer l'importance de ces dernières dans le mécanisme de reconnaissance de l'antigène par le TCR public.

De plus, nous avons obtenu la structure à moyenne résolution (4.7Å) d'un nouveau complexe ternaire entre un autre TCR public (RA15) et NLV/HLA-A2 (voir Figure 11). Cette structure ainsi que la conservation des aminoacides clés du TCR RA14 contactant la surface de l'antigène chez d'autres TCR spécifiques de NLV/A2 montre clairement que ces TCR de forte avidité peuvent adopter un mode de reconnaissance similaire. En parallèle, les mesures d'affinité de différents TCR anti-NLV/HLA-A2 nous ont permis de classifier ces TCR suivant leur pouvoir d'interaction, de corrélérer ce dernier avec l'avidité des clones T associés et avec le nombre d'acides aminés conservés avec RA14. Nous avons pu ainsi suggérer que le TCR RA14 représente la solution optimale d'interaction avec l'épitope. L'ensemble de ces données nous a ainsi permis de démontrer que l'émergence d'une solution publique peut aussi être basée sur un TCR présentant une complémentarité structurale favorable avec l'ensemble de la surface peptide/CMH.

Deux articles décrivant ce travail en détail ont été publiés dans *The Journal of Immunology* (publication P8) et dans le journal *Acta Crystallography section F* (publication P9).

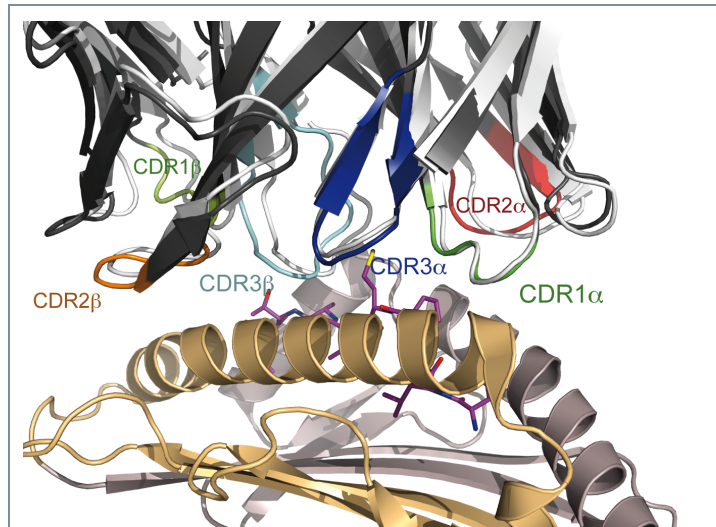


Figure 11 : Superposition des sites antigéniques des TCR RA14 et RA15

Représentation en mode ruban des TCR $\alpha\beta$ RA14 (dégradé de gris, code PDB : 3GSN) et RA15 (coloré). L'interface entre les 2 TCR et le même antigène NLV/HLA-A2 est quasiment conservée.

4. Projet à venir

Pour l'avenir, nous souhaitons continuer la mise en évidence des paramètres affectant la fréquence et diversité du répertoire naïf de cellules T spécifiques d'un antigène donné. En effet, les données immunologiques et structurales obtenues ces dernières années nous montrent que ces mécanismes restent très complexes et difficilement appréhendables. Des données préliminaires immunologiques obtenues par nos collaborateurs suggèrent des contributions intriquées à la fois de la réactivité croisée pour différents antigènes et inhérentes au répertoire T, mais aussi des caractéristiques résultant des processus de sélection intra-thymique. Nous continuerons nos travaux complexes de combinaison des données structurales et analytiques des répertoires T issus de patients.

Nous souhaitons ainsi focaliser notre attention sur l'importance du polymorphisme et de la restriction CMH, mais aussi l'implication de peptides du soi sur les événements de sélections intra-thymiques, pour identifier l'impact de ces antigènes sur la formation du répertoire T spécifique antiviral et anti-tumoral. Cette question pourra être résolue par exemple par la détermination de structures et la caractérisation (i) de TCR issus de clones spécifiques d'une molécule CMH et d'un antigène tumoral (HLA-A2/Melan-A) mais sélectionnés dans des contextes CMH différents (d'individus HLA-A2 négatifs) et (ii) de TCR issus de répertoires spécifiques du peptide tumoral (Melan-A) et d'un peptide du soi dont la séquence est similaire (PGT).

Enfin, nous nous focalisons aussi sur une série de lymphocytes T et leurs TCR spécifiques d'un nouveau peptide tumoral (MELOE-1) présenté par HLA-A2. Ces lymphocytes T présentent la particularité de prévenir les rechutes après réinjection chez des patients atteints du mélanome [53, 85]. MELOE-1 représente ainsi une cible thérapeutique très intéressante pour le développement de l'immunothérapie contre le mélanome. Par la résolution de la structure de complexes binaires HLA-A2/MELOE-1 et ternaires avec les TCR spécifiques, nous tenterons de proposer des analogues de ce peptide qui pourraient activer de façon optimale une réponse anti-MELOE.

Chapitre IV

Une nouvelle voie de signalisation de la phagocytose

La phagocytose de particules, et en particulier la phagocytose par les cellules spécialisées de l'immunité comme les macrophages ou les cellules dendritiques, joue un rôle essentiel dans le contrôle des réactions inflammatoires et dans l'activation d'une réponse immunitaire adaptative. Cependant, la réponse cellulaire à la phagocytose est modulée selon le type de l'élément internalisé. Lors de l'infection, les phagocytes sont responsables de l'élimination du pathogène, d'une réponse inflammatoire et de la stimulation de la réponse immunitaire qui en résulte. Chez un individu sain, les nombreux événements de développement, de remodelage et de maintenance des tissus et organes ou encore la réponse immunitaire conduisent continuellement à la formation de cellules dont la mort a été programmée, aussi appelée apoptose. La phagocytose et l'élimination des corps apoptotiques sont donc aussi indispensables pour assurer l'équilibre de l'organisme de manière tolérogène. Les différentes formes de réponse du système immunitaire chez les mammifères sont ainsi le résultat de l'activation de voies de signalisation spécifiques.

Les acteurs moléculaires impliqués dans la signalisation associée à la phagocytose restent à ce jour très mal connus et controversés dans les modèles mammifères. Deux familles diversifiées de protéines, DOCK et ELMO, ont été identifiées pour être au cœur de processus signalétiques qui conduisent au réarrangement du cytosquelette d'actine. En effet, le couple DOCK/ELMO est à l'origine de l'activation de petites protéines G de la famille Rho (*i.e.* Rac1 et Cdc42) connues pour réguler les processus de remodelage dans différents types cellulaires. Cette activation conduit ainsi à la modification morphologique de la cellule pour l'internalisation de particules et sa motilité.

Mon arrivée à l'Institut de Biologie Structurale a coïncidé avec le démarrage d'un nouveau projet de recherche ayant pour l'objectif d'élucider les voies de signalisation allant de la reconnaissance des cellules apoptotiques jusqu'à l'internalisation par les phagocytes professionnels aussi bien aux niveaux physiologiques qu'aux niveaux moléculaires et structuraux. Ce projet de recherche est développé par la synergie de plusieurs équipes de l'institut pour mettre en commun leurs expertises en immunologie, biologie moléculaire, biologie cellulaire, biochimie et biologie structurale.

1. Informations générales

1.1. Laboratoire et équipe d'accueil

Institut de Biologie Structurale Jean-Pierre Ebel, Grenoble

Laboratoire de cristallogénèse et cristallographie des protéines (2006-2011)

Groupe Réponse immunitaire aux pathogènes et soi altéré (2012-présent)

Equipe Housset

Partnership for Structural Biology, Grenoble

1.2. Compétences, synergies et financements

Ce projet de recherche s'inscrit dans un large projet du groupe « Réponse immunitaire aux pathogènes et soi altéré » qui s'intéresse aux mécanismes d'élimination des cellules en mort programmée, depuis leur reconnaissance jusqu'à la signalisation intracellulaire. Les aspects de signalisation auxquels nous nous intéressons tout particulièrement sont développés par la synergie de plusieurs équipes de l'IBS pour mettre en commun leurs expertises en immunologie, biologie moléculaire, biologie cellulaire, biochimie et biologie structurale.

Deux grandes approches sont actuellement utilisées par les différentes équipes :

(i) Une approche cellulaire par l'équipe de Philippe Frachet et Jean-Philippe Kleman pour l'étude fonctionnelle. Elle a pour but le développement de modèles cellulaires humains qui permettent d'identifier les protéines ciblées et leurs partenaires aussi bien au niveau des ARN messagers (qPCR), qu'au niveau des protéines (TAP-Tag, immuno-précipitation, biochimie des protéines). En parallèle, ces modèles permettent aussi la localisation des protéines cibles par des méthodes de microscopie optique grâce à une plateforme de microscopie développée à l'IBS par Jean-Philippe Kleman.

(ii) Une approche biochimique et structurale avec le clonage, l'expression, la purification, le marquage isotopique et la résolution de structure de cibles déjà connues (équipe de Dominique Housset et moi-même, groupe RMN), mais aussi la caractérisation des interactions entre les différents partenaires protéiques par résonance des plasmons de surface (SPR - Biacore) disponible dans le groupe et par RMN.

Pour ce projet, notre équipe a déjà obtenu pour l'année 2010 un financement de l'ANRS dont le porteur a été D. Housset. Différentes demandes de financement dont je suis le porteur ont été sollicitées ces derniers mois notamment auprès du Sidaction, de la fondation FINOVI ou de la région Rhône-Alpes.

1.3. Contribution scientifique

Publications : P12.

Communications : C13., C14.

2. Contexte scientifique

2.1. La phagocytose des corps apoptotiques : éliminer les cellules mortes

Chez les organismes multicellulaires, les nombreux événements de développement, de remodelage et de maintenance des tissus et organes ou encore la réponse immunitaire conduisent continuellement à la formation de cellules dont la mort a été programmée ou induite, aussi appelée apoptose. Chaque jour de la vie d'un homme, ces mécanismes provoquent la mort de milliards de cellules endommagées, âgées ou en excès. Leur élimination est donc cruciale pour maintenir l'homéostasie et l'équilibre de l'organisme [86].

L'événement cellulaire, assurément clé dans les processus d'élimination des cellules apoptotiques, est la reconnaissance et l'ingestion de celles-ci par les cellules phagocytaires. Il est assuré par les phagocytes professionnels, que sont les macrophages et les cellules dendritiques, mais aussi par des cellules voisines non-professionnelles comme les cellules épithéliales ou endothéliales et les fibroblastes. Outre le fait d'apporter la fonction d'« éboueurs cellulaires », les processus de phagocytose sont essentiels à la régulation immunitaire et au contrôle des réactions inflammatoires, notamment quand ils sont conduits par les phagocytes professionnels qui sont capables de sécréter des molécules anti- ou pro-inflammatoires et qui sont présentateurs d'antigènes. En effet, la réponse cellulaire à la phagocytose est modulée selon le type de l'élément internalisé. Lors de l'infection par un pathogène, les phagocytes sont aussi responsables de son élimination, d'une réponse inflammatoire ainsi que de la stimulation de la réponse immunitaire qui en résulte. Au contraire, dans le contexte d'un individu sain, la phagocytose des cellules apoptotiques reste tolérogène. Les différentes formes de réponse du système immunitaire face à l'objet internalisé chez les mammifères sont ainsi et certainement le résultat de l'activation de voies spécifiques de signalisation [87-89].

En condition normale, les processus d'élimination sont rapides et peu de cellules en apoptose sont observées dans les différents tissus et organes. Mais les défauts d'élimination des cellules sont associés à différentes pathologies (pour revue [90]). Par exemple, des conditions auto-immunes, comme le lupus érythémateux systémique, trouvent leur origine dans un défaut d'élimination de cellules de l'immunité [91]. Des maladies respiratoires, comme les maladies pulmonaires obstructives chroniques, la mucoviscidose ou l'asthme, présentent une accumulation de cellules en apoptose dans les tissus pulmonaires [92]. Dans les maladies cardiovasculaires issues de la formation de plaques d'athérosclérose, les macrophages présentent un ralentissement de l'ingestion de cellules apoptotiques [93]. L'excès des cellules en mort programmée est associé aux maladies neurodégénératives de type Parkinson [94] ou Alzheimer [95]. Les mécanismes de régulation de l'inflammation induits par l'apoptose et la phagocytose pourraient jouer un rôle dans le développement de tumeurs cancéreuses [96]. Enfin, certains pathogènes utilisent leurs facteurs de virulence pour usurper les mécanismes de phagocytose afin de

permettre leur invasion (*Shigella*) [97] ou moduler l'action des cellules immunes (HIV) [98].

2.2. La reconnaissance et la signalisation : les acteurs moléculaires

L'élimination des cellules apoptotiques se déroule par une série d'étapes caractéristiques. Initialement, le phagocyte se lie à la cellule pour une reconnaissance spécifique de la cible puis la signalisation intracellulaire conduit au remodelage du réseau d'actine et à la formation de la coupe phagocytaire. De façon ultime, la cible est complètement internalisée par la formation du phagosome et dégradée par sa machinerie enzymatique.

2.2.1. La reconnaissance : entre signaux « mangez-moi » et récepteurs

Pendant l'apoptose, les cellules exposent à leur surface des signaux moléculaires spécifiques de leur état. Avec la dénomination d'ACAMP (pour Apoptotic Cell-Associated Molecular Pattern), ils sont aussi qualifiés de signaux « mangez-moi » ou « eat-me ». Ces signaux incluent un répertoire complexe de molécules qui apparaissent à la surface, ou qui sont déjà présentes, et qui sont altérées par des mécanismes d'oxydation ou de modifications des chaînes de glycosylation (pour revue [88]). Parmi ces molécules « mangez-moi », la phosphatidylserine (PS) est le marqueur emblématique. Cette molécule est translocalisée depuis la face interne vers la face externe de la membrane de la cellule en apoptose [99].

Une fois à la surface, ces signaux « mangez-moi » sont reconnus par la cellule phagocytaire soit directement grâce à différents récepteurs membranaires spécifiques, soit indirectement par divers récepteurs solubles dits molécules pontantes (« bridging molecules ») ou « opsonines » qui sont à leur tour reconnues par d'autres récepteurs du phagocyte (pour revue [87, 88, 100]).

L'exposition et la reconnaissance de la PS ont fait l'objet de la grande majorité des études sur les mécanismes d'élimination des cellules apoptotiques comme nous le montre la littérature très foisonnante. Pendant longtemps, il a été cru qu'un seul récepteur de la PS pouvait être universel à toutes les cellules de type phagocytaire. Mais depuis plusieurs années, il apparaît que de nombreuses molécules, membranaires ou solubles, sont capables de se lier à la PS. Les récepteurs à la surface du phagocyte et capables d'interagir directement avec la PS incluent les membres de la famille TIM [101], BAI1 (brain angiogenesis inhibitor 1, [102]) ou de famille de la Stabiline [103]. Plusieurs molécules solubles pontantes ont aussi été identifiées. Sont à citer: l'Annexine-1 [104], MFG-E8 (milk-fat-globule-EGF-factor 8, [105]), Gas6 (growth-arrest-specific 6), la Protéine S [106], β 2-GPI (β 2-glycoprotein 2, [107]) ou encore des molécules du complément comme C1q [108]. Ces molécules sont capables de se fixer à la PS du côté de la cellule en apoptose et aux récepteurs du phagocyte comme au récepteur PS (Annexin-1, [104]), à certaines intégrines (MFG-E8, [109]), aux récepteurs de la famille TAM (Gas6 ou Protéine S ; [106]), au récepteur de β 2-GPI [110] ou au complexe entre la Calreticuline et CD91 (C1q, [111, 112]). D'autres sites incluant des sucres ou des phospholipides modifiés sont aussi la cibles de molécules

pontantes comme la TSP-1 (Thrombospondine-1) [113], certaines molécules du complément C3b/bi [114], les lectines MBL (mannose binding lectin) ou SP-A et D (lung surfactant protein A et D) [115]. Les récepteurs phagocytaires qui leur sont associés sont les récepteurs scavenger de classe A (SR-A) [116], lectin-like oxLDL-recepteur 1 (LOX1) [117], CD68 [118] ou CD36 [119] et le récepteur des collectines, le complexe Calreticuline/CD91 [115] (voir Figure 12).

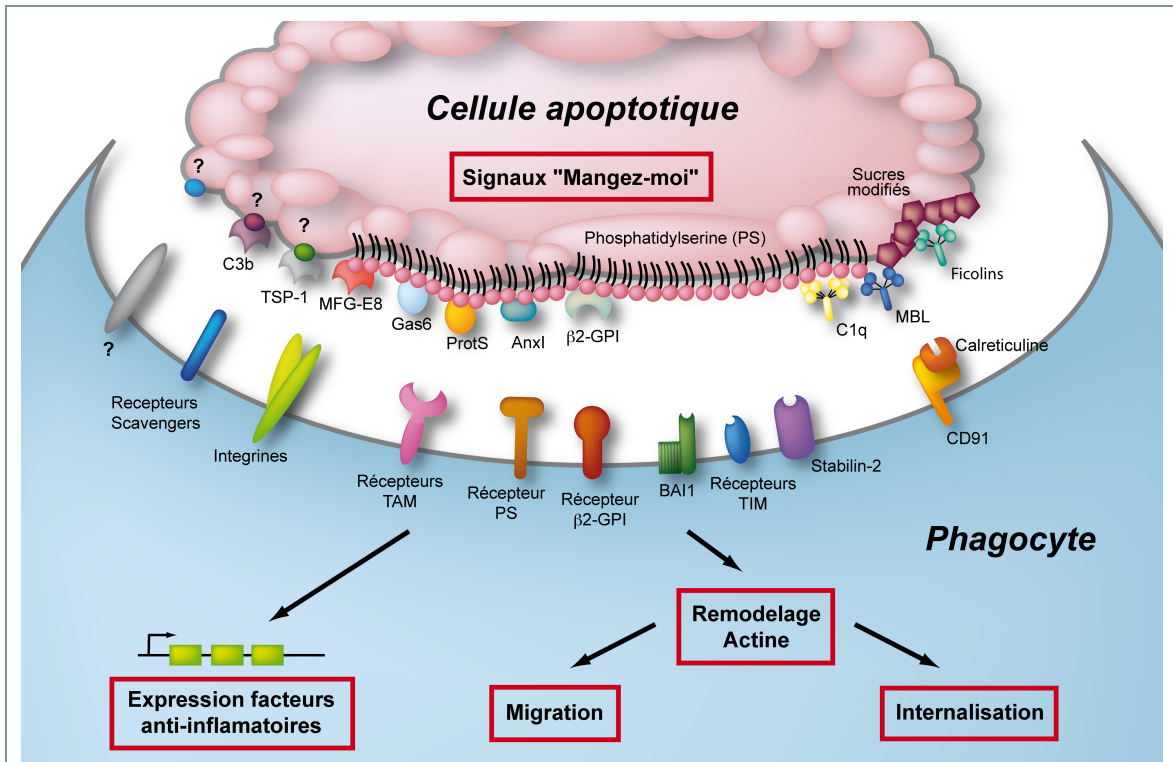


Figure 12 : Représentation schématique des molécules connues pour être impliquées dans la reconnaissance des cellules apoptotiques par les cellules phagocytaires

Les cellules en mort programmée exposent à leur surface des molécules « mangez-moi » comme la Phosphatidylsérine (PS) qui sont reconnues soit directement par les récepteurs du phagocyte (BAI1, TIM, Stabilin-2), soit indirectement par l'intermédiaire de molécules pontantes (MFG-E8, Annexin-1, C1q...) puis de récepteurs spécifiques (Intégrines, TAM, Calreticuline/CD91).

Il est à noter l'implication de nombreuses opsonines de l'immunité innée (molécules du complément C1q, C3b/bi et les Lectines) qui sont connues pour intervenir dans la réponse immunitaire sous des conditions de stress pathogéniques et pour induire une réponse pro-inflammatoire contre les microorganismes invasifs [120]. Ceci montre le parallèle et la similarité entre la reconnaissance des cellules altérées, aussi caractérisées de « soi-altéré », et la reconnaissance du « non-soi » pathogénique. Néanmoins, comme l'élimination des cellules apoptotiques n'induit pas de réponse immune mais, au contraire, stimule la sécrétion de facteurs anti-inflammatoires [121, 122], il doit exister des différences cruciales lors de la signalisation des deux processus qui sont encore aujourd'hui à déterminer.

Le nombre important de molécules impliquées dans la reconnaissance des cellules apoptotiques est le reflet de la complexité des mécanismes moléculaires mis en jeu pour leur élimination efficace, rapide et tolérogène. Il est aussi le reflet d'une

part, d'une redondance dans les signaux de signalisation nécessaires pour induire la phagocytose ou/et d'autre part, d'une hiérarchisation dans ces signaux. En effet, tous les récepteurs ne sont pas exprimés en même temps par toutes les cellules capables de phagocytose. Par exemple, les récepteurs de la famille TIM sont beaucoup plus abondants que BAI1 chez les macrophages, alors que les cellules de Sertoli expriment BAI1 de façon plus importante. Ainsi, suivant le type de cellules phagocytaires, un type de récepteur associé à son mécanisme de reconnaissance peut être nécessaire plutôt qu'un autre. De plus, il apparaît que la majorité des signaux « mangez-moi » et des récepteurs associés pris individuellement sont capables de promouvoir l'adhésion entre partenaires cellulaires sans toutefois induire un fort taux d'internalisation, la PS restant nécessaire à l'induction de l'internalisation [123]. Le modèle actuel proposé par Henson et ses collaborateurs suggère ainsi que dans le contexte d'engagements multiples de différents récepteurs phagocytaires et différents signaux « mangez-moi », certains servent à l'adhésion intercellulaire, alors que d'autres ont pour fonction de promouvoir la signalisation de reconnaissance et l'induction de la phagocytose ; l'interaction entre la PS et ses récepteurs étant alors l'interaction centrale [124]. D'après ce modèle, un parallèle évident peut être fait avec la synapse immunologique (voir Chapitre I, section 2.4.3) et le point de contact entre la cellule apoptotique et le phagocyte peut être alors vu comme une synapse d'internalisation avec un arrangement hiérarchique entre des différentes interactions intermoléculaires.

2.2.2. La signalisation intracellulaire : une machinerie conservée

Les événements qui suivent la reconnaissance des cellules apoptotiques et la signalisation intracellulaire qui conduisent à son internalisation et qui préviennent les réactions inflammatoires par la cellule phagocytaire restent encore à ce jour mal connus. Les premières informations sur la machinerie de signalisation proviennent des études génétiques sur le ver *Caenorhabditis elegans* chez qui plusieurs gènes ont été identifiés comme étant essentiels à la fonction d'élimination des cellules en mort programmée au cours du développement [125]. Par la suite, les gènes orthologues chez les mammifères ont aussi été caractérisés. Deux voies de signalisation sont aujourd'hui les objets de nos attentions. Elles sont partiellement redondantes et fonctionnent en amont des GTPases de la famille Rho (CED10 chez *C. elegans*) qui sont connues pour être des régulateurs clés du cytosquelette d'actine. Toutefois, les données restent aujourd'hui très parcellaires et extrêmement confuses. Cela est certainement dû à la complexité des mécanismes mis en jeu pour réguler efficacement la phagocytose et à l'implication de nombreuses molécules aux fonctions cellulaires multiples et redondantes ; mais aussi à la multitude des modèles cellulaires étudiés et des techniques expérimentales utilisées.

La voie GULP (CED-6)

La première voie de signalisation implique le récepteur Calreticuline/CD91 (CED-1), le transporteur ABC1 (CED-7) et l'adaptateur GULP (CED-6, voir Figure 13).

Le détail des mécanismes de leurs interactions reste encore à être clarifié. Mais aujourd'hui, les différentes études permettent de dresser un scénario suivant lequel des transporteurs ABC seraient responsables de l'externalisation de la phosphatidylsérine à la surface des cellules apoptotiques [126], à son tour reconnue par le phagocyte via le complexe Calreticuline/CD91 [127], la Stabiline [128] ou le récepteur scavenger SR-BI [129]. Le rôle précis des transporteurs ABC reste une énigme dans ce scénario car leur présence à la fois côté cellule cible et côté phagocyte est nécessaire pour promouvoir la phagocytose [130]. Le signal de reconnaissance est ensuite transduit par la protéine adaptatrice GULP qui interagit directement avec la partie cytoplasmique des récepteurs de surface CD91 [131], Stabiline [132, 133] ou SR-BI [129] pour promouvoir l'activation de la petite protéine G Rac1 et le réarrangement du réseau d'actine d'une façon encore inconnue. De façon intéressante, GULP a aussi été décrit dans des mécanismes de régulation et trafic du cholestérol [134] ou dans la régulation d'une GTPase de la famille Ras (Arf6) impliquée dans de nombreux processus comme l'endocytose, la sécrétion, l'adhésion cellulaire et la migration [135].

La voie DOCK/ELMO (CED-5/CED-12)

Au cœur de la seconde voie de signalisation identifiée aujourd'hui et qui va nous intéresser par la suite, se trouve un complexe entre les protéines de la famille DOCK et les protéines de la famille ELMO. Les protéines DOCK sont des Facteurs d'Echange de Guanine (GEF) atypiques pour les petites protéines G de la famille Rho et les protéines ELMO, leurs régulateurs (voir Figure 13) [87, 136, 137]. Les archétypes de ces deux familles, DOCK180 (CED-5) et ELMO1, (CED-12) ont été rapidement identifiés comme essentiels pour promouvoir la phagocytose chez le ver et les mammifères [125, 138, 139].

Toutefois, la cascade moléculaire et les mécanismes de recrutement de ce complexe DOCK180/ELMO1 sont encore mal compris et font aujourd'hui l'objet de nombreuses études (voir Figure 13). Plusieurs récepteurs semblent impliqués directement ou indirectement dans la cascade de signalisation de la reconnaissance qui conduit à l'activation de la GTPase Rac1 par DOCK180/ELMO1. Les études chez l'homme et le ver ont identifié le récepteur PS [123, 140], l'opsonine MGF-E8 et l'intégrine $\alpha_V\beta_5$ [109, 138] ou la molécule pontante Gas6 et son récepteur tyrosine kinase MER de la famille TAM [141] comme acteurs de la reconnaissance et de la transduction du signal vers la voie DOCK180/ELMO1. Le récepteur BAI1 serait aussi impliqué en interagissant directement avec la phosphatidylsérine via un domaine de type Thrombospondine et avec ELMO1 via son domaine cytoplasmique [102].

Les partenaires et les mécanismes de recrutement du couple DOCK180/ELMO1 à la membranaire plasmique apparaissent encore confus [142]. Les premières études suggèrent que son recrutement peut se faire par l'intermédiaire d'une protéine adaptatrice, CrkII [125] suite à l'engagement de l'intégrine $\alpha_V\beta_5$ et/ou du récepteur tyrosine kinase MER et l'activation de protéines kinases de la famille Src. Toutefois, bien que CrkII peut interagir avec DOCK/ELMO, cette interaction ne

semble pas obligatoire pour une phagocytose efficace [143]. De plus, il apparaît que DOCK180 et ELMO1 peuvent agir indépendamment de CrkII et par l'intermédiaire de voies régulées par la petite GTPase RhoG et sa GEF Trio [144]. En effet, RhoG sous sa forme liée au GTP est capable d'interagir avec ELMO1 [145] et leur surexpression dans des modèles cellulaires induit l'ingestion de cellules apoptotiques [146]. Les protéines DOCK possèdent aussi un domaine qui interagit avec les phospholipides membranaires et qui permettrait le recrutement direct du complexe DOCK/ELMO à la membranaire [142]. Enfin, ELMO1 peut être phosphorylé par la protéine kinase Hck de la famille Src [147] qui est connue pour être essentiellement exprimée chez les phagocytes et impliquée dans de nombreux processus cellulaires liés au remodelage de l'actine [148]. Pour le détail des mécanismes connus de recrutement du complexe DOCK/ELMO, nous pouvons nous référer à la section 2.5

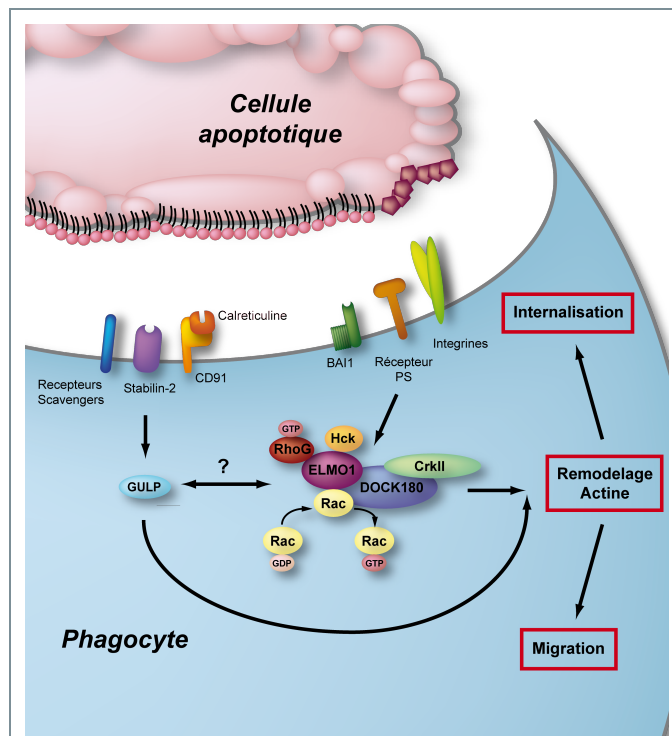


Figure 13 : Représentation schématique des molécules connues pour être impliquées dans la signalisation de phagocytose des cellules apoptotiques

2 voies distinctes ont été identifiées pour activer la petite GTPase Rac1 et le remodelage du cytosquelette d'actine qui mène à l'internalisation ou à la migration cellulaire : la voie GULP et la voie DOCK/ELMO

Le fait que les voies GULP et DOCK/ELMO conduisent simultanément à l'activation de Rac1 et du remodelage d'actine laisse entrevoir des mécanismes de redondance ou de coordination des différents effecteurs qui conduisent à une phagocytose efficace. La dernière hypothèse semble être privilégiée. En effet, Arf6 et sa GEF Arno, qui seraient régulés par la voie GULP (voir paragraphe précédent), pourraient aussi servir de médiateurs dans le recrutement du complexe DOCK/ELMO à la membrane [149, 150]. De plus, l'interaction entre la Stabiline 2, inducteur de la voie GULP, et l'intégrine $\alpha_v\beta_5$, qui active la voie DOCK/ELMO, suggère un mécanisme de coopération entre les différents signaux de reconnaissance pour orchestrer l'ingestion des cellules apoptotiques [151].

2.3. La famille DOCK : nouveaux facteurs d'échange de guanine

Comme toutes les petites protéines G (ou petites GTPases), les membres de la famille Rho cyclent entre deux états : un état inactif lié au GDP et un état actif lié au GTP, tous deux caractérisés par des conformations différentes. La conformation à l'état actif permet l'interaction avec des effecteurs spécifiques et l'induction d'une

cascade de signalisation (voir Chapitre V, section 5.1). Ces GTPases participent largement à la régulation de divers processus de modification morphologique, de polarisation, de migration, d'apoptose, de phagocytose, de transport ou encore de transcription de gènes [152, 153]. D'un point de vue classique, l'activité et la fonction d'interrupteur moléculaire de ces GTPases sont fermement régulées par trois classes de molécules : (i) les protéines activatrices des GTPases (GAP pour GTPase Activating Proteins) qui en se fixant aux GTPases sous forme GTP vont activer la fonction d'hydrolyse et favoriser la forme inactive ; (ii) les facteurs d'échanges de guanine (GEF pour Guanine Exchange Factor) qui en favorisation la forme libre de la GTPase permettent l'échange entre le GDP et GTP ; (iii) les protéines inhibitrices de GTPase (GDI pour Guanosine dissociation inhibitors) qui maintiennent les protéines G dans un état inactif (voir Chapitre V, section 5.1.4). Jusqu'à récemment, la conversion de la forme liée au GDP vers la forme active liée au GTP des protéines de la famille Rho n'a été crue possible que seulement par des GEF, dits classiques, présentant l'association de deux domaines universellement conservés dénommés Dbl-homologue (ou DH) et Pleckstrin-homologue (ou PH) [154, 155].

DOCK180 (CED-5) a été initialement identifié comme facteur d'échange de guanine pour Rac1 (CED-10) pour promouvoir la phagocytose chez le ver. Or, il s'avère que de nombreux orthologues de DOCK180 qui partagent les mêmes fonctions ont été identifiés chez les organismes eucaryotes évolués, notamment chez les mammifères et chez l'Homme [156]. L'ensemble de ces protéines forme une nouvelle famille de GEF atypiques pour les petites protéines G de la famille Rho. Appelée famille CDM (pour CED-5/DOCK/Myoblast City) [157], CZH [158] ou encore protéines « DOCK-related » (Dedicator Of CytoKinesis), elles stimulent spécifiquement l'échange GDP/GTP et l'activation des GTPases Rac1 et Cdc42, qui sont largement impliquées dans la régulation du cytosquelette d'actine. Les protéines DOCK sont atypiques au sens où elles ne présentent pas de domaine DH mais une nouvelle forme de domaines GEF, sans homologie aucune avec les domaines GEF classiques des GTPases Rho.

2.3.1. Les différents DOCK chez l'Homme

Les études bioinformatiques et phylogéniques ont permis d'identifier chez l'Homme 11 protéines DOCK qui peuvent être classées en 4 sous-familles en fonction de leur homologie de séquence (entre 50 et 65%) et de la présence des différents domaines spécifiques de la famille (voir Figure 14) [156]. De façon générale et chez les mammifères, les protéines DOCK sont très diversifiées. Et bien que les informations dont nous disposons sur les différents membres de cette famille soient très disparates, il apparaît qu'elles sont au cœur de nombreux processus signalétiques dans différents types cellulaires.

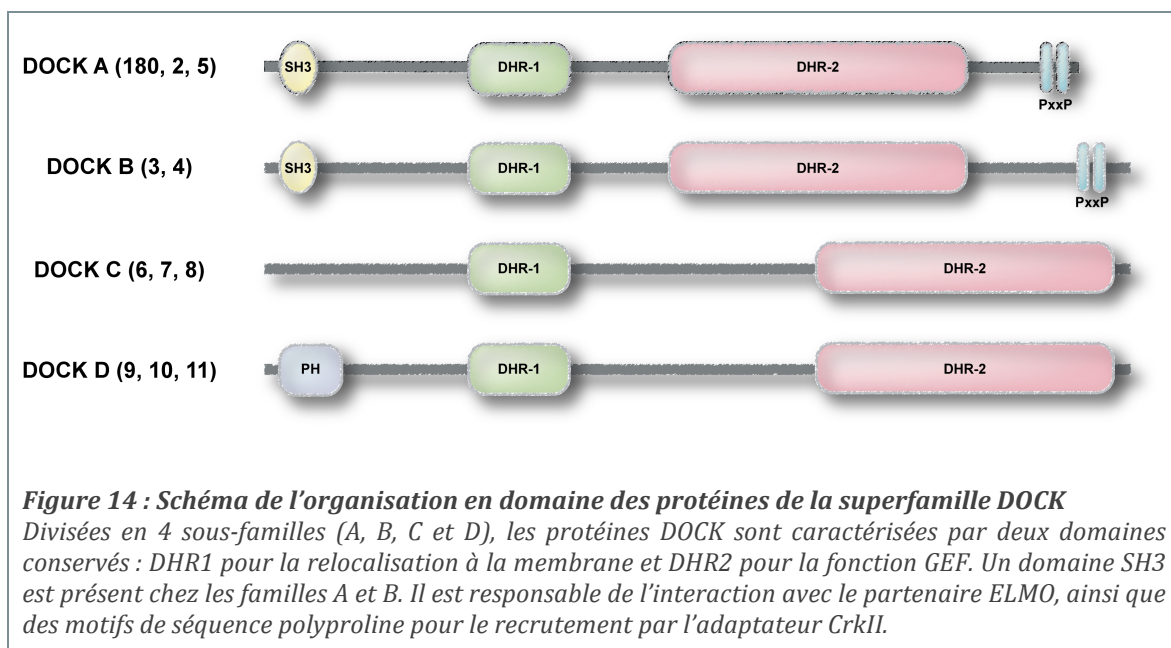
La famille A comprend DOCK180, DOCK2 et DOCK5 :

DOCK180 est largement ubiquitaire et spécifique de Rac1 et, comme nous l'avons vu précédemment, elle intervient dans la signalisation intracellulaire de la phagocytose des cellules apoptotiques (voir section 2.2.2). Mais cet archétype de la

famille joue aussi un rôle important dans les différents processus de mobilité cellulaire. En effet, surexprimée dans des cellules en culture, DOCK180 induit la formation de lamellipodes, augmente l'étalement et la migration cellulaire [130, 139, 142, 159, 160].

DOCK2 est exprimée sélectivement par les cellules hématopoïétiques et notamment par les lymphocytes et les cellules dendritiques, au contraire de DOCK180 qui y est absent [161, 162]. Différentes études ont montré le rôle important de DOCK2 dans le développement, l'activation et la migration des lymphocytes [157]. Par exemple, DOCK2 intervient dans la translocation des TCR et la formation de la synapse immunologique [163, 164].

DOCK5 est la protéine de cette famille la moins bien connue. Il a été suggéré qu'elle soit spécifique de Rac1 et tout comme DOCK180 impliquée dans la signalisation conduisant à la migration cellulaire. DOCK5 pourrait ainsi jouer un rôle complémentaire ou redondant à DOCK180. En effet, les déplétions de DOCK5 ou de DOCK180 conduisent à des phénotypes similaires de déficience de migration [165] et récemment il a été montré que DOCK5 peut former un hétérodimère *in vitro* avec DOCK180 [166].



La famille B est constituée par DOCK3 et DOCK4 :

DOCK3 (ou MOCA) et DOCK4 interviennent essentiellement dans les cellules de types neuronales. DOCK3 participerait à la croissance des axones [167, 168]. DOCK4 serait une GEF pour les GTPases Rap1 ou Rac1 et serait capable de réguler la croissance dendritique et la migration cellulaire dans les tissus neuronaux [169, 170].

DOCK6, DOCK7 et DOCK8 constituent la famille C :

DOCK6 et DOCK7 semblent aussi spécifiques des cellules neuronales. DOCK6 serait un GEF pour Rac1 et Cdc42 et il jouerait un rôle dans la croissance des neurites [171]. DOCK7 serait impliquée dans la régulation des microtubules dans les neurones

et dans la formation des axones [172] mais aussi dans le réarrangement du réseau d'actine dans ces cellules [173]. DOCK8, activateur de Cdc42, intervient dans les cellules immunes comme dans les cellules dendritiques, les cellules B ou T. Plusieurs études montrent que la déficience en DOCK8 compromet la signalisation dans le lymphocyte B ainsi que la production d'anticorps [174, 175], la survie des lymphocytes T [176] et que DOCK8 est essentielle pour la migration des cellules dendritiques [177].

DOCK9, DOCK10 et DOCK11 forme la dernière famille D :

DOCK9 (ou zizimine-1) est capable d'activer Cdc42 et intervient, elle aussi, dans la croissance des neurites [178]. De façon intéressante, DOCK9 est capable de dimériser via son domaine GEF [179]. La fonction de DOCK10 reste à ce jour inconnue. L'expression de son gène peut être induite suite à l'activation des lymphocytes B par l'interleukine 4 [180]. Enfin, la seule information dont nous disposons sur DOCK11 est qu'elle serait spécifique de Cdc42 et préférentiellement exprimée chez les lymphocytes [181].

2.3.2. Les domaines fonctionnels

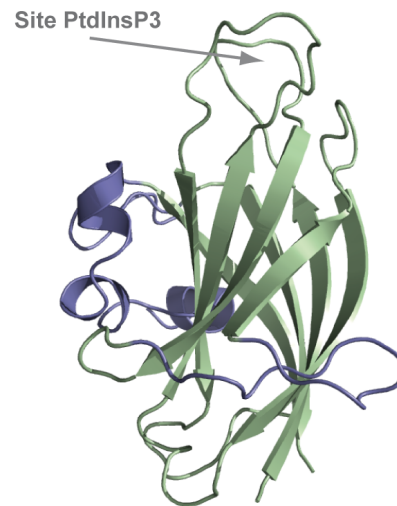
Les protéines DOCK sont de grands polypeptides de 1800 à 3100 acides aminés et elles sont composées de plusieurs domaines fonctionnels qui sont caractéristiques et spécifiques de la famille (voir figure 14). Elles ont tout d'abord été identifiées par bioinformatique au début des années 2000 [156] et leur fonction ainsi que les mécanismes de recrutement et de régulation sont depuis quelques années largement étudiés. La famille DOCK se distingue par deux domaines conservés, DHR1 et DHR2 (pour DOCK Homology Region). De façon intéressante, la comparaison des séquences des domaines DHR1 et DHR2 issus de différents organismes laisse penser que ces deux domaines ont co-évolué dans le contexte de la protéine entière [158].

Le domaine DHR1 :

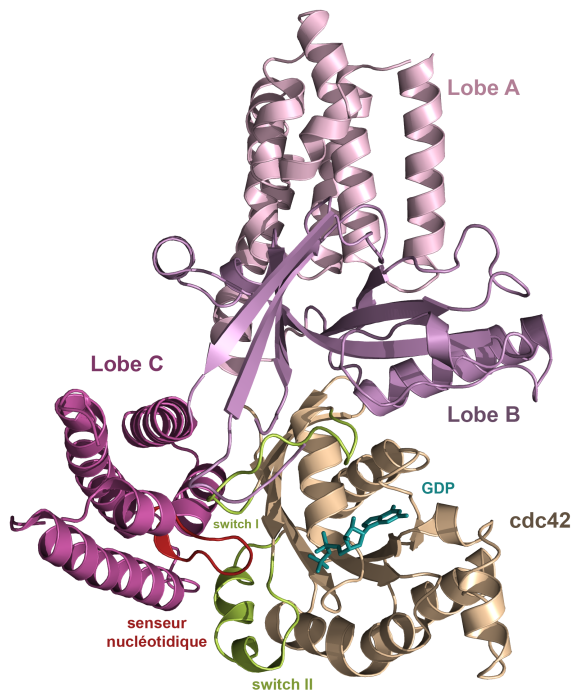
D'environ 200 acides aminés, le domaine DHR1 est invariablement associé en amont du domaine DHR2 chez tous les membres de la famille DOCK (voir Figure 14). Pour certains d'entre eux, il présente une faible homologie avec la famille de domaines C2, connus pour être des domaines d'interaction avec les lipides [156]. Le domaine DHR1 de DOCK180 est d'ailleurs capable d'interagir avec un des produits de la membrane : le PtdInsP3 (phosphatidylinositol(3,4,5)-triphosphate) et il est nécessaire pour la localisation de DOCK180 à la membrane plasmique et pour permettre l'élongation et la migration cellulaire [142]. La structure du domaine DHR1 de DOCK180, résolue en 2010, révèle un repliement en sandwich β similaire à celui des domaines C2, bien que l'identité de séquences avec les domaines structurellement les plus proches ne dépassent pas 16% (voir Figure 15). Le domaine DHR1 présente des insertions structurales non-triviales et une série de boucles qui formerait un site de fixation des phospholipides. Confirmé par des expériences de mutagenèse et de modélisation, ce site s'avère être différent du site de fixation qui est classiquement trouvé dans les domaines C2 [182].

Figure 15 : Structure du domaine DHR1 (à droite)

Représentation en mode ruban de la structure du domaine DHR1 de DOCK180 (code PDB : 3L4C). Son cœur (en vert) possède un repliement similaire aux domaines C2, connu pour interagir avec les phospholipides. Toutefois, il présente des insertions importantes (en bleu) et le site de fixation du ligand phospholipidique est différent.

**Figure 16 : Structure du complexe entre un domaine DHR2 et une petite GTPase Rho (à gauche)**

Représentation en mode ruban de la structure du complexe entre le domaine DHR2 de DOCK9 et de la GTPase cdc42 (code PDB : 2WM9). Cdc42 est en beige et ses régions switches en vert. Le nucléotide GDP en bleu. Les 3 lobes du domaine DHR2 sont en rose du plus clair (A) au plus foncé (C). La boucle active (« senseur nucléotidique ») est colorée en rouge. Elle vient s'insérer dans le site actif de la GTPase pour permettre le désengagement de l'ion magnésium et déstabiliser les interactions nucléotide/GTPase suivant un mode propre aux domaines GEF des protéines DOCK.

**Le domaine DHR2 :**

La fonction de GEF pour les GTPases Rho est portée par le domaine DHR2 (voir Figure 14). Contrairement au domaine DH des GEF classiques qui contient environ 240 acides aminés, les domaines DHR2 sont plus larges avec 450 à 500 acides aminés suivant la définition des extrémités. De façon similaire aux GEF classiques, il a été montré que le domaine DHR2 de plusieurs membres de la superfamille est capable d'interagir avec la forme vide de la GTPase pour laquelle il est spécifique [139, 156, 171, 179, 181, 183, 184]. Ainsi, les domaines DHR2 des familles A et B sont spécifiques de Rac1, la famille D et DOCK8 de la famille C interagissent avec Cdc42, et DOCK6 semble bénéficier d'une double spécificité pour Rac1 et Cdc42. La grande majorité des études fonctionnelles ont été focalisée sur DOCK180 et en accord avec les observations précédentes, elles montrent que son domaine DHR2 est nécessaire et suffisant pour promouvoir l'échange GDP/GTP et pour activer Rac1 à la fois *in vitro* et *in cellulo* [156, 185, 186]. Toutefois, des études cellulaires suggèrent que DOCK180 ne serait pas fonctionnelle en elle-même *in vivo*. En effet et dans certains cas, l'activation de Rac1 *in cellulo* est observée seulement si DOCK180 et ELM01, son partenaire, sont

tous les deux surexprimés dans les modèles cellulaires [139, 159, 187]. Cette apparente contradiction peut trouver son origine dans les mécanismes complexes de régulations et de recrutement du couple DOCK/ELMO que nous verrons par la suite (voir section 2.5).

Les structures des domaines DHR2 de DOCK2 et DOCK9 en complexe avec leur GTPase spécifique, respectivement Rac1 et Cdc42, ont permis de lever le voile aussi bien sur la structure de ce domaine GEF atypique, le mécanisme qu'ils utilisent pour permettre l'échange du GDP en GTP que sur les différences qui confèrent leur spécificité à la GTPase (voir Figure 16). De façon très intéressante, le domaine DHR2 présente une architecture différente des domaines DH retrouvés chez les Rho-GEF classiques. Il est constitué de 3 lobes de taille équivalente qui possèdent un repliement unique à la famille DOCK ; le site de fixation de la petite protéine G et d'échange du nucléotide étant formé par 2 des 3 lobes. De manière générale, les Rho-GEF en se liant à leur partenaire catalytique induisent des changements de conformations des deux boucles « switches » de la GTPase qui permettent la déstabilisation des interactions entre le nucléotide et son site de fixation. Notamment, la formation du complexe GEF/GTPase rompt des liaisons importantes entre le cycle guanine et le switch I ainsi que des interactions entre les groupements phosphates et la boucle P. De plus, la présence de la GEF impose un déplacement du switch II qui à son tour chasse l'ion magnésium présent dans le site catalytique (voir chapitre V, section 5.1.4) [155]. Il s'avère que les protéines DOCK et leur domaine DHR2 utilisent un mécanisme différent des Rho-GEF classiques pour permettre le relargage du nucléotide diphosphate de la poche catalytique de la GTPase. En effet, alors que le switch I de la GTPase subit un réarrangement similaire à celui qui est observé classiquement dans l'interaction entre les domaines DH et les GTPases, le switch II ne présente aucun changement structural. Le domaine DHR2 utilise plutôt une boucle, baptisée « senseur nucléotidique », pour exclure l'ion Mg^{2+} du site catalytique [188, 189].

Le domaine SH3 et le domaine hélical :

Les membres des familles A et B possèdent en leur extrémité amino-terminale un domaine de type SH3 (Src homology 3) (voir Figure 14). Les domaines SH3 sont extrêmement répandus chez de nombreuses familles de protéines. Ils sont connus pour interagir avec des séquences riches en Prolines, dites « polyproline » (PxxP) et pour être impliqués dans de nombreuses voies de signalisation [190]. Les premières caractérisations de la fonction du domaine SH3 de DOCK180 ont montré *in vitro* sa spécificité pour le motif polyproline d'ELMO1 (voir section 2.4) [191]. Elles ont aussi démontré le rôle *in vivo* de l'interaction SH3/PxxP dans la promotion de la phagocytose et de la migration cellulaire [192, 193]. Toutefois et de façon surprenante, cette interaction n'est pas indispensable *in vitro* pour la formation d'un complexe DOCK180/ELMO1 qui serait alors possible par l'intermédiaire de domaines hélicaux situés en aval du domaine SH3 de DOCK et autour du domaine PH d'ELMO (voir section 2.4) [193]. Le domaine SH3 est aussi capable d'interagir avec le domaine DHR2 de façon indépendante d'un motif polyproline absent dans la séquence de ce

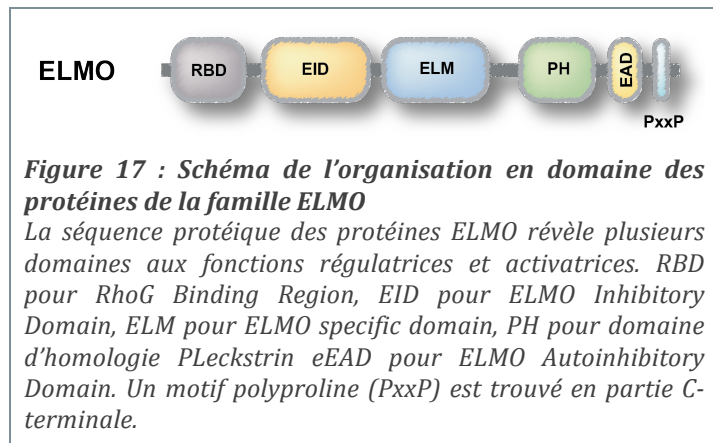
dernier. Cette interaction suggère un mécanisme d'auto-inhibition des protéines DOCK des familles A et B (voir section 2.5.2) [194].

Les séquences PxxP :

Enfin, la séquence des DOCK180 et DOCK2 à 5 présente plusieurs motifs de type PxxP en partie carboxy-terminale (voir Figure 14). Par ces motifs, DOCK180 est capable d'interagir avec la protéine adaptatrice CrkII et son domaine SH3 [195], ce qui permettrait son recrutement au niveau des récepteurs membranaires.

2.4. La famille ELMO : un régulateur ?

La famille de protéines ELMO (pour Engulfment and Motility) a été initialement identifiée chez *C. elegans* (CED-12) pour participer à la même cascade de signalisation que les protéines DOCK et pour activer les GTPases Rho conduisant aux mécanismes de remodelage de l'actine [159, 191]. Cette famille est conservée parmi tous les organismes étudiés jusqu'alors et on dénombre 3 membres chez les mammifères : ELMO1, 2 et 3. De façon surprenante ELMO1 et 2 présentent une identité de séquence remarquable et supérieure à 75% et elles peuvent facilement être interchangeables. De nombreuses études ont démontré l'interaction entre ELMO1/2 et DOCK180, DOCK2, DOCK3 et DOCK4 [139, 162, 196]. Bien que génétiquement très proche, le rôle d'ELMO3 reste aujourd'hui encore élué. Comme entraperçu précédemment, la fonction exacte des protéines ELMO dans le complexe avec les protéines DOCK reste très discutée. *In vitro*, ELMO1 stimule la fonction GEF de DOCK180 envers Rac1 de façon modérée et semble être dispensable, alors que dans certains modèles cellulaires, ELMO est requis pour favoriser l'activation de Rac1 et stimuler la phagocytose et la migration cellulaire [139, 159, 187]. Le mécanisme d'interaction entre DOCK et ELMO semble complexe et reste donc à être complètement élucidé. Toutefois, les fonctions de régulateur et d'activateur des protéines DOCK émergent : elles reposent sur les différents domaines identifiés.



Le domaine PH :

La première caractéristique découverte est la présence d'un domaine de type PH (Pleckstrin Homology) dans la partie C-terminale des protéines ELMO (voir Figure 17). Les domaines PH sont connus pour leur spécificité pour les phospholipides, mais aussi pour les phosphotyrosines et les peptides polyproline. Ils permettent en général le recrutement des protéines contenant ce type de domaines vers différentes localisations cellulaires [197]. Une première structure du domaine PH d'ELMO1 a été résolue en 2008 [193]. Pendant longtemps, il a été cru que ce domaine pouvait être le

pendant des domaines PH des Rho-GEF classiques et pouvait être ainsi responsable de la localisation à la membrane du complexe DOCK/ELMO. Mais, la structure a révélé qu'il est structuralement plus proche des domaines PH trouvés dans certaines phospholipases (*i.e.* PLC γ) ou protéines adaptatrices (*i.e.* Dual Adaptator of Phosphotyrosine and 3-Phosphoinositides – DAPP1) que des domaines PH typiquement trouvés chez les Rho-GEF classiques, certainement à cause des insertions différentes trouvées chez ces protéines. Mais de façon surprenante, le domaine PH d'ELMO1 ne contient pas les motifs structuraux nécessaires à la fixation des phospholipides et il a été confirmé par différentes méthodes que ce domaine est incapable de se lier aux différents lipides testés [193].

Mécanistiquement, la fonction du domaine PH d'ELMO reste évasive et est décrite comme coopérative à l'activité GEF de DOCK. En effet, le domaine PH ne se fixe ni aux protéines DOCK ni à Rac1 mais la surexpression d'ELMO augmente l'activité GEF *in cellulo* et la présence du domaine PH augmente l'activité GEF de DOCK *in vitro*. Il a ainsi été proposé que ce domaine coopère « *in trans* » à l'activité GEF du domaine DHR2 pour former un complexe tripartite DOCK/GTPase/ELMO qui stabiliserait l'interaction DOCK/GTPase [187]. Toutefois le rôle précis du domaine PH reste à être clarifié car certaines études montrent que la surexpression de DOCK180 seule est capable d'activer Rac1 et que le domaine DHR2 seul *in vitro* est efficace pour l'échange GDP/GTP.

Le domaine ELM :

Une séquence spécifique de la famille ELMO, appelée ELM, est trouvée au centre de la séquence (voir Figure 17). A l'heure actuelle aucune fonction n'est décrite pour ce domaine. Des domaines homologues ont été découverts chez des protéines activatrices des GTPases de la famille Ras (ELMOD1/2 pour la GAP et Arf-like pour les GTPases associées [198]). Toutefois, ce type de fonction reste à être démontré pour ELM01, 2 ou 3.

Le domaine RBD :

Les premières études de la séquence des protéines ELMO n'avaient pas identifié jusqu'à récemment de domaines spécifiques ou ressemblant à des domaines connus. Seules des séquences répétitives similaires aux séquences de type Armadillo ont été rapidement pressenties dans la première moitié N-terminale [146]. Une première zone a notamment été identifiée pour interagir avec différents partenaires. Cette séquence spécifique contient un domaine dénommé RBD (pour RhoG Binding Region) et ne présente aucune homologie avec des séquences protéiques connues (voir Figure 17). Elle est la cible de la GTPase RhoG sous sa forme active liée au GTP [145] ou encore, elle permet l'interaction directe entre ELMO1 et le récepteur membranaire BAI1 [102]. Ces deux effecteurs pourraient jouer un rôle important dans les mécanismes de recrutement et de régulation du complexe DOCK/ELMO à la membrane (voir section 2.5.1). Cette même région pourrait aussi être impliquée dans le recrutement de protéines de la famille ERM (Erzine-Radixine-Moesine) jouant aussi un rôle dans la formation et la régulation du cytosquelette [199].

Les domaines EID et EAD :

Récemment, des études bioinformatiques plus approfondies des homologies de séquence avec des domaines connus laissent entrevoir une fonction possible pour deux domaines différents. Ainsi certaines parties de la séquence d'ELMO présentent une homologie lointaine avec des domaines trouvés dans une famille de protéines, les Formines relatives à Dia1 (DRF). Ces protéines sont des acteurs majeurs de la nucléation de l'actine et sont nécessaires à la formation des filaments constituant le cytosquelette d'actine [200]. Les protéines de cette famille ont la particularité de présenter deux domaines d'autoinhibition : DID (Diaphanous Inhibitory Domain) au repliement Armadillo et DAD (Diaphanous Auto-regulatory Domain). ELMO présente deux domaines similaires naturellement dénommés EID (ELMO Inhibitory Domain) et EAD (ELMO Autoregulatory Domain) (voir Figure 17). Les études fonctionnelles ont permis de suggérer un modèle d'auto-inhibition basée sur l'interaction entre ces deux domaines (voir section 2.5.1) [166, 201].

La séquence PxxP et hélices α :

La séquence des protéines ELMO révèle aussi la présence d'une séquence riche en prolines située en partie C-terminale (voir Figure 17). Comme nous l'avons vu précédemment, ce motif interagit avec le domaine SH3 des protéines DOCK (voir section 2.3.2). Dans un modèle d'autoinhibition de DOCK, cette interaction ELMO/SH3 permettrait de rompre l'interaction intramoléculaire entre les domaines de DOCK et d'activer l'activité GEF du complexe tripartite (voir section 2.5.2). Une première structure entre la partie N-terminale de DOCK2 et la partie C-terminale d'ELMO1 a permis de révéler que le domaine SH3 et le polyproline d'ELMO forment une interaction canonique SH3/PxxP [202]. Toutefois, des expériences de co-précipitation montrent que cette interaction est dispensable à la formation du complexe. De plus, des mutants de délétion de DOCK180 et ELMO1 ne possédant pas ni le domaine SH3 ni le motif PxxP sont capables de promouvoir la migration cellulaire [193]. Ces mêmes études, ainsi que la structure du complexe N-terminal de DOCK2/C-terminal d'ELMO1, ont permis d'identifier les domaines minimaux de complexation : côté ELMO, le domaine d'interaction inclut le domaine PH et deux hélices qui l'encadrent en N- et C-terminal ; et côté DOCK, la région d'interaction est une région hélicale située après le domaine SH3 N-terminal (voir Figure 18).

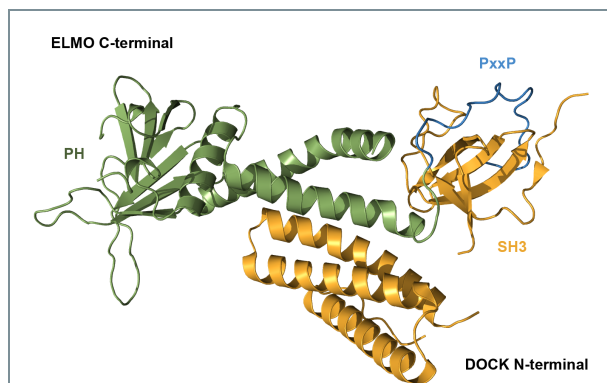


Figure 18 : Structure du complexe entre la partie N-terminale de DOCK et la partie C-terminale d'ELMO

Représentation en mode ruban de la structure du complexe entre les 167 premiers acides aminés de DOCK2 (orangé) et les 197 derniers acides aminés d'ELMO1 (vert et bleu). Le motif riche en Prolines d'ELMO est en bleu (20 derniers résidus). Code PDB : 3A98

2.5. Une régulation complexe

Les premières études ont suggéré un modèle de facteur d'échange de guanine atypique formé par le complexe DOCK/ELMO. En effet, certaines de ces études montrent que DOCK180 est catalytiquement active envers Rac1 seulement en complexe avec ELMO1 [139]. En accord, avec cette hypothèse, des mutants de délétion de DOCK180 incapables de se fixer à ELMO1 n'activent pas Rac1, même surexprimées en modèle cellulaire [159]. Une des premières explications a été l'élaboration d'un modèle bipartite de GEF dont les deux domaines fonctionnels équivalents aux domaines Dbl et PH des Rho-GEF classiques sont portés par les deux protéines différentes : DOCK avec son domaine DHR2 et ELMO avec son domaine PH. Aujourd'hui, l'ensemble des données obtenues pour les protéines DOCK et ELMO montrent que le recrutement, l'activation et les mécanismes régissant la fonction de facteur d'échange de guanine du couple de protéines DOCK/ELMO est beaucoup plus complexe que pressenti au premier abord. Ces protéines, en présentant des domaines spécifiques et uniques, montrent une extrême complexité dont les études commencent seulement à apporter quelques éléments de réponses.

2.5.1. Le recrutement et l'autorégulation d'ELMO

Plusieurs mécanismes pourraient aujourd'hui expliquer le recrutement de la protéine ELMO à la membrane plasmique où elle peut à son tour interagir avec ses partenaires et notamment former un complexe avec les protéines DOCK.

Par interaction avec la GTPase RhoG :

Différents éléments tendent à faire penser que le recrutement du complexe DOCK/ELMO par la GTPase RhoG pourrait être un des mécanismes pour permettre son activation et sa localisation au niveau de la membrane plasmique. Les études récentes ont permis de démontrer que la région située en partie N-terminale d'ELMO1 (domaine RBD) est capable d'interagir avec la GTPases RhoG sous sa forme active, liée au GTP (voir section 2.4) [145]. De plus, l'abrogation du complexe DOCK/ELMO ou de RhoG réduit de façon importante et similaire la migration cellulaire ainsi que l'étalement cellulaire [203] et il a été montré que RhoG recrute DOCK/ELMO pour promouvoir l'ingestion de billes de latex dans un modèle simple de phagocytose [146]. Toutefois, la pertinence physiologique de cette interaction reste encore à être pleinement déterminée. En effet, la délétion de RhoG dans des modèles murins ne conduit pas à des phénotypes particuliers [204].

L'auto-inhibition EID/EAD

L'homologie des domaines EID et EAD avec des domaines d'autorégulation présents chez les protéines homologue à la protéine DIA1 a conduit à étudier plus finement la possibilité de mécanisme similaire chez ELMO1 (voir section 2.4). DIA1 ou Diaphanous 1 et ses homologues forment une sous-famille des protéines Formines qui sont des effecteurs de la petite GTPase RhoA. Elles sont activées par l'interaction avec la petite protéine G qui vient rompre une interaction intramoléculaire entre deux domaines terminaux, DID et DAD [200]. Ainsi, il est observé que ces domaines

d'ELMO1 isolés sont capables d'interagir l'un avec l'autre, et que des mutations dans ces domaines qui pourraient abolir l'interaction EID/EAD conduisent à une augmentation de la motilité cellulaire [201]. Ces mêmes études montrent aussi que RhoG et le domaine EAD présentent une compétition dans l'interaction *in vitro* avec un fragment contenant les domaines EID et RBD, laissant entrevoir un possible mécanisme de levée d'auto-inhibition d'ELMO par la GTPase sous sa forme active similaire au mécanisme observé pour les Formines.

Par des récepteurs PS :

Les récepteurs à la surface des cellules phagocytaires impliqués dans la reconnaissance des signaux « mangez-moi » pourraient être directement engagés dans la localisation du complexe DOCK/ELMO. Les premières études génétiques chez *C. elegans* ont identifié le récepteur PS dans la même voie de signalisation que CED-2, CED-5 et CED-12 lors de la phagocytose des cellules en apoptose (voir section 2.2.2). De plus, il a été montré une interaction *in vitro* de CED-5 et CED-12 avec ce récepteur [140]. Chez les mammifères, l'interaction entre les molécules orthologues à celles du ver reste encore à être prouvée. Toutefois, un autre récepteur transmembranaire de la phosphatidylserine, BAI1, est capable à la fois d'interagir avec la PS grâce à son domaine Thrombospondine et avec une région N-terminale d'ELMO1 grâce à sa partie cytoplasmique (voir section 2.2.2) [102]. Ces mêmes études démontrent aussi que BAI1 coopère avec DOCK/ELMO pour induire l'ingestion des cellules apoptotiques.

Par phosphorylation :

ELMO1 peut être phosphorylé sur différentes Tyrosines par la protéine kinase Hck de la famille de protéines Src. Ces protéines kinases sont largement impliquées dans différents processus cellulaires liés au cytosquelette. La mutation de des Tyrosines d'ELMO1 réduit aussi la phagocytose et la migration de fibroblastes [147]. Toutefois, la preuve de l'existence de formes endogènes d'ELMO phosphorylées et leur rôle lors de la phagocytose ou de la migration cellulaire reste encore à être apportée.

2.5.2. Le recrutement et l'auto-inhibition de DOCK

En parallèle de l'activation et du recrutement des protéines ELMO, les protéines DOCK pourraient, elles-aussi, posséder leurs propres mécanismes de recrutement et d'activation sous-membranaires, garantissant ainsi une régulation fine de l'activation des GTPases Rho et du remodelage de l'actine.

Par crkII

DOCK180 a été originellement identifié comme un partenaire de la protéine adaptatrice CrkII [125] suite à l'engagement de l'intégrine $\alpha_v\beta_5$ et/ou du récepteur tyrosine kinase MER et à l'activation de protéines kinases de la famille Src. Il a été rapidement montré que cette interaction fait intervenir un domaine SH3 de CrkII et un motif PxxP de DOCK180 (voir sections 2.2.2 et 2.3.2). Cette interaction lie élégamment les différentes voies de signalisation régulées par CrkII et la voie

DOCK/ELMO dans la phagocytose des cellules apoptotiques. Toutefois, cette interaction ne semble pas obligatoire pour une phagocytose efficace, ce qui suggère l'existence éventuelle de voies alternatives [143].

Par les phospholipides

Plusieurs observations laissent à penser que le recrutement des protéines DOCK à la membrane par les phospholipides comme PtdIns(3,4,5)P₃ est un mécanisme majeur de régulation de l'activité des Rho-GTPases. Bien avant la caractérisation de l'interaction entre le domaine DHR1 et ces lipides (voir section 2.3.2), il avait été observé que des formes de DOCK180 ne possédant pas ce domaine conservé étaient incapables de promouvoir la migration cellulaire par l'intermédiaire de l'activation de Rac1 [142]. En accord avec la fonction maintenant connue du DHR1, le remplacement du domaine DHR1 par un domaine PH canonique d'adressage aux phospholipides dans une chimère de DOCK180 permet de restaurer la fonction GEF et la motilité cellulaire [142]. Bien que pour le moment aucune spécificité pour ces phospholipides n'ait été observée pour le domaine PH d'ELMO, il serait intéressant d'investiguer une possible synergie entre ce domaine d'ELMO et le domaine DHR1 de DOCK pour une localisation du complexe aux sites de production des PtdIns(3,4,5)P₃.

Par protection à l'ubiquitylation

DOCK180 est ubiquitynilé dans les modèles cellulaires, ce qui conduit rapidement à sa dégradation par le protéasome. Cette ubiquitynylation est en partie bloquée par la co-surexpression d'ELMO1 ou des fragments d'ELMO1 qui sont connus pour interagir avec DOCK180 [205]. Une des hypothèses envisagées est qu'ELMO protège DOCK180 de la dégradation par le protéasome d'une façon encore inconnue. Cette hypothèse est d'autant plus attractive qu'elle permet d'expliquer en partie pourquoi il est observé une activité accrue envers Rac1 de DOCK180 quand elle est surexprimée avec ELMO1. Dans ce contexte, l'accroissement d'activité peut être mise sur le compte d'une augmentation de la stabilité de DOCK180 dans la cellule en présence d'ELMO1 plutôt que sur le compte de l'augmentation des activités catalytiques.

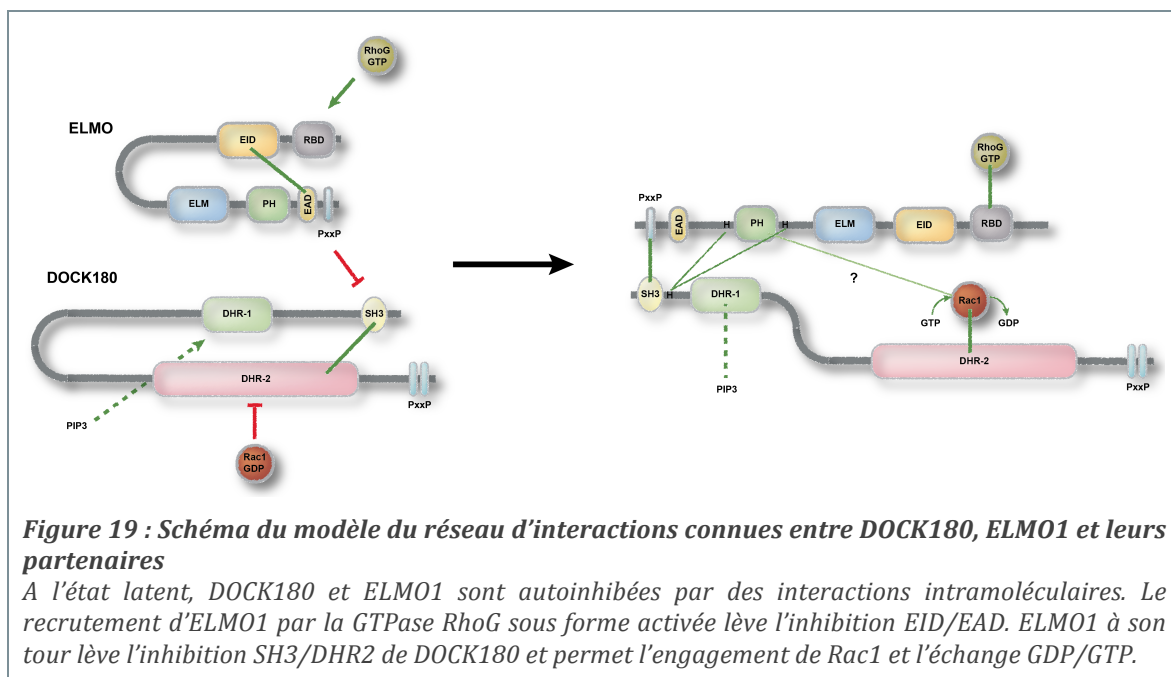
L'autoinhibition SH3/DHR2

Plusieurs observations laissent à penser qu'un mécanisme d'autorégulation est aussi présent pour les protéines DOCK possédant un domaine SH3 (famille A et B). En effet, si nous comparons à l'activation de Rac1 dans les mêmes conditions avec DOCK180 entier, la surexpression de mutants de délétion de DOCK180 pour lesquels il manque le domaine SH3 conduit à une augmentation relative de Rac1 sous forme GTP liée [194]. De plus, il a pu être montré une interaction directe entre les domaines SH3 et DHR2 de DOCK180 et cela sans intervention d'un motif polyproline absent de la séquence du domaine GEF. Ainsi, il a été suggéré que cette interaction SH3/DHR2 est à la base de l'inhibition de la fonction GEF de DOCK et prévient la fixation de la GTPase au domaine DHR2. Cette autoinhibition pourrait être levée grâce à la compétition par l'interaction entre le domaine SH3 de DOCK et la partie C-terminale d'ELMO (voir section 2.4).

Dimérisation ?

Différentes observations ont mené à l'idée récente qu'une oligomérisation des protéines DOCK peut intervenir et jouer un rôle dans l'activation de la fonction d'échange. Premièrement, il a été observé que DOCK9 se trouve sous forme de dimère. Cette dimérisation intervient par l'intermédiaire du domaine DHR2 ; chaque monomère fixant une molécule de la GTPase Cdc42 [179]. La structure du complexe a permis de démontrer que l'interface de dimérisation est indépendante du site de fixation de la GTPase [188]. Meller et ses collaborateurs ont aussi montré que l'on pouvait trouver DOCK180 sous forme de dimères *in cellulo*, sans toutefois montrer la pertinence fonctionnelle. Un rapport récent du groupe de Coté apporte aussi l'observation que DOCK5 peut être co-immunoprécipité avec DOCK180 depuis des extraits cellulaires [166], suggérant ainsi l'idée de la formation *in cellulo* de complexes dimériques de différentes protéines DOCK qui permettrait de moduler l'activité GEF.

2.5.3. Un modèle d'activation et de régulation



L'ensemble de ces données permet de dresser un scénario de l'activation du couple DOCK/ELMO bien qu'il soit encore entaché d'ombres et d'interrogations (voir Figure 19). A l'état latent, les deux protéines seraient sous des formes autoinhibées par l'intermédiaire d'interactions intramoléculeuses : ELMO1 par ses domaines EID et EAD et DOCK180 par ses domaines SH3 et DHR2. Grâce à différents recruteurs (CrkII, RhoG sous forme GTP, récepteur BAI1, phospholipides, phosphorylation), ELMO1 et DOCK180 migraient dans la région sous-membranaire. Certaines de ces molécules seraient alors capables en premier lieu de lever l'autoinhibition d'ELMO1 (RhoG, BAI1). ELMO1 ainsi activée interagirait avec DOCK180 par l'intermédiaire du domaine SH3 de DOCK180, du motif PxxP d'ELMO1 et d'un groupement de différentes hélices α des deux

partenaires. Cette interaction complexe lèverait l'interaction SH3/DHR2 et libérerait la fonction GEF du DHR2 pour Rac1 qui serait en partie stimulée par le domaine PH.

2.6. Les pirates Nef et IpgB1

Certains pathogènes sont capables de détourner les voies de signalisation qui mènent à l'activation des GTPases de la famille Rho et du remodelage du réseau d'actine pour favoriser leur entrée dans la cellule hôte et leur virulence [206, 207]. Pour arriver à ces fins, ils produisent des toxines qui miment ou activent les GTPases eucaryotes pour conduire à la réorganisation du cytosquelette. Par exemple, une nouvelle famille d'effecteurs bactériens et sécrétés par le système de sécrétion de type III, la famille WxxE, a été proposée comme inhibitrice ou activatrice des petites protéines G ou comme mime de leur états actifs [208, 209]. Dans cette famille, une sous-famille est formée par les protéines IpgB1 et 2 de *Shigella*, Map d'*E. coli* et les protéines SifA et B de *Salmonella*. IpgB1 est d'autant plus intéressante qu'elle est capable d'induire des modifications morphologiques de la membrane similaires à celles observées lors de la phagocytose. IpgB1 a été proposée pour jouer un rôle de mime de la forme active de la GTPase RhoG pour recruter le complexe DOCK/ELMO [97]. Toutefois le mécanisme de détournement des voies de signalisation par IpgB1 au profit de la bactérie reste discuté. En effet, IpgB1 pourrait plutôt avoir une activité GEF intrinsèque pour Rac1 similaire à IpgB2 avec qui elle partage 25% d'identité de séquence [210, 211].

La protéine Nef de HIV-1 pourrait être aussi un pirate des voies de signalisation régulées par DOCK et ELMO. Les virus HIV-1, HIV-2 et SIV utilisent leur protéine Nef pour modifier de façon indépendante le fonctionnement de plusieurs voies de transport et de signalisation cellulaire de l'hôte (Présentation des récepteurs CD4, molécules CMH, altération de la signalisation des cellules T...) [212]. Récemment, il a été mis en évidence une interaction entre Nef et le complexe DOCK2/ELMO1, inhibant la chimiotaxie des lymphocytes T pour un effet négatif sur le déclenchement d'une réponse cellulaire adaptée [98]. Le site de myristoylation et le motif polyproline de Nef ont été montrés comme essentiels pour l'activation de Rac1 et ils pourraient être des motifs de recrutement du complexe DOCK/ELMO.

3. Résultats marquants

Comme nous l'avons vu, les informations fonctionnelles et structurales du couple ELMO/DOCK et ses mécanismes de régulation dans les différentes lignées cellulaires sont à ce jour encore très parcellaires et restent un champ d'investigation très vaste. En combinant différentes approches de biologie structurale, de biochimie et de biologie cellulaire, nous nous attachons dans un premier temps à répondre à des questions ciblées autour du mode d'interaction entre DOCK180 et ELMO1 qui sont les archétypes des deux familles de protéines. Nous essayons notamment de comprendre quelles sont la nature et la fonction des différents domaines d'ELMO1 et comment la protéine ELMO1 régule l'activité du facteur d'échange de guanine DOCK180. De plus, la manière dont les facteurs de virulence de pathogènes, comme la protéine Nef de HIV ou la protéine IpgB1 de *Shigella*, recrutent le complexe DOCK/ELMO pour détourner la voie d'activation Rac1 dans le lymphocyte ou le macrophage reste aussi une question de premier ordre. Elucider le mode de détournement est essentielle pour notre compréhension des mécanismes de piratage des voies de signalisation qui permettent aux pathogènes d'échapper à la surveillance immunitaire et d'accroître leur virulence.

Au laboratoire, nous avons mis en place une stratégie de clonage et production en bactérie des protéines entières et des domaines fonctionnels pour les études de caractérisation structurales et fonctionnelles des interactions protéine-protéine. Nous avons ainsi pu obtenir différents domaines d'ELMO, de DOCK et leurs partenaires connus (i.e. Hck, Rac1, Nef). Depuis octobre 2008, une grande partie des travaux présentés ici ont fait l'objet de la thèse de Marion Sévajol que j'ai pu co-encadrer. Nous avons rencontré plusieurs difficultés majeures qui nous ont obligés à ajuster nos stratégies régulièrement. D'une part, la définition des domaines fonctionnels de chaque protéine reste encore à ce jour débattue et ce malgré les études de biochimie et de prédiction menées par notre équipe et nos concurrents. Nous avons donc dû tester de nombreuses constructions et différentes définitions de domaines afin d'obtenir des fragments solubles et d'une pureté suffisante pour les études structurales et biophysiques. D'autre part, les difficultés d'élaboration des protocoles pour caractériser les interactions des différents complexes restent des obstacles non-triviaux et ont accaparé une grande partie du temps que Marion a consacré à ces études.

3.1. L'interaction entre ELMO et le domaine SH3 de DOCK

Nous nous sommes dans un premier temps intéressés à comprendre les mécanismes de recrutement de DOCK180 par ELMO1. Comme nous l'avons vu précédemment, plusieurs études suggèrent des importances différentes dans la promotion de la phagocytose et de la migration cellulaire pour l'interaction entre le domaine SH3 situé en partie N-terminale des protéines DOCK et le motif polyproline situé en partie C-terminale des protéines ELMO. En effet les premières études ont montré *in vitro* la spécificité du domaine SH3 de DOCK180 pour le motif polyproline

d'ELMO1 [191] et elles ont démontré *in vivo* le rôle de l'interaction SH3/PxxP [192, 193]. Toutefois et comme le révèle une série d'expériences de co-précipitation avec différents mutants de délétion, la formation d'un complexe DOCK180/ELMO1 ne semble pas reposer sur cette interaction mais plutôt sur une interaction entre deux domaines hélicaux trouvés en aval du domaine SH3 et autour du domaine PH [193]. Pourtant, la structure entre la partie N-terminale de DOCK2 et la partie C-terminale d'ELMO1 montre bien une interaction canonique entre le domaine SH3 et le motif polyproline (voir section 2.4) [202].

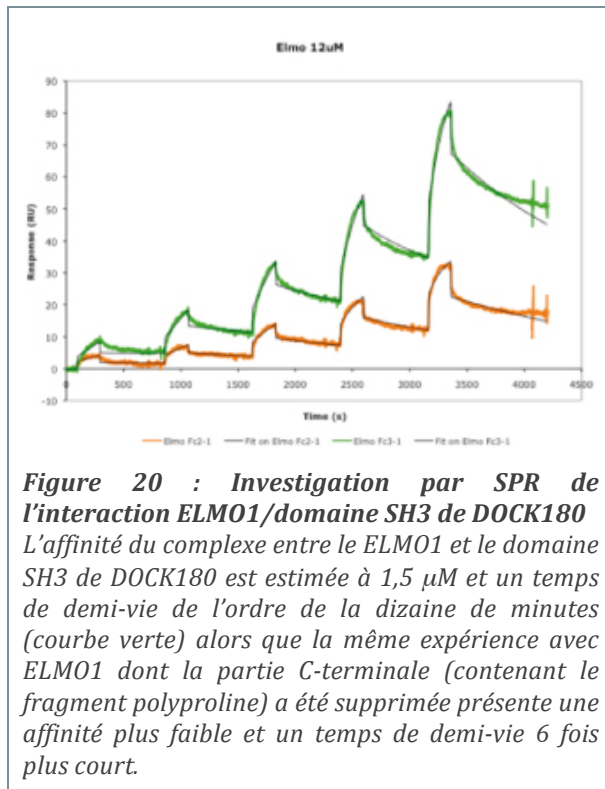


Figure 20 : Investigation par SPR de l'interaction ELMO1/domaine SH3 de DOCK180
L'affinité du complexe entre le ELMO1 et le domaine SH3 de DOCK180 est estimée à $1,5 \mu\text{M}$ et un temps de demi-vie de l'ordre de la dizaine de minutes (courbe verte) alors que la même expérience avec ELMO1 dont la partie C-terminale (contenant le fragment polyproline) a été supprimée présente une affinité plus faible et un temps de demi-vie 6 fois plus court.

micromolaire. De plus, et comme cela avait été pressenti, il apparaît que le motif polyproline d'ELMO1 n'est pas essentiel pour la formation du complexe SH3/ELMO. En effet, les différents mutants de délétion d'ELMO1 pour lesquels nous avons supprimé notamment le motif PxxP montrent toujours une affinité non négligeable pour le fragment SH3 de DOCK180 (voir Figure 20). En accord avec nos résultats, la structure du complexe SH3-DOCK2/C-terminal-ELMO1 révèle une interface de contact entre le domaine SH3 et le bundle d'hélices d'ELMO1 qui pourrait représenter l'interface minimale d'interaction.

Toutefois et bien que nos mesures SPR montrent que la présence du motif PxxP n'augmente pas significativement l'affinité des différents complexes testés par SPR, elle tend tout de même à augmenter leur temps de demi-vie. Ainsi nous pouvons proposer une formation du complexe DOCK/ELMO en deux temps. Un premier contact entre les deux partenaires se ferait par l'interaction primordiale entre un domaine de DOCK comprenant le domaine SH3 et le domaine hélicale qui le suit et un domaine structural d'ELMO composé par les deux hélices encadrant le domaine PH. Dans un second temps, la stabilisation et la fixation de la séquence polyproline

Dans le but d'apporter une réponse à cette apparente contradiction, nous avons réalisé lors de la thèse de Marion un ensemble d'expériences par résonance des plasmons de surface (SPR) et par résonance magnétique nucléaire (RMN, collaboration avec Pierre Gans). Elles nous ont permis de caractériser en détail l'interaction entre le domaine SH3 de DOCK180 et la partie C-terminale d'ELMO1 comprenant le domaine PH et le motif polyproline. En particulier et contrairement à ce qui avait pu être montré précédemment, nos résultats montrent que le domaine SH3 est capable d'interagir seul avec le domaine C-terminal d'ELMO1 et avec une affinité de l'ordre du

d'ELMO par le domaine SH3 renforcerait le complexe pour probablement permettre l'activation de la protéine DOCK et faciliter le recrutement du complexe DOCK/ELMO et sa fonction GEF.

Nous avons publié ces résultats dans un article paru dans *Biochimie* (publication P12).

3.2. Caractérisation structurale d'ELMO1

Nous avons aussi l'objectif d'élucider au niveau structural les mécanismes de régulation des protéines DOCK par ELMO. Afin d'avoir une image précise, nous utilisons deux approches expérimentales différentes : (i) la cristallographie nous permettra d'avoir une image à l'échelle atomique des interactions mises en jeu pour la formation des complexes et (ii) la diffusion des rayons X aux petits angles (SAXS) apportera des informations à basse résolution sur les changements conformationnels et les réarrangements éventuels de domaines.

Pendant la thèse de Marion, de nombreux essais de cristallisation d'ELMO1 entier et de ses différents domaines seuls ou en complexes avec le domaine SH3 de DOCK180 ont été réalisés. Toutefois et malgré plus de 15 000 conditions testées avec les différents domaines et complexes produits, aucune piste exploitable n'a été obtenue pour cette protéine qui semble réfractaire à la cristallisation.

Des expériences SAXS avec des échantillons de la protéine ELMO1 entière et des fragments contenant les domaines N-terminaux (ELMO Δ C) ou C-terminaux (ELMO Δ N) ont été réalisés. Ces données dans le domaine des bases résolutions nous ont permis d'obtenir un premier modèle *ab initio* de l'enveloppe tridimensionnelle de ces domaines. Grâce à la comparaison et la superposition de ces enveloppes, une première localisation des domaines a pu être faite dans le modèle d'ELMO1 entier. Toutefois, différentes mesures et d'autres méthodes de modélisation soulèvent plusieurs interrogations.

Tout d'abord et contrairement à tous les autres échantillons, les données obtenues pour le fragment ELMO Δ N, qui contient le domaine PH et le domaine hélical de la partie C-terminale, n'a pas pu être modélisé sous la forme d'un monomère mais sous la forme d'un dimère. Pourtant, les autres méthodes biophysiques utilisées pour caractériser l'état oligomérique de tous les échantillons nous suggèrent la présence exclusive de monomères lors de l'expérience SAXS. Notamment, nous avons utilisé la chromatographie d'exclusion de taille couplée à la diffusion de lumière (SEC-MALLS). Mais la présence de dimères de faible interaction et séparables par chromatographie ou la formation d'oligomères par dommage radiatif dû à l'exposition aux rayons X ne sont pas à exclure. La question de leur relevance physiologique reste donc ouverte. Vraisemblablement, cette dimérisation de ce fragment d'ELMO1 pourrait être en relation avec l'exposition de ses zones hélicales qui sont responsables de la complexation avec DOCK et qui pourraient induire une dimérisation artéfactuelle. Elle est peut être aussi à mettre en rapport avec l'observation de possibles

homodimères ou hétérodimères des protéines DOCK (voir section 2.5.2) et pourrait jouer un rôle dans la régulation du complexe.

Ensuite, une autre méthode de modélisation a été appliquée pour obtenir un modèle d'ELMO1 entier à partir de la structure connue du fragment qui contient le domaine PH et le domaine hélical et qui correspond à peu de chose près à notre fragment ELMO Δ N (voir section 2.4). Cette méthode, qui utilise les structures cristallographiques obtenues sur des domaines seuls et les données SAXS enregistrées sur la protéine entière, permet non seulement de reconstruire un modèle de la protéine entière mais aussi de déduire la position du domaine utilisé et de construire une enveloppe du fragment « manquant ». Une des premières observations que nous pouvons faire est que ce nouveau modèle d'ELMO1 correspond au modèle *ab initio* obtenu précédemment. De plus, la localisation du domaine ELMO Δ N est en accord avec notre première estimation de sa localisation dans l'enveloppe globale. Par contre, le modèle d'enveloppe obtenu pour le fragment ELMO Δ C à partir de cette dernière méthode ne correspond pas au modèle *ab initio* obtenu pour ce fragment N-terminal produit seul en solution (voir Figure 21). La comparaison de ces deux modèles nous suggère ainsi que ce fragment posséderait un repliement différent suivant qu'il est produit seul ou dans le contexte de la protéine entière. Il reste maintenant à déterminer si cette différence reflète un changement de conformation correspondant au modèle d'auto-inhibition ou un artefact dû à la production du fragment seul. En effet, ELMO Δ C contient le domaine EID mais pas le domaine EAD qui se trouve dans la partie C-terminale (ELMO Δ N, voir section 2.4). Dans le contexte où ce fragment est seul en solution, il pourrait ne pas subir les contraintes du domaine EAD comme dans le contexte de la protéine entière.

Enfin cette dernière enveloppe autorise un modèle de complexe grâce à une superposition avec la structure cristalline du complexe entre la partie C-terminale d'ELMO1 et la partie N-terminale de DOCK2 (voir section 2.4 et Figure 21). Ce modèle nous indique que la complexation du domaine SH3 et du domaine hélical serait possible sans changement de conformation d'ELMO, mais il n'exclue pas des changements quaternaires dus aux autres domaines de DOCK.

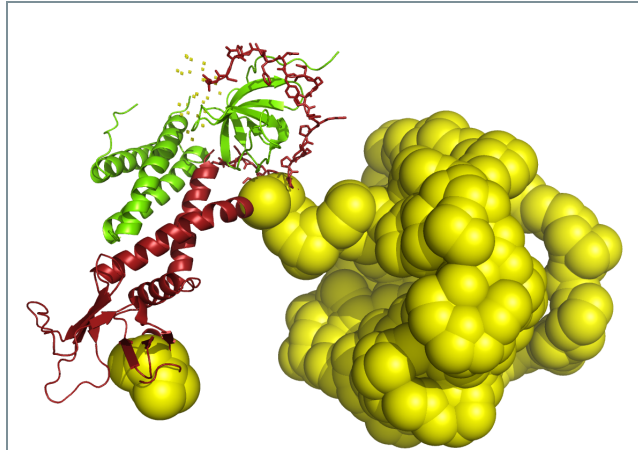


Figure 21 : Modèle combiné de la structure d'un complexe entre ELMO1 et la partie N-terminale de DOCK180

Le modèle du complexe a été construit par superposition du modèle d'ELMO1 entier (en jaune et rouge) et la structure cristalline du complexe entre la partie N-terminale de DOCK2 (en vert) et la partie C-terminale d'ELMO1 (en rouge). Le modèle d'ELMO1 entier a été obtenu en utilisant les données SAXS enregistrées sur un échantillon d'ELMO1, la structure cristalline de la partie C-terminale contenant le domaine PH et le groupement d'hélices α (en rouge) et après reconstruction du fragment N-terminal (en jaune).

3.3. Caractérisation préliminaire de l'interaction entre DOCK, ELMO et Nef

Récemment, l'interaction entre Nef du virus HIV-1 et le complexe DOCK2/ELMO1 a été montrée comme jouant un rôle dans l'inhibition de la chimiotaxie des lymphocytes T avec effet négatif sur le déclenchement d'une réponse cellulaire adaptée lors de l'infection par le virus du SIDA [98]. Nef est une protéine de 25 à 35kDa qui interagit avec un grand nombre d'effecteurs de la cellule hôte pour moduler les voies de signalisation et promouvoir la pathogénicité du virus. Bien que largement documentés, les mécanismes précis d'interférence utilisés par Nef restent encore aujourd'hui très mal compris et apparaissent d'une grande complexité. Plusieurs régions de cette protéine virale comprenant des motifs similaires à des motifs trouvés dans les protéines de l'hôte ont été identifiées: (i) un site de myristoylation est situé en partie N-terminale et permet les différentes localisations de la protéine (ii) une région cœur contient un motif polyproline pour ses interactions avec les effecteurs contenant des domaines SH3 comme les protéines kinases de la famille Src et (iii) deux régions flexibles : une première région située dans la partie C-terminale avec un motif di-leucine et une seconde région contenant un cluster acide [212-214].

Janardhan et ses collaborateurs ont montré que le site de myristoylation et le motif polyproline de Nef sont essentiels pour l'activation de Rac1 et qu'ils pourraient être des motifs de recrutement du complexe DOCK/ELMO [98]. Dans le but de vérifier une première hypothèse de recrutement par l'intermédiaire du domaine SH3 des protéines DOCK et du motif PxxP de Nef, j'ai mené des expériences préliminaires d'interactions par SPR entre différents domaines SH3 de protéines DOCK (DOCK180 et DOCK2) et différentes constructions de Nef (souche LAI) contenant ou non la partie N-terminale. Ces premières données montrent clairement une interaction entre le cœur de Nef et les domaines SH3, bien qu'elle soit faible et avec des affinités de l'ordre de la centaine de micromolaires (voir Figure 22).

Une de nos hypothèses est que Nef se substituerait ou amplifierait le rôle d'ELMO1 en interagissant directement avec le domaine SH3 de DOCK grâce à son motif PxxP. Mais les séquences différentes entre les domaines SH3 pourraient conduire à des affinités et des spécificités différentes pour Nef. En effet, bien que les bases de l'interaction Nef/ SH3 reposent sur l'interaction entre le motif PxxP et la région spécifique du fragment SH3, la spécificité de Nef peut aussi dépendre d'autres motifs. Les interactions caractérisées et les plus documentées sont entre Nef et les protéines kinases de la famille Src (Src, Fyn, Hck, Lck) [214]. Ces études ont montré

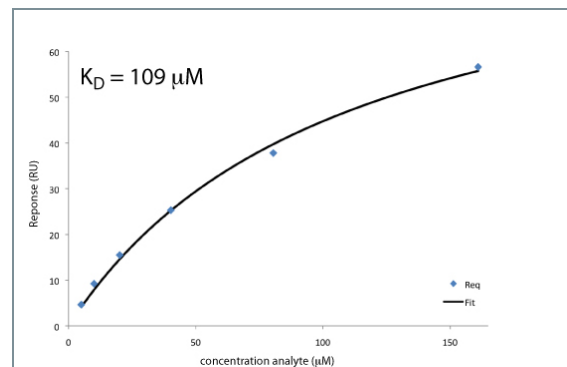


Figure 22 : Première estimation par SPR de l'affinité de Nef pour les domaines SH3 de DOCK

L'affinité d'interaction entre le domaine SH3 de DOCK2 et la région cœur de Nef a été déterminée par SPR.

une remarquable sélectivité de Nef pour les différents domaines SH3 dont les affinités peuvent varier de la centaine de micromolaires au nanomolaire. Les quelques structures (seulement 4 structures de Nef en complexe sont présentes dans la PDB) ont révélé non seulement l'interaction canonique entre le motif polyproline de Nef et les domaines SH3, mais aussi le rôle d'une autre région des domaines SH3, appelée boucle RT, et dont les variations de séquence sont responsables des différentes affinités observées [215-217].

Enfin, ELMO1 semble obligatoire pour l'action de Nef sur le complexe DOCK/ELMO et le détournement des voies de signalisation qui mènent à l'activation de Rac1 dans les cellules T [98]. Nef pourrait donc aussi interagir avec ELMO1 d'une façon encore inconnue. Pour vérifier cette hypothèse, des expériences préliminaires de co-précipitation menées par l'étudiante en thèse dirigée par Jean-Philippe Kleman, Rida Awad, avec les protéines disponibles au laboratoire suggèrent une interaction entre la région cœur de Nef et ELMO1. Toutefois, elle devra être confirmée par d'autres techniques de biochimie et de biophysique pour permettre de caractériser sa pertinence physiologique.

4. Projet à venir

Dans un contexte très concurrentiel, nous allons continuer nos efforts de caractérisation des différents domaines du complexe DOCK/ELMO ainsi que des interactions entre DOCK, ELMO et les différents partenaires connus. Ces données nous permettront d'apporter quelques réponses aux questions qui restent nombreuses et ouvertes sur le mode de régulation du complexe DOCK/ELMO.

Dans un premier temps, la fin de la thèse de Marion a été consacrée à la caractérisation de l'interaction entre la partie N-terminal de DOCK180 et la partie C-terminale d'ELMO1. En effet, nous avons vu que le domaine d'extension hélicale au domaine SH3 de DOCK180 serait fortement impliqué dans l'interaction avec la partie C-terminale d'ELMO1. Grâce aux différentes données obtenues par Marion au début de l'été grâce aux méthodes de SAXS et de SPR, nous tâcherons d'explicitier cette interaction. Nous avons notamment recueilli des données sur de nouvelles constructions protéiques de DOCK180, contenant le domaine SH3 et le domaine hélicale, et les mutants de délétion d'ELMO1 avec ou sans le motif polyproline. Nous serons ainsi à même de caractériser l'importance de l'interaction entre les domaines hélicaux en les comparant avec les données que nous avons déjà en notre possession avec le domaine SH3 seul.

L'équipe de Philippe Frachet et Jean-Philippe Kleman, en particulier leur étudiante en thèse Rida Awad, travaillent aussi à développer différentes méthodes pour identifier et caractériser *in vitro* et *in cellulo* les interactions possibles entre ELMO1 et ses différents partenaires connus ou encore à découvrir. A l'heure actuelle, ils s'attachent à caractériser les mécanismes d'interaction entre ELMO1 et la protéine kinase Hck, et notamment son domaine SH3. Les premières données semblent suggérer un mode d'interaction entre ce dernier et ELMO1. Ce mode serait différent du mode d'interaction ELMO/DOCK en ne faisant pas intervenir le motif PxxP d'ELMO mais plutôt sa région N-terminale. De façon étonnante et à l'encontre de ce qui est décrit dans la littérature, Marion a pu observer par SPR une interaction de faible affinité entre cette même région et le domaine SH3 de DOCK. Nous tâcherons d'explicitier son rôle. En parallèle, les nouveaux partenaires du couple ELMO/DOCK qui pourront être identifiés à partir des méthodes *in cellulo* seront aussi au cœur de notre travail de caractérisation structurale.

Par les méthodes que nous employons depuis cinq ans, nos objectifs seront aussi de développer une nouvelle thématique de caractérisation thermodynamique et structurale des interactions entre les domaines fonctionnels de DOCK ou ELMO et des partenaires pathogéniques. Nous souhaitons notamment clarifier (i) comment la protéine Nef de HIV-1 interagit avec le couple DOCK/ELMO ou encore (ii) comment la protéine IpgB1 de *Shigella* intervient dans le recrutement du couple DOCK/ELMO.

Pour Nef, il semble que son motif polyproline soit impliqué dans le recrutement du complexe DOCK/ELMO et l'activation de Rac1. A l'heure actuelle, les premières expériences de SPR que j'ai menées pour caractériser l'interaction entre Nef et les domaines SH3 de DOCK180 et DOCK2 montrent clairement une affinité de

faible ampleur, confirmant notre première hypothèse. Pour autant, le modèle d'activation du couple DOCK/ELMO suggère une interaction entre le domaine C-terminal d'ELMO et le domaine SH3 de DOCK. Par comparaison avec les données que nous possédons du complexe DOCK/ELMO, les données complémentaires de SPR et de structure du complexe SH3-DOCK/Nef nous permettront d'avoir une première approche afin d'identifier le mode de complexation que peut former Nef avec DOCK et ELMO. D'autre part, la présence obligatoire d'ELMO1 pour l'action de Nef laisse penser que cette dernière pourrait aussi interagir avec ELMO1 lors de la formation d'un éventuel complexe ternaire. Nous souhaitons donc profiter des avancées que nous avons faites sur ELMO et DOCK pour analyser thermodynamiquement et structurellement les interactions que Nef pourrait former aussi avec les protéines DOCK180, DOCK2 et ELMO1.

Le virus HIV est connu pour sa persistance dans l'organisme hôte par divers mécanismes de réplication aux différents stades de l'infection mais aussi pour sa réémergence depuis des réservoirs cellulaires. Très connu pour infecter les lymphocytes T CD4⁺, le virus est aussi capable d'infecter de nombreux types cellulaires de la lignée myéloïde et notamment les macrophages et les cellules dendritiques. Après infection de ces cellules, HIV-1 est capable de détourner les fonctions de migration cellulaire, de phagocytose et de production de cytokines pour promouvoir une latence et sa pathogénicité. Dans ce contexte, nous nous poserons aussi la question de la pertinence *in vivo* des interactions entre DOCK, ELMO et Nef. En effet, à l'heure actuelle, il a seulement été montré le possible détournement de la migration cellulaire des lymphocytes T par Nef et par l'intermédiaire du complexe DOCK2/ELMO1. Mais l'expression et l'implication des protéines DOCK diffèrent d'un type cellulaire à l'autre. Notamment, DOCK180 n'est pas exprimé dans les lignées cellulaires de type lymphoïde mais plutôt dans les lignées phagocytaires au contraire de DOCK2. Nos données préliminaires laissent entrevoir une possible interaction entre DOCK180, ELMO1 et Nef. Avec l'équipe de Philippe Frachet et Jean-Philippe Kleman, nous testerons par différentes méthodes la pertinence du recrutement du complexe spécifique des phagocytes par Nef dans un modèle cellulaire simple de phagocyte. Ces données seront une première approche pour évaluer l'importance du détournement des voies de signalisation de la phagocytose par le virus du SIDA.

En ce qui concerne IpgB1, il a été montré que cette protéine du facteur de sécrétion de type III de *Shigella* joue un rôle prépondérant dans le remodelage du réseau d'actine chez la cellule hôte pour favoriser son internalisation. Toutefois, la fonction d'IpgB1 reste controversée : d'un côté, il a été proposé que IpgB1 joue un rôle de mime de la GTPase RhoG sous sa forme activée pour recruter le complexe DOCK/ELMO en interagissant directement avec un domaine N-terminal d'ELMO. De l'autre, IpgB1 pourrait aussi avoir une activité de GEF pour Rac1. Nous nous proposons donc d'élucider la fonction d'IpgB1 en caractérisant structurellement et thermodynamiquement les interactions entre IpgB1 et ELMO (et/ou son domaine N-terminal) et aussi de comparer ces données avec celle du complexe RhoG/ELMO que nous pourrions aussi obtenir.

Chapitre V

Recherche de nouveaux agents anticancéreux et antibactériens

La recherche médicamenteuse et son application pharmaceutique contre les différentes maladies ont prouvé sans contestation leur efficacité dans l'augmentation extraordinaire de l'espérance et de la qualité de la vie humaine durant ce dernier siècle. Les efforts considérables de la recherche permettent aujourd'hui d'identifier chez l'Homme de plus en plus de mutations de gènes ou de dérégulations des voies de signalisation cellulaire associées à de très diverses pathologies. Depuis les années 80, la recherche et l'industrie pharmaceutique se sont ainsi focalisées dans la recherche de molécules bioactives ciblant les protéines incriminées et nous pouvons compter à l'heure actuelle plus de deux cents cibles protéiques pour deux mille molécules administrables afin de lutter contre des maladies aussi bien cancéreuses qu'auto-immunes ou inflammatoires. De même, la recherche de thérapies médicamenteuses contre les organismes microbiens a permis l'identification de nombreuses molécules antibiotiques qui sont aujourd'hui largement utilisées dans la médecine humaine et vétérinaire. Toutefois, les règles de plus en plus strictes issues des agences nationales de santé et les coûts de plus en plus importants pour qu'une molécule soit mise sur le marché poussent les industriels à réfléchir sur les critères de choix des cibles et d'évaluation des molécules. Dans le registre scientifique, les problèmes liés à la toxicité ou à la résistance aux molécules utilisées montrent aussi la nécessité toujours croissante de trouver des cibles protéiques pertinentes et de développer de nouvelles techniques de caractérisation des molécules.

Les protéines capables de fixer et d'utiliser les nucléotides représentent une des plus vastes familles de protéines dans le monde du vivant. Elles interviennent dans pratiquement tous les processus biologiques pour leurs capacités d'hydrolyse des nucléotides triphosphates. L'énergie dégagée permet ainsi le fonctionnement de machineries moléculaires, de catalyser des réactions chimiquement défavorables ou de modifier chimiquement des substrats moléculaires. Les fonctions de ces protéines sont extrêmement diverses et incluent des fonctions d'hydrolase, de synthétase, de ligase, de transférase... Et parce que leur implication est souvent vitale pour la cellule et parce que les nucléotides sont de petites molécules imitables, les protéines fixant les nucléotides représentent des cibles thérapeutiques de premiers choix.

Dans ce chapitre, nous allons nous intéresser à deux familles de protéines : les protéines kinases qui catalysent le transfert d'un groupement phosphate vers un peptide substrat et les GTPases, véritables interrupteurs moléculaires qui adoptent deux conformations suivant le nucléotide présent dans leur site.

1. Informations générales

1.1. Laboratoire et équipe d'accueil

Institut de Biologie Structurale Jean-Pierre Ebel, Grenoble

Laboratoire de cristallogénèse et cristallographie des protéines (2006-2011)

Groupe Réponse immune aux pathogènes et soi altéré (2012-présent)

Partnership for Structural Biology (2006-présent)

1.2. Compétences, synergies et financements

Ce projet de recherche est développé en deux volets par ma collaboration avec, pour une première partie, un consortium d'équipes européennes pour les aspects de recherche de molécules anticancéreuses et une collaboration entre notre équipe et l'équipe de Jean-Michel Jault de l'IBS pour le projet de caractérisation de GTPases bactériennes.

1.2.1. Recherche de nouveaux agents anticancéreux

Le premier volet concernant la recherche et l'optimisation de nouvelles molécules inhibitrices de protéines kinases aux pouvoirs anticancéreux est le fruit de la collaboration entre de nombreuses équipes pour associer des compétences en biologie, enzymologie, chimie, modélisation, biologie structurale. Pour cette recherche de nouveaux inhibiteurs spécifiques, l'équipe de Claude Cochet (CEA/Inserm, iRTSV, Grenoble) s'est entourée d'équipe de chimistes de l'Institut Curie de Paris (équipes de Jean-Claude Florent et Chi-Hung N'Guyen), de développement de criblage à haut débit (équipe de Laurence Lafanechère), et de biologistes structuraux (moi-même et l'équipe de Stefan Knapp, Université d'Oxford, United Kindom).

L'équipe de Claude Cochet est depuis longtemps soutenue par différents organismes et organisations non-gouvernementales (i.e. Ligue contre le cancer, Association pour la recherche contre le cancer...). Pour les aspects structuraux que je m'attache à développer, une nouvelle demande de financement a été récemment déposée en mon nom auprès du comité départemental de la Ligue contre le cancer.

1.2.2. Caractérisation de nouvelles GTPases bactériennes

Le projet de caractérisation de nouvelles GTPases bactériennes est basé sur la collaboration entre notre équipe, en charge des aspects biostructuraux, l'équipe de Jean-Michel Jault (groupe protéines membranaire de l'IBS), en charge des études de biochimie et d'enzymologie, et le groupe d'Ahcène Boumendjel (Département de pharmacochimie moléculaire de l'Université Joseph Fourier, Grenoble), en charge de la synthèse des bibliothèques de petites molécules.

La caractérisation structurale et la recherche d'inhibiteurs sont financées pour les années 2011 et 2012 par un financement PEPII du CNRS qui a été obtenu

récemment et dont J.-M. Jault est le coordinateur. Ce financement nous a permis de recruter un étudiant de master 2 (Marc Heitz) dont je supervise le travail.

1.3. Contribution scientifique

Publications : P6., P10., P11., P13.

Communications : C14.

Structures déposées (code PDB) :30WK, 30WJ, 30WL, 3MB6, 3MA3, 4DCS, 4DCT, 4DCU, 4DCV

2. Contexte scientifique : les protéines kinases

2.1. Les protéines kinases

2.1.1. Quelques généralités

Il est aujourd'hui généralement admis que les protéines kinases sont d'importants régulateurs des processus cellulaires en contrôlant de très nombreux évènements intracellulaires de transduction du signal. Ces protéines, qui intègrent, amplifient et transmettent les signaux vers leurs cibles, forment une superfamille d'enzymes. Elles sont retrouvées dans pratiquement toutes les voies de signalisation intracellulaires connues à ce jour et sont donc impliquées dans tous les aspects de la vie cellulaire comme le développement, la motilité, le transport, l'immunité, *etc.* De nombreux travaux ont montré le rôle essentiel des protéines kinases. En effet, l'inactivation génétique de nombre de ces enzymes donne lieu à des phénotypes fortement altérés et elle peut même être létale pour la cellule [218]. L'ensemble des gènes codant pour les protéines kinases chez l'Homme, aussi appelé kinome, forme une des familles de gènes les plus répandues dans le génome humain avec 518 gènes identifiés (environ 1,7% du génome) [219]. Et il est estimé que plus de 30% des protéines cellulaires peuvent être modifiées par leur activité [220] ; révélant ainsi la grande implication de ces enzymes dans la vie de la cellule et de l'organisme.

Biochimiquement, les protéines kinases catalysent la phosphorylation de substrats protéiques, c'est-à-dire le transfert d'un groupement phosphate depuis l'ATP vers un acide aminé phospho-accepteur d'une protéine (Sérine, Thréonine, Tyrosine ou Histidine) [221]. Historiquement découverte par G. Burnett et E. Kennedy [222], cette modification, qui est une des modifications post-traductionnelles les plus fréquentes, permet de modifier les propriétés physico-chimiques et/ou les conformations des protéines ainsi phosphorylées et leur permet d'interagir avec des partenaires de façons phospho-dépendantes.

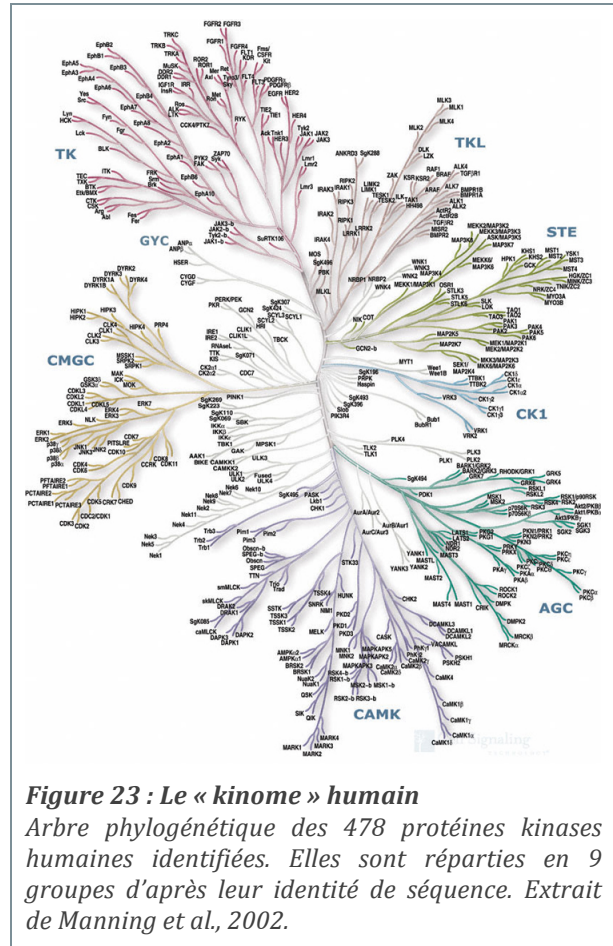


Figure 23 : Le « kinome » humain

Arbre phylogénétique des 478 protéines kinases humaines identifiées. Elles sont réparties en 9 groupes d'après leur identité de séquence. Extrait de Manning et al., 2002.

La phosphorylation des protéines est un phénomène dynamique, réversible et rapide qui est fortement régulé. La réaction inverse de déphosphorylation catalysée par les protéines phosphatases ramène notamment les substrats à leur état de base. Le cycle phosphorylation/déphosphorylation permet ainsi une régulation sensible et rapide des voies de signalisation.

2.1.2. Les protéines kinases : cibles thérapeutiques

Par leur fonction et leur implication à toutes les étapes de la vie cellulaire, leur dérégulation est à la source de nombreuses pathologies cancéreuses, cardiovasculaires, neurologiques, métaboliques ou inflammatoires (quelques exemples : [223-228]). Il est estimé que plus de 200 gènes de kinases sont reliés à des maladies et à de nombreux cancers. Il est même considéré que la mutation ou la surexpression de protéines kinases est une caractéristique des cellules cancéreuses [229, 230]. En étant dérégulées, l'expression de ces gènes, aussi appelés oncogènes, perturbe les voies métaboliques. Ces derniers procurent ainsi aux cellules cancéreuses un avantage adaptatif pour leur survie, leur prolifération et l'invasion des tissus environnants, conduisant au développement et/ou à la progression de cellules cancéreuses. Quelques exemples de ces gènes les plus étudiés ces dernières années sont à citer : le gène BCR-ABL pour la leucémie myéloïde chronique ; BRAF pour les cancers de la peau, ovariens, thyroïdiens, colorectaux ou pancréatiques ; c-RET pour des cancers de la thyroïde ; ALK pour les neuroblastomes, EGFR pour des tumeurs mammaires ou ovariennes ; c-Kit pour des tumeurs du stroma gastro-intestinal.

L'importance des protéines kinases dans la progression de nombreuses maladies a ainsi conduit à les considérer comme cibles thérapeutiques de premiers choix et font aujourd'hui l'objet de nombreux développements d'inhibiteurs pharmacologiques [231, 232]. Il est actuellement estimé que 30% des programmes de recherches pharmaceutiques sont tournés vers le développement de molécules ciblant les protéines kinases. Ainsi, une centaine de molécules sont recensées en essais cliniques à différentes phases et les inhibiteurs administrables oralement, utilisés essentiellement dans les thérapies anticancéreuses et commercialisés actuellement sont au nombre de 11 : Gleveec[®]/Fusadil[®]/Imatinib, Iressa[®]/Gefitinib, Tarceva[®]/Erlotinib, Nexavar[®]/Sorafenib, Sutent[®]/Sunitinib, Tassigna[®]/Nilotinib, Spryrcel[®]/Dasatinib, Tykerb[®]/Lapatinib, Rapamune[®]/Rapamycin.

Toutefois, avant ce succès grandissant des inhibiteurs de kinases dans la pharmacopée, l'industrie et la recherche pharmaceutiques ont longtemps été réticents à la recherche de molécules ciblant cette classe d'oncoprotéines, certainement à cause des limitations et des difficultés qu'imposent les propriétés de la famille des protéines kinases. D'une part, la majorité des inhibiteurs découverts à ce jour ont la capacité de se fixer au site ATP des kinases cibles. Afin d'être efficaces, ils doivent présenter une affinité importante pour contrecarrer les fortes concentrations intracellulaires d'ATP (entre 1 à 5mM) [233]. D'autre part, la forte conservation du site ATP des protéines kinases présage des difficultés au développement de molécules spécifiques pour une protéine ou une famille de

protéines. Des tests d'affinité sur des représentants des différentes familles de kinases montrent en effet que les profils de sélectivité des quelques dizaines d'inhibiteurs testés à ce jour peuvent être très opposés : certaines molécules sont capables d'inhiber efficacement un nombre restreint de cibles, d'autres sont très peu sélectives et inhibent la majorité du kinome avec des infinités de l'ordre de la centaine de nanomolaires [234]. Ce manque de sélectivité pourrait être à l'origine d'effets secondaires graves et freine le développement vers des molécules thérapeutiques efficaces et sûres.

2.1.3. La structure du domaine catalytique des protéines-kinases

L'analyse des séquences des 518 protéines kinases humaines montre que 478 d'entre elles possèdent au moins un domaine catalytique fortement conservé, dénommé domaine catalytique eucaryote (ePK) ; les 40 autres sont qualifiés d'atypiques (aPK) car présentant des séquences non-caractéristiques de l'activité kinase, bien que pour certaines d'entre elles, leur fonction catalytique ait été démontrée *in vitro* [219]. Les études structurales sont depuis plusieurs années très fécondes en termes de connaissances des mécanismes de phosphorylation des substrats par cette superfamille de molécules et en termes des modes d'inhibition par les petites molécules. Depuis la première structure d'une protéine kinase, cAPK en 1991 [235], près de 1000 structures de complexes entre différentes kinases et divers ligands naturels ou synthétiques ont été déposées dans la PDB. Ce nombre impressionnant de structures cristallographiques est un atout inestimable dans la compréhension fine des interactions intermoléculaires entre ces enzymes et les molécules inhibitrices. En combinant informations structurales et connaissances d'activité inhibitrice, il est possible d'exploiter ainsi des détails structuraux précis de l'enzyme et d'optimiser les composés dans le but d'augmenter leur affinité et leur spécificité [232, 236].

Les premières structures obtenues dans les années 90 ont confirmé la conservation de l'organisation structurale du domaine catalytique ePK pressentie par la similarité des séquences. Elles ont aussi montré la remarquable conservation du site actif. Le domaine kinase est constitué d'environ 270 acides aminés constituant deux lobes distincts reliés par une région charnière. Plusieurs motifs structuraux caractéristiques des domaines ePK participent à la régulation et à l'activité enzymatique : la boucle A (ou « A-loop »), la région charnière, la boucle P (ou « P-loop »), l'hélice C, le motif DFG et la boucle catalytique. La phosphorylation du substrat se produit dans le site catalytique à l'intersection des deux lobes et comprend deux poches juxtaposées : le site de fixation de l'ATP et le site de fixation du peptide (voir Figure 24).

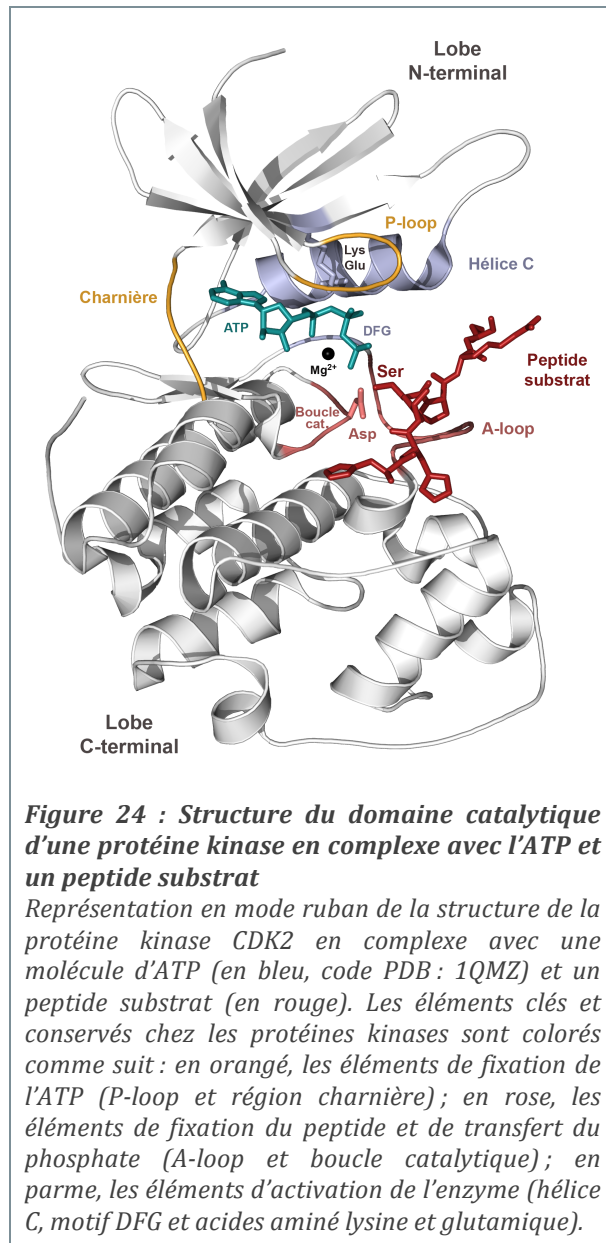
Le site de reconnaissance du peptide substrat est formé essentiellement par la boucle A portée par le lobe C-terminal. Sa séquence et sa structure constituent deux des rares éléments qui diffèrent d'une protéine kinase à l'autre et confèrent à l'enzyme sa spécificité pour la cible protéique.

Le site de fixation de l'ATP est formé par des acides aminés provenant des deux lobes. La base adénine est prise en étau dans une poche formée de différents acides aminés au caractère majoritairement hydrophobe. Elle est par ailleurs orientée par deux liaisons hydrogènes avec la chaîne principale de la région charnière. Le sucre ribose est stabilisé par des liaisons hydrogènes avec des acides aminés du lobe C-terminal. Les trois groupements phosphates sont quant à eux alignés par la boucle P, riche en glycines (voir section 2.2.4 et Figure 26).

Quatre acides aminés conservés ont été identifiés dans pratiquement tous les domaines ePK et sont essentiels à la catalyse du transfert du groupement phosphate γ de l'ATP vers le phospho-accepteur (voir Figure 24). Une première Lysine située dans un brin central du lobe N-terminal est stabilisée grâce à une liaison ionique avec un résidu Glutamate provenant de l'hélice C activatrice (voir paragraphe suivant). Cette Lysine participe à l'orientation correcte des groupements phosphates α et β de l'ATP. Un troisième acide aminé important est l'acide Aspartique du motif conservé DFG qui est situé en amont de la boucle A. Il permet la coordination d'ions magnésium nécessaires au positionnement des phosphates β et γ . Enfin, un dernier acide Aspartique appartenant à la boucle catalytique et à son motif conservé HRD sert à orienter le groupement hydroxyle du résidu phospho-accepteur.

Un dernier acide aminé appelé « gatekeeper » est à mentionner. Situé à la limite entre le feuillet $\beta 5$ et la région charnière, il varie en séquence entre les différentes protéines kinases.

Pourtant, les différentes études de recherches d'inhibiteurs ont identifié ce résidu comme jouant un rôle important dans la sélectivité. En effet, suivant l'encombrement stérique présenté par l'acide aminé à cette position, une nouvelle poche au caractère hydrophobe apparaît proche de la poche de fixation de l'adénine, permettant ainsi d'étendre la sélectivité du site pour différentes molécules spécifiquement inhibitrices.



D'une manière générale, les protéines kinases transitent entre deux conformations catalytiquement actives et inactives sous l'effet de divers modes de régulation qui peuvent être très différents d'une kinase à l'autre : par exemple une phosphorylation (dans le cas de cAPK) ou l'interaction d'une sous-unité régulatrice (dans le cas de CDK2). A chaque fois, ils font intervenir l'hélice C et le motif DFG qui se trouvent dans différentes conformations suivant l'état de l'enzyme. Dans les formes inactives, l'hélice C est généralement dans une position éloignée, positionnant le résidu Glutamate hors du site catalytique. Dans les formes actives, l'hélice C est dans une orientation plus proche et permet le positionnement du Glutamate dans le site et l'orientation de la Lysine. Le motif DFG peut aussi de manière indépendante osciller entre deux états. Dans les formes inactives, il adopte une conformation incompatible avec la chélation de l'ion Mg^{2+} et déstabilisante pour la liaison entre la Lysine et le Glutamate. Il est aussi à noter que l'ensemble des variations structurales entre les conformations actives et inactives est aussi associé à des changements de conformation de boucle A de fixation du substrat. Enfin, certaines études suggèrent aussi que la boucle P, trouvée dans plusieurs états dont certains incompatibles avec la stabilisation de l'ATP, pourrait jouer un rôle régulateur, bien que cela reste encore très débattu.

2.1.4. Quelques modes d'action des inhibiteurs

Dans la recherche de molécules actives contre les protéines kinases, trois types d'inhibiteurs sont à distinguer :

- Les inhibiteurs de types I sont des mimes du groupement adénine et sont capables d'interagir avec l'environnement du site de fixation du nucléotide sous sa forme active de façon similaire à l'ATP, c'est-à-dire qu'ils sont pris en étau par les résidus hydrophobes de la cavité et qu'ils forment une combinaison d'interactions hydrogènes avec la région charnière. La majorité des inhibiteurs développés à ce jour sont de type I, en partie du fait que les formes actives de kinases sont souvent utilisées dans les essais biochimiques de criblage [237]. Nous pouvons citer par exemple le Sunitinib utilisé dans le traitement contre le carcinome des cellules rénales ou contre certaines tumeurs du tube digestif (PDB : 3G0F [238]). Toutefois, les profils de sélectivités de ce type d'inhibiteur sont souvent très larges et peu ciblés sur une famille restreinte de kinases. Un sous-groupe d'inhibiteurs de type I prend avantage de la poche hydrophobe supplémentaire trouvée chez les protéines kinases qui possèdent un acide aminé gatekeeper de petite taille [237]. Ces inhibiteurs présentent en général une sélectivité plus restreinte. Le Lapatinib, utilisé dans les traitements contre le cancer du sein, est un bel exemple d'inhibiteur occupant à la fois la poche adénine et la poche additionnelle du site actif (PDB : 1XKK, [239]).

- Les inhibiteurs de type II sont allostériques et compétitifs de l'ATP du fait qu'ils ciblent le site actif des protéines kinases sous leur forme inactive. Mais ils restent encore rares. L'Imatinib, un des tous premiers inhibiteurs identifiés et commercialisés pour des thérapies contre les cancers gastro-intestinaux et la leucémie myéloïde chronique, fait parti de cette famille (PDB : 1IEP, [240]).

- Enfin, les inhibiteurs de type III sont apparus très récemment grâce aux différentes campagnes de criblage de banques de petites molécules. Elles ont permis d'identifier des molécules capables d'inhiber l'activité phospho-transférase sans empêcher la fixation de l'ATP. Ces nouveaux inhibiteurs sont capables de se fixer dans différentes poches à la surface des kinases comme le site de fixation du substrat ou comme des sites provoquant l'allostérie de l'hélice C et de la boucle A. En exemple, la Rapamycine (PDB : 2FAP) s'avère être à la fois une molécule utilisée contre le carcinome des cellules rénales et une molécule au pouvoir immunosuppresseur.

2.2. La protéine kinase 2 : CK2

Initialement appelée Caséine Kinase 2, la protéine kinase 2 (ou CK2) est une des premières enzymes de cette famille de protéines à avoir été identifiée suite aux travaux de Burnett et Kennedy sur la phosphorylation de la caséine [222]. Cependant, il est aujourd'hui admis que la caséine n'est pas un substrat physiologique de CK2 et le nom définitif de protéine kinase 2 a été adopté en 1994. Pour plusieurs raisons que nous allons détailler par la suite, CK2 est une protéine kinase toute particulière.

2.2.1. Une « Master kinase »

Contrairement à beaucoup de homologues, CK2 possède une structure quaternaire avec deux domaines catalytiques, CK2 α ou CK2 α' codés par 2 gènes distincts, et un homodimère de régulation, CK2 β . Le domaine α , qui fait parti du sous-groupe des kinases ayant l'appellation « autres », possède une activité catalytique intrinsèque très élevée, comparée aux différentes kinases. L'affinité entre les domaines catalytiques et de régulation (de l'ordre de 5nM) conduit à la formation de « l'holoenzyme » (2 CK2 α /2 CK2 β) qui est elle aussi catalytiquement active. Toutefois, la localisation et les déplacements indépendants de chacune des sous-unités observées dans des cellules vivantes [241] et les données structurales qui montrent une flexibilité de l'interface entre les sous-unités α et β suggèrent que la formation du complexe serait transitoire, réversible et non-obligatoire malgré sa remarquable stabilité *in vitro* [242].

CK2 est exprimée chez tous les organismes eucaryotes et elle est sûrement la protéine kinase la plus conservée. A titre d'exemple, le domaine catalytique des protéines kinases 2 de l'Homme et des plantes présentent près de 70% d'identité de séquence et l'identité est totale entre les protéines issues des mammifères.

L'expression de CK2 est quasi-ubiquitaire, c'est-à-dire que son expression est observée dans pratiquement tous les tissus et cellules de l'organisme et CK2 est retrouvé aussi bien dans le cytosol que dans le noyau [243].

CK2 est une Sérine/Thréonine kinase capable d'utiliser aussi bien l'ATP que le GTP pour phosphoryler ses cibles protéiques. Elle est aussi le prototype d'une enzyme pléiotropique qui régule de multiples voies de signalisation intracellulaire de façon simultanée. CK2 est même qualifiée de « master kinase » tant le nombre de substrats identifiés est important. A ce jour, plus de trois cents substrats protéiques

sont connus pour être des cibles de CK2 et il est estimé que plus de 20% du phosphoprotéome, c'est-à-dire les protéines susceptibles d'être phosphorylées, le serait par CK2 [244-247]. Il serait trop long d'explicitier en détails toutes les voies de signalisation et de régulation pour lesquelles CK2 joue un rôle. Mais en résumé, elle intervient dans la transduction du signal, le contrôle de l'expression génomique, le cycle cellulaire, la promotion de la survie et de la prolifération, le contrôle de l'apoptose ou encore la réparation des lésions de l'ADN [248, 249]. Par son rôle multiple, il apparaît aujourd'hui que CK2 est au centre d'un contrôle fin entre les signaux de survie et de mort de la cellule.

Alors que la grande majorité des protéines kinases ne sont actives qu'en réponse à un stimulus spécifique, CK2 fait parti d'une sous-classe de protéines kinases qui sont constitutivement actives. A ce jour, aucun messenger unique ou autre événement comme une phosphorylation ou interaction avec un partenaire protéique n'est connu pour réguler ou activer l'activité de CK2. Bien qu'aucun mécanisme général et classique de type « on/off » n'ait été caractérisé pour CK2, plusieurs processus de régulation ont été suggérés et il semble aujourd'hui qu'ils pourraient être spécifiques de la voie de signalisation dans laquelle CK2 intervient. Ainsi, nous pouvons citer en exemple l'assemblage et l'expression des différentes sous-unités de CK2 ou la formation d'oligomères, les modifications chimiques comme sa phosphorylation, son autophosphorylation ou son ubiquitinylation, son inhibition par des petites molécules naturelles comme l'héparine, les polyamines ou des phospholipides, certains effecteurs protéiques ou encore sa localisation intracellulaire [246, 249, 250].

A ce jour, l'ensemble des études sur le domaine catalytique CK2 α et sur ses substrats permet d'avoir un des répertoires les plus étendus parmi les répertoires connus de sites phosphorylés par une protéine kinase. Ainsi CK2 α reconnaît plus de 300 sites différents qui sont des peptides majoritairement acides avec un site de fixation consensus du type S/T-X-X-D/E-pS [245, 247]. Toutefois, bien que CK2 α soit classée parmi la famille des Sérine/Thréonine kinase, il existe des exemples de phosphorylation de substrat possédant un Tyrosine [251-254]. De façon très intéressante, il apparaît que la sous-unité catalytique seule et l'holoenzyme présentent des spécificités de substrats différentes, mais il apparaît aussi que la complexation de CK2 α à CK2 β change la spécificité du domaine catalytique pour ses nombreux partenaires, certainement en modulant les sites potentiels de fixation [255]. Ainsi, les substrats de CK2 peuvent être classés en 3 catégories distinctes [256] : (i) les substrats complètement indépendants de CK2 β qui peuvent à la fois être phosphorylés par le domaine α isolé ou l'holoenzyme ; (ii) les substrats dépendants de la sous-unité de régulation β qui sont exclusivement phosphorylés par le complexe CK2 α /CK2 β , (iii) et les substrats exclusivement indépendant de CK2 β qui sont phosphorylés par CK2 α mais pas par l'holoenzyme.

2.2.2. CK2 : une cible thérapeutique

L'implication de CK2 dans de très nombreuses voies de signalisation intracellulaire contrôlant la survie et l'apoptose de la cellule laisse entrevoir que sa dérégulation pourrait être à l'origine de différentes pathologies. L'implication de CK2 a été identifiée dans de nombreuses maladies inflammatoires, respiratoires (mucoviscidose), parasitaires (paludisme), virales (HPV, EBV, Influenza, HCV) ou de dégénérescence (maculaire) (pour revue [257]). Des niveaux aberrants d'expression et d'activité kinase de CK2 sont observés dans différents types de cellules cancéreuses comme celles de la prostate, du sein, des poumons, du rein, de lymphome, faisant de CK2 un marqueur de mauvais pronostic cancéreux et suggérant son rôle dans le tumorigenèse. La surexpression de CK2 dans des modèles cellulaires provoque en effet l'apparition de caractères cancéreux. Différentes études suggèrent aussi le potentiel thérapeutique que représente l'inhibition de CK2 dans les cellules cancéreuses. Par exemple, la régulation négative de l'activité catalytique de CK2 a montré son efficacité dans la réduction de la prolifération cellulaire mais aussi dans l'induction de l'apoptose des cellules cancéreuses. Des stratégies à base de petites molécules, mais aussi d'expression de mutants négativement dominants ou d'ARN anti-sens ou intercalant ont prouvé leur efficacité dans des modèles de xénogreffes cancéreuses chez la souris [249, 250, 257, 258]. Les études extensives sur la recherche de petites molécules inhibitrices de l'activité kinase de CK2 conduit aujourd'hui à la mise en essai thérapeutique du premier inhibiteur ciblant CK2 et administrable oralement pour lutter contre la progression de tumeurs solides ou du myélome [259].

2.2.3. La structure de CK2

L'activité catalytique de CK2 est portée par la sous-unité α dont la première structure a été résolue en 1998 [260, 261]. Aujourd'hui plus de quatre-vingts structures de complexes entre CK2 α issue du maïs ou de l'Homme et des ligands sont déposées dans la PDB. Globalement, l'architecture est celle d'un domaine ePK classique et tous les éléments structuraux et fonctionnels qui sont caractéristiques des protéines kinases sont retrouvés. Toutefois quelques singularités sont à souligner.

En premier lieu, la région charnière reliant les deux lobes présente une séquence et différentes structures qui permettent à CK2 α d'accommoder aussi bien l'ATP que le GTP comme source de phosphate [262]. Une autre des caractéristiques de CK2 α est le volume du site de fixation du nucléotide. Possédant des acides aminés plus encombrants que la plupart des autres kinases et une phénylalanine comme gatekeeper, la poche ATP est plus petite que celle de la majorité des protéines kinases [236, 263]. Enfin, l'activité constitutive du domaine α s'explique par deux éléments structuraux. Le premier est la modification du motif DFG par la substitution de la Phénylalanine en Tryptophane (motif DWG). Cette modification a pour conséquence d'apporter de nouvelles interactions entre ce motif et le reste du domaine. Ces

dernières pourraient à leur tour contribuer à maintenir ce segment dans une forme active permanente [261]. La seconde contribution à la forme active de CK2 est la contrainte sur l'hélice C qui est toujours trouvée dans une conformation active. Ce segment est en effet stabilisé par des interactions avec un segment supplémentaire présent en partie N-terminale de la séquence de CK2 et structuralement en contact avec l'hélice activatrice [262, 264].

Bien qu'aucune structure de CK2 complexée à un substrat protéique ne soit connue, certainement à cause des compétitions entre le peptide chargé et les espèces ioniques utilisées pour la cristallisation [265], les structures du domaine catalytique de CK2 ont révélé la nature du site potentiel de fixation du peptide substrat. Il est essentiellement composé d'acides aminés chargés positivement et préfigure des interactions ioniques que pourrait faire CK2 α avec ses substrats acides.

Enfin, les structures de la sous-unité catalytique seule et dans le contexte de l'holoenzyme ont permis de localiser le site d'interaction α/β à la surface du lobe N-terminal. Composé essentiellement d'acides aminés hydrophobes, ce site peut adopter différentes conformations et peut ainsi accommoder, non seulement la boucle d'interaction de CK2 β , mais aussi de petites molécules organiques [266, 267].

La détermination de la structure de l'holoenzyme en 2001 a été une avancée majeure dans la connaissance de la régulation de CK2. Elle se présente sous la forme d'un « papillon » composé d'un dimère de la sous-unité régulatrice β au centre et de deux sous-unités catalytiques α en périphérie sans contact entre elles (voir Figure 25) [242]. La structure cristallographique de la sous-unité β résolue précédemment par une équipe de l'IBS en 1999 avait révélé la forme du dimère qui s'opère par une interaction homotypique et un motif en doigt de zinc [268]. La

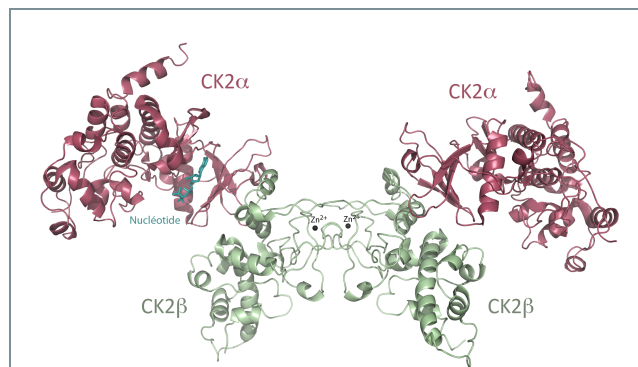


Figure 25 : Structure de l'« holoenzyme » CK2
Représentation en mode ruban de la structure de la protéine kinase CK2 sous sa forme hétérotétramérique (2 CK2 α / 2 CK2 β). Les unités catalytiques CK2 α indépendantes sont en rose foncé, le dimère régulateur 2CK2 β est en vert pale. Un analogue de l'ATP (en bleu) est observé dans la structure d'un des deux domaines alpha. Chaque domaine régulateur bêta possède un ion zinc. Code PDB : 1JHW

comparaison des structures de chaque domaine seul et en complexe ne montre aucun changement structural majeur autre que ceux observés à l'interface α/β et laisse différentes questions non résolues, notamment celles sur les mécanismes de modulation des substrats par CK2 β ou d'autophosphorylation. De façon inattendue, l'interface entre les deux sous-unités α et β est réduite (moins de 900Å²) et est à mettre en comparaison avec la forte constante d'affinité qui est de l'ordre de 5nM.

Enfin, CK2 peut former *in vitro* des assemblages supramoléculaires dont certains présentent une activité accrue comparée à celle du tétramère ou de l'unité

catalytique seule, alors que d'autres assemblages perdent l'activité de phosphorylation [269]. Toutefois, leur existence *in vivo* et leur rôle physiologique restent encore à être démontrés mais ces assemblages pourraient jouer un rôle dans les mécanismes de régulation de l'activité de CK2.

2.2.4. Les bases structurales des inhibiteurs ATP-compétitifs

Le potentiel oncogénique de CK2 a conduit à de nombreuses études dans le but d'identifier des inhibiteurs spécifiques aussi bien par les laboratoires académiques que par l'industrie pharmaceutique. La littérature sur la caractérisation aussi bien biochimique, cellulaire et structurale d'inhibiteurs de CK2 est très foisonnante et de nombreux inhibiteurs sont décrits. Les études les plus complètes qui combinent études structurales et relations structure-activité d'inhibiteurs regroupent les familles moléculaires des Anthraquinones et Xanthenones [270, 271], des Indoloquinoline [272, 273], des dérivées Tetralogénobenzos [274-281], des Coumarines [282, 283] ou encore des Pyrazolo-triazines [284, 285]. Il est d'une part intéressant de noter que ces inhibiteurs sont majoritairement des inhibiteurs ATP-compétitifs, c'est à dire qu'ils vont se fixer dans le site ATP, et d'autre part, il est frappant de constater leur grande diversité chimique.

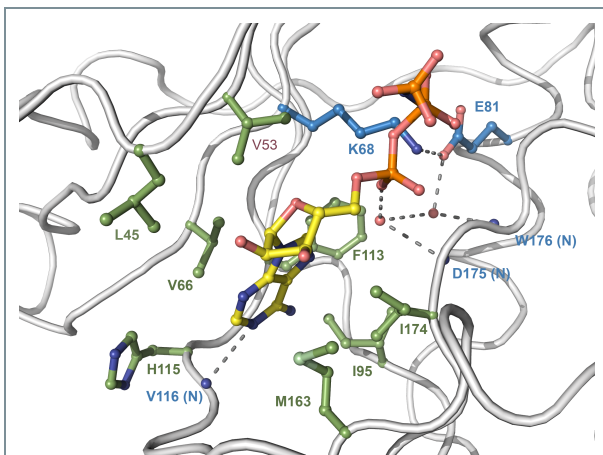


Figure 26 : Structure d'un analogue de l'ATP dans le site actif de la protéine kinase 2

Représentation en mode billes et bâtons d'une molécule analogue à l'ATP (AMPPNP) dans le site actif de CK2 α . Le groupement adénine et le sucre ribose sont en jaune, les groupements phosphate en orange. Les inhibiteurs de CK2 ciblent plusieurs zones distinctes du site : (i) les poches adénine et ribose formées par des acides aminés hydrophobes (vert clair) (ii) la lysine 68 et l'acide glutamique 81 (en bleu) (iii) une poche électrostatiquement positive occupée par 2 molécules d'eau conservées (sphères rose) (iv) la région charnière (113-117) qui forme des liaisons hydrogène avec le groupement adénine de l'ATP. Code PDB : 1PJK

Vu le nombre important de structures de complexes CK2 α /inhibiteur, il serait mal venu de présenter dans le détail les caractéristiques de chacun et il est possible de se référer à de nombreuses et excellentes revues [250, 286-290]. Cependant et bien que très variables en terme de classes chimiques, il est essentiel de mentionner que ces inhibiteurs possèdent des caractéristiques communes dictées par les impératifs inhérents au site ATP de CK2. En effet, les différentes structures cristallographiques de complexes CK2 α /inhibiteurs ont permis de révéler plusieurs éléments clés nécessaires à un inhibiteur pour être efficacement en compétition avec l'ATP. Tout d'abord, pour nombre de molécules décrites dans la littérature, la contribution énergétique la plus importante pour leur fixation au site

ATP est certainement celle due aux interactions apolaires entre les acides aminés composant la surface de la poche adénine et l'inhibiteur (voir Figure 26). Pour

certaines d'entre eux comme les dérivés Tetrabromobenzoides [275], de Coumarine [282] ou de Pyrazolo-triazine [284, 285], les analyses des relations entre structure et activité ont confirmé le rôle important joué par ces interactions. De plus, cette zone de la poche ATP semble être déterminante pour la sélectivité des inhibiteurs spécifiques de CK2. Comme mentionné précédemment, la poche ATP de CK2 présente un volume plus petit, notamment par la présence de deux acides aminés encombrants (Valine 66 et Isoleucine 174) qui sont en général plus petits dans les autres kinases (souvent une Alanine, Thréonine ou Leucine). Par exemple, le site ATP de la protéine kinase CDK2 est plus large que celle de CK2 à cause de la présence de deux Alanines au lieu du couple Valine/Isoleucine. Une conséquence directe est un mode de fixation de l'inhibiteur TBB différent entre les deux kinases et une sélectivité remarquable du TBB pour CK2 [270]. De même, la mutation de ces deux résidus dans CK2 réduit la capacité inhibitrice des dérivés du TBB, confirmant leur importance [290]. Un autre exemple est celui de l'inhibiteur Emodin qui est capable d'inhiber l'activité de CK2 mais qui possède une activité inhibitrice réduite pour la kinase PKA. Ceci est à mettre aussi en rapport avec une cavité ATP plus large pour PKA [271].

Un autre point essentiel pour l'interaction efficace entre CK2 et une molécule inhibitrice de son activité est la présence d'une zone à la charge électrostatique positive enfouie au fond de la cavité ATP. Cette zone est proche du pont salin conservé entre la Lysine 68 et l'acide Glutamique 81 (voir Figure 26). En effet, la comparaison de l'ensemble des structures CK2/inhibiteur montre que les groupements chargés négativement se retrouvent essentiellement dans cette zone spécifique du site ATP. Ceci est vrai pour plusieurs molécules comme IQA, MNA, MNX, TBB, DMAT ou DRB mais aussi pour l'ATP et le GTP qui y présentent leur groupement phosphate α . Un exemple de l'influence de l'interaction électrostatique et de la charge de la molécule inhibitrice vient de la différence entre le TBB et le TBA. Le TBB avec un pKa proche de 5 se lie à CK2 dans cette zone positivement chargée alors que son dérivé TBI avec un pKa de 9 se trouve éloigné vers la région charnière [291].

Les derniers éléments importants sont les interactions de type liaison hydrogène que peut former l'inhibiteur avec les atomes du site ATP. Il apparaît que les ligands portant une fonction acide ont tendance à venir interagir avec le fond du site soit directement avec la Lysine catalytique en position 68 en remplaçant les molécules d'eau trouvées dans cette zone, soit indirectement par la conservation d'une ou plusieurs de ces molécules d'eau. Les autres molécules préfèrent se lier à la région charnière et en particulier avec la chaîne principale des acides aminés en position 114 et 116 (voir Figure 26).

Les études structurales ont ainsi permis de montrer les caractéristiques essentielles d'une molécule pour inhiber efficacement l'activité de CK2 de façon ATP-compétitive : une hydrophobicité adéquate, une complémentarité de forme pour s'accommoder de l'étroitesse du site et la capacité d'établir des interactions polaires avec au moins un des deux points d'ancrage que sont la région charnière et le couple conservé Lysine/Glutamate. Les dérivés Pyrazolo-triazines en sont la meilleure

illustration : ils présentent toutes les caractéristiques pour interagir avec tous ces éléments structuraux et sont les molécules les plus efficaces [284, 285].

2.2.5. Les inhibiteurs non-ATP-compétitifs

Les informations biochimiques et structurales concernant des inhibiteurs non-ATP-compétitifs, c'est-à-dire des molécules capables d'inhiber l'activité kinase sans modifier la fixation du nucléotide, sont beaucoup plus restreintes. Plusieurs sites de fixation d'inhibiteurs situés à la surface de CK2 α sont potentiellement visés.

Le site de fixation de l'unité régulatrice CK2 β est une cible intéressante. La stratégie d'inhiber la formation de l'holoenzyme a le mérite de vouloir cibler spécifiquement des voies de signalisation reposant sur la régulation par l'unité β . De petites molécules organiques comme le glycérol ont déjà été observées dans les structures du domaine catalytique. Mais à l'heure actuelle, un seul inhibiteur chimique, le DRB, est observé dans ce site secondaire [250]. Plus récemment, un peptide cyclique, qui a été élaboré par nos collaborateurs et qui contient une séquence similaire à la boucle de CK2 β venant en contact avec CK2 α , a montré sa capacité à inhiber l'assemblage de l'enzyme et à affecter sa préférence de substrat [292]. Toutefois, ni le mode exact d'interaction avec CK2 α , ni sa capacité d'inhibition ne sont connus. Dans la même perspective, un peptide intéressant est celui dérivé d'une protéine membranaire responsable du flux d'ion chlorure, CFTR, et contenant la séquence consensus de phosphorylation par CK2 α [293]. Il a été montré que ce peptide inhibe l'activité kinase par un mécanisme non-ATP-compétitif et pourrait donc se fixer à la poche du substrat. Il est clair que la recherche et la caractérisation de molécules de cette classe d'inhibiteurs est aujourd'hui un champ d'investigation très ouvert.

3. Résultats marquants sur CK2

Comme nous venons de le voir, de nombreuses molécules de différentes natures capables d'inhiber *in vitro* l'activité kinase de CK2 de manière efficace ont déjà été découvertes, offrant ainsi des outils de choix pour les études biochimiques, structurales ou cellulaires. Toutefois, très peu de ces molécules se sont avérées être de bons candidats pour le développement de molécules à visée thérapeutique ; essentiellement parce que beaucoup d'entre elles n'ont pas démontré leur efficacité dans des modèles cellulaires ou chez l'animal. Dans ce contexte, l'équipe de Claude Cochet (iRTSV/CEA, Grenoble) a développé une stratégie basée sur le criblage de nombreuses bibliothèques de composés chimiques, ou chimiothèques, avec l'objectif ultime d'identifier et de caractériser de nouvelles molécules inhibitrices de l'activité kinase de CK2 qui seraient efficaces à la fois *in vitro* mais aussi sur des modèles animaux.

La stratégie employée s'appuie sur la combinaison de techniques de criblage automatisé à haut débit et de caractérisation biochimique, cellulaire et structurale :

- Grâce à la plateforme de Criblage pour les Molécules BioActives (Laurence Lafanechère et Caroline Barette, CMBA/CEA Grenoble), la première étape consiste au criblage de larges chimiothèques (NCI structural diversity set, NCI Mechanistic set, Prestwick chemical library, Chimiothèque nationale, BioMol kinase inhibitors library). Ce criblage est basé sur un test d'activité kinase luminescent non-radioactif qui a été miniaturisé par l'équipe de Claude Cochet.

- A l'issue du criblage et à l'identification de « touches », celles-ci sont caractérisées biochimiquement par un test radioactif permettant de mesurer l'activité inhibitrice des molécules identifiées, d'identifier leur mode d'inhibition (ATP-compétitif, substrat compétitif...) et de les valider en tant que réels inhibiteurs *in vitro*. Les composés ainsi trouvés sont aussi testés sur un ensemble de protéines kinases afin d'évaluer leur sélectivité.

- En parallèle, des séries de dérivés de la « touche » primaire sont synthétisées par différentes équipes de chimistes avec lesquelles nous collaborons (Equipes de Jean-Claude Florent et de Marie-Paule Teulade-Fichou de l'Institut Curie, Paris) et caractérisées biochimiquement. Les études de relation entre la structure des dérivés et leur activité inhibitrice permettent d'identifier les meilleurs dérivés mais aussi d'optimiser leur composition en identifiant les groupements chimiques favorables à l'inhibition.

- Une fois validée, l'efficacité des inhibiteurs est testée *in cellulo* grâce à un test rapporteur de l'activité kinase de CK2 dans un modèle cellulaire simple. Développé par l'équipe du CEA, ce test est basé sur un peptide phosphorylable par CK2 et fusionné à une protéine fluorescente. Le plasmide contenant la protéine fusion est transfecté dans des cellules vivantes et le taux de phosphorylation est mesuré après surexpression du peptide rapporteur et incubation de la culture cellulaire avec l'inhibiteur. Ce test permet non seulement de mesurer l'inhibition intracellulaire mais

aussi dévaluer la perméabilité ou la séquestration par la membrane, la métabolisation ou encore la toxicité de l'inhibiteur sur la viabilité cellulaire [294].

- Enfin, les inhibiteurs les plus prometteurs sont validés dans un modèle animal. Ces études qui ont été menées chez la souris ont pour but d'estimer la toxicité et l'efficacité des molécules *in vivo*. En résumé, elles consistent à inoculer des cellules cancéreuses humaines (cellules de gliomes) à des souris saines. La régression de la tumeur et la survie des populations de souris sont ensuite mesurées suite à l'injection de l'inhibiteur.

L'équipe de Claude Cochet a pu ainsi identifier plusieurs nouvelles familles de molécules possédant une action inhibitrice à des degrés divers et spécifiques de CK2, dont certaines ont montré *in vivo* une efficacité dans la régression de cellules tumorales. Les inhibiteurs identifiés sont aussi bien ATP-compétitifs comme les Salicylaldéhydes [295], les acides Difuranedicarboxyliques [296] et les Pyridocarbazoles [297] ou ATP non-compétitifs comme les Polyoxometallates [298] et les Azonaphthalènes [299].

Comme nous avons pu le comprendre précédemment, l'information obtenue par la résolution des structures cristallographiques de complexes entre les protéines kinases et leurs inhibiteurs est essentielle pour identifier les interactions atomiques clés entre l'enzyme et son ligand. Elle permet par la suite l'optimisation rationnelle et la conception de nouvelles molécules plus efficaces. En parallèle du travail de l'équipe de Claude Cochet et de nos collaborateurs chimistes, nous tentons d'obtenir les structures de complexes entre le domaine catalytique de CK2 et les différentes molécules identifiées. Bien que des nombreux cristaux aient été obtenus, la première difficulté réside ici dans l'obtention de complexe avec une population d'inhibiteurs ATP-compétitifs suffisante dans chaque cristal pour pouvoir les identifier sans ambiguïté dans les cartes de densités électroniques. L'autre difficulté concerne les inhibiteurs non-ATP-compétitifs dont le site de fixation parfois inconnu peut être aussi un site d'interaction entre les molécules du cristal qui participe à la cohésion cristalline. Dans ce cas la forme cristalline est incompatible avec la formation du complexe et de nouvelles formes cristallines doivent être trouvées. Avec l'aide de Renaud Prudent pour qui j'ai encadré le travail cristallographique pendant sa thèse, nous avons pu résoudre la structure de composés appartenant à deux familles d'inhibiteurs ATP-compétitifs en complexe avec CK2 α .

3.1. Les acides Difuranedicarboxyliques

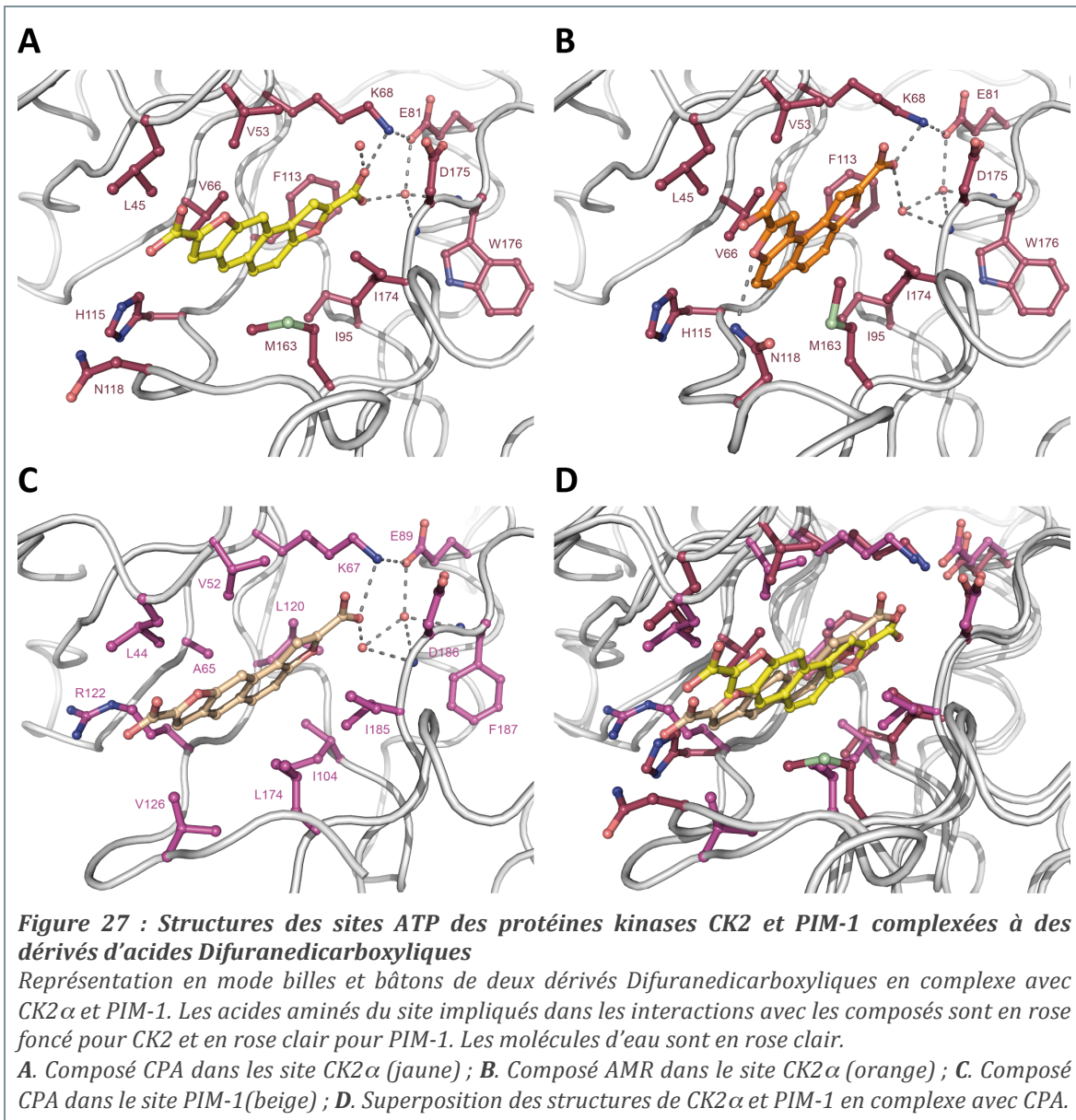
Les acides Difuranedicarboxyliques ont été initialement identifiés après un large criblage de la librairie de petites molécules de l'Institut Curie. Ces composés présentent la particularité structurale de posséder un cœur plan, rigide et hydrophobe avec deux groupements polaires en périphérie (voir Figure 27). Les caractérisations biochimiques et l'expansion chimique du répertoire de dérivés, toutes deux associées aux analyses de relation structure/activité ont permis d'identifier trois composés qui font parti des inhibiteurs les plus puissants de CK2 α .

L'autre intérêt des acides Difuranedicarboxyliques réside dans leur sélectivité atypique et très restreinte. En effet, ces inhibiteurs sont non seulement spécifiques de CK2 α mais aussi de la famille des protéines kinases PIM. De façon très intéressante, ces dernières possèdent des fonctions similaires à CK2 : elles sont aussi constitutivement actives, elles ont des profils d'expression similaires à CK2 et elles interviennent dans des voies de signalisation cellulaire comparables. De part leur rôle anti-apoptotique et leur implication dans des pathologies cancéreuses similaires, les kinases PIM sont aussi des cibles thérapeutiques [300]. Ces composés Difuranedicarboxyliques pourraient ainsi présenter l'avantage de cibler deux voies différentes de développement des cellules cancéreuses et déboucher sur de nouvelles perspectives thérapeutiques, notamment dans le cancer de la prostate où la dérégulation de ces deux kinases est corrélée avec l'expansion de la maladie.

Dans le but de rationaliser cette sélectivité restreinte et d'obtenir une image précise des interactions mises en jeu pour l'inhibition de l'activité des deux kinases, nous avons résolu différentes structures de complexes entre les acides Difuranedicarboxyliques et CK2 α . La structure d'un complexe avec PIM-1 a aussi été déterminée en collaboration avec le groupe de Stephan Knapp du Genomic Structural Consortium (Oxford).

Une première modélisation des complexes avait été entreprise par une équipe collaboratrice de l'Institut Curie (équipe de Jean-Claude Florent). Elle avait permis d'expliquer en partie les relations structure/activité de la librairie de dérivés. Toutefois, les structures cristallographiques ont révélé des interactions qui n'avaient pas été prédites. Les acides Difuranedicarboxyliques occupent le site ATP de CK2 α et PIM-1 à une position similaire à celle de tous les inhibiteurs ATP-compétitifs de ces protéines kinases, c'est à dire avec le cœur hydrophobe pris en étau par les résidus formant la poche adénine (voir Figure 27). Toutefois, la modélisation n'avait pas prédit la configuration des interactions de type liaison hydrogène que forment les groupements acides de ces composés. D'une part, ils contactent le couple Lysine 68/Glutamate 81 grâce à une liaison hydrogène directe formée avec la Lysine, mais aussi par l'intermédiaire de plusieurs molécules d'eau qui sont très souvent conservées dans les structures de CK2 α et PIM-1 (voir section 2.2.4 et Figure 27). D'autre part, la configuration des cycles du cœur peut aussi être favorable à une interaction électrostatique avec la région charnière qui adopte alors une conformation qui n'avait pas été prédite (voir Figure 27).

Le rationnel de la sélectivité restreinte de ces composés pour CK2 α et les kinases PIM réside dans la conservation en séquence des acides aminés formant le site ATP. Les poches adénines des deux kinases sont en effet structurellement très proches. Comme nous l'avons vu précédemment (voir section 2.2.4), la sélectivité peut reposer sur les caractéristiques de taille et de forme. La plus grande différence de CK2 α et de PIM-1 comparée aux autres protéines kinases est l'étroitesse de leur site ATP. La structure plane des Difuranedicarboxylates semble ici très adéquate pour se fixer avec une forte affinité à ce type de site ATP. Les kinases possédant des sites



plus spacieux présentent pour ces inhibiteurs des affinités plus faibles qui sont certainement dues à des interactions plus faibles avec le cœur de ces composés.

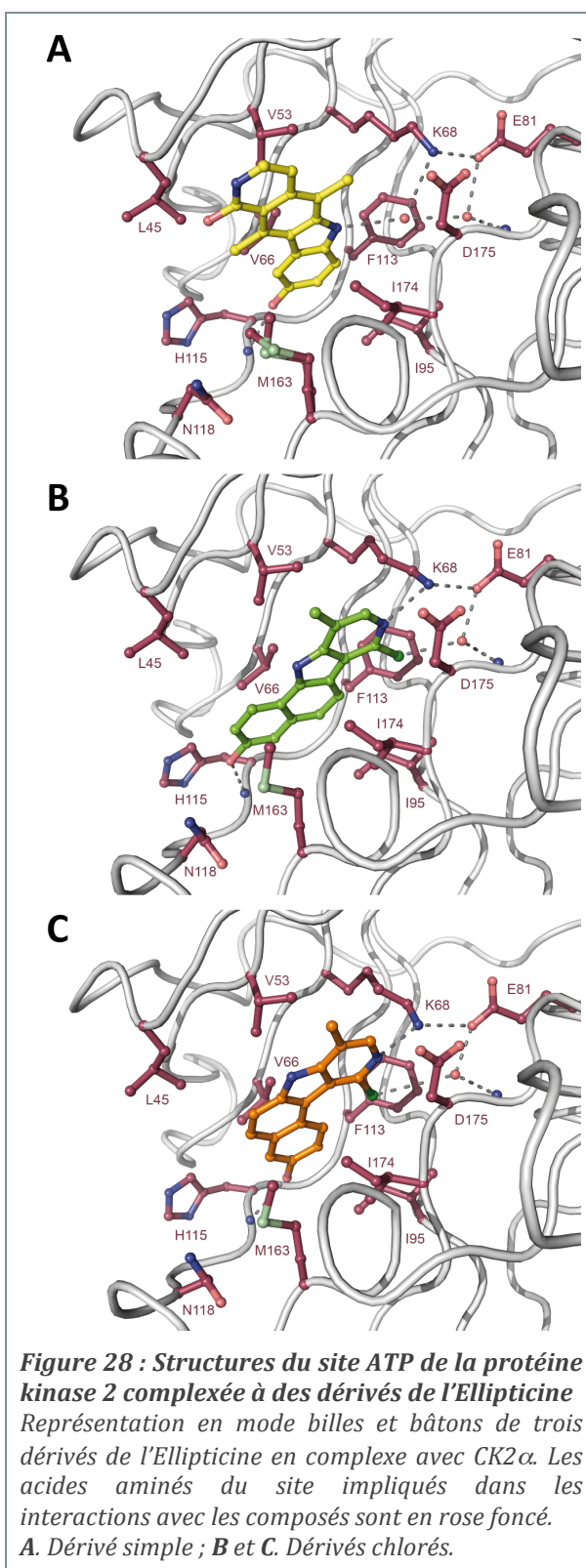
Ainsi, nous avons pu montrer les caractéristiques structurales qui font des Difuranes des inhibiteurs duaux potentiellement attractifs pour le développement de nouveaux composés cliniques. Toutefois, l'équipe de Claude Cochet a pu montrer l'inefficacité de ces composés sur des cellules mis en culture bien que l'activité inhibitrice *in cellulo* a pu être observée. Ceci suggère un mécanisme de neutralisation des groupements acides qui reste encore inconnu. Tout récemment, une nouvelle famille d'inhibiteurs duaux de CK2 et des kinases PIM a aussi été aussi identifiée aux Etats-Unis. Ces inhibiteurs montrent une activité antiprolifératives sur des cultures cellulaires [301]. Ce travail a fait l'objet d'un article publié dans FASEB Journal en 2010.

3.2. Les dérivés d'Ellipticine

L'Ellipticine ou Pyridocarbazole et ses dérivés sont connus pour leur efficacité thérapeutique dans la lutte contre le cancer essentiellement en intervenant comme intercalant de l'ADN [302]. Quelques études montrent aussi que ces composés sont inhibiteurs de cibles protéiques comme la Topoisomérase II [303] ou les protéines kinases c-Kit [304], Cdk2 [305] et Aurora [306]. L'équipe de Claude Cochet a montré que les Ellipticines ont aussi un potentiel inhibiteur de type ATP-compétitif pour la kinase CK2. Un point particulièrement attrayant de ces composés souligné par les études des relations structure/activité est la possibilité de moduler la sélectivité des composés en modifiant des groupements chimiques précis. Toutefois, le contrôle précis des effets inhibiteurs sur une kinase ou un groupe de kinases précis reste difficile. Mais la compréhension mécanistique et structurale pourrait à terme permettre de concevoir des composés ciblant spécifiquement les kinases cibles.

Une première étape a été franchie par notre équipe grâce à la résolution de la structure CK2 α en complexe avec trois dérivés différents de l'Ellipticine. Comme beaucoup de composés précédemment identifiés, ces inhibiteurs miment la base adénine de l'ATP en occupant une position similaire à cette dernière dans le site et en formant une liaison hydrogène avec la région charnière (voir Figure 28). Toutefois, les similitudes s'arrêtent ici,

notamment parce que les différents dérivés d'Ellipticine présentent des orientations et des géométries d'interaction avec le site ATP différentes et dépendantes de la



géométrie du composé. Pour la première fois, ces structures ont montré un nouveau mode d'interaction faisant intervenir un réseau de liaisons hydrogène entre le groupement indole d'un inhibiteur et le fond de la cavité ATP composé d'une molécule d'eau conservée et du couple Lysine/Glutamate (voir Figure 28).

Malgré des spécificités jugées moyennes, c'est-à-dire peu restreintes, les tests de régression tumorale sur les modèles murins de xénogreffe ont révélé un effet positif anti-tumoral et non-toxique de certains dérivés d'Ellipticine. L'ensemble de ces données nous prouve le potentiel thérapeutique de ces composés inhabituels dans le développement de nouvelles molécules actives. L'article décrivant ces résultats a été publié dans *Cancer Research* en 2010.

4. Projet à venir sur les inhibiteurs des protéines kinases

Les résultats structuraux obtenus ces dernières années sur les complexes entre CK2 α et ses inhibiteurs ATP-compétitifs ont démontré sans contestation la nécessité de caractériser dans le détail les interactions entre l'unité catalytique de cette protéine kinase et les molécules potentiellement candidates pour le développement d'agents anticancéreux.

Dans un premier temps, les informations déjà obtenues grâce aux structures cristallines permettent aujourd'hui à nos collaborateurs chimistes d'élaborer de nouvelles séries d'inhibiteurs. En adaptant les groupements chimiques, ils essayent de synthétiser des nouvelles molécules soit pour optimiser les interactions avec le site actif, soit pour rendre les composés biologiquement plus actifs, c'est-à-dire perméables à la membrane cellulaire et/ou résistants à la dégradation *in cellulo*, sans modifier les interactions favorables à l'inhibition. Dans le but de valider les dernières molécules synthétisées, nous continuerons nos efforts dans la caractérisation de ces inhibiteurs de CK2 aussi bien aux niveaux biochimique, cellulaire, animal ou structural.

Ensuite, la découverte de molécules inhibitrices de CK2 atypiques qui s'avèrent être non-ATP-compétitifs nous pousse à caractériser leur mode d'interaction avec l'unité catalytique. Une série d'inhibiteurs à base métallique comme les Polyoxymetallates a notamment été caractérisée par l'équipe de Claude Cochet. Le mode d'interaction de ces complexes métalliques reste à être déterminé. Des méthodes de criblage virtuel ont aussi permis d'identifier de nouveaux inhibiteurs allostériques qui semblent cibler l'interface entre les sous-unités α et β de l'holoenzyme. La difficulté pour la caractérisation structurale de ces nouvelles molécules est ici de trouver une forme cristalline compatible avec la formation d'un complexe avec CK2 α . Par exemple, les cristaux et l'empilement cristallin qui sont obtenus au laboratoire présentent CK2 α avec son site β occupé par des atomes d'une molécule symétrique. Si nous voulons obtenir une image d'un inhibiteur en complexe, il faudra obtenir d'autres formes cristallines similaires à celles déjà été mises en évidence par d'autres équipes.

Enfin, comme nous l'avons vu précédemment, la sélectivité des inhibiteurs identifiés reste un enjeu important pour inhiber de façon spécifique la protéine kinase cible. Pour être capables dans le futur de moduler cette sélectivité, il est nécessaire de comprendre dans le détail les différences et les points communs des différentes kinases qui pourraient être inhibées par une même série de molécules. Il est donc nécessaire de résoudre les structures de différentes kinases sensibles aux inhibiteurs testés comme nous l'avons déjà fait avec CK2 et Pim1.

5. Contexte scientifique : les GTPases bactériennes

5.1. Les GTPases

5.1.1. Quelques généralités

Historiquement, les premières investigations biochimiques sur les GTPases viennent de recherches sur des enzymes utilisant le GTP au lieu de l'ATP pour réguler différents processus cellulaires puis de l'identification de l'oncogène *ras* chez l'Homme [307-309]. Pendant ces quarante dernières années, l'identification et l'analyse fonctionnelle d'un grand nombre de protéines capables de fixer et d'hydrolyser le GTP ont permis de comprendre et de découvrir la grande diversité de cette superfamille d'enzymes. Il apparaît qu'elles jouent un rôle prépondérant dans la vie de la cellule. En effet, retrouvées dans tous les règnes du monde du vivant, de la bactérie à l'Homme, elles interviennent dans des mécanismes cellulaires très diversifiés comme la transduction du signal, la biosynthèse des protéines, le contrôle et la différenciation cellulaire ou encore le transport intracellulaire. Ces enzymes sont considérées comme de véritables interrupteurs moléculaires car elles oscillent entre différents états suivant la forme du nucléotide fixé : un état actif (« on ») sous la forme liée au GTP et un état inactif (« off ») sous la forme liée au GDP. La forme GTP des GTPases interagit en général avec des effecteurs moléculaires pour activer les processus cellulaires. L'hydrolyse du nucléotide triphosphate en nucléotide diphosphate entraîne ensuite l'inactivation de l'enzyme et le relargage de ce dernier permet à la protéine d'entrer dans un nouveau cycle (pour revue : [310-313])

Depuis la découverte de la première GTPase, de nombreuses familles de ces enzymes ont été identifiées. A partir de la similarité de séquence et des similitudes structurales, les études récentes ont montré que les différentes familles de GTPases peuvent être regroupées en différentes classes. La famille SIMIBI (pour Signal recognition particule MInd and Blod) contient par exemple des GTPases intervenant dans les signaux de reconnaissance comme les SRP (pour Signal Recognition Particule) mais aussi des ATPases assimilées aux familles de MinD, de la dithiobiotin synthétase ou de la formyltetrahydrofolate synthétase. La famille TRAFAC (pour TRAnslation FACtors) inclut des protéines impliquées dans la translation, la transduction du signal, la motilité cellulaire ou le transport intracellulaire. Cette famille d'enzymes regroupe les GTPases les plus connues comme des facteurs de translation (*i.e.* EF-Tu), les protéines G hétérotrimériques, les petites protéines G monomériques (*i.e.* famille Ras), les protéines des familles des Dynamines, des Myosines et des Kinésines ou encore la majorité des protéines G procaryotes qui nous intéressent ici [314]. Enfin, une dernière classe de GTPases est à signaler et regroupe la Tubuline et ses homologues qui sont des constituants majeurs du cytosquelette en formant les microtubules.

Les GTPases eucaryotes sont connues pour intervenir dans de très nombreux et divers processus cellulaires comme la synthèse des protéines, la signalisation transmembranaire, la translocation des molécules, le trafic cellulaire, l'organisation

du cytosquelette, la différenciation ou la prolifération cellulaire. Les GTPases bactériennes semblent avoir un rôle plus restreint à la synthèse protéique et représentent une des plus grandes familles de facteurs d'assemblage du ribosome bactérien [315]. La révolution du séquençage génomique dans les années 90 a permis d'identifier un certain nombre de gènes chez la bactérie codant pour des protéines comparables aux protéines G eucaryotes. Précédant la découverte de leur implication dans la biosynthèse du ribosome, plusieurs études ont rapidement démontré le rôle essentiel de ces protéines dans la survie et la croissance bactérienne [316]. Toutefois et bien qu'il existe beaucoup d'études de caractérisation biochimique, fonctionnelle et structurale, peu d'informations sont connues à ce jour sur la fonction exacte de ces GTPases chez la bactérie et sur le détail de leurs mécanismes de régulation dans l'assemblage du ribosome (pour revue : [317]). Leurs rôles probables sont le recrutement et le déplacement des protéines constituant les sous-unités ribosomales, le recrutement et la régulation de facteurs d'assemblage ou encore l'induction de réarrangements conformationnels du ribosome naissant. En plus de leur rôle dans l'assemblage ribosomique, certaines de ces GTPases seraient aussi impliquées dans d'autres processus cellulaires comme la réplication de l'ADN, la division cellulaire, la réponse au stress ou la pathogénicité. Enfin, la plupart de ces GTPases procaryotes possèdent des homologues chez les eucaryotes, prouvant leur conservation tout au long de l'évolution. Une compréhension de leur mécanisme est donc cruciale pour comprendre les mécanismes généraux de la biogenèse du ribosome [315, 316].

5.1.2. La structure du domaine G

Malgré la grande diversité des rôles des GTPases dans les processus cellulaires, les protéines de cette superfamille partagent un domaine structural commun : le domaine G qui porte la fonction d'hydrolyse du nucléotide. Cette similarité structurale s'appuie sur une identité de séquence significative qui suggère une origine commune [311, 314]. L'objectif d'élucider les mécanismes d'hydrolyse et de régulation des différentes voies de signalisation par les familles de GTPases a conduit à de très nombreuses études structurales et une littérature très abondante. Parmi les GTPases les plus étudiées, la petite GTPase Ras de l'Homme représente l'archétype du domaine G [312].

De façon remarquable, le site de fixation du nucléotide à base de guanine est composé de cinq boucles qui sont les éléments les plus conservés entre les GTPases. Ils permettent ainsi de définir la superfamille. Ces cinq boucles sont désignées motifs G1 à G5 (voir Figure 29 et Figure 30) [311, 312] :

- Le motif G1 correspond à la boucle P (« P-loop ») ou motif de Walker A. Avec une séquence consensus GXXXXGKT/S, ce motif contient notamment une Lysine nécessaire à la stabilisation des groupements phosphate du nucléotide.

- Le motif G2, qui peut être très conservé au sein d'une même famille de GTPases mais qui varie en séquence d'une famille à l'autre, peut être difficilement identifié. Ce motif est également appelé « switch I » car il présente des conformations différentes suivant l'état de la GTPase et le nucléotide fixé. Souvent observé dans un

état non-ordonné à l'état « off », ce switch qui est structuré à l'état « on » permet de stabiliser le groupement phosphate γ du nucléotide triphosphate. Il participe aussi à la coordination d'un ion magnésium qui est nécessaire à la catalyse. Ce segment est aussi impliqué dans l'interaction des protéines G monomériques avec leurs effecteurs.

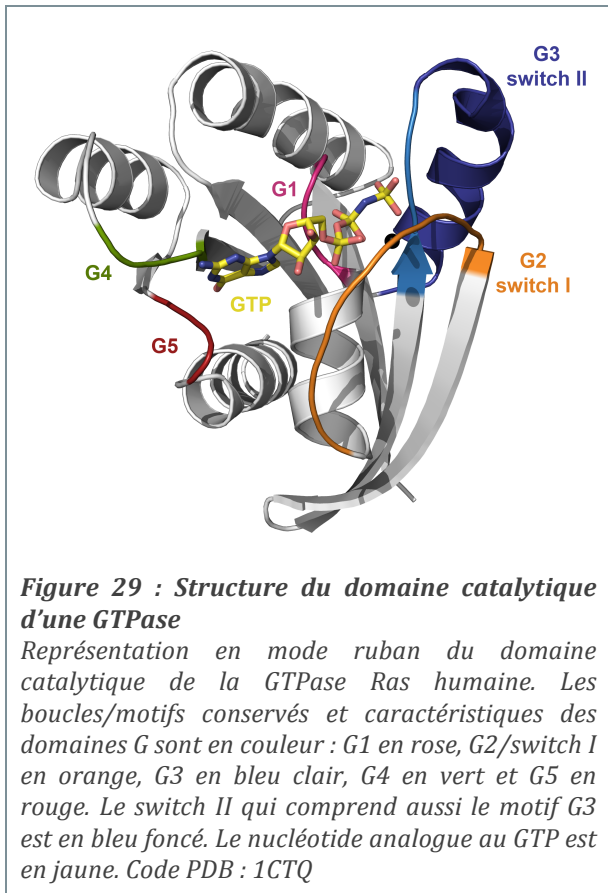
- Le motif G3 est l'équivalent du motif de Walker B. De séquence consensus DXXG, il adopte en général une conformation spécifique des GTPases et différente des protéines à boucle P. L'Aspartate conservé de ce motif intervient, par l'intermédiaire d'une molécule d'eau, dans la fixation de l'ion magnésium. La Glycine conservée stabilise le groupement phosphate γ du nucléotide. Dans certaines familles de GTPases, le motif G3 contient un acide aminé polaire supplémentaire (une Glutamine dans le cas de Ras) qui est impliqué dans le mécanisme d'hydrolyse du GTP (voir section 5.1.3). Le motif G3 est inclus dans la seconde région flexible des GTPases, le « switch II », qui adopte des structures différentes, voire désordonnées. Comme pour le switch I, ses conformations sont influencées par le nucléotide fixé. La région switch II est un site important de reconnaissance des différents régulateurs des GTPases et les changements de conformation jouent un rôle primordial dans l'échange GDP/GTP (voir section 5.1.4) [312, 318, 319].

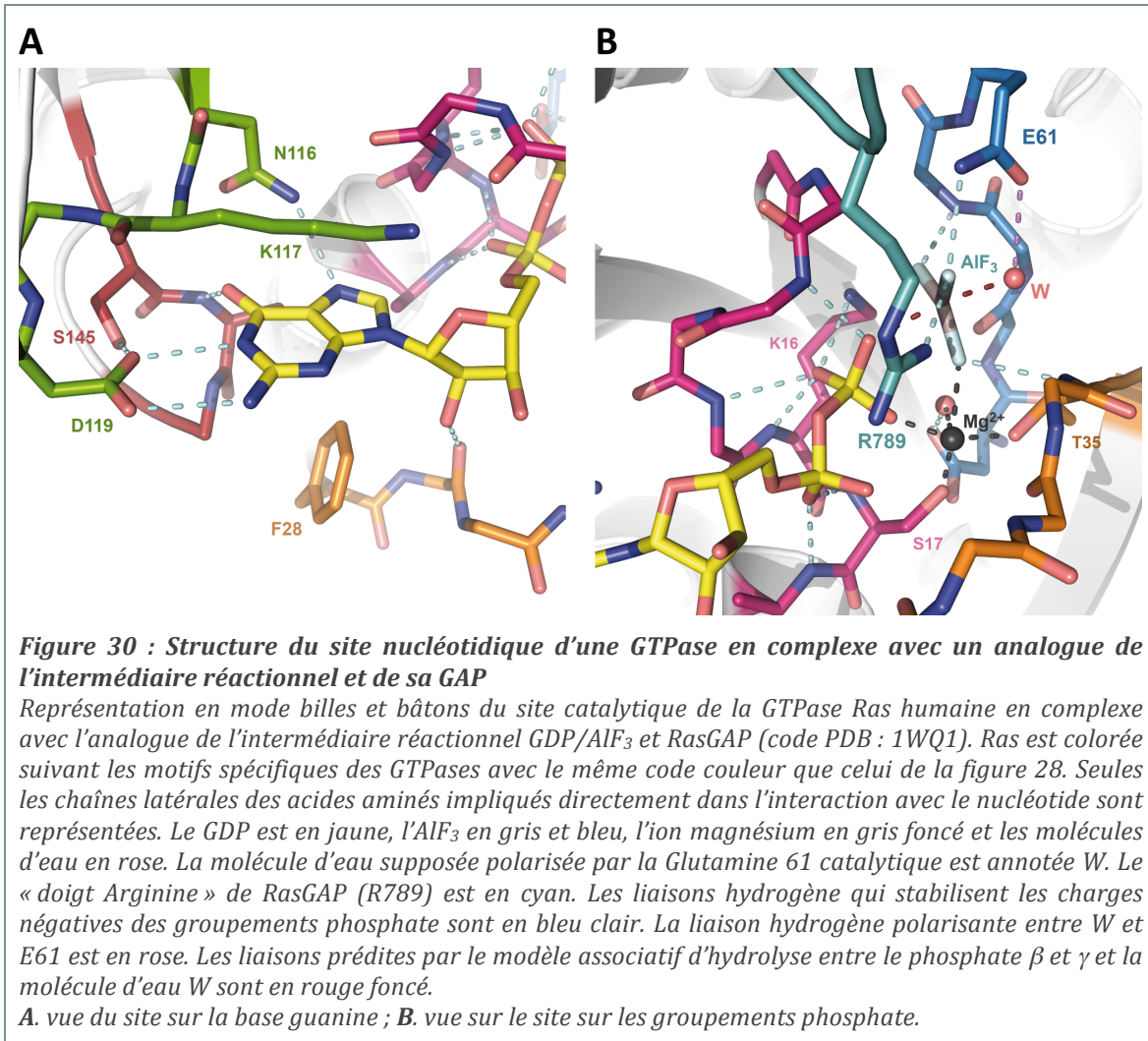
- Le motif G4, de séquence N/TKXD, confère la spécificité des GTPases pour le nucléotide et permet la stabilisation de la base Guanine. Ce motif oriente la base en formant avec cette dernière à la fois un pont salin grâce à son acide Aspartique, une liaison hydrogène grâce à son Asparagine et des contacts hydrophobes grâce à sa Lysine.

- Le motif G5 n'est pas présent dans toutes les GTPases. Sa séquence de type SAL/K est aussi responsable de la sélectivité des GTPases pour le GTP et le GDP grâce à une interaction spécifique avec l'atome d'oxygène O6 de la base Guanine mais absent de la base Adénine.

5.1.3. Le mécanisme d'hydrolyse

Le cycle hydrolytique des GTPases peut être vu comme un mécanisme universel de commutateur permettant le contrôle et la régulation des fonctions





cellulaires. A part quelques exemptions dont les GTPases bactériennes impliquées dans la synthèse du ribosome [316], les GTPases présentent une forte affinité pour le nucléotide, typiquement de l'ordre du picomolaire au nanomolaire. Par conséquent, ces enzymes alternent entre leurs deux états actifs et inactifs et sont rarement observées sous forme libre. Ces commutateurs moléculaires peuvent aussi être vus comme des minuteurs dont le temps de vie des formes liées au GTP peut déterminer l'activation des processus régulés : l'hydrolyse du GTP interrompt le minuteur et le signal de régulation positive [320].

Le mécanisme fondamental d'hydrolyse de GTP en GDP par les domaines G a suscité de nombreuses interrogations depuis la découverte des GTPases et reste encore aujourd'hui très discuté. Deux modèles extrêmes de l'état de transition du nucléotide permettent d'imaginer la réaction d'hydrolyse (pour revue [312, 321]).

Le modèle dit associatif a pendant longtemps prévalu. Ce modèle décrit l'état de transition par un intermédiaire réactionnel pentavalent qui est induit par la formation d'une liaison entre une molécule d'eau nécessaire à la réaction et le groupement phosphate γ du nucléotide. Cette liaison est associée au transfert des charges négatives vers les atomes d'oxygène de ce groupement phosphate. Dans ce

cas, il est nécessaire que le site catalytique de la GTPase possède une base pour extraire un proton de la molécule d'eau alors « activée » et pour favoriser son attaque nucléophile du GTP et de la liaison phosphoanhydride. L'acide aminé supplémentaire, qui peut être polaire ou acide et qui est trouvé dans le switch II/motif G3 de certaines familles de GTPases (i.e. Glutamine 61 dans Ras, voir Figure 28), est considéré comme l'activateur de la molécule d'eau « catalytique ». Toutefois, plusieurs études et observations ne sont pas en accord avec ce modèle. Notamment, la mutation de la Glutamine « base » de la GTPase Ras en un analogue isoélectronique (NGln) ne montre aucune influence de la réduction de charge du site sur le taux d'hydrolyse [322]. De plus, certaines familles de GTPases, les HAS-GTPases (pour Hydrophobic Amino acid Substitution-GTPases) possèdent un acide aminé hydrophobe à la place de l'acide aminé polaire, suggérant un mécanisme hydrolytique différent [323]. Certaines études suggèrent même que le GTP en lui-même pourrait faire office de base [324, 325].

D'autres observations faites par des études cinétiques, structurales et de spectroscopies supportent un modèle d'hydrolyse plutôt dissociatif. Ce modèle prévoit la formation d'un ion métaphosphate résultant de la rupture de la liaison entre les deux groupements phosphate β et γ et le déplacement des charges négatives sur le groupement β . Toutefois, cet intermédiaire métaphosphate n'a jamais été observé en tant que tel [321]. Dans ce contexte, l'acide aminé polaire du switch II (i.e. Gln61 de Ras) est vu comme un stabilisateur de l'état de transition en orientant la molécule d'eau [326].

Quel que soit le modèle d'hydrolyse, les structures cristallographiques de différentes GTPases seules ou en complexe soit avec les nucléotides, soit avec des analogues d'intermédiaires réactionnels ont permis d'identifier les atomes et les acides aminés catalytiques (voir Figure 30). Ces derniers sont nécessaires à l'orientation du substrat et de la molécule d'eau ainsi qu'à la stabilisation des charges négatives produites lors de la réaction [312, 321]. Le site catalytique possède ainsi plusieurs espèces de charge positive. Il est à noter la présence de l'ion magnésium et d'atomes donneurs d'hydrogène issus de la chaîne principale de la boucle P (i.e. Gly13, Val 14, Gly15 de Ras) et issus de sa Lysine (i.e. Lys16 de Ras). Une charge positive supplémentaire peut aussi venir compléter le site actif et fortement accélérer la réaction d'hydrolyse. Cette charge, généralement portée par une Arginine, peut être apportée *in trans* par une GAP (« doigt Arginine », voir section 5.1.4 et Figure 30) ou *in cis* par un autre domaine de la GTPase.

5.1.4. Les régulateurs GEF et GAP

Le cycle GDP/GTP des GTPases est en général fortement régulé par différentes familles de partenaires : Les Facteurs d'Echange de Guanine (GEF) qui catalysent le relargage du GDP qui sera remplacé par le GTP et les Protéines Activatrices des GTPases (GAP) qui apportent habituellement un groupement chargé essentiel à l'hydrolyse du GTP (voir section 5.1.3).

Parce que l'affinité de la plupart des GTPases pour les nucléotides est importante et par conséquent la vitesse de dissociation du GDP très lente (temps de demi-vie de l'ordre de l'heure), l'activité d'échange des GEF est nécessaire pour un échange GDP/GTP rapide et pour assurer les processus biologiques dont la durée est de l'ordre de la minute. Ainsi, les GEF sont capables d'augmenter la vitesse d'échange de plusieurs ordres de grandeur lors de leur complexation avec les GTPases. Pour cela, le facteur d'échange modifie le site de fixation du nucléotide et diminue l'affinité du GDP pour la GTPase. La concentration plus importante du GTP dans le milieu cellulaire permet alors la substitution [155, 327, 328]. Les séquences et les différentes structures de GEF seul ou en complexe avec leur GTPase cible ont révélé des structures différentes des facteurs d'échange et des approches différentes du site catalytique. La majorité des études structurales et biochimiques concernent les GEF des petites protéines G humaines [329-333]. L'ensemble de ces données montre toutefois que ces GEF utilisent des mécanismes similaires pour déformer le site de fixation du nucléotide en imposant des changements de conformation des régions switches et de la boucle P. Il en résulte des perturbations des interactions entre la GTPase et le GDP, ainsi qu'un déplacement de l'ion magnésium, qui permettent la libération du GDP [155, 327, 334].

Bien que les GTPases soient capables d'hydrolyse du GTP, cette réaction est en général très lente et la réaction de catalyse est rendue efficace par les GAP qui accélèrent la vitesse de réaction de plusieurs ordres de grandeur. L'exemple le plus connu est la GAP de la petite protéine G Ras humaine (RasGAP) : sa présence permet d'augmenter par un facteur 100 000 la vitesse d'hydrolyse du GTP. En se complexant à la GTPase, la protéine activatrice amène dans le site actif une Arginine supplémentaire, appelée aussi « doigt Arginine ». La charge supplémentaire portée par cet acide aminé compense les charges négatives de l'intermédiaire réactionnel (voir section 5.1.3 et Figure 30) [335, 336]. D'autres mécanismes d'activation sont aussi observés mais restent sur le même principe de stabilisation du site catalytique de la GTPase et par l'insertion *in trans* d'acides aminés issus de la GAP [155, 327].

Il existe une dernière classe de régulateurs des GTPases : les GDI (pour GTP Binding Domain Inhibitors). A l'heure actuelle, seuls des GDI spécifiques des GTPases de la famille Rho chez les mammifères ont été identifiés. Leur fonction est de bloquer la GTPase dans une forme inactive [337].

Les GTPases bactériennes semblent posséder un cycle régulé différemment qui reste aujourd'hui à être déterminé. Aucun GEF homologue aux facteurs d'échange eucaryotes n'a encore été révélé. Ceci est certainement dû à une affinité moindre des GTPases procaryotes pour le nucléotide (de l'ordre du micromolaire). Elle pourrait permettre un échange spontané sans l'aide de partenaire [317]. Récemment, une protéine conservée chez les bactéries (YihI) a été identifiée comme stimulateur de l'activité d'hydrolyse de la GTPase Der. Malgré une activité doublement stimulée, il a été proposé que YihI joue un rôle de GAP [338]. D'autres mécanismes d'activation de l'activité GTPases pourraient aussi avoir lieu. Par exemple, il a été montré que les interactions avec les effecteurs ribosomaux ou nucléiques comme dans le cas de la

GTPase Era [339] ou la présence d'ion potassium comme dans le cas de MnmE ou YqeH [340, 341] permettent d'augmenter de façon significative l'activité de ces GTPases procaryotes.

5.1.5. Les GTPases : cibles antibactériennes ?

Les recherches de molécules inhibitrices de GTPases sont encore aujourd'hui très confidentielles et peu nombreuses. Parce qu'impliquées dans des pathologies cancéreuses [342], les petites protéines G humaines ont fait l'objet de recherche de molécules à visée thérapeutique. Contrairement au succès de la recherche d'inhibiteurs qui ciblent les protéines kinases de manière compétitive à la fixation du nucléotide (voir section 2.1.2), il n'existe que de rares exemples de ce type d'inhibiteurs pour les GTPases [343]. Aujourd'hui, les molécules identifiées permettent surtout de bloquer le recrutement des protéines G à la membrane en inhibant les modifications post-traductionnelles comme la prénylation ou en inhibant leurs régulateurs [310]. Par exemple, sont à citer les efforts faits sur le développement d'inhibiteurs de la farnesyltransferase. Ils permettent de bloquer l'association à la membrane de mutants de la GTPase Ras [344]. Un autre inhibiteur est la Brefeldine A, un composé naturel qui bloque l'activation des GTPases Arf. En stabilisant le complexe entre ces enzymes et leurs GEF dans une conformation particulière, il ne permet pas l'échange nucléotidique [345, 346]. Des petites molécules ciblant la surface de la GTPases Rac1 ont aussi été identifiées. Elles interviennent dans l'interaction avec les facteurs d'échange, bloquant ainsi l'activation de Rac1 par ses GEF spécifiques comme TRIO ou Tiam1 [347, 348].

Peu de GTPases bactériennes sont la cible de molécules antibiotiques. Les meilleurs exemples sont certainement les inhibiteurs des facteurs de translation bactériens EF-Tu [349]. Différentes molécules fonctionnant sur deux mécanismes distincts ont été identifiés : des inhibiteurs de la fixation des ARNt [350] et des inhibiteurs comme les dérivés de la Kiromycine qui, en se fixant à l'interface entre le domaine G et un autre domaine, bloquent EF-Tu dans une conformation similaire à celle de la forme active. Ces composés empêchent ainsi le relargage du GDP de la protéine et la dissociation de la protéine du ribosome [351, 352].

Par ces exemples, les GTPases, leurs effecteurs et leurs régulateurs semblent être des cibles potentielles dans la recherche de nouveaux agents thérapeutiques. Toutefois, d'un point de vue « drogabilité », c'est-à-dire d'un point de vue de l'identification, de la caractérisation, de l'optimisation et des essais précliniques, les études sur les molécules inhibitrices des fonctions de ces protéines ne sont qu'au stade naissant quand nous les comparons à la découverte des inhibiteurs de kinases, pour ne prendre qu'un seul exemple. En ce qui concerne les GTPases bactériennes, beaucoup restent donc à découvrir. Leurs rôles et leurs fonctions exacts dans la physiologie de la cellule restent élusifs bien que quantité de données soient déjà publiées sur nombre de ces enzymes. Et contrairement aux GTPases eucaryotes pour lesquelles les régulations par les partenaires GEF, GAP ou GDI et les interactions avec leurs effecteurs sont largement documentées, la régulation du cycle des GTPases

bactériennes est encore inconnue. Il sera donc nécessaire de comprendre le rôle de ces enzymes dans la physiologie de la cellule, d'identifier leurs facteurs de régulation et de caractériser structuralement leurs mécanismes d'activation et d'hydrolyse qui apparaissent significativement différents des GTPases eucaryotes. Ces données permettront dans le futur le développement de nouvelles classes d'antibiotiques pour lutter contre les infections bactériennes.

5.2. EngA

Parmi les gènes de GTPases bactériennes conservées et impliquées dans la biogenèse du ribosome, le gène *yphC* a été décrit pour la première fois en 2000 comme un gène essentiel à la survie de la bactérie. Identifié chez *Neisseria gonorrhoeae*, le gène a été préalablement baptisé *engA* (pour Essential Neisserial GTPase A, [353]). Les études d'inactivation de ce gène qui ont suivi chez différentes espèces bactériennes comme *Escherichia* [354], *Salmonella* [355], *Bacillus* [356], *Streptococcus*, *Staphylococcus* ou *Haemophilus* [357] ont confirmé son rôle essentiel chez les procaryotes et il a été proposé qu'il fasse parti du génome minimal nécessaire à la vie bactérienne [358].

Le produit d'expression du gène *yphC* est aussi bien appelé Der, YphC, EngA ou YfgK suivant l'organisme. Nous l'appellerons EngA pour plus de clarté dans le texte. La protéine EngA est décrite comme appartenant à une famille de GTPases dénommée suivant ses différents membres : Era/Obg/EngA/MnmE/YchF/YihA. Il est intéressant de noter que ce gène est largement conservé chez les eubactéries mais pas chez les archaebactéries, ni chez les eucaryotes à l'exception d'un exemplaire chez *Arabidopsis* qui pourrait provenir d'une symbiose bactérienne ancestrale [359].

5.2.1. La structure de EngA

L'étude de la séquence montre que les protéines EngA présentent la particularité de posséder deux fois les motifs spécifiques des GTPases. EngA est donc unique au sens où elle possède deux domaines G en tandem, dénommé GD1 et GD2. La ressemblance entre les deux domaines suggère une duplication de gène et il a été proposé que EngA soit issue de la duplication du gène d'Era, une GTPase de la même famille impliquée dans le métabolisme du ribosome ; d'où son autre nom : Der pour « double Era » [360]. L'homologie de séquence des motifs G2 entre les deux domaines suggère aussi l'existence d'effecteurs similaires voire identiques. Les deux domaines G sont reliés entre eux par une séquence acide dont la longueur varie d'une bactérie à l'autre. Enfin, la partie C-terminale montre un domaine avec un repliement similaire à celui des domaines KH (pour hnRNP K Homology). Ces domaines sont connus pour leur interaction avec les acides nucléiques [361]. Cependant, EngA ne possède pas le motif spécifique de fixation retrouvé dans ces domaines.

Notre démarche structurale se base sur deux structures de EngA issue de *B. subtilis* [362] ou *T. maritima* [360] et publiées précédemment à notre travail. Ces structures cristallines ont révélé l'architecture de EngA et l'arrangement de ses trois domaines. Les domaines G adoptent un repliement tertiaire caractéristique des

domaines G de la famille TRAFAC et comparable aux petites protéines G eucaryotes. De façon inattendue, ils se positionnent de part et d'autre du domaine KH d'une manière telle qu'il n'existe pas d'interaction directe entre eux (voir Figure 31).

Toutefois, l'arrangement quaternaire des domaines est différent entre les deux structures publiées et il a été proposé qu'il soit influencé par la nature des nucléotides fixés et par l'état on/off des domaines G. En effet, la structure de EngA de *T. maritima* a été résolue avec le domaine GD2 contenant du GDP dans son site actif et le domaine GD1 possédant deux ions phosphate. Les positions de ces derniers sont proches des positions supposées des groupements phosphates β et γ du GTP. Il a alors été suggéré que cette structure du domaine GD1 représenterait la structure d'un

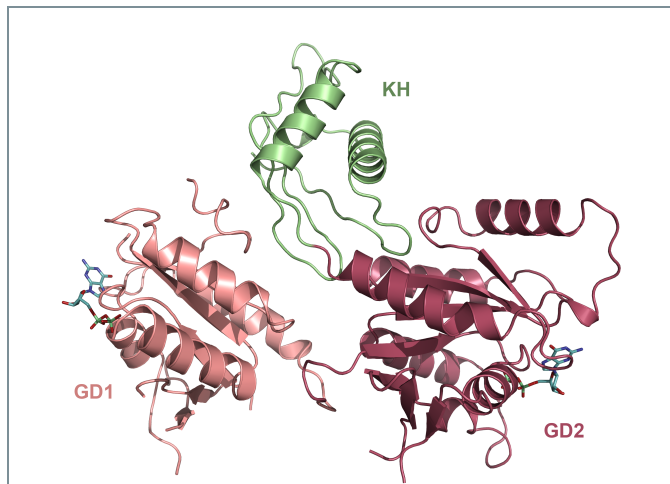


Figure 31 : Structure de la GTPase EngA de *B. subtilis*
 Représentation en mode ruban de la structure de la GTPase EngA de *B. subtilis* en complexe avec le GDP. La structure présente deux domaines G en tandem, GD1 en rose clair et GD2 en rose foncé. Ils s'arrangent autour d'un domaine central qui possède une topologie similaire aux domaines d'homologie K. Code PDB : 2HJG

état similaire à son état « on » (GTP lié) [360]. La structure de EngA de *B. subtilis* a quant à elle été résolue avec du GDP dans les deux domaines G, suggérant ainsi une image de leur état « off » [362]. Pourtant, les deux structures montrent une différence très importante de position du domaine GD1 par rapport aux domaines GD2 et KH qui restent inchangés. Lorsque les couples GD2/KH sont superposés, les domaines GD1 montrent en effet une différence d'orientation de plus de 60°. Ce déplacement drastique est de plus associé à un déplacement d'une hélice α qui fait partie du switch II du domaine GD1. A l'état « on », la position de ce switch permet de stabiliser le domaine dans une première position. L'hydrolyse du GTP en GDP par le domaine GD1 et son passage de l'état « on » à l'état « off » imposerait alors un changement de conformation de l'hélice α pour la libérer des interactions que forme cette dernière avec le domaine central. Ce changement de conformation serait à son tour à l'origine des modifications de l'arrangement quaternaire [362].

La différence de position du domaine GD1 permet aussi d'exposer des surfaces moléculaires aux caractéristiques électrostatiques différentes. Dans le cas de l'état supposé « on » du domaine GD1, sa position permet de mettre en évidence une large surface positivement chargée. Elle a été proposée pour être un site d'interaction avec les ARN ribosomiaux. Ainsi, les changements de conformation d'EngA pourraient jouer un rôle crucial dans la régulation des interactions que pourrait faire cette GTPase avec ses partenaires ribosomiaux [362].

5.2.2. Les fonctions de EngA

Peu de données sont disponibles concernant l'expression de EngA *in vivo*. Il a été estimé une expression constitutive de 7000 molécules par bactérie chez *B. subtilis* [356] et une localisation cytoplasmique chez *E. coli* [363]. La première indication de l'implication de cette GTPase dans les mécanismes d'assemblage du ribosome bactérien vient d'une étude de complémentation de gène chez *E. coli*. Sa surexpression compense la mutation et la délétion d'un gène de modification de l'ARN ribosomal 23S qui provoque un défaut de croissance bactérienne et une accumulation des sous-unités 30S et 50S du ribosome au détriment du ribosome 70S [364]. Il a ensuite été démontré que la suppression d'EngA chez *B. subtilis* et *E. coli* entraîne une altération du profil ribosomique issue d'un défaut de maturation de la sous-unité 50S du ribosome [365, 366]. Ces mêmes études montrent aussi que la GTPase interagit spécifiquement avec la sous-unité 50S du ribosome *in vitro* de manière nucléotide dépendante. Plus récemment, une étude chez *E. coli* montre que EngA est aussi capable de se lier aux particules ribosomales 30S et 70S. En plus d'être dépendantes du nucléotide, ces interactions interviennent par l'intermédiaire des différents domaines G de la GTPase. La condition nécessaire à la fixation au ribosome serait la forme GTP liée du domaine GD2. Dans cet état, EngA présente une faible association à la particule 50S qui serait alors stabilisée par l'addition de GTP lié au domaine GD1. L'hydrolyse du GTP en GDP et les différentes formes des domaines G permettraient donc de réguler la fixation spécifique de la GTPase aux différentes formes et sous-unités du ribosome [367]. Cette interaction spécifique avec le ribosome est confortée par la mise en évidence d'une interaction avec des protéines ribosomales. EngA d'*E. coli* est co-purifiée avec différentes protéines ribosomales [365, 368] et les interactions entre les protéines structurales S7 ou S9 et EngA de *Salmonella* ont été caractérisées *in vitro* [369]. Enfin le domaine KH semble aussi essentiel pour l'interaction avec la sous-unité 50S [370].

5.2.3. La biochimie de EngA

Les études biochimiques du cycle GTPase d'EngA sont encore plus succinctes dans la littérature. Les tests enzymatiques de différents homologues de EngA ont confirmé qu'il s'agit bien d'une GTPase active capable de fixer aussi bien le GTP que le GDP mais pas le GMP et les nucléotides à base Adénine et capable d'hydrolyser le GTP *in vitro*. Il apparaît par les études chez *E. coli*, *T. maritima* et *S. typhimurium* que EngA peut être considérée comme une GTPase à forte activité intrinsèque avec des activités catalytiques k_{cat} comprises entre 0,1 et 1,2 min^{-1} [365, 369, 371]. Cette forte activité est à mettre en rapport avec l'activité de hRas seule ($5 \cdot 10^{-3} \text{ min}^{-1}$, [372]) ou Era (10^{-2} min^{-1} , [373]). Chaque domaine G possède une activité GTPase supposée indépendante. La vitesse d'hydrolyse par le domaine GD1 isolé d'EngA de *T. maritima* est similaire à celle de la protéine entière et son domaine GD2 a une activité deux fois moindre. De façon intéressante, la mutation du motif G4 (Asparagine en Aspartate) du domaine GD1 induit une diminution de l'activité de la protéine entière alors que la même mutation dans le domaine GD2 une augmentation.

Ainsi contrairement aux GTPases eucaryotes qui généralement nécessitent des partenaires activateurs comme les GAP (voir section 5.1.4), il semble que les GTPases bactériennes et notamment EngA hydrolysent le GTP sans cofacteurs protéiques. Toutefois, il peut s'avérer que des cofacteurs chimiques additionnels améliorent l'efficacité hydrolytique. C'est le cas pour le potassium qui augmente l'activité de certaines GTPases bactériennes, notamment celles de la même famille comme MnmE [340, 374].

L'étude de EngA de *S. typhimurium* par calorimétrie montre que les deux sites nucléotidiques ne sont pas équivalents en terme d'affinité : en absence de magnésium, EngA de *Salmonella* fixe le GDP avec une affinité vingt fois supérieure à celle du GTP, le GD2 ayant une affinité cinq fois supérieure à celle du GD1 pour le GDP [369]. De l'ordre du micromolaire, cette affinité pour le nucléotide est à mettre en relation avec celle des GTPases qui sont en général de l'ordre du nanomolaire. Cette faible affinité est aussi observée pour les autres GTPases bactériennes comme MnmE [340, 374] ou Era [375].

Enfin, l'étude de mutants des boucles P de EngA de *E. coli* ont confirmé l'implication du motif G1 dans l'hydrolyse du nucléotide mais aussi suggère une coopération entre les deux domaines : la fixation et/ou l'hydrolyse par un des domaines stimulerait ou régulerait le second [365].

6. Résultats marquants sur EngA

L'ensemble des données préliminaires sur la GTPase bactérienne EngA chez différents organismes a irrémédiablement montré son implication dans la biogenèse du ribosome. Son caractère essentiel à la vie de la bactérie, son occurrence exclusive dans le règne bactérien et son rôle illustrent l'intérêt d'étudier en détails son cycle d'hydrolyse du nucléotide et les mécanismes qui sont associés à ce dernier. Sa conservation dans le règne procaryote et son absence chez les eucaryotes, notamment chez les mammifères, font de EngA une cible attrayante pour le développement de nouveaux agents antibactériens. Nous avons commencé notre étude en nous focalisant sur EngA de *B. subtilis* dont une première structure en complexe avec le GDP a été résolue précédemment [362].

Les premiers objectifs sont les caractérisations enzymatiques et structurales de cette protéine ainsi que de ses domaines G indépendamment et dans le but de comprendre le rôle respectif de chaque domaine.

Les résultats obtenus par nos collaborateurs ont permis de démontrer l'effet stimulateur du potassium sur l'activité GTPase de EngA de *Bacillus*. La présence de cet ion augmente de vingt à trente fois l'efficacité d'hydrolyse. L'étude sur les domaines G séparés nous montre aussi que ces deux domaines sont tous les deux stimulés. L'activité GTPase des deux domaines est néanmoins différente : le domaine GD1 est beaucoup plus actif que le domaine GD2, confirmant ce qui avait déjà été observé pour l'homologue de *Thermotoga*. En effet, l'activité du domaine GD1 est comparable à la protéine entière et est toujours supérieure à celle du domaine GD2. Cet effet stimulateur du potassium a aussi été observé pour d'autres GTPases bactériennes de la famille comme MnmE [340] ou FeoB [376]. Les analyses phylogénétiques suggèrent aussi que plusieurs autres membres de cette famille pourraient aussi être sensibles au potassium [317]. Les structures des domaines G de MnmE et FeoB en complexe avec le potassium montrent même que l'ion K^+ , qui est chélaté par le switch I, se place dans le site actif à une position identique à celle du « doigt Arginine » observé dans la structure du complexe Ras/RasGAP (voir section 5.1.3). Ces observations soulignent le caractère général de l'activation de ces domaines G et il a été proposé que l'ion potassium agisse comme un facteur d'activation chimique (GAF pour GTPase Activator Factor), à l'instar des GAP.

L'affinité des nucléotides pour le domaine G peut influencer l'activité GTPase. Précédemment, l'étude de microcalorimétrie sur EngA de *Salmonella* suggère que cette GTPase possède deux classes distinctes de sites, l'un de forte affinité et l'autre de faible affinité [369]. Toutefois, Les mesures faites par nos collaborateurs vont à l'encontre des observations précédentes. En effet, par des méthodes de fluorescence les affinités des deux domaines isolés pour les nucléotides sont comparables.

Avec l'objectif d'obtenir une image des différentes étapes du cycle réactionnel de cette GTPase à double domaine G, nous avons entrepris une étude par cristallographie des différents complexes que peut former EngA de *Bacillus* avec les nucléotides ou des analogues supposés non-hydrolysables. Nous avons ainsi résolu la

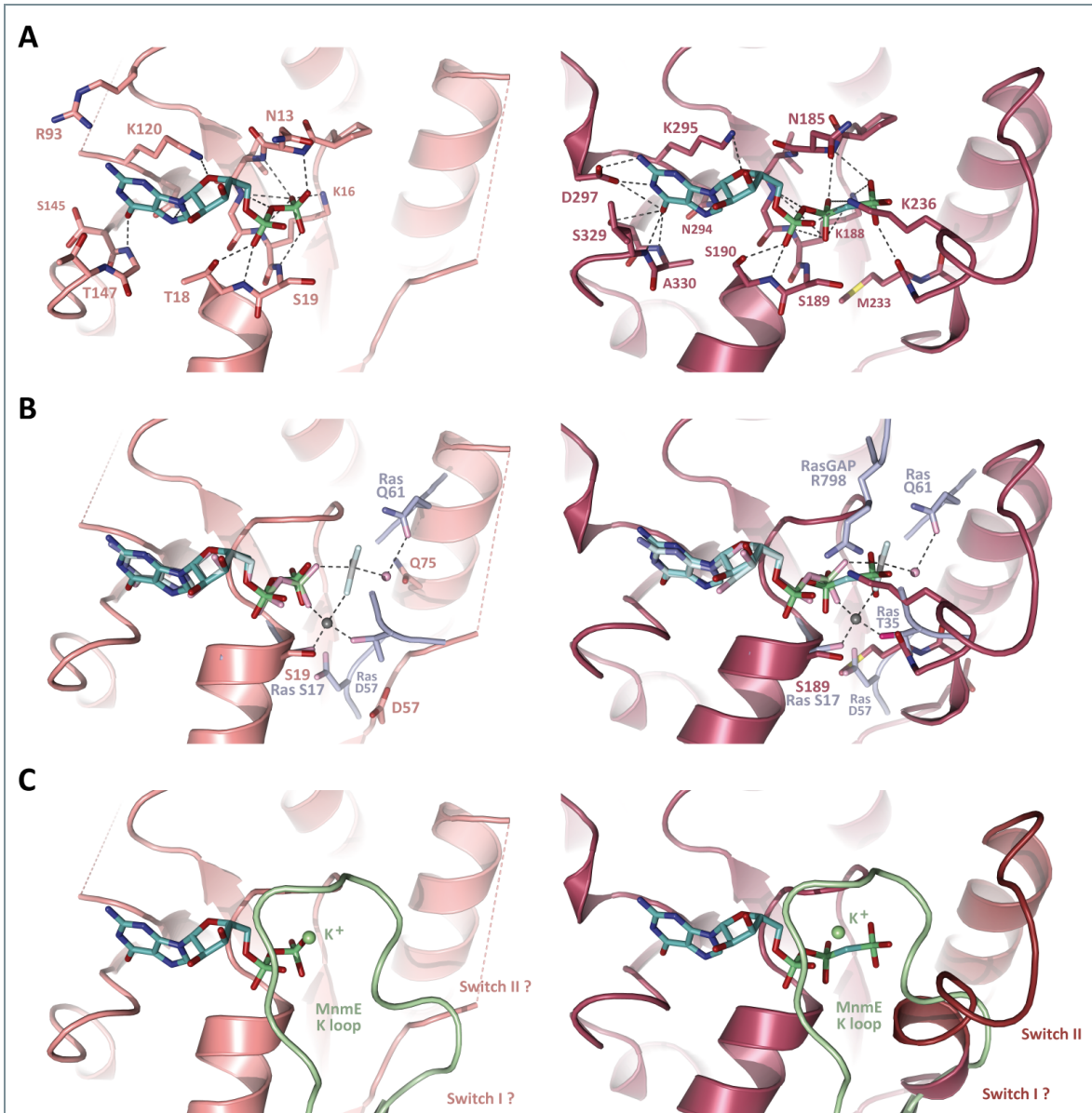


Figure 32 : Structure du site actif des domaines GD1 et GD2 de EngA

Représentation en mode mixte du site catalytique des domaines GD1 en complexe avec du GDP (à gauche en rose clair) et du domaine GD2 en complexe avec du GMPPCP (à droite en rose foncé). Les switches désordonnés sont représentés par des lignes pointillées.

A. vue des sites actifs avec les acides aminés des domaines G impliqués dans la fixation du nucléotide. Les liaisons hydrogènes sont représentées par des lignes pointillées grises.

B. superposition avec les acides aminés de la GTPase Ras (en parme), l'ion magnésium (en gris) et la molécule d'eau catalytique (en rose) ; la conformation du switch II du domaine GD1 doit changer pour permettre le positionnement d'acides aminés (Q75 et D57) qui pourraient être impliqués dans le mécanisme d'hydrolyse. le site catalytique du domaine GD2 est incomplet et encombré par une Lysine (K236) qui prend la place supposée de l'ion K^+ ou du « doigt Arginine » de rasGAP (R798).

C. superposition avec la boucle de chélation du potassium de MnME (en vert), une GTPase activée par cet ion. Les switches I des domaines G désordonnés pourraient adopter une conformation similaire. Un changement de conformation du switch II du domaine GD2 est néanmoins nécessaire pour éviter un conflit stérique.

structure de EngA de *B. subtilis* en présence de nucléotides di-phosphates ou triphosphates. Nous avons ainsi à notre disposition les structures de la protéine entière avec le domaine GD1 sous une forme vide, avec du GDP ou du GTP et le domaine GD2 avec du GDP, du GTP ou du GMPPCP (voir Figure 32). Toutefois, malgré de nombreux essais de cristallisation en présence des ions cofacteurs, nous n'observons jamais d'ions magnésium ou potassium dans les cartes de densité électronique. De plus, toutes les structures obtenues sont similaires à la structure de EngA contenant du GDP dans les deux domaines. Les régions précédemment absentes dans la structure de EngA publiée, notamment certains switches des domaines G, restent inobservables dans les nombreuses structures que nous avons résolues. Ceci vient certainement du fait inévitable de leur caractère désordonné et flexible. Seul le switch II du domaine GD2 est structuré.

Aucun changement de conformation quaternaire et de réarrangement de domaine n'est donc observé. Notamment, la structure avec le GTP complexé au domaine GD1 ne montre pas d'arrangement des domaines comparables à celui observé dans la structure de EngA de *T. maritima* (voir section 5.2.1) [360]. La question de savoir comment la nature du nucléotide dans le site du domaine GD1 influence l'arrangement quaternaire reste donc ouverte. Les structures du domaine GD2 contenant du GTP ou son analogue GMPPCP ne montrent pas non plus de différences drastiques si nous la comparons à sa structure en complexe avec le GDP. Pour les deux domaines G, il n'est pas exclure que la fixation du nucléotide triphosphate ne soit pas capable d'induire à elle seule les changements de conformation et la stabilisation des régions switches. Par exemple et comme il a été observé pour Ras ou MnmE, la présence du facteur d'activation (l'Arginine de la GAP ou l'ion potassium) peut être nécessaire à l'obtention de la forme active du domaine G. Comme nous n'observons pas de GAF chimique dans nos structures malgré sa présence dans les conditions de cristallisation, les domaines G pourraient rester dans un état similaire à l'état « off », même en complexe avec le nucléotide triphosphate seul (voir Figure 32).

Bien que nous ayons à l'heure actuelle une image incomplète des sites actifs de EngA, les activités différentielles, les séquences et structures disponibles nous suggèrent tout de même des mécanismes d'hydrolyse différents pour les deux domaines G. La séquence du domaine GD1 nous montre que ce domaine possède tous les éléments nécessaires à une catalyse efficace en présence de potassium. Une séquence spécifique de la boucle de chélation de l'ion K^+ est bien retrouvée dans la séquence du domaine GD1. Il est donc probable que la région switch I adopte une conformation similaire à la conformation active observée dans les structures de MnmE et FeoB (voir paragraphe précédent). La séquence du switch II présente des acides aminés plutôt acides. Il est donc vraisemblable que les acides aminés au niveau du motif B de Walker soient impliqués d'une manière similaire à celle qui est généralement observée pour les GTPases (voir section 5.1.3 et Figure 32). A l'inverse, le domaine GD2 présente certaines caractéristiques pour une hydrolyse altérée. La séquence de son switch II présente un caractère basique et les structures révèlent un

site GTPase incomplet. La position inhabituelle du switch II et la présence d'une Lysine placée dans la zone supposée de fixation de l'ion potassium sont aussi incompatibles avec la présence de cet ion et la conformation active que pourrait adopter le switch I (voir Figure 32). Un changement de conformation drastique et coûteux énergétiquement est donc nécessaire au switch II pour permettre l'orientation correcte des régions switches et la fixation du facteur d'activation. Ce coût pourrait être à l'origine des différences d'efficacité catalytique observées entre les deux domaines.

Ces résultats obtenus pendant la thèse d'Anne-Emmanuelle Foucher font l'objet d'un article soumis au journal PlosOne (Publication P13).

7. Projet à venir sur les inhibiteurs de EngA

L'objectif ultime de ce projet est d'élaborer des candidats moléculaires et de rassembler des données structurales complètes pour la conception de nouvelles générations d'antibiotiques. Ces molécules que nous recherchons seront dirigées pour la première fois contre l'assemblage du ribosome bactérien. En effet, jusqu'à maintenant, beaucoup de molécules ont été développées pour inhiber directement les mécanismes ribosomiaux mais jamais les mécanismes en amont de son assemblage qui est pourtant vital pour la bactérie.

Notre projet consiste à rechercher des inhibiteurs capables d'interagir avec EngA et d'inhiber soit son activité GTPase, soit ses interactions avec ses partenaires. Grâce à la modélisation de la structure complète des domaines G et à un criblage virtuel de librairies chimiques, des molécules capables de réduire à la fois l'activité de GTPases EngA issues de différentes bactéries et la croissance de ces dernières ont récemment été identifiées [377]. Notre approche sera sensiblement différente et utilisera des techniques expérimentales pour tester directement la faculté de petites molécules à se fixer ou non à la GTPase. Le test utilisé pour cribler ces dernières sera le TSA (pour "Thermal Shift Assay"). Cette technique est basée sur la fixation d'une sonde (le Sypro Orange) sur les zones hydrophobes d'une protéine qui sont progressivement exposées à la suite d'une dénaturation thermique, elle-même induite par une augmentation de la chaleur. Il a été montré pour de nombreuses protéines que l'addition de ligands permet de les stabiliser et de les protéger contre cette dénaturation. Une différence dans la température de dénaturation, qui est directement reliée à l'affinité de la protéine pour ses ligands, peut alors être observée. Cette technique permet un criblage à moyen débit afin d'identifier de nouveaux inhibiteurs de diverses protéines. Elle a notamment été utilisée avec succès pour identifier des inhibiteurs de protéines kinases eucaryotes. Dans notre cas, nos collaborateurs ont validé l'utilisation de ce test avec différents nucléotides. Par cette approche, ils cribleront plusieurs chimiothèques dans un premier temps restreintes à des composés inhibiteurs de NTPases ou à des composés plus variés mais dont une activité biologique est connue.

En parallèle, il sera nécessaire de connaître dans le détail les mécanismes hydrolytiques de EngA et les interactions avec ses partenaires qui permettent à la GTPase de réguler le cycle de biogenèse du ribosome bactérien. Un des premiers objectifs de notre étude sera donc de continuer nos efforts dans la caractérisation structurale des états de EngA de *B. subtilis* afin d'avoir une image complète du cycle réactionnel particulier à cette famille de GTPases. En complément, différents mutants de la protéine entière ou des domaines G seuls seront aussi étudiés afin d'élucider les différences entre ces deux domaines et de comprendre leur rôle respectif dans le contexte de la protéine entière. Notamment, l'étude de différents mutants par nos collaborateurs laisse entrevoir une dépendance des domaines G contrairement à ce qui est supposé par leur arrangement structural. Une première mutation dans la boucle P du domaine GD1 (K16A) réduit l'activité de la protéine entière à un niveau

comparable à celle du domaine GD2 isolé, suggérant que l'activité résiduelle mesurée correspond à celle du domaine GD2. Par contre, la protéine entière possédant la même mutation dans la boucle P du domaine GD2 (K118A) montre une activité bien inférieure à celle du domaine GD1 isolé. L'absence d'hydrolyse du domaine GD2 affecte certainement l'hydrolyse par le domaine GD1 d'une manière encore inconnue.

L'identification des partenaires ribosomaux, qu'ils soient protéiques ou nucléiques, et la caractérisation de leurs interactions avec EngA est aussi indispensable. A l'heure actuelle, les données recensées ne sont que préliminaires et sont insuffisantes pour la recherche d'inhibiteurs, notamment pour la mise au point d'un test utilisable en criblage à haut débit. Une première approche sera de travailler avec des ribosomes entiers pour mesurer l'interaction de EngA avec la particule prise dans son ensemble. L'autre approche sera de travailler avec les constituants du ribosome avec lesquels EngA est susceptible d'interagir. Ils sont néanmoins encore à découvrir. Les premières pistes seront les protéines ribosomales qui ont déjà été identifiées par co-purification chez *Escherichia* [365, 368] ou *Salmonella* [369]. Nous avons commencé le travail par la purification de ribosomes entiers issus de *B. subtilis* (par l'équipe de J.M. Jault) et la production recombinante de la protéine structurale S7 (par notre équipe).

Dans le cas où cette approche s'avère fructueuse, la détermination des structures de complexes entre EngA, ses partenaires et les nouveaux ligands moléculaires identifiés pour inhiber leur interaction sera réalisée pour obtenir une image des interactions et apporter une aide précieuse à la conception rationnelle d'inhibiteurs. Ils pourront alors être synthétisés par des méthodes de chimie médicinale (Collaboration avec l'équipe d'Ahcène Boumendjel).

Bibliographie

1. Janeway, C.A., Jr. and R. Medzhitov, *Innate immune recognition*. Annu Rev Immunol, 2002. **20**: p. 197-216.
2. Ferrandon, D., J.L. Imler, and J.A. Hoffmann, *Sensing infection in Drosophila: Toll and beyond*. Semin Immunol, 2004. **16**(1): p. 43-53.
3. Hoffmann, J.A., *The immune response of Drosophila*. Nature, 2003. **426**(6962): p. 33-8.
4. Dziarski, R., *Peptidoglycan recognition proteins (PGRPs)*. Mol Immunol, 2004. **40**(12): p. 877-86.
5. Reiser, J.B., L. Teyton, and I.A. Wilson, *Crystal structure of the Drosophila peptidoglycan recognition protein (PGRP)-SA at 1.56 Å resolution*. J Mol Biol, 2004. **340**(4): p. 909-17.
6. Kim, M.S., M. Byun, and B.H. Oh, *Crystal structure of peptidoglycan recognition protein LB from Drosophila melanogaster*. Nat Immunol, 2003. **4**(8): p. 787-93.
7. Chang, C.I., et al., *Structure of the ectodomain of Drosophila peptidoglycan-recognition protein LCa suggests a molecular mechanism for pattern recognition*. Proc Natl Acad Sci U S A, 2005. **102**(29): p. 10279-84.
8. Chang, C.I., et al., *A Drosophila pattern recognition receptor contains a peptidoglycan docking groove and unusual L,D-carboxypeptidase activity*. PLoS Biol, 2004. **2**(9): p. E277.
9. Guan, R., et al., *Crystal structure of human peptidoglycan recognition protein I alpha bound to a muramyl pentapeptide from Gram-positive bacteria*. Protein Sci, 2006. **15**(5): p. 1199-206.
10. Guan, R., et al., *Crystal structure of the C-terminal peptidoglycan-binding domain of human peptidoglycan recognition protein I alpha*. J Biol Chem, 2004. **279**(30): p. 31873-82.
11. Guan, R., et al., *Structural basis for peptidoglycan binding by peptidoglycan recognition proteins*. Proc Natl Acad Sci U S A, 2004. **101**(49): p. 17168-73.
12. Guan, R., et al., *Crystal structure of a peptidoglycan recognition protein (PGRP) in complex with a muramyl tripeptide from Gram-positive bacteria*. J Endotoxin Res, 2005. **11**(1): p. 41-6.
13. Guan, R., et al., *Crystal structure of human peptidoglycan recognition protein S (PGRP-S) at 1.70 Å resolution*. J Mol Biol, 2005. **347**(4): p. 683-91.
14. Sharma, P., et al., *Multiligand specificity of pathogen-associated molecular pattern-binding site in peptidoglycan recognition protein*. J Biol Chem, 2011. **286**(36): p. 31723-30.
15. Sharma, P., et al., *Crystal structure of the peptidoglycan recognition protein at 1.8 Å resolution reveals dual strategy to combat infection through two independent functional homodimers*. J Mol Biol, 2008. **378**(4): p. 923-32.
16. Weber, A.N., et al., *Binding of the Drosophila cytokine Spatzle to Toll is direct and establishes signaling*. Nat Immunol, 2003. **4**(8): p. 794-800.
17. Hoffmann, A., et al., *Biophysical characterization of refolded Drosophila Spatzle, a cystine knot protein, reveals distinct properties of three isoforms*. J Biol Chem, 2008. **283**(47): p. 32598-609.
18. Nagata, S., *Fas ligand-induced apoptosis*. Annu Rev Genet, 1999. **33**: p. 29-55.

19. Orlinick, J.R., A.K. Vaishnav, and K.B. Elkon, *Structure and function of Fas/Fas ligand*. *Int Rev Immunol*, 1999. **18**(4): p. 293-308.
20. Screaton, G. and X.N. Xu, *T cell life and death signalling via TNF-receptor family members*. *Curr Opin Immunol*, 2000. **12**(3): p. 316-22.
21. Wallach, D., et al., *Tumor necrosis factor receptor and Fas signaling mechanisms*. *Annu Rev Immunol*, 1999. **17**: p. 331-67.
22. Mongkolsapaya, J., et al., *Structure of the TRAIL-DR5 complex reveals mechanisms conferring specificity in apoptotic initiation*. *Nat Struct Biol*, 1999. **6**(11): p. 1048-53.
23. Jones, E.Y., *The tumour necrosis factor receptor family: life or death choices*. *Curr Opin Struct Biol*, 2000. **10**(6): p. 644-8.
24. Jensen, P.E., *Recent advances in antigen processing and presentation*. *Nat Immunol*, 2007. **8**(10): p. 1041-8.
25. Neefjes, J., et al., *Towards a systems understanding of MHC class I and MHC class II antigen presentation*. *Nat Rev Immunol*, 2011. **11**(12): p. 823-36.
26. Sette, A., et al., *The relationship between class I binding affinity and immunogenicity of potential cytotoxic T cell epitopes*. *J Immunol*, 1994. **153**(12): p. 5586-92.
27. van der Burg, S.H., et al., *Immunogenicity of peptides bound to MHC class I molecules depends on the MHC-peptide complex stability*. *J Immunol*, 1996. **156**(9): p. 3308-14.
28. Jones, E.Y., *MHC class I and class II structures*. *Curr Opin Immunol*, 1997. **9**(1): p. 75-9.
29. Madden, D.R., *The three-dimensional structure of peptide-MHC complexes*. *Annu Rev Immunol*, 1995. **13**: p. 587-622.
30. Maenaka, K. and E.Y. Jones, *MHC superfamily structure and the immune system*. *Curr Opin Struct Biol*, 1999. **9**(6): p. 745-53.
31. Fink, P.J. and M.J. Bevan, *Positive selection of thymocytes*. *Adv Immunol*, 1995. **59**: p. 99-133.
32. Nossal, G.J., *Negative selection of lymphocytes*. *Cell*, 1994. **76**(2): p. 229-39.
33. Ignatowicz, L., J. Kappler, and P. Marrack, *The repertoire of T cells shaped by a single MHC/peptide ligand*. *Cell*, 1996. **84**(4): p. 521-9.
34. Tourne, S., et al., *The influence of invariant chain on the positive selection of single T cell receptor specificities*. *Eur J Immunol*, 1995. **25**(7): p. 1851-6.
35. Zerrahn, J., W. Held, and D.H. Raulet, *The MHC reactivity of the T cell repertoire prior to positive and negative selection*. *Cell*, 1997. **88**(5): p. 627-36.
36. Bouneaud, C., P. Kourilsky, and P. Bousso, *Impact of negative selection on the T cell repertoire reactive to a self-peptide: a large fraction of T cell clones escapes clonal deletion*. *Immunity*, 2000. **13**(6): p. 829-40.
37. Arstila, T.P., et al., *A direct estimate of the human alphabeta T cell receptor diversity*. *Science*, 1999. **286**(5441): p. 958-61.
38. Even, J., et al., *T-cell repertoires in healthy and diseased human tissues analysed by T-cell receptor beta-chain CDR3 size determination: evidence for oligoclonal expansions in tumours and inflammatory diseases*. *Res Immunol*, 1995. **146**(2): p. 65-80.
39. Mason, D., *A very high level of crossreactivity is an essential feature of the T-cell receptor*. *Immunol Today*, 1998. **19**(9): p. 395-404.
40. Wilson, D.B., et al., *Specificity and degeneracy of T cells*. *Mol Immunol*, 2004. **40**(14-15): p. 1047-55.

41. Nagy, Z.A., *Alloreactivity: an old puzzle revisited*. Scand J Immunol, 2012. **75**(5): p. 463-70.
42. Blattman, J.N., et al., *Estimating the precursor frequency of naive antigen-specific CD8 T cells*. J Exp Med, 2002. **195**(5): p. 657-64.
43. Kotturi, M.F., et al., *Naive precursor frequencies and MHC binding rather than the degree of epitope diversity shape CD8+ T cell immunodominance*. J Immunol, 2008. **181**(3): p. 2124-33.
44. Legoux, F., et al., *Impact of TCR reactivity and HLA phenotype on naive CD8 T cell frequency in humans*. J Immunol, 2010. **184**(12): p. 6731-8.
45. Moon, J.J., et al., *Tracking epitope-specific T cells*. Nat Protoc, 2009. **4**(4): p. 565-81.
46. Moon, J.J., et al., *Naive CD4(+) T cell frequency varies for different epitopes and predicts repertoire diversity and response magnitude*. Immunity, 2007. **27**(2): p. 203-13.
47. Obar, J.J., K.M. Khanna, and L. Lefrancois, *Endogenous naive CD8+ T cell precursor frequency regulates primary and memory responses to infection*. Immunity, 2008. **28**(6): p. 859-69.
48. Nikolich-Zugich, J., M.K. Slifka, and I. Messaoudi, *The many important facets of T-cell repertoire diversity*. Nat Rev Immunol, 2004. **4**(2): p. 123-32.
49. Venturi, V., et al., *The molecular basis for public T-cell responses?* Nat Rev Immunol, 2008. **8**(3): p. 231-8.
50. Rudolph, M.G., R.L. Stanfield, and I.A. Wilson, *How TCRs bind MHCs, peptides, and coreceptors*. Annu Rev Immunol, 2006. **24**: p. 419-66.
51. Borbulevych, O.Y., et al., *T cell receptor cross-reactivity directed by antigen-dependent tuning of peptide-MHC molecular flexibility*. Immunity, 2009. **31**(6): p. 885-96.
52. Tynan, F.E., et al., *The immunogenicity of a viral cytotoxic T cell epitope is controlled by its MHC-bound conformation*. J Exp Med, 2005. **202**(9): p. 1249-60.
53. Godet, Y., et al., *Frequent occurrence of high affinity T cells against MELOE-1 makes this antigen an attractive target for melanoma immunotherapy*. Eur J Immunol, 2010. **40**(6): p. 1786-94.
54. Scott-Browne, J.P., et al., *Evolutionarily conserved features contribute to alphabeta T cell receptor specificity*. Immunity, 2011. **35**(4): p. 526-35.
55. Yin, L., et al., *A single T cell receptor bound to major histocompatibility complex class I and class II glycoproteins reveals switchable TCR conformers*. Immunity, 2011. **35**(1): p. 23-33.
56. Marrack, P., et al., *T cell receptor specificity for major histocompatibility complex proteins*. Curr Opin Immunol, 2008. **20**(2): p. 203-7.
57. Reiser, J.B., et al., *CDR3 loop flexibility contributes to the degeneracy of TCR recognition*. Nat Immunol, 2003. **4**(3): p. 241-7.
58. Reiser, J.B., et al., *A T cell receptor CDR3beta loop undergoes conformational changes of unprecedented magnitude upon binding to a peptide/MHC class I complex*. Immunity, 2002. **16**(3): p. 345-54.
59. Garcia, K.C., et al., *Structural basis of plasticity in T cell receptor recognition of a self peptide-MHC antigen*. Science, 1998. **279**(5354): p. 1166-72.
60. Ding, Y.H., et al., *Four A6-TCR/peptide/HLA-A2 structures that generate very different T cell signals are nearly identical*. Immunity, 1999. **11**(1): p. 45-56.

61. Kjer-Nielsen, L., et al., *A structural basis for the selection of dominant alphabeta T cell receptors in antiviral immunity*. Immunity, 2003. **18**(1): p. 53-64.
62. Reiser, J.B., et al., *Crystal structure of a T cell receptor bound to an allogeneic MHC molecule*. Nat Immunol, 2000. **1**(4): p. 291-7.
63. Luz, J.G., et al., *Structural comparison of allogeneic and syngeneic T cell receptor-peptide-major histocompatibility complex complexes: a buried alloreactive mutation subtly alters peptide presentation substantially increasing V(beta) Interactions*. J Exp Med, 2002. **195**(9): p. 1175-86.
64. Stewart-Jones, G.B., et al., *A structural basis for immunodominant human T cell receptor recognition*. Nat Immunol, 2003. **4**(7): p. 657-63.
65. Hahn, M., et al., *Unconventional topology of self peptide-major histocompatibility complex binding by a human autoimmune T cell receptor*. Nat Immunol, 2005. **6**(5): p. 490-6.
66. Bridgeman, J.S., et al., *Structural and biophysical determinants of alphabeta T-cell antigen recognition*. Immunology, 2012. **135**(1): p. 9-18.
67. Sykulev, Y., et al., *Kinetics and affinity of reactions between an antigen-specific T cell receptor and peptide-MHC complexes*. Immunity, 1994. **1**(1): p. 15-22.
68. McKeithan, T.W., *Kinetic proofreading in T-cell receptor signal transduction*. Proc Natl Acad Sci U S A, 1995. **92**(11): p. 5042-6.
69. Dustin, M.L. and D. Depoil, *New insights into the T cell synapse from single molecule techniques*. Nat Rev Immunol, 2011. **11**(10): p. 672-84.
70. Rodriguez-Fernandez, J.L., L. Riol-Blanco, and C. Delgado-Martin, *What is an immunological synapse?* Microbes Infect, 2010. **12**(6): p. 438-45.
71. Krummel, M.F. and J.P. Allison, *CD28 and CTLA-4 have opposing effects on the response of T cells to stimulation*. J Exp Med, 1995. **182**(2): p. 459-65.
72. Springer, T.A., *Adhesion receptors of the immune system*. Nature, 1990. **346**(6283): p. 425-34.
73. Reth, M., *Antigen receptor tail clue*. Nature, 1989. **338**(6214): p. 383-4.
74. Nika, K., et al., *Constitutively active Lck kinase in T cells drives antigen receptor signal transduction*. Immunity, 2010. **32**(6): p. 766-77.
75. Deindl, S., et al., *Stability of an autoinhibitory interface in the structure of the tyrosine kinase ZAP-70 impacts T cell receptor response*. Proc Natl Acad Sci U S A, 2009. **106**(49): p. 20699-704.
76. Mingueneau, M., et al., *Loss of the LAT adaptor converts antigen-responsive T cells into pathogenic effectors that function independently of the T cell receptor*. Immunity, 2009. **31**(2): p. 197-208.
77. Smith, A., et al., *A talin-dependent LFA-1 focal zone is formed by rapidly migrating T lymphocytes*. J Cell Biol, 2005. **170**(1): p. 141-51.
78. Nolz, J.C., et al., *WAVE2 regulates high-affinity integrin binding by recruiting vinculin and talin to the immunological synapse*. Mol Cell Biol, 2007. **27**(17): p. 5986-6000.
79. Cannon, J.L. and J.K. Burkhardt, *Differential roles for Wiskott-Aldrich syndrome protein in immune synapse formation and IL-2 production*. J Immunol, 2004. **173**(3): p. 1658-62.
80. Kim, S.T., et al., *The alphabeta T cell receptor is an anisotropic mechanosensor*. J Biol Chem, 2009. **284**(45): p. 31028-37.
81. Goulmy, E., et al., *Mismatches of minor histocompatibility antigens between HLA-identical donors and recipients and the development of graft-versus-host*

- disease after bone marrow transplantation. *N Engl J Med*, 1996. **334**(5): p. 281-5.
82. Goulmy, E., *Minor histocompatibility antigens: allo target molecules for tumor-specific immunotherapy*. *Cancer J*, 2004. **10**(1): p. 1-7.
 83. Fujii, N., et al., *Expression of minor histocompatibility antigen, HA-1, in solid tumor cells*. *Transplantation*, 2002. **73**(7): p. 1137-41.
 84. Turner, S.J., et al., *Lack of prominent peptide-major histocompatibility complex features limits repertoire diversity in virus-specific CD8+ T cell populations*. *Nat Immunol*, 2005. **6**(4): p. 382-9.
 85. Godet, Y., et al., *MELOE-1 is a new antigen overexpressed in melanomas and involved in adoptive T cell transfer efficiency*. *J Exp Med*, 2008. **205**(11): p. 2673-82.
 86. Strasser, A., L. O'Connor, and V.M. Dixit, *Apoptosis signaling*. *Annu Rev Biochem*, 2000. **69**: p. 217-45.
 87. Kinchen, J.M. and K.S. Ravichandran, *Journey to the grave: signaling events regulating removal of apoptotic cells*. *J Cell Sci*, 2007. **120**(Pt 13): p. 2143-9.
 88. Lauber, K., et al., *Clearance of apoptotic cells: getting rid of the corpses*. *Mol Cell*, 2004. **14**(3): p. 277-87.
 89. Ren, Y. and J. Savill, *Apoptosis: the importance of being eaten*. *Cell Death Differ*, 1998. **5**(7): p. 563-8.
 90. Elliott, M.R. and K.S. Ravichandran, *Clearance of apoptotic cells: implications in health and disease*. *J Cell Biol*, 2010. **189**(7): p. 1059-70.
 91. Gaipal, U.S., et al., *Defects in the disposal of dying cells lead to autoimmunity*. *Curr Rheumatol Rep*, 2004. **6**(6): p. 401-7.
 92. Henson, P.M. and R.M. Tuder, *Apoptosis in the lung: induction, clearance and detection*. *Am J Physiol Lung Cell Mol Physiol*, 2008. **294**(4): p. L601-11.
 93. Schrijvers, D.M., et al., *Phagocytosis of apoptotic cells by macrophages is impaired in atherosclerosis*. *Arterioscler Thromb Vasc Biol*, 2005. **25**(6): p. 1256-61.
 94. Mochizuki, H., et al., *Histochemical detection of apoptosis in Parkinson's disease*. *J Neurol Sci*, 1996. **137**(2): p. 120-3.
 95. Su, J.H., et al., *Immunohistochemical evidence for apoptosis in Alzheimer's disease*. *Neuroreport*, 1994. **5**(18): p. 2529-33.
 96. Solinas, G., et al., *Tumor-associated macrophages (TAM) as major players of the cancer-related inflammation*. *J Leukoc Biol*, 2009. **86**(5): p. 1065-73.
 97. Handa, Y., et al., *Shigella IpgB1 promotes bacterial entry through the ELMO-Dock180 machinery*. *Nat Cell Biol*, 2007. **9**(1): p. 121-8.
 98. Janardhan, A., et al., *HIV-1 Nef binds the DOCK2-ELMO1 complex to activate rac and inhibit lymphocyte chemotaxis*. *PLoS Biol*, 2004. **2**(1): p. E6.
 99. Fadok, V.A., et al., *Exposure of phosphatidylserine on the surface of apoptotic lymphocytes triggers specific recognition and removal by macrophages*. *J Immunol*, 1992. **148**(7): p. 2207-16.
 100. Ravichandran, K.S., *Find-me and eat-me signals in apoptotic cell clearance: progress and conundrums*. *J Exp Med*, 2010. **207**(9): p. 1807-17.
 101. Wong, K., et al., *Phosphatidylserine receptor Tim-4 is essential for the maintenance of the homeostatic state of resident peritoneal macrophages*. *Proc Natl Acad Sci U S A*, 2010. **107**(19): p. 8712-7.
 102. Park, D., et al., *BAI1 is an engulfment receptor for apoptotic cells upstream of the ELMO/Dock180/Rac module*. *Nature*, 2007. **450**(7168): p. 430-4.

103. Park, S.Y., et al., *Rapid cell corpse clearance by stabilin-2, a membrane phosphatidylserine receptor*. Cell Death Differ, 2008. **15**(1): p. 192-201.
104. Arur, S., et al., *Annexin I is an endogenous ligand that mediates apoptotic cell engulfment*. Dev Cell, 2003. **4**(4): p. 587-98.
105. Hanayama, R., et al., *Identification of a factor that links apoptotic cells to phagocytes*. Nature, 2002. **417**(6885): p. 182-7.
106. Lemke, G. and C.V. Rothlin, *Immunobiology of the TAM receptors*. Nat Rev Immunol, 2008. **8**(5): p. 327-36.
107. Balasubramanian, K. and A.J. Schroit, *Characterization of phosphatidylserine-dependent beta2-glycoprotein I macrophage interactions. Implications for apoptotic cell clearance by phagocytes*. J Biol Chem, 1998. **273**(44): p. 29272-7.
108. Paidassi, H., et al., *C1q binds phosphatidylserine and likely acts as a multiligand-bridging molecule in apoptotic cell recognition*. J Immunol, 2008. **180**(4): p. 2329-38.
109. Akakura, S., et al., *The opsonin MFG-E8 is a ligand for the alphavbeta5 integrin and triggers DOCK180-dependent Rac1 activation for the phagocytosis of apoptotic cells*. Exp Cell Res, 2004. **292**(2): p. 403-16.
110. Maiti, S.N., et al., *Beta-2-glycoprotein 1-dependent macrophage uptake of apoptotic cells. Binding to lipoprotein receptor-related protein receptor family members*. J Biol Chem, 2008. **283**(7): p. 3761-6.
111. Paidassi, H., et al., *Investigations on the C1q-calreticulin-phosphatidylserine interactions yield new insights into apoptotic cell recognition*. J Mol Biol, 2011. **408**(2): p. 277-90.
112. Duus, K., et al., *Direct interaction between CD91 and C1q*. FEBS J, 2010. **277**(17): p. 3526-37.
113. Stern, M., J. Savill, and C. Haslett, *Human monocyte-derived macrophage phagocytosis of senescent eosinophils undergoing apoptosis. Mediation by alpha v beta 3/CD36/thrombospondin recognition mechanism and lack of phlogistic response*. Am J Pathol, 1996. **149**(3): p. 911-21.
114. Takizawa, F., S. Tsuji, and S. Nagasawa, *Enhancement of macrophage phagocytosis upon iC3b deposition on apoptotic cells*. FEBS Lett, 1996. **397**(2-3): p. 269-72.
115. Vandivier, R.W., et al., *Role of surfactant proteins A, D, and C1q in the clearance of apoptotic cells in vivo and in vitro: calreticulin and CD91 as a common collectin receptor complex*. J Immunol, 2002. **169**(7): p. 3978-86.
116. Platt, N., et al., *Role for the class A macrophage scavenger receptor in the phagocytosis of apoptotic thymocytes in vitro*. Proc Natl Acad Sci U S A, 1996. **93**(22): p. 12456-60.
117. Oka, K., et al., *Lectin-like oxidized low-density lipoprotein receptor 1 mediates phagocytosis of aged/apoptotic cells in endothelial cells*. Proc Natl Acad Sci U S A, 1998. **95**(16): p. 9535-40.
118. Erdosova, B., et al., *Part of CD68+ macrophages in the clearance of apoptotic bodies in human metanephros*. Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub, 2002. **146**(2): p. 41-5.
119. Savill, J., et al., *Thrombospondin cooperates with CD36 and the vitronectin receptor in macrophage recognition of neutrophils undergoing apoptosis*. J Clin Invest, 1992. **90**(4): p. 1513-22.
120. Dunkelberger, J.R. and W.C. Song, *Complement and its role in innate and adaptive immune responses*. Cell Res, 2010. **20**(1): p. 34-50.

121. Fadok, V.A., et al., *Macrophages that have ingested apoptotic cells in vitro inhibit proinflammatory cytokine production through autocrine/paracrine mechanisms involving TGF-beta, PGE2, and PAF.* J Clin Invest, 1998. **101**(4): p. 890-8.
122. Voll, R.E., et al., *Immunosuppressive effects of apoptotic cells.* Nature, 1997. **390**(6658): p. 350-1.
123. Hoffmann, P.R., et al., *Phosphatidylserine (PS) induces PS receptor-mediated macropinocytosis and promotes clearance of apoptotic cells.* J Cell Biol, 2001. **155**(4): p. 649-59.
124. Henson, P.M. and D.A. Hume, *Apoptotic cell removal in development and tissue homeostasis.* Trends Immunol, 2006. **27**(5): p. 244-50.
125. Gumienny, T.L. and M.O. Hengartner, *How the worm removes corpses: the nematode C. elegans as a model system to study engulfment.* Cell Death Differ, 2001. **8**(6): p. 564-8.
126. Marguet, D., et al., *Engulfment of apoptotic cells involves the redistribution of membrane phosphatidylserine on phagocyte and prey.* Nat Cell Biol, 1999. **1**(7): p. 454-6.
127. Gardai, S.J., et al., *Cell-surface calreticulin initiates clearance of viable or apoptotic cells through trans-activation of LRP on the phagocyte.* Cell, 2005. **123**(2): p. 321-34.
128. Park, S.Y., et al., *Epidermal growth factor-like domain repeat of stabilin-2 recognizes phosphatidylserine during cell corpse clearance.* Mol Cell Biol, 2008. **28**(17): p. 5288-98.
129. Osada, Y., et al., *Signalling pathway involving GULP, MAPK and Rac1 for SR-BI-induced phagocytosis of apoptotic cells.* J Biochem, 2009. **145**(3): p. 387-94.
130. Wu, Y.C. and H.R. Horvitz, *The C. elegans cell corpse engulfment gene ced-7 encodes a protein similar to ABC transporters.* Cell, 1998. **93**(6): p. 951-60.
131. Su, H.P., et al., *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, 2002. **277**(14): p. 11772-9.
132. Park, S.Y., et al., *Requirement of adaptor protein GULP during stabilin-2-mediated cell corpse engulfment.* J Biol Chem, 2008. **283**(16): p. 10593-600.
133. Park, S.Y., et al., *Adaptor protein GULP is involved in stabilin-1-mediated phagocytosis.* Biochem Biophys Res Commun, 2010. **398**(3): p. 467-72.
134. Landry, Y.D., et al., *ATP-binding cassette transporter A1 expression disrupts raft membrane microdomains through its ATPase-related functions.* J Biol Chem, 2006. **281**(47): p. 36091-101.
135. Ma, Z., et al., *Regulation of Arf6 and ACAP1 signaling by the PTB-domain-containing adaptor protein GULP.* Curr Biol, 2007. **17**(8): p. 722-7.
136. Cote, J.F. and K. Vuori, *GEF what? Dock180 and related proteins help Rac to polarize cells in new ways.* Trends Cell Biol, 2007. **17**(8): p. 383-93.
137. Reddien, P.W. and H.R. Horvitz, *The engulfment process of programmed cell death in caenorhabditis elegans.* Annu Rev Cell Dev Biol, 2004. **20**: p. 193-221.
138. Albert, M.L., J.I. Kim, and R.B. Birge, *alpha5beta5 integrin recruits the CrkII-Dock180-rac1 complex for phagocytosis of apoptotic cells.* Nat Cell Biol, 2000. **2**(12): p. 899-905.
139. Brugnera, E., et al., *Unconventional Rac-GEF activity is mediated through the Dock180-ELMO complex.* Nat Cell Biol, 2002. **4**(8): p. 574-82.

140. Wang, X., et al., *Cell corpse engulfment mediated by C. elegans phosphatidylserine receptor through CED-5 and CED-12*. Science, 2003. **302**(5650): p. 1563-6.
141. Wu, Y., et al., *A role for Mer tyrosine kinase in alphavbeta5 integrin-mediated phagocytosis of apoptotic cells*. J Cell Sci, 2005. **118**(Pt 3): p. 539-53.
142. Cote, J.F., et al., *A novel and evolutionarily conserved PtdIns(3,4,5)P3-binding domain is necessary for DOCK180 signalling*. Nat Cell Biol, 2005. **7**(8): p. 797-807.
143. Tosello-Trampont, A.C., et al., *Identification of two signaling submodules within the CrkII/ELMO/Dock180 pathway regulating engulfment of apoptotic cells*. Cell Death Differ, 2007. **14**(5): p. 963-72.
144. Henson, P.M., *Engulfment: ingestion and migration with Rac, Rho and TRIO*. Curr Biol, 2005. **15**(1): p. R29-30.
145. Katoh, H. and M. Negishi, *RhoG activates Rac1 by direct interaction with the Dock180-binding protein Elmo*. Nature, 2003. **424**(6947): p. 461-4.
146. deBakker, C.D., et al., *Phagocytosis of apoptotic cells is regulated by a UNC-73/TRIO-MIG-2/RhoG signaling module and armadillo repeats of CED-12/ELMO*. Curr Biol, 2004. **14**(24): p. 2208-16.
147. Yokoyama, N., et al., *Identification of tyrosine residues on ELMO1 that are phosphorylated by the Src-family kinase Hck*. Biochemistry, 2005. **44**(24): p. 8841-9.
148. Guiet, R., et al., *Hematopoietic cell kinase (Hck) isoforms and phagocyte duties - from signaling and actin reorganization to migration and phagocytosis*. Eur J Cell Biol, 2008. **87**(8-9): p. 527-42.
149. Santy, L.C., K.S. Ravichandran, and J.E. Casanova, *The DOCK180/Elmo complex couples ARNO-mediated Arf6 activation to the downstream activation of Rac1*. Curr Biol, 2005. **15**(19): p. 1749-54.
150. Chen, E.H., et al., *Control of myoblast fusion by a guanine nucleotide exchange factor, loner, and its effector ARF6*. Cell, 2003. **114**(6): p. 751-62.
151. Kim, S., et al., *Cross Talk between Engulfment Receptors Stabilin-2 and Integrin alphavbeta5 Orchestrates Engulfment of Phosphatidylserine-Exposed Erythrocytes*. Mol Cell Biol, 2012. **32**(14): p. 2698-708.
152. Aznar, S., et al., *Rho GTPases: potential candidates for anticancer therapy*. Cancer Lett, 2004. **206**(2): p. 181-91.
153. Etienne-Manneville, S. and A. Hall, *Rho GTPases in cell biology*. Nature, 2002. **420**(6916): p. 629-35.
154. Erickson, J.W. and R.A. Cerione, *Structural elements, mechanism, and evolutionary convergence of Rho protein-guanine nucleotide exchange factor complexes*. Biochemistry, 2004. **43**(4): p. 837-42.
155. Rossman, K.L., C.J. Der, and J. Sondek, *GEF means go: turning on RHO GTPases with guanine nucleotide-exchange factors*. Nat Rev Mol Cell Biol, 2005. **6**(2): p. 167-80.
156. Cote, J.F. and K. Vuori, *Identification of an evolutionarily conserved superfamily of DOCK180-related proteins with guanine nucleotide exchange activity*. J Cell Sci, 2002. **115**(Pt 24): p. 4901-13.
157. Reif, K. and J. Cyster, *The CDM protein DOCK2 in lymphocyte migration*. Trends Cell Biol, 2002. **12**(8): p. 368-73.
158. Meller, N., S. Merlot, and C. Guda, *CZH proteins: a new family of Rho-GEFs*. J Cell Sci, 2005. **118**(Pt 21): p. 4937-46.

159. Grimsley, C.M., et al., *Dock180 and ELMO1 proteins cooperate to promote evolutionarily conserved Rac-dependent cell migration*. J Biol Chem, 2004. **279**(7): p. 6087-97.
160. Lu, M. and K.S. Ravichandran, *Dock180-ELMO cooperation in Rac activation*. Methods Enzymol, 2006. **406**: p. 388-402.
161. Fukui, Y., et al., *Haematopoietic cell-specific CDM family protein DOCK2 is essential for lymphocyte migration*. Nature, 2001. **412**(6849): p. 826-31.
162. Sanui, T., et al., *DOCK2 regulates Rac activation and cytoskeletal reorganization through interaction with ELMO1*. Blood, 2003. **102**(8): p. 2948-50.
163. Nishihara, H., et al., *DOCK2 mediates T cell receptor-induced activation of Rac2 and IL-2 transcription*. Biochem Biophys Res Commun, 2002. **296**(3): p. 716-20.
164. Sanui, T., et al., *DOCK2 is essential for antigen-induced translocation of TCR and lipid rafts, but not PKC-theta and LFA-1, in T cells*. Immunity, 2003. **19**(1): p. 119-29.
165. Sanders, M.A., D. Ampasala, and M.D. Basson, *DOCK5 and DOCK1 regulate Caco-2 intestinal epithelial cell spreading and migration on collagen IV*. J Biol Chem, 2009. **284**(1): p. 27-35.
166. Patel, M., A. Pelletier, and J.F. Cote, *Opening up on ELMO regulation: New insights into the control of Rac signaling by the DOCK180/ELMO complex*. Small Gtpases, 2011. **2**(5): p. 268-275.
167. Namekata, K., et al., *MOCA induces membrane spreading by activating Rac1*. J Biol Chem, 2004. **279**(14): p. 14331-7.
168. Namekata, K., et al., *Dock3 regulates BDNF-TrkB signaling for neurite outgrowth by forming a ternary complex with Elmo and RhoG*. Genes Cells, 2012.
169. Yajnik, V., et al., *DOCK4, a GTPase activator, is disrupted during tumorigenesis*. Cell, 2003. **112**(5): p. 673-84.
170. Ueda, S., et al., *Dock4 regulates dendritic development in hippocampal neurons*. J Neurosci Res, 2008. **86**(14): p. 3052-61.
171. Miyamoto, Y., et al., *Dock6, a Dock-C subfamily guanine nucleotide exchanger, has the dual specificity for Rac1 and Cdc42 and regulates neurite outgrowth*. Exp Cell Res, 2007. **313**(4): p. 791-804.
172. Watabe-Uchida, M., et al., *The Rac activator DOCK7 regulates neuronal polarity through local phosphorylation of stathmin/Op18*. Neuron, 2006. **51**(6): p. 727-39.
173. Majewski, L., et al., *Dock7: A GEF for Rho-family GTPases and a novel myosin VI-binding partner in neuronal PC12 cells*. Biochem Cell Biol, 2012. **90**(4): p. 565-74.
174. Su, H.C., H. Jing, and Q. Zhang, *DOCK8 deficiency*. Ann N Y Acad Sci, 2011. **1246**: p. 26-33.
175. Jabara, H.H., et al., *DOCK8 functions as an adaptor that links TLR-MyD88 signaling to B cell activation*. Nat Immunol, 2012. **13**(6): p. 612-20.
176. Randall, K.L., et al., *DOCK8 deficiency impairs CD8 T cell survival and function in humans and mice*. J Exp Med, 2011. **208**(11): p. 2305-20.
177. Harada, Y., et al., *DOCK8 is a Cdc42 activator critical for interstitial dendritic cell migration during immune responses*. Blood, 2012. **119**(19): p. 4451-61.
178. Meller, N., et al., *Function of the N-terminus of zizimin1: autoinhibition and membrane targeting*. Biochem J, 2008. **409**(2): p. 525-33.

179. Meller, N., et al., *The novel Cdc42 guanine nucleotide exchange factor, zizimin1, dimerizes via the Cdc42-binding CZH2 domain.* J Biol Chem, 2004. **279**(36): p. 37470-6.
180. Yelo, E., et al., *Dock10, a novel CZH protein selectively induced by interleukin-4 in human B lymphocytes.* Mol Immunol, 2008. **45**(12): p. 3411-8.
181. Nishikimi, A., et al., *Zizimin2: a novel, DOCK180-related Cdc42 guanine nucleotide exchange factor expressed predominantly in lymphocytes.* FEBS Lett, 2005. **579**(5): p. 1039-46.
182. Premkumar, L., et al., *Structural basis of membrane targeting by the Dock180 family of Rho family guanine exchange factors (Rho-GEFs).* J Biol Chem, 2010. **285**(17): p. 13211-22.
183. Hiramoto, K., M. Negishi, and H. Katoh, *Dock4 is regulated by RhoG and promotes Rac-dependent cell migration.* Exp Cell Res, 2006. **312**(20): p. 4205-16.
184. Lin, Q., et al., *Identification of a DOCK180-related guanine nucleotide exchange factor that is capable of mediating a positive feedback activation of Cdc42.* J Biol Chem, 2006. **281**(46): p. 35253-62.
185. Cote, J.F. and K. Vuori, *In vitro guanine nucleotide exchange activity of DHR-2/DOCKER/CZH2 domains.* Methods Enzymol, 2006. **406**: p. 41-57.
186. Kiyokawa, E., et al., *Activation of Rac1 by a Crk SH3-binding protein, DOCK180.* Genes Dev, 1998. **12**(21): p. 3331-6.
187. Lu, M., et al., *PH domain of ELMO functions in trans to regulate Rac activation via Dock180.* Nat Struct Mol Biol, 2004. **11**(8): p. 756-62.
188. Yang, J., et al., *Activation of Rho GTPases by DOCK exchange factors is mediated by a nucleotide sensor.* Science, 2009. **325**(5946): p. 1398-402.
189. Kulkarni, K., et al., *Multiple factors confer specific Cdc42 and Rac protein activation by dedicator of cytokinesis (DOCK) nucleotide exchange factors.* J Biol Chem, 2011. **286**(28): p. 25341-51.
190. Reebye, V., et al., *A perspective on non-catalytic Src homology (SH) adaptor signalling proteins.* Cell Signal, 2012. **24**(2): p. 388-92.
191. Gumienny, T.L., et al., *CED-12/ELMO, a novel member of the CrkII/Dock180/Rac pathway, is required for phagocytosis and cell migration.* Cell, 2001. **107**(1): p. 27-41.
192. Zhou, Z., et al., *The C. elegans PH domain protein CED-12 regulates cytoskeletal reorganization via a Rho/Rac GTPase signaling pathway.* Dev Cell, 2001. **1**(4): p. 477-89.
193. Komander, D., et al., *An alpha-helical extension of the ELMO1 pleckstrin homology domain mediates direct interaction to DOCK180 and is critical in Rac signaling.* Mol Biol Cell, 2008. **19**(11): p. 4837-51.
194. Lu, M., et al., *A Steric-inhibition model for regulation of nucleotide exchange via the Dock180 family of GEFs.* Curr Biol, 2005. **15**(4): p. 371-7.
195. Matsuda, M., et al., *Interaction between the amino-terminal SH3 domain of CRK and its natural target proteins.* J Biol Chem, 1996. **271**(24): p. 14468-72.
196. Biersmith, B., et al., *The DOCK protein sponge binds to ELMO and functions in Drosophila embryonic CNS development.* PLoS One, 2011. **6**(1): p. e16120.
197. Scheffzek, K. and S. Welti, *Pleckstrin Homology (PH) like domains - versatile modules in protein-protein interaction platforms.* FEBS Lett, 2012.
198. Bowzard, J.B., et al., *ELMOD2 is an Arl2 GTPase-activating protein that also acts on Arfs.* J Biol Chem, 2007. **282**(24): p. 17568-80.

199. Grimsley, C.M., et al., *Characterization of a novel interaction between ELMO1 and ERM proteins*. J Biol Chem, 2006. **281**(9): p. 5928-37.
200. Schonichen, A. and M. Geyer, *Fifteen formins for an actin filament: a molecular view on the regulation of human formins*. Biochim Biophys Acta, 2010. **1803**(2): p. 152-63.
201. Patel, M., et al., *An evolutionarily conserved autoinhibitory molecular switch in ELMO proteins regulates Rac signaling*. Curr Biol, 2010. **20**(22): p. 2021-7.
202. Hanawa-Suetsugu, K., et al., *Structural basis for mutual relief of the Rac guanine nucleotide exchange factor DOCK2 and its partner ELMO1 from their autoinhibited forms*. Proc Natl Acad Sci U S A, 2012. **109**(9): p. 3305-10.
203. Katoh, H., K. Hiramoto, and M. Negishi, *Activation of Rac1 by RhoG regulates cell migration*. J Cell Sci, 2006. **119**(Pt 1): p. 56-65.
204. Vigorito, E., et al., *Immunological function in mice lacking the Rac-related GTPase RhoG*. Mol Cell Biol, 2004. **24**(2): p. 719-29.
205. Makino, Y., et al., *Elmo1 inhibits ubiquitylation of Dock180*. J Cell Sci, 2006. **119**(Pt 5): p. 923-32.
206. Boquet, P., *Small GTP binding proteins and bacterial virulence*. Microbes Infect, 2000. **2**(7): p. 837-43.
207. Kahn, R.A., H. Fu, and C.R. Roy, *Cellular hijacking: a common strategy for microbial infection*. Trends Biochem Sci, 2002. **27**(6): p. 308-14.
208. Alto, N.M., *Mimicking small G-proteins: an emerging theme from the bacterial virulence arsenal*. Cell Microbiol, 2008. **10**(3): p. 566-75.
209. Alto, N.M., et al., *Identification of a bacterial type III effector family with G protein mimicry functions*. Cell, 2006. **124**(1): p. 133-45.
210. Huang, Z., et al., *Structural insights into host GTPase isoform selection by a family of bacterial GEF mimics*. Nat Struct Mol Biol, 2009. **16**(8): p. 853-60.
211. Klink, B.U., et al., *Structure of Shigella IpgB2 in complex with human RhoA: implications for the mechanism of bacterial guanine nucleotide exchange factor mimicry*. J Biol Chem, 2010. **285**(22): p. 17197-208.
212. Foster, J.L. and J.V. Garcia, *HIV-1 Nef: at the crossroads*. Retrovirology, 2008. **5**: p. 84.
213. Foster, J.L., et al., *Mechanisms of HIV-1 Nef function and intracellular signaling*. J Neuroimmune Pharmacol, 2011. **6**(2): p. 230-46.
214. Geyer, M., O.T. Fackler, and B.M. Peterlin, *Structure--function relationships in HIV-1 Nef*. EMBO Rep, 2001. **2**(7): p. 580-5.
215. Breuer, S., et al., *A cleavage enzyme-cytometric bead array provides biochemical profiling of resistance mutations in HIV-1 Gag and protease*. Biochemistry, 2011. **50**(20): p. 4371-81.
216. Horenkamp, F.A., et al., *Conformation of the dileucine-based sorting motif in HIV-1 Nef revealed by intermolecular domain assembly*. Traffic, 2011. **12**(7): p. 867-77.
217. Lee, C.H., et al., *Crystal structure of the conserved core of HIV-1 Nef complexed with a Src family SH3 domain*. Cell, 1996. **85**(6): p. 931-42.
218. Graves, J.D. and E.G. Krebs, *Protein phosphorylation and signal transduction*. Pharmacol Ther, 1999. **82**(2-3): p. 111-21.
219. Manning, G., et al., *The protein kinase complement of the human genome*. Science, 2002. **298**(5600): p. 1912-34.
220. Cohen, P., *The origins of protein phosphorylation*. Nat Cell Biol, 2002. **4**(5): p. E127-30.

221. Adams, J.A., *Kinetic and catalytic mechanisms of protein kinases*. Chem Rev, 2001. **101**(8): p. 2271-90.
222. Burnett, G. and E.P. Kennedy, *The enzymatic phosphorylation of proteins*. J Biol Chem, 1954. **211**(2): p. 969-80.
223. Bennett, B.L., *c-Jun N-terminal kinase-dependent mechanisms in respiratory disease*. Eur Respir J, 2006. **28**(3): p. 651-61.
224. Dhillon, A.S., et al., *MAP kinase signalling pathways in cancer*. Oncogene, 2007. **26**(22): p. 3279-90.
225. Lugo, T.G., et al., *Tyrosine kinase activity and transformation potency of bcr-abl oncogene products*. Science, 1990. **247**(4946): p. 1079-82.
226. Santarius, T., et al., *A census of amplified and overexpressed human cancer genes*. Nat Rev Cancer, 2010. **10**(1): p. 59-64.
227. Wang, Y., *Mitogen-activated protein kinases in heart development and diseases*. Circulation, 2007. **116**(12): p. 1413-23.
228. Pandya, N., D. Santani, and S. Jain, *Role of mitogen-activated protein (MAP) kinases in cardiovascular diseases*. Cardiovasc Drug Rev, 2005. **23**(3): p. 247-54.
229. Greenman, C., et al., *Patterns of somatic mutation in human cancer genomes*. Nature, 2007. **446**(7132): p. 153-8.
230. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer*. Cell, 2000. **100**(1): p. 57-70.
231. Cohen, P., *Protein kinases--the major drug targets of the twenty-first century?* Nat Rev Drug Discov, 2002. **1**(4): p. 309-15.
232. Zhang, J., P.L. Yang, and N.S. Gray, *Targeting cancer with small molecule kinase inhibitors*. Nat Rev Cancer, 2009. **9**(1): p. 28-39.
233. Knight, Z.A. and K.M. Shokat, *Features of selective kinase inhibitors*. Chem Biol, 2005. **12**(6): p. 621-37.
234. Karaman, M.W., et al., *A quantitative analysis of kinase inhibitor selectivity*. Nat Biotechnol, 2008. **26**(1): p. 127-32.
235. Knighton, D.R., et al., *Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase*. Science, 1991. **253**(5018): p. 407-14.
236. Liao, J.J., *Molecular recognition of protein kinase binding pockets for design of potent and selective kinase inhibitors*. J Med Chem, 2007. **50**(3): p. 409-24.
237. Zuccotto, F., et al., *Through the "gatekeeper door": exploiting the active kinase conformation*. J Med Chem, 2010. **53**(7): p. 2681-94.
238. Gajiwala, K.S., et al., *KIT kinase mutants show unique mechanisms of drug resistance to imatinib and sunitinib in gastrointestinal stromal tumor patients*. Proc Natl Acad Sci U S A, 2009. **106**(5): p. 1542-7.
239. Wood, E.R., et al., *A unique structure for epidermal growth factor receptor bound to GW572016 (Lapatinib): relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells*. Cancer Res, 2004. **64**(18): p. 6652-9.
240. Nagar, B., et al., *Crystal structures of the kinase domain of c-Abl in complex with the small molecule inhibitors PD173955 and imatinib (STI-571)*. Cancer Res, 2002. **62**(15): p. 4236-43.
241. Filhol, O., et al., *Live-cell fluorescence imaging reveals the dynamics of protein kinase CK2 individual subunits*. Mol Cell Biol, 2003. **23**(3): p. 975-87.

242. Niefind, K., et al., *Crystal structure of human protein kinase CK2: insights into basic properties of the CK2 holoenzyme*. EMBO J, 2001. **20**(19): p. 5320-31.
243. Ahmad, K.A., et al., *Protein kinase CK2--a key suppressor of apoptosis*. Adv Enzyme Regul, 2008. **48**: p. 179-87.
244. Gyenis, L. and D.W. Litchfield, *The emerging CK2 interactome: insights into the regulation and functions of CK2*. Mol Cell Biochem, 2008. **316**(1-2): p. 5-14.
245. Meggio, F. and L.A. Pinna, *One-thousand-and-one substrates of protein kinase CK2? Faseb J*, 2003. **17**(3): p. 349-68.
246. Olsten, M.E., J.E. Weber, and D.W. Litchfield, *CK2 interacting proteins: emerging paradigms for CK2 regulation? Mol Cell Biochem*, 2005. **274**(1-2): p. 115-24.
247. Salvi, M., et al., *Extraordinary pleiotropy of protein kinase CK2 revealed by weblogo phosphoproteome analysis*. Biochim Biophys Acta, 2009. **1793**(5): p. 847-59.
248. Litchfield, D.W., *Protein kinase CK2: structure, regulation and role in cellular decisions of life and death*. Biochem J, 2003. **369**(Pt 1): p. 1-15.
249. St-Denis, N.A. and D.W. Litchfield, *Protein kinase CK2 in health and disease: From birth to death: the role of protein kinase CK2 in the regulation of cell proliferation and survival*. Cell Mol Life Sci, 2009. **66**(11-12): p. 1817-29.
250. Battistutta, R., *Protein kinase CK2 in health and disease: Structural bases of protein kinase CK2 inhibition*. Cell Mol Life Sci, 2009. **66**(11-12): p. 1868-89.
251. Chardot, T., H. Shen, and J.C. Meunier, *Dual specificity of casein kinase II from the yeast Yarrowia lipolytica*. C R Acad Sci III, 1995. **318**(9): p. 937-42.
252. Donella-Deana, A., et al., *Autocatalytic tyrosine-phosphorylation of protein kinase CK2 alpha and alpha' subunits: implication of Tyr182*. Biochem J, 2001. **357**(Pt 2): p. 563-7.
253. Marin, O., et al., *Tyrosine versus serine/threonine phosphorylation by protein kinase casein kinase-2. A study with peptide substrates derived from immunophilin Fpr3*. J Biol Chem, 1999. **274**(41): p. 29260-5.
254. Vilk, G., et al., *Protein kinase CK2 catalyzes tyrosine phosphorylation in mammalian cells*. Cell Signal, 2008. **20**(11): p. 1942-51.
255. Filhol, O., J.L. Martiel, and C. Cochet, *Protein kinase CK2: a new view of an old molecular complex*. EMBO Rep, 2004. **5**(4): p. 351-5.
256. Buchou, T. and C. Cochet, *[Protein kinase CK2: an enzyme that likes to be different]*. Med Sci (Paris), 2003. **19**(6-7): p. 709-16.
257. Duncan, J.S. and D.W. Litchfield, *Too much of a good thing: the role of protein kinase CK2 in tumorigenesis and prospects for therapeutic inhibition of CK2*. Biochim Biophys Acta, 2008. **1784**(1): p. 33-47.
258. Trembley, J.H., et al., *Protein kinase CK2 in health and disease: CK2: a key player in cancer biology*. Cell Mol Life Sci, 2009. **66**(11-12): p. 1858-67.
259. Battistutta, R., et al., *Unprecedented selectivity and structural determinants of a new class of protein kinase CK2 inhibitors in clinical trials for the treatment of cancer*. Biochemistry, 2011. **50**(39): p. 8478-88.
260. Guerra, B., et al., *Expression, purification and crystallization of the catalytic subunit of protein kinase CK2 from Zea mays*. Acta Crystallogr D Biol Crystallogr, 1998. **54**(Pt 1): p. 143-5.
261. Niefind, K., et al., *Crystal structure of the catalytic subunit of protein kinase CK2 from Zea mays at 2.1 Å resolution*. EMBO J, 1998. **17**(9): p. 2451-62.
262. Niefind, K., et al., *GTP plus water mimic ATP in the active site of protein kinase CK2*. Nat Struct Biol, 1999. **6**(12): p. 1100-3.

263. Blencke, S., et al., *Characterization of a conserved structural determinant controlling protein kinase sensitivity to selective inhibitors*. Chem Biol, 2004. **11**(5): p. 691-701.
264. Ermakova, I., et al., *Crystal structure of a C-terminal deletion mutant of human protein kinase CK2 catalytic subunit*. J Mol Biol, 2003. **330**(5): p. 925-34.
265. Niefind, K., et al., *Evolved to be active: sulfate ions define substrate recognition sites of CK2alpha and emphasise its exceptional role within the CMGC family of eukaryotic protein kinases*. J Mol Biol, 2007. **370**(3): p. 427-38.
266. Raaf, J., et al., *The interaction of CK2alpha and CK2beta, the subunits of protein kinase CK2, requires CK2beta in a preformed conformation and is enthalpically driven*. Protein Sci, 2008. **17**(12): p. 2180-6.
267. Raaf, J., et al., *The CK2 alpha/CK2 beta interface of human protein kinase CK2 harbors a binding pocket for small molecules*. Chem Biol, 2008. **15**(2): p. 111-7.
268. Chantalat, L., et al., *Crystal structure of the human protein kinase CK2 regulatory subunit reveals its zinc finger-mediated dimerization*. EMBO J, 1999. **18**(11): p. 2930-40.
269. Valero, E., et al., *Quaternary structure of casein kinase 2. Characterization of multiple oligomeric states and relation with its catalytic activity*. J Biol Chem, 1995. **270**(14): p. 8345-52.
270. De Moliner, E., et al., *Inhibition of protein kinase CK2 by anthraquinone-related compounds. A structural insight*. J Biol Chem, 2003. **278**(3): p. 1831-6.
271. Yim, H., et al., *Emodin, an anthraquinone derivative isolated from the rhizomes of Rheum palmatum, selectively inhibits the activity of casein kinase II as a competitive inhibitor*. Planta Med, 1999. **65**(1): p. 9-13.
272. Sarno, S., et al., *Biochemical and three-dimensional-structural study of the specific inhibition of protein kinase CK2 by [5-oxo-5,6-dihydroindolo-(1,2-a)quinazolin-7-yl]acetic acid (IQA)*. Biochem J, 2003. **374**(Pt 3): p. 639-46.
273. Vangrevelinghe, E., et al., *Discovery of a potent and selective protein kinase CK2 inhibitor by high-throughput docking*. J Med Chem, 2003. **46**(13): p. 2656-62.
274. Battistutta, R., et al., *Structural features underlying selective inhibition of protein kinase CK2 by ATP site-directed tetrabromo-2-benzotriazole*. Protein Sci, 2001. **10**(11): p. 2200-6.
275. Battistutta, R., et al., *Inspecting the structure-activity relationship of protein kinase CK2 inhibitors derived from tetrabromo-benzimidazole*. Chem Biol, 2005. **12**(11): p. 1211-9.
276. Duncan, J.S., et al., *An unbiased evaluation of CK2 inhibitors by chemoproteomics: characterization of inhibitor effects on CK2 and identification of novel inhibitor targets*. Mol Cell Proteomics, 2008. **7**(6): p. 1077-88.
277. Golub, A.G., et al., *Evaluation of 4,5,6,7-tetrahalogeno-1H-isoindole-1,3(2H)-diones as inhibitors of human protein kinase CK2*. Biochim Biophys Acta, 2008. **1784**(1): p. 143-9.
278. Pagano, M.A., et al., *Tetrabromocinnamic acid (TBCA) and related compounds represent a new class of specific protein kinase CK2 inhibitors*. Chembiochem, 2007. **8**(1): p. 129-39.
279. Raaf, J., O.G. Issinger, and K. Niefind, *Insights from soft X-rays: the chlorine and sulfur sub-structures of a CK2alpha/DRB complex*. Mol Cell Biochem, 2008. **316**(1-2): p. 15-23.

280. Sarno, S., et al., *Selectivity of 4,5,6,7-tetrabromobenzotriazole, an ATP site-directed inhibitor of protein kinase CK2 ('casein kinase-2')*. FEBS Lett, 2001. **496**(1): p. 44-8.
281. Zien, P., et al., *Tetrabromobenzotriazole (TBBt) and tetrabromobenzimidazole (TBBz) as selective inhibitors of protein kinase CK2: evaluation of their effects on cells and different molecular forms of human CK2*. Biochim Biophys Acta, 2005. **1754**(1-2): p. 271-80.
282. Chilin, A., et al., *Coumarin as attractive casein kinase 2 (CK2) inhibitor scaffold: an integrate approach to elucidate the putative binding motif and explain structure-activity relationships*. J Med Chem, 2008. **51**(4): p. 752-9.
283. Cozza, G., et al., *Identification of ellagic acid as potent inhibitor of protein kinase CK2: a successful example of a virtual screening application*. J Med Chem, 2006. **49**(8): p. 2363-6.
284. Nie, Z., et al., *Structure-based design and synthesis of novel macrocyclic pyrazolo[1,5-a] [1,3,5]triazine compounds as potent inhibitors of protein kinase CK2 and their anticancer activities*. Bioorg Med Chem Lett, 2008. **18**(2): p. 619-23.
285. Nie, Z., et al., *Structure-based design, synthesis, and study of pyrazolo[1,5-a][1,3,5]triazine derivatives as potent inhibitors of protein kinase CK2*. Bioorg Med Chem Lett, 2007. **17**(15): p. 4191-5.
286. Bortolato, A., G. Cozza, and S. Moro, *Protein kinase CK2 inhibitors: emerging anticancer therapeutic agents?* Anticancer Agents Med Chem, 2008. **8**(7): p. 798-806.
287. Mazzorana, M., L.A. Pinna, and R. Battistutta, *A structural insight into CK2 inhibition*. Mol Cell Biochem, 2008. **316**(1-2): p. 57-62.
288. Sarno, S. and L.A. Pinna, *Protein kinase CK2 as a druggable target*. Mol Biosyst, 2008. **4**(9): p. 889-94.
289. Sarno, S., et al., *Development and exploitation of CK2 inhibitors*. Mol Cell Biochem, 2005. **274**(1-2): p. 69-76.
290. Sarno, S., et al., *Features and potentials of ATP-site directed CK2 inhibitors*. Biochim Biophys Acta, 2005. **1754**(1-2): p. 263-70.
291. Battistutta, R., et al., *The ATP-binding site of protein kinase CK2 holds a positive electrostatic area and conserved water molecules*. Chembiochem, 2007. **8**(15): p. 1804-9.
292. Laudet, B., et al., *Identification of chemical inhibitors of protein-kinase CK2 subunit interaction*. Mol Cell Biochem, 2008. **316**(1-2): p. 63-9.
293. Pagano, M.A., et al., *Modulation of protein kinase CK2 activity by fragments of CFTR encompassing F508 may reflect functional links with cystic fibrosis pathogenesis*. Biochemistry, 2008. **47**(30): p. 7925-36.
294. Prudent, R., et al., *In vitro and in vivo assays of protein kinase CK2 activity*. Methods Enzymol, 2010. **485**: p. 597-610.
295. Prudent, R., et al., *Salicylaldehyde derivatives as new protein kinase CK2 inhibitors*. Biochim Biophys Acta, 2008. **1780**(12): p. 1412-20.
296. Lopez-Ramos, M., et al., *New potent dual inhibitors of CK2 and Pim kinases: discovery and structural insights*. Faseb J, 2010. **24**(9): p. 3171-85.
297. Prudent, R., et al., *Antitumor activity of pyridocarbazole and benzopyridindole derivatives that inhibit protein kinase CK2*. Cancer Res, 2010. **70**(23): p. 9865-74.

298. Prudent, R., et al., *Identification of polyoxometalates as nanomolar noncompetitive inhibitors of protein kinase CK2*. Chem Biol, 2008. **15**(7): p. 683-92.
299. Moucadel, V., et al., *Antitumoral activity of allosteric inhibitors of protein kinase CK2*. Oncotarget, 2011. **2**(12): p. 997-1010.
300. Merkel, A.L., E. Meggers, and M. Ocker, *PIM1 kinase as a target for cancer therapy*. Expert Opin Investig Drugs, 2012. **21**(4): p. 425-36.
301. Pierre, F., et al., *Novel potent dual inhibitors of CK2 and Pim kinases with antiproliferative activity against cancer cells*. Bioorg Med Chem Lett, 2012. **22**(9): p. 3327-31.
302. Rivalle, C., et al., *Antitumor amino-substituted pyrido[3',4':4,5]pyrrolo[2,3-g]isoquinolines and pyrido[4,3-b]carbazole derivatives: synthesis and evaluation of compounds resulting from new side chain and heterocycle modifications*. J Med Chem, 1983. **26**(2): p. 181-5.
303. Fosse, P., et al., *Sequence requirements for mammalian topoisomerase II mediated DNA cleavage stimulated by an ellipticine derivative*. Nucleic Acids Res, 1991. **19**(11): p. 2861-8.
304. Vendome, J., et al., *Molecular modeling of wild-type and D816V c-Kit inhibition based on ATP-competitive binding of ellipticine derivatives to tyrosine kinases*. J Med Chem, 2005. **48**(20): p. 6194-201.
305. Ohashi, M., E. Sugikawa, and N. Nakanishi, *Inhibition of p53 protein phosphorylation by 9-hydroxyellipticine: a possible anticancer mechanism*. Jpn J Cancer Res, 1995. **86**(9): p. 819-27.
306. Hoang, T.M., et al., *Benzo[e]pyridoindoles, novel inhibitors of the aurora kinases*. Cell Cycle, 2009. **8**(5): p. 765-72.
307. Journak, F., *Structure of the GDP domain of EF-Tu and location of the amino acids homologous to ras oncogene proteins*. Science, 1985. **230**(4721): p. 32-6.
308. Kabsch, W., et al., *Low resolution structure of partially trypsin-degraded polypeptide elongation factor, EF-TU, from Escherichia coli*. J Mol Biol, 1977. **117**(4): p. 999-1012.
309. McCormick, F., et al., *A model for the tertiary structure of p21, the product of the ras oncogene*. Science, 1985. **230**(4721): p. 78-82.
310. Cherfils, J. and M. Zeghouf, *Chronicles of the GTPase switch*. Nat Chem Biol, 2011. **7**(8): p. 493-5.
311. Paduch, M., F. Jelen, and J. Otlewski, *Structure of small G proteins and their regulators*. Acta Biochim Pol, 2001. **48**(4): p. 829-50.
312. Sprang, S.R., *G protein mechanisms: insights from structural analysis*. Annu Rev Biochem, 1997. **66**: p. 639-78.
313. Schweins, T. and A. Wittinghofer, *GTP-binding proteins. Structures, interactions and relationships*. Curr Biol, 1994. **4**(6): p. 547-50.
314. Leipe, D.D., et al., *Classification and evolution of P-loop GTPases and related ATPases*. J Mol Biol, 2002. **317**(1): p. 41-72.
315. Karbstein, K., *Role of GTPases in ribosome assembly*. Biopolymers, 2007. **87**(1): p. 1-11.
316. Britton, R.A., *Role of GTPases in bacterial ribosome assembly*. Annu Rev Microbiol, 2009. **63**: p. 155-76.
317. Verstraeten, N., et al., *The universally conserved prokaryotic GTPases*. Microbiol Mol Biol Rev, 2011. **75**(3): p. 507-42, second and third pages of table of contents.

318. Milburn, M.V., et al., *Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic ras proteins.* Science, 1990. **247**(4945): p. 939-45.
319. Stouten, P.F., et al., *How does the switch II region of G-domains work?* FEBS Lett, 1993. **320**(1): p. 1-6.
320. Li, G. and H. Qian, *Kinetic timing: a novel mechanism that improves the accuracy of GTPase timers in endosome fusion and other biological processes.* Traffic, 2002. **3**(4): p. 249-55.
321. Li, G. and X.C. Zhang, *GTP hydrolysis mechanism of Ras-like GTPases.* J Mol Biol, 2004. **340**(5): p. 921-32.
322. Chung, H.H., et al., *Probing the role of loop 2 in Ras function with unnatural amino acids.* Proc Natl Acad Sci U S A, 1993. **90**(21): p. 10145-9.
323. Mishra, R., et al., *Analysis of GTPases carrying hydrophobic amino acid substitutions in lieu of the catalytic glutamine: implications for GTP hydrolysis.* Proteins, 2005. **59**(2): p. 332-8.
324. Schweins, T., et al., *Substrate-assisted catalysis as a mechanism for GTP hydrolysis of p21ras and other GTP-binding proteins.* Nat Struct Biol, 1995. **2**(1): p. 36-44.
325. Pasqualato, S. and J. Cherfils, *Crystallographic evidence for substrate-assisted GTP hydrolysis by a small GTP binding protein.* Structure, 2005. **13**(4): p. 533-40.
326. Kleuss, C., et al., *Mechanism of GTP hydrolysis by G-protein alpha subunits.* Proc Natl Acad Sci U S A, 1994. **91**(21): p. 9828-31.
327. Bos, J.L., H. Rehmann, and A. Wittinghofer, *GEFs and GAPs: critical elements in the control of small G proteins.* Cell, 2007. **129**(5): p. 865-77.
328. Vetter, I.R. and A. Wittinghofer, *The guanine nucleotide-binding switch in three dimensions.* Science, 2001. **294**(5545): p. 1299-304.
329. Boriack-Sjodin, P.A., et al., *The structural basis of the activation of Ras by Sos.* Nature, 1998. **394**(6691): p. 337-43.
330. Goldberg, J., *Structural basis for activation of ARF GTPase: mechanisms of guanine nucleotide exchange and GTP-myristoyl switching.* Cell, 1998. **95**(2): p. 237-48.
331. Itzen, A., et al., *Nucleotide exchange via local protein unfolding--structure of Rab8 in complex with MSS4.* EMBO J, 2006. **25**(7): p. 1445-55.
332. Renault, L., et al., *Structural basis for guanine nucleotide exchange on Ran by the regulator of chromosome condensation (RCC1).* Cell, 2001. **105**(2): p. 245-55.
333. Worthylake, D.K., K.L. Rossman, and J. Sondek, *Crystal structure of Rac1 in complex with the guanine nucleotide exchange region of Tiam1.* Nature, 2000. **408**(6813): p. 682-8.
334. Cherfils, J. and P. Chardin, *GEFs: structural basis for their activation of small GTP-binding proteins.* Trends Biochem Sci, 1999. **24**(8): p. 306-11.
335. Scheffzek, K., et al., *The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants.* Science, 1997. **277**(5324): p. 333-8.
336. Wittinghofer, A., K. Scheffzek, and M.R. Ahmadian, *The interaction of Ras with GTPase-activating proteins.* FEBS Lett, 1997. **410**(1): p. 63-7.
337. DerMardirossian, C. and G.M. Bokoch, *GDI: central regulatory molecules in Rho GTPase activation.* Trends Cell Biol, 2005. **15**(7): p. 356-63.

338. Hwang, J. and M. Inouye, *A bacterial GAP-like protein, YihI, regulating the GTPase of Der, an essential GTP-binding protein in Escherichia coli*. J Mol Biol, 2010. **399**(5): p. 759-72.
339. Meier, T.I., et al., *Era GTPase of Escherichia coli: binding to 16S rRNA and modulation of GTPase activity by RNA and carbohydrates*. Microbiology, 2000. **146 (Pt 5)**: p. 1071-83.
340. Scrima, A. and A. Wittinghofer, *Dimerisation-dependent GTPase reaction of MnmE: how potassium acts as GTPase-activating element*. EMBO J, 2006. **25**(12): p. 2940-51.
341. Anand, B., P. Surana, and B. Prakash, *Deciphering the catalytic machinery in 30S ribosome assembly GTPase YqeH*. PLoS One, 2010. **5**(4): p. e9944.
342. Karnoub, A.E. and R.A. Weinberg, *Ras oncogenes: split personalities*. Nat Rev Mol Cell Biol, 2008. **9**(7): p. 517-31.
343. Agola, J.O., et al., *A Competitive Nucleotide Binding Inhibitor: In Vitro Characterization of Rab7 GTPase Inhibition*. ACS Chem Biol, 2012. **7**(6): p. 1095-108.
344. Blum, R., A.D. Cox, and Y. Kloog, *Inhibitors of chronically active ras: potential for treatment of human malignancies*. Recent Pat Anticancer Drug Discov, 2008. **3**(1): p. 31-47.
345. Renault, L., B. Guibert, and J. Cherfils, *Structural snapshots of the mechanism and inhibition of a guanine nucleotide exchange factor*. Nature, 2003. **426**(6966): p. 525-30.
346. Mossessova, E., R.A. Corpina, and J. Goldberg, *Crystal structure of ARF1*Sec7 complexed with Brefeldin A and its implications for the guanine nucleotide exchange mechanism*. Mol Cell, 2003. **12**(6): p. 1403-11.
347. Gao, Y., et al., *Rational design and characterization of a Rac GTPase-specific small molecule inhibitor*. Proc Natl Acad Sci U S A, 2004. **101**(20): p. 7618-23.
348. Ferri, N., et al., *Virtual screening approach for the identification of new Rac1 inhibitors*. J Med Chem, 2009. **52**(14): p. 4087-90.
349. Noble, C.G. and H. Song, *Structural studies of elongation and release factors*. Cell Mol Life Sci, 2008. **65**(9): p. 1335-46.
350. Parmeggiani, A., et al., *Structural basis of the action of pulvomycin and GE2270 A on elongation factor Tu*. Biochemistry, 2006. **45**(22): p. 6846-57.
351. Vogeley, L., et al., *Conformational change of elongation factor Tu (EF-Tu) induced by antibiotic binding. Crystal structure of the complex between EF-Tu.GDP and aurodox*. J Biol Chem, 2001. **276**(20): p. 17149-55.
352. Parmeggiani, A., et al., *Enacyloxin IIa pinpoints a binding pocket of elongation factor Tu for development of novel antibiotics*. J Biol Chem, 2006. **281**(5): p. 2893-900.
353. Mehr, I.J., et al., *A homologue of the recombination-dependent growth gene, rdcC, is involved in gonococcal pilin antigenic variation*. Genetics, 2000. **154**(2): p. 523-32.
354. Hwang, J. and M. Inouye, *An essential GTPase, der, containing double GTP-binding domains from Escherichia coli and Thermotoga maritima*. J Biol Chem, 2001. **276**(33): p. 31415-21.
355. Knuth, K., et al., *Large-scale identification of essential Salmonella genes by trapping lethal insertions*. Mol Microbiol, 2004. **51**(6): p. 1729-44.
356. Morimoto, T., et al., *Six GTP-binding proteins of the Era/Obg family are essential for cell growth in Bacillus subtilis*. Microbiology, 2002. **148**(Pt 11): p. 3539-52.

357. Zalacain, M., et al., *A global approach to identify novel broad-spectrum antibacterial targets among proteins of unknown function*. J Mol Microbiol Biotechnol, 2003. **6**(2): p. 109-26.
358. Gil, R., et al., *Determination of the core of a minimal bacterial gene set*. Microbiol Mol Biol Rev, 2004. **68**(3): p. 518-37.
359. Mittenhuber, G., *Comparative genomics of prokaryotic GTP-binding proteins (the Era, Obg, EngA, ThdF (TrmE), YchF and YihA families) and their relationship to eukaryotic GTP-binding proteins (the DRG, ARF, RAB, RAN, RAS and RHO families)*. J Mol Microbiol Biotechnol, 2001. **3**(1): p. 21-35.
360. Robinson, V.L., et al., *Domain arrangement of Der, a switch protein containing two GTPase domains*. Structure, 2002. **10**(12): p. 1649-58.
361. Valverde, R., L. Edwards, and L. Regan, *Structure and function of KH domains*. FEBS J, 2008. **275**(11): p. 2712-26.
362. Muench, S.P., et al., *The essential GTPase YphC displays a major domain rearrangement associated with nucleotide binding*. Proc Natl Acad Sci U S A, 2006. **103**(33): p. 12359-64.
363. Watt, R.M., et al., *Visualizing the proteome of Escherichia coli: an efficient and versatile method for labeling chromosomal coding DNA sequences (CDSs) with fluorescent protein genes*. Nucleic Acids Res, 2007. **35**(6): p. e37.
364. Tan, J., U. Jakob, and J.C. Bardwell, *Overexpression of two different GTPases rescues a null mutation in a heat-induced rRNA methyltransferase*. J Bacteriol, 2002. **184**(10): p. 2692-8.
365. Bharat, A., et al., *Cooperative and critical roles for both G domains in the GTPase activity and cellular function of ribosome-associated Escherichia coli EngA*. J Bacteriol, 2006. **188**(22): p. 7992-6.
366. Schaefer, L., et al., *Multiple GTPases participate in the assembly of the large ribosomal subunit in Bacillus subtilis*. J Bacteriol, 2006. **188**(23): p. 8252-8.
367. Tomar, S.K., et al., *Distinct GDP/GTP bound states of the tandem G-domains of EngA regulate ribosome binding*. Nucleic Acids Res, 2009. **37**(7): p. 2359-70.
368. Butland, G., et al., *Interaction network containing conserved and essential protein complexes in Escherichia coli*. Nature, 2005. **433**(7025): p. 531-7.
369. Lamb, H.K., et al., *Functional analysis of the GTPases EngA and YhbZ encoded by Salmonella typhimurium*. Protein Sci, 2007. **16**(11): p. 2391-402.
370. Hwang, J. and M. Inouye, *Interaction of an essential Escherichia coli GTPase, Der, with the 50S ribosome via the KH-like domain*. J Bacteriol, 2010. **192**(8): p. 2277-83.
371. Hwang, J. and M. Inouye, *RelA functionally suppresses the growth defect caused by a mutation in the G domain of the essential Der protein*. J Bacteriol, 2008. **190**(9): p. 3236-43.
372. Gibbs, J.B., et al., *Intrinsic GTPase activity distinguishes normal and oncogenic ras p21 molecules*. Proc Natl Acad Sci U S A, 1984. **81**(18): p. 5704-8.
373. Chen, S.M., et al., *Expression and characterization of RNase III and Era proteins. Products of the rnc operon of Escherichia coli*. J Biol Chem, 1990. **265**(5): p. 2888-95.
374. Yamanaka, K., J. Hwang, and M. Inouye, *Characterization of GTPase activity of TrmE, a member of a novel GTPase superfamily, from Thermotoga maritima*. J Bacteriol, 2000. **182**(24): p. 7078-82.

375. Sullivan, S.M., et al., *Analysis of guanine nucleotide binding and exchange kinetics of the Escherichia coli GTPase Era*. J Bacteriol, 2000. **182**(12): p. 3460-6.
376. Ash, M.R., et al., *Potassium-activated GTPase reaction in the G Protein-coupled ferrous iron transporter B*. J Biol Chem, 2010. **285**(19): p. 14594-602.
377. Hwang, J., et al., *Structure-based design and screening of inhibitors for an essential bacterial GTPase, Der*. J Antibiot (Tokyo), 2012. **65**(5): p. 237-43.



Institut de Biologie Structurale



Annexes au Mémoire

Pour obtenir le diplôme d'

HABILITATION A DIRIGER LES RECHERCHES

Spécialité : **biologie**

Présenté par

Jean-Baptiste REISER

préparé à l'Institut de Biologie Structurale Jean-Pierre Ebel, Grenoble

Quelques exemples de biologie structurale Publications



Chapitre II

Travaux postdoctoraux

Crystal Structure of the *Drosophila* Peptidoglycan Recognition Protein (PGRP)-SA at 1.56 Å Resolution

Jean-Baptiste Reiser¹, Luc Teyton² and Ian A. Wilson^{1,3*}

¹Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla CA 92037 USA

²Department of Immunology The Scripps Research Institute 10550 North Torrey Pines Road La Jolla CA 92037, USA

³The Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road La Jolla CA 92037, USA

Peptidoglycan recognition proteins (PGRPs) form a recently discovered protein family, which is conserved from insect to mammals and is implicated in the innate immune system by interacting with/or degrading microbial peptidoglycans (PGNs). *Drosophila* PGRP-SA is a member of this family of pattern recognition receptors and is involved in insect Toll activation. We report here the crystal structure of PGRP-SA at 1.56 Å resolution, which represents the first example of a “recognition” PGRP. Comparison with the catalytic *Drosophila* PGRP-LB reveals an overall structure conservation with an L-shaped hydrophilic groove that is likely the PGN carbohydrate core binding site, but further suggests some possible functional homology between recognition and catalytic PGRPs. Consistent with sequence analysis, PGRP-SA does not contain the canonical zinc-binding residues found in catalytic PGRPs. However, substitution of the zinc-binding cysteine residue by serine, along with an altered coordinating histidine residue, assembles a constellation of residues that resembles a modified catalytic triad. The serine/histidine juxtaposition to a threonine residue and a carbonyl oxygen atom, along with conservation of the catalytic water molecule found in PGRP-LB, tantalizingly suggests some hydrolytic function for this member of receptor PGRPs.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: pattern recognition receptor; innate immunity; Toll pathway; serine hydrolase; zinc hydrolase

*Corresponding author

Introduction

Innate immunity represents the first line of defense against microorganisms. In contrast to the adaptive response, innate immunity receptors are able to recognize conserved motifs that are present in microorganisms but absent from the host organism, in order to discriminate between pathogens and self. These conserved motifs, called pathogen associated molecular patterns (PAMPs) include, for example, peptidoglycans (PGNs), lipopolysaccharides (LPSs), and glycans, and were postulated to be recognized by germ-line encoded pattern recognition receptors (PRRs).¹

PGNs are an essential cell-wall component of Gram-positive bacteria and are found in lesser abundance in Gram-negative bacteria. They are composed of glycan chains of alternating units of *N*-acetylglucosamine (GluNAc) and *N*-acetyl-

muramic acid (MurNAc) that are cross-linked by short peptides of which the composition depends of the bacterial species.² Several immune proteins are known to interact with PGNs: CD14,³ Toll-like receptor (TLR)-2⁴ and intracellular nucleotide-binding oligomerization domain (Nod)-containing proteins.^{5,6} The peptidoglycan recognition proteins (PGRPs) form another family of PRRs which are related to the bacteriophage T7 lysosyme, a zinc-dependent, *N*-acetylmuramoyl-L-alanine amidase (NAMLAA).^{7,8} PGRPs are found in a wide range of organisms (*Bombix mori*,⁹ *Drosophila melagonaster*,¹⁰ *Homo sapiens*,¹¹ *Mus musculus*¹²), are highly conserved from insects to mammals,⁷ and are able to bind PGNs with a high affinity.¹³ The different PGRPs are expressed in various innate immune cells, such as neutrophils¹³ and macrophage.^{10,11}

In *Drosophila*, microbial infections activate two distinct intracellular pathways called imd/relish and Toll, respectively, that lead to secretion of antimicrobial peptides.¹⁴ The imd/relish cascade involves PGRP-LC¹⁵ or PGRP-LE¹⁶ in adult and larvae, respectively, which constitute essential components for recognition of Gram-negative bacteria

Abbreviations used: PGRP, peptidoglycan recognition protein; PGN, peptidoglycan.

E-mail address of the corresponding author: wilson@scripps.edu

through interaction with *meso*-diaminopimelic acid (DAP)-containing PGNs. The Toll pathway mediates response to fungi and Gram-positive bacteria. Circulating PGRP-SA is a key upstream component of this pathway by interacting with Gram-positive bacteria, or their derived lysine-type PGNs, and by activating a downstream effector, which is still unknown and that triggers the proteolytic cleavage of the Toll receptor ligand, Spätzle.¹⁷

This large family of PGRPs can be functionally divided into "catalytic" PGRPs with amidase activity and "recognition" PGRPs that bind peptidoglycans. The crystal structure of the PGRP-LB from *Drosophila* revealed the first structural insights into the catalytic role of PGRP.¹⁸ To provide further knowledge of pattern recognition proteins and to gain insights into the structural characteristics of recognition PGRPs, the crystal structure of *Drosophila* PGRP-SA was determined to 1.56 Å resolution.

Results

Structure overview and overall differences

Peptidoglycan recognition protein SA (PGRP-SA) from *D. melanogaster* was successfully cloned and expressed at a high level (1 mg per liter of cell culture) in *D. melanogaster* SC2 cells. PGRP-SA was then purified, crystallized and its crystal structure was determined by molecular replacement using PGRP-LB¹⁸ as a starting model (see Materials and Methods, and Table 1). The final refined model consists of residues 11–177 of the cloned sequence corresponding to the peptidoglycan-binding domain (Figure 1). No electron density was observed for the ten first residues, as for PGRP-LB, suggesting flexibility of the N terminus in PGRP domains.

The PGRP-SA structure adopts the unique fold of the PGRP family with a mixed β -sheet of five strands (four parallel, β 3, β 4, β 6 and β 7, and one anti-parallel, β 5) and three α -helices (α 1 to α 3) that form the protein core and two additional β -strands (β 1 and β 2) at the N terminus. As expected, the PGRP-SA overall structure is very similar to that of PGRP-LB¹⁸ with a root-mean-square deviation (r.m.s.d.) of 1.19 Å for 166 α -carbon atoms. Nevertheless, a significant difference with PGRP-LB is noted in the β 3- α 1 loop/ α 1 helix region (Figure 2). Indeed the β 3- α 1 loop is shorter in PGRP-SA due to a single residue deletion (Figure 3) and has a slightly different conformation. Because of the conserved disulfide bond between Cys48 and Cys54, which anchors the β 3- α 1 loop to the α 1-helix, the PGRP-SA α 1-helix tilts about 5° around a pivot at residue 65. However, as these two PGRPs crystallize in different space groups, the position and conformation of the β 3- α 1 loop/ α 1-helix group may also be influenced by crystal packing. Other structural differences

Table 1. Data collection, structure determination and refinement statistics

<i>Data collection</i>	
Space group	C2
Unit cell parameters	
<i>a</i> (Å)	136.29
<i>b</i> (Å)	64.42
<i>c</i> (Å)	45.87
β (deg.)	100.33
No. molecules in asymm. unit	2
Resolution range (Å)	45.13–1.56 (1.62–1.56) ^a
Completeness (%)	99.4 (94.0)
No. unique reflections	54,845 (5148)
Redundancy	3.6 (2.5)
Mean <i>I</i> / σ	10.7 (2.0)
<i>R</i> _{sym} (%) ^b	5.6 (45.1)
<i>Molecular replacement</i>	
Resolution range (Å)	15.00–3.00
Structure factor corr. coeff. (%)	50.2
Structure factor <i>R</i> (%)	47.5
<i>Refinement</i>	
Resolution range (Å)	45.00–1.56 (1.60–1.56) ^a
No. reflections (<i>F</i> > 0)	52,381 (2635)
No. test set reflections	5275 (317)
No. non-hydrogen protein atoms	2737
No. water atoms	389
No. sulfate atoms	10
No. ethylene glycol atoms	16
Deviations from ideality (r.m.s.)	
Bond length (Å)	0.010
Bond angle (deg.)	1.299
Ramachandran statistics (%)	
Most favored	91.0
Favored	7.9
Allowed	0.3
Disallowed ^c	0.7
Average <i>B</i> value (Å ²)	18.4
<i>R</i> _{free} ^d (%)	21.1 (29.0)
<i>R</i> _{work} ^d (%)	17.7 (23.3)
<i>R</i> _{cryst} ^d (%)	18.1 (23.6)

^a Numbers in parentheses indicate statistics for the highest resolution shell.

^b $R_{\text{sym}} = \sum \langle |I_{hkl}| - I_{hkl}| / |I_{hkl}| \rangle$, where $\langle I_{hkl} \rangle$ is the average intensity over symmetry-related reflections and I_{hkl} is the observed intensity.

^c Val44 was found in disallowed region for both molecules in the asymmetric unit ($\Phi = 71^\circ$, $\psi = 118^\circ$), but its density is well defined.

^d $R_{\text{value}} = \sum ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum |F_{\text{obs}}|$, where F_{obs} and F_{calc} are the observed and model-based calculated structure factors. For R_{free} , the sum is done on the test set reflections (10% of total reflections), for R_{work} on the remaining reflections, and for R_{cryst} on all reflections included in the resolution range.

include a shift of the N terminus of PGRP-SA by about 10 Å towards the protein core, as compared to PGRP-LB, that enables Cys11 (corresponding to Thr12 in PGRP-LB) to form a disulfide bond with Cys134 that stabilizes interaction of the β 1-strand with the protein core. Some distal loops show also minor shifts less than 1.5 Å.

A conserved binding groove

Comparison of PGRP-SA with PGRP-LB and the bacteriophage T7 lysosyme, a zinc-dependent amidase related to the PGRP domain,¹⁹ reveals a

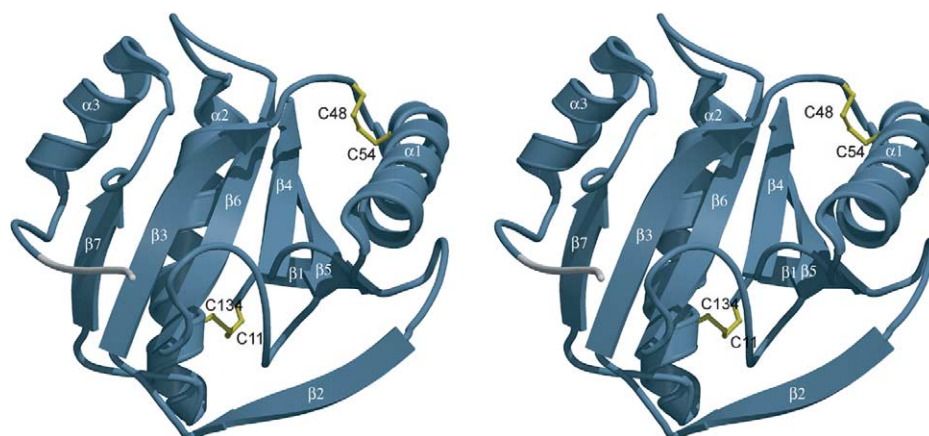


Figure 1. Stereoview of PGRP-SA crystal structure. The ribbon structure of PGRP-SA from *D. melanogaster* is represented and the secondary structures are labeled in order from N to C terminus. The disulfide bonds between Cys11 and Cys134, and between Cys48 and Cys54 are drawn in yellow ball-and-sticks. The six-histidine tag at the C terminus is shown in white.

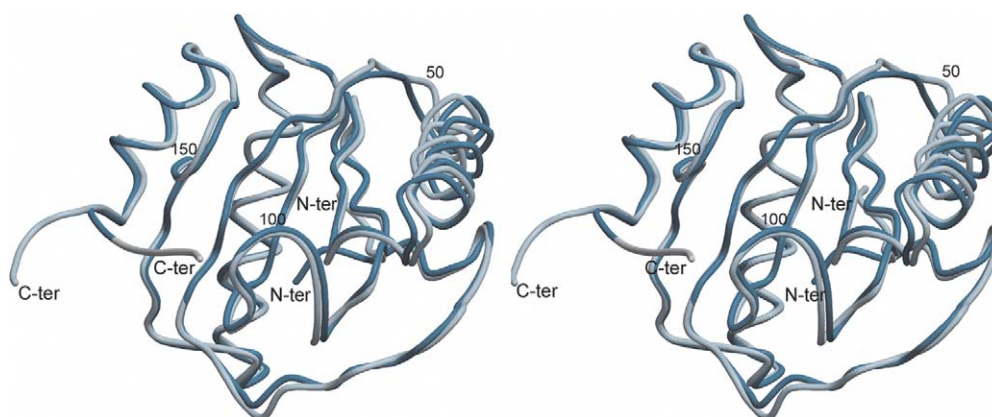


Figure 2. Stereoview of superimposition of PGRP-SA and PGRP-LB. PGRP-SA and PGRP-LB backbone structures are represented in coils, oriented as in Figure 1. The PGRP-SA structure is in blue and the PGRP-LB structure is in gray. The N and C termini are labeled, as well as every 50 residues.

conserved L-shaped groove which likely represents the peptidoglycan-binding site¹⁸ (Figures 2 and 4). This groove is delineated by the β -sheet, the α 1 helix and the β 3- α 1, α 1- β 4, β 5- β 6 and β 7- α 3 loops (Figures 1 and 4) and is lined by residues that are more than 70% conserved in PGRPs (Figures 3 and 4). Among these residues (with at least 15 Å² of solvent-exposed surface), the side-chains of Met61, Tyr64, Phe71, Tyr76 and Thr156 point towards the groove and are well positioned for ligand interactions. Other possible binding residues, which could utilize either their side-chains or even their main chain, that would then require no sequence conservation, include residues 44, 46, 69, 70, 97 and 155.

Three highly conserved residues, Cys48, Tyr76 and Thr156, have been already identified as essential for PGN recognition. Although Cys48 presents no solvent-accessible surface in the putative PGN-binding groove, its mutation abolishes Toll receptor activation by PGRP-SA¹⁷ and the amidase activity of human PGRP-L.²⁰ As this cysteine residue is involved in a disulfide bond with Cys54 at

the edge of the putative PGN-binding groove, it likely participates in the structural integrity of the groove. Tyr76 is important for amidase activity in bacteriophage T7 lysozyme¹⁹ and human PGRP-L²⁰ and is highly conserved in PGRP domains. The structure of PGRP-SA shows that the tyrosine hydroxyl group points towards the base of the binding groove and suggests likely interaction with PGN either directly or indirectly *via* a water molecule (see below). Mutation of Thr156 to lysine in PGRP-LB also abolishes PGN binding,¹⁸ which again suggests critical interactions with PGN.

In this unliganded PGRP-SA structure, the PGN binding groove is filled with water molecules and comparison of the three related amidase domains identifies a conserved network of water molecules in the PGN-binding groove. In bacteriophage T7 lysozyme, the water molecule that completes the zinc coordination sphere was proposed to initiate nucleophilic attack on the substrate carbonyl group, as for other Zn²⁺ proteases, such as carboxypeptidase A¹⁹ and thermolysin.²¹ An oxygen atom from a glutamic acid of a symmetry-related

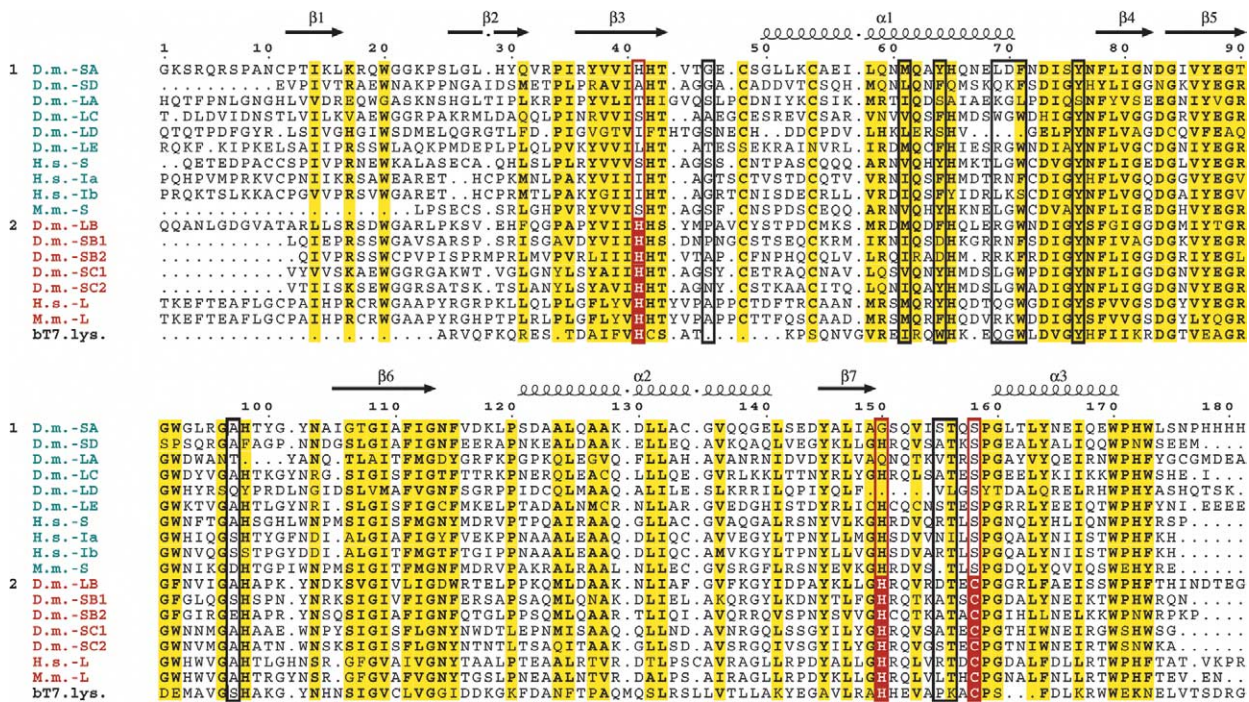


Figure 3. Sequence alignment of PGRPs and secondary structure assignment. The PGRPs sequences from *D. melanogaster* (D.m.), *H. sapiens* (H.s.) and *M. musculus* (M.m.) and the lysozyme sequence from the bacteriophage T7 were aligned with Multialign program.⁴⁷ The secondary structure elements from PGRP-SA are indicated above the sequences. Sequences are sorted according whether they belong to (1) recognition PGRP (in blue) or (2) catalytic PGRP (in red) subgroups. In the latter subgroup, the conserved amino acid residues that bind the catalytic zinc ion, as in T7 lysozyme, are highlighted in red boxes and white letters. The corresponding residues in recognition PGRPs are boxed in red. The yellow columns indicate residues with a sequence homology greater than 70% and black boxes show residues that line the PGN-binding groove and that have a relative solvent accessibility greater than 40%; half of them are conserved.

molecule is found at this equivalent position in the zinc ion coordination sphere of PGRP-LB, confirming the crucial role of a coordinating oxygen atom in the amidase catalytic mechanism.¹⁸ In both structures, the conserved oxygen atom directly interacts with the zinc ion and a conserved tyrosine residue (Tyr46 in T7 lysozyme and Tyr78 in PGRP-

LB, Figure 5). The PGRP-SA structure surprisingly reveals conservation of this catalytic water molecule (Figure 5). Indeed, it forms hydrogen bonds to the conserved tyrosine residue (position 76 in PGRP-SA) and to His41, where the ND1 atom replaces the zinc ion (see below). By masking the zinc ion in PGRP-LB or His41 in PGRP-SA from

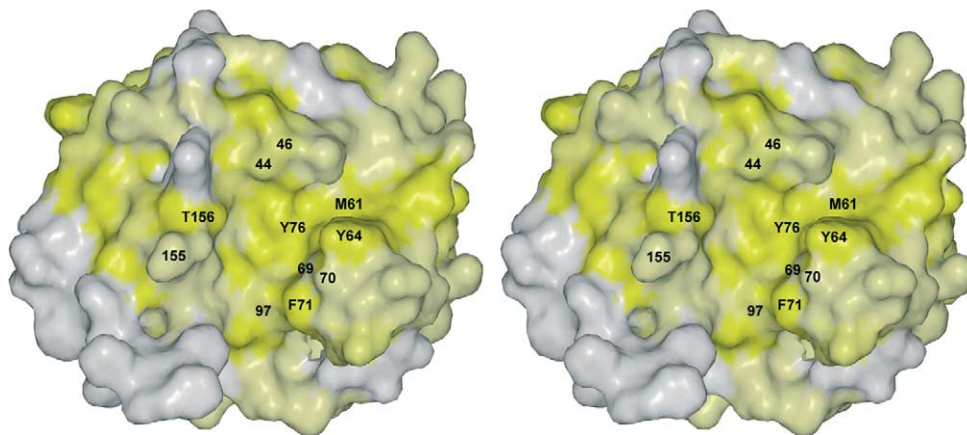


Figure 4. Sequence conservation represented on the PGRP-SA molecular surface. The sequence homology is shown on a stereoview of the PGRP-SA molecular surface looking directly down to the PGN-binding groove. The surface is colored according to the sequence homology from white for surface residues less than 30% conserved to yellow for more than 70% conserved. The name and number of the residues lining the groove are reported in bold characters; the residues found to be mostly conserved in PGRPs are indicated with both letter and sequence number, while non-conserved residues are indicated only with the sequence number.

solvent, this structurally conserved water molecule allows conservation of the shape and properties of this region of the PGN-binding groove. This water molecule conservation also suggests an important role for water in PGN binding.

Structural differences between recognition and catalytic PGRPs

According to PGRP sequence analyses, two main subgroups in this protein family can be designated as catalytic and recognition PGRPs (Figure 3). The first subgroup refers to PGRPs with amidase activity against PGNs that have conserved residues to coordinate a zinc ion (two histidine residues and a cysteine) and for amidase activity (a threonine and a tyrosine). Eight PGRPs are implicated in that subgroup. Mammalian long PGRPs (PGRP-L)^{20,22} and *Drosophila* PGRP-SC1B²³ have an NAMLAA activity similar to bacteriophage T7 lysozyme;¹⁹ whereas PGRP-LB shows lytic activity against sensitized *Escherichia coli*, but is also proposed to have such an amidase activity.¹⁸ By contrast, the recognition subgroup refers to PGRPs that lack zinc-binding residues, but still retain the ability to bind PGNs, as shown for short mammals PGRPs¹¹ and *Drosophila* PGRP-SA.¹⁰

The PGRP-LB active site structure reveals a catalytic zinc ion coordinated in a tetrahedral geometry by two histidine residues (residues 42 and 152 in PGRP-LB sequence numbering), a cysteine residue (160) and a carboxyl oxygen atom from a symmetry-related glutamate residue, which likely substitutes for a water molecule (Figure 5).¹⁸ The PGRP-SA sequence lacks two of the three Zn-coordinating residues; His152 and Cys160 of PGRP-LB are replaced by Gly150 and Ser158 in PGRP-SA that abolish zinc binding (Figure 5). X-ray absorp-

tion measurement at the zinc edge on a PGRP-SA crystal confirmed that no metal ion is present in this recognition PGRP (data not shown). Although this groove region can be closely superimposed between both PGRPs, some significant differences can be correlated with the presence or the absence of a zinc ion (Figure 5). Compared to the PGRP-LB structure, PGRP-SA His41 (corresponding to His42 of PGRP-LB) is rotated about 30° around X₁ so that its ND1 atom is adjacent to the former zinc location and its Thr156 (corresponding to Thr158 of PGRP-LB) at the β 7- α 3 loop apex is shifted about 0.7 Å (C $^{\alpha}$) towards the center of the groove. However, despite the absence of a catalytic Zn²⁺, the structural differences observed between PGRP-SA and PGRP-LB are mainly confined to the base of the groove.

Comparison with serine hydrolases

The crystal structure of PGRP-SA reveals that Ser158 is now in close proximity *via* hydrogen bonds to His41 (OG-ND1 distance = 2.76 Å, Figure 6). The juxtaposition of serine and histidine residues is common in serine hydrolases. Members of this protein superfamily also represent prototype examples of convergent evolution by having very low sequence homology and adopting different folds that can still converge to similar active sites.^{24–26} Although no data are currently available to suggest that PGRP-SA from *Drosophila* can catalyze any such hydrolytic reaction, some structural features of the catalytic triad motifs of serine hydrolases are found in PGRP-SA.

First, the PGRP-SA structure revealed a constellation of residues that resembles the catalytic triad of serine hydrolases (Figure 6). In general, the catalytic triad is composed of a nucleophilic serine that

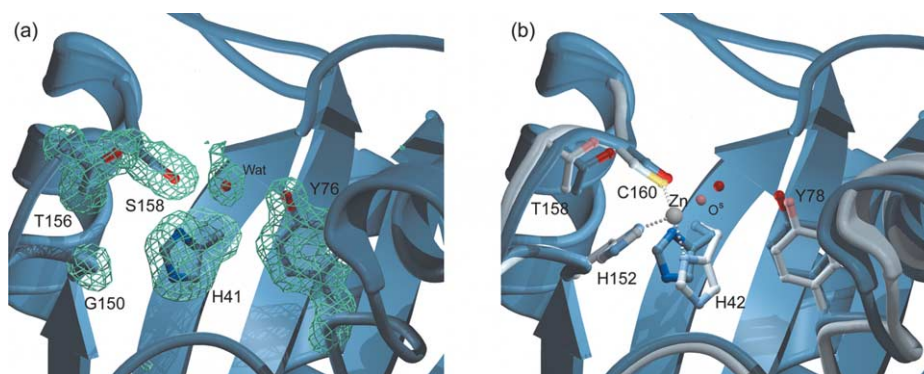


Figure 5. The PGRP-SA and PGRP-LB binding and catalytic site. (a) The omit $F_o - F_c$ electron density map is shown around PGRP-SA side-chains that correspond to PGRP-LB zinc-coordinating residues. The density is represented by light green mesh and contoured at 3σ . The residues are labeled and numbered as in PGRP-SA sequence. Atoms are colored according to their type: oxygen in red, nitrogen in blue and carbon as backbone atoms. No density is present for a zinc ion as expected in the PGRP-SA structure. (b) Superposition of the PGRP-LB catalytic site (in light gray) and the corresponding PGRP-SA binding site (in blue color). The residues are labeled and numbered as in the PGRP-LB sequence. Only PGRP-LB structure elements that show significant deviation from the PGRP-SA structure are shown for clarity. The zinc ion found in PGRP-LB and its interactions with coordinating residues are colored in light gray. The O^{*} atom corresponds to the carboxyl oxygen atom of the PGRP-LB symmetry-related molecule Glu182 assumed to substitute for the normal coordinating water molecule. A water molecule is found in the PGRP-SA structure adjacent to the catalytic water molecule.

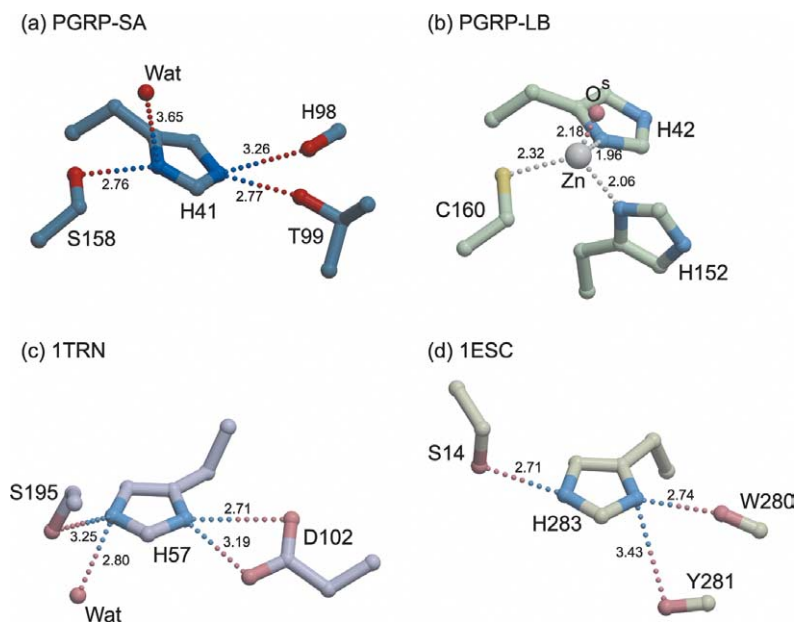


Figure 6. PGRP-SA Ser/His juxtaponition in comparison with the catalytic center of PGRP-LB and serine hydrolases. (a) Histidine 41, serine 158, threonine 99 and the carbonyl group of residue 98 of PGRP-SA are represented in ball and sticks. Atoms are colored according to their type. Hydrogen bonds are represented in broken lines and distances in Å. (b) Although structural changes are observed compared to PGRP-LB, the location of the presumed catalytic water molecule O^s is structurally conserved. (c) The Ser/His/Asp triad of bovine trypsin (PDB access code: 1TRN) and (d) the modified triad of a bacterial esterase (PDB access code: 1ESC) are shown for comparison. The configuration of PGRP-SA S158/H41/T99/H98-carbonyl resembles the catalytic center of serine hydrolases, especially the modified triad of bacterial esterase.

attacks the substrate carbonyl *via* a covalently bonded tetrahedral intermediate, a histidine residue that acts as a base to activate the serine and an acidic residue (aspartate or glutamate) to orient and polarize the base²⁶. In PGRP-SA, His41 is hydrogen bonded to Ser158 (N^{δ1}-O^s distance = 2.76 Å) and is appropriately oriented to act as a base and to activate the serine. Moreover, His41 hydrogen bonds to the hydroxyl group of Thr99 (N^{ε2}-O^{γ1} distance = 2.77 Å) and to the carbonyl group of His98 (N^{ε2}-O distance = 3.26 Å). When the PGRP-SA His41 is superimposed on the classical catalytic histidine residue of serine hydrolases (the bovine trypsin structure has been taken as the reference, PDB access code 1TRN²⁷), both oxygen-containing groups (Thr99 OH^γ and His98 CO) occupy an equivalent position of the aspartate (Figure 6). Secondly, an oxyanion hole composed of proton donors is necessary to stabilize the negative charge of the transition state.²⁸ The hydrophilic character of the PGRP-SA groove makes it ideal for such an interaction. For example, residues that could act as an oxyanion hole, include the backbone amide of Ser158, as well as the side-chain nitrogen atom of His42 or His41. The hydroxyl groups of Thr156 or Tyr76 could also act as hydrogen bond donors and replace main-chain amide nitrogen interactions, as already observed in the bacterial cocaine esterase²⁹ or the porcine muscle oligopeptidase,²⁸ where a tyrosine residue forms the oxyanion hole. Finally, a water molecule, activated by the histidine residue, is essential to complete the catalytic reaction.²⁶ In PGRP-SA, the water molecule, described above, that is conserved in amidase domains and interacts with His41 in PGRP-SA (Figure 6), could be a good candidate to be such a component of any hydrolytic mechanism.

Discussion

The crystal structure of PGRP-SA presented here represents the second structure of a PGRP and is the first structure of a recognition PGRP known to be directly involved in bacterial recognition through their cell-wall PGNs. This structure allows us to address some key structural elements in bacterial recognition by the immune system. Both PGRP domain structures, PGRP-SA and PGRP-LB, have substantial similarities (Figure 2) and confirm the proposal from the high sequence homology that all PGRP domains have an identical fold.¹⁸

Most PGRPs, including PGRP-SA, bind PGN with a high affinity,¹⁰⁻¹³ suggesting common structural features for binding PGNs. PGRP-SA and PGRP-LB have a conserved L-shaped groove that is likely the PGN-binding site. However, the question of how PGRP domains specifically bind PGNs still remains to be determined. Half of the residues that line the groove are highly conserved, especially at the base of the groove (Figures 3 and 4), indicating that the conformation and chemical properties of the groove are maintained in all PGRPs. These residues are mostly hydrophilic, which suggests that the interaction with the PGN core is through hydrogen bonds and electrostatic interactions. The tyrosine residue (at positions 76 in PGRP-SA and 78 in PGRP-LB) and the threonine residue (at position 156 in PGRP-SA and 158 in PGRP-LB), that are conserved in all PGRPs (Figure 3), along with the structurally conserved water molecule, are good candidates as essential residues for PGN binding.

However, each protein of the PGRP family has its own specificity for PGNs.⁷ Residues found at

the edges of the groove are more variable in sequence (Figure 4) and this sequence variation may tailor the PGRP surface for each PGN ligand.¹⁸ Furthermore, compared to PGRP-LB, the different PGRP-SA α 1 helix/ β 3- α 1 conformation (Figure 2) could suggest a direct role in PGN recognition by providing a slightly different topology to each PGRP groove, or that its flexibility may play a role in accommodating each specific ligand. This dual nature of conserved and unconserved features of the PGRP molecular surface may then be related to the dual properties of peptidoglycans. PGNs are composed of a common core of alternating β -1,4 linked GluNAc and MurNAc strands, which have little variation in bacteria, but are cross-linked by short peptides that are highly variable.² Thus, the conserved structural features of PGRPs are likely to be involved in the recognition of the common PGN carbohydrates, while the more variable parts of the PGRP groove and its vicinity interact with the cross-linked variable peptides. A liganded complex structure between a PGRP and its specific ligand is clearly necessary to resolve this issue

Comparison between the structures of PGRP-LB and PGRP-SA, which represent the catalytic and recognition PGRPs, respectively, indicates that the main structural differences between these two PGRP subgroups are located at the bottom of the groove around the zinc-binding site in catalytic PGRPs (Figure 5). Nevertheless, the observed substitutions in PGRPs do not grossly affect the groove shape and apparent binding properties. Furthermore, a water molecule is conserved and may aid in conserving the structural complementarity for ligand binding. Catalytic and recognition PGRPs likely share a similar PGN-binding mode, which remains unperturbed by the presence or absence of the bound zinc ion. A PGRP-SC1B mutation study in which the coordinating cysteine residue was replaced by a serine or an alanine residue supports this assertion. Although these mutants are enzymatically inactive, certainly due to the absence of zinc, they retained their ability to bind PGNs.²³ However, the ability to bind zinc may not be the only difference between catalytic and recognition PGRPs. Wang *et al.* mutated residues in the binding site to furnish zinc-binding capability to human and mouse PGRP-S that should have changed the proteins into an "active" configuration; yet no amidase activity was observed, suggesting that other residues are also required for the amidase properties.²⁰

PGRP-SA appears to be the key in *Drosophila* Toll pathway activation by interacting with either Gram-positive bacteria or their derived PGNs, which then induces an unknown proteolytic cascade that leads to the cleavage of the Toll ligand, Spätzle.¹⁷ The observation that PGRP-SA has some distant structural similarities with the superfamily of serine hydrolases leads to the question of whether there is indeed any hydrolytic capability for this supposed recognition PGRP. X-ray crystal-

lography and biochemical studies have demonstrated how the conventional catalytic triad, which provides the hydrolase activity, can vary among different families of enzymes that cleave amide or ester bonds.²⁶ For example, the structure of *Streptomyces scabies* esterase (PDB access code 1ESC) revealed that the acidic residue can be replaced by two main-chain carbonyl groups (Figure 6).³⁰ Type I signal peptidases SipW³¹ and SipS³² hydrolyze their substrates without the requirement of a third acidic member of the catalytic triad, and Craik and colleagues demonstrated that mutation of the aspartate to asparagine does not totally disrupt the protease activity of trypsin.³³ A hydrolase mechanism *via* a modified catalytic triad is then not excluded. If this hypothesis for PGRP-SA is later confirmed by experimental approaches, degradation of any ligand would then be a key step in the pathogen response.

Recently, surprising reports have emerged for a potential cooperative interaction of PGRP-SA with the Gram-negative binding protein (GNBP-1) for induction of the signal cascade. These studies suggest that the GNBP-1/PGRP-SA complex formation either may act as a first step in Gram-positive recognition cascade, or may be directly involved in PGN recognition.^{34,35} Compared to the T7 lysozyme structure, PGRP-LB shows a distinctive back face through addition of two β -strands in its N-terminal region.¹⁸ PGRP-SA shows structural conservation of this specific region of the PGRP domain. This unique back face has been proposed as a hot spot for protein-protein interaction for downstream signal recognition, but it remains to be confirmed if and how this segment is indeed involved in any interaction between GNBP-1 and PGRP-SA and whether the complex is involved in any hydrolytic events.

Materials and Methods

Expression and purification

The full-length cDNA of *Drosophila* PGRP-SA was modified by PCR by using overhang primers to allow the addition of a C-terminal six-histidine tag (5' primer GAATTCGAAATGCAGCCGGTTCGATTCCG, 3' primer GGATCCTTAGTGATGGTGATGGTGATGGGGATTGA GAGCCAGTGC). The fragment was cloned between the EcoRI and BamHI sites of the expression vector pRMHa3 and co-transfected with a G418 resistance gene into *D. melanogaster* SC2 cells by calcium phosphate precipitation. Stable cell lines were derived from G418 selection and kept under selection in Schneider's medium (GIBCO/BRL) containing 500 μ g/ml of G418. Stable transfectants of PGRP-SA were then expanded in serum-free medium (Insect Xpress, BioWhittaker) and protein expression was induced over a period of three to five days by addition of 0.7 mM copper sulfate. After collection, concentration and dialysis of the cell supernatant, the recombinant protein was recovered by metal affinity chromatography on a Ni-NTA resin (Qiagen) by elution with 200 mM imidazole (pH 8.0)

and phosphate-buffered saline (PBS). The PGRP-SA-containing fractions were pooled and purified to homogeneity by size-exclusion chromatography (Superdex 200 10/30, Pharmacia) in 20 mM Tris-HCl (pH 8.0), 50 mM NaCl on FPLC (Pharmacia). Minor contaminants were then removed by anion-exchange chromatography (Mono Q 5/5, Pharmacia) in 20 mM Tris-HCl (pH 8.0); PGRP-SA was then eluted with addition of about 220 mM NaCl. Protein-containing fractions were analyzed by reducing 4%–20%-gradient SDS/PAGE (Bio-rad) and the PGRP-SA fractions were pooled and concentrated by centrifugation to about 15 mg/ml.

Crystallography

PGRP-SA crystals were grown by the sitting-drop method at 22 °C by mixing 2 µl of protein solution with 2 µl of reservoir solution containing 1.3 M to 1.5 M Li₂SO₄ and 0.1 M Mes (pH 5.5 to 6.0). A two-step seeding method was used to obtain crystals of a suitable size for diffraction experiments. Crystal clusters were crushed and seeds transferred to clear drops at lower sulfate concentrations. Small single, plate-shaped crystals (less than 30 µm) were obtained and then transferred to freshly made protein drops where they attained a final size of about 100 µm in the largest dimension. The crystal used for the X-ray data collection was cooled directly in liquid nitrogen after a short transfer in a cryo-protecting solution containing 1.8 M Li₂SO₄, 0.1 M Mes (pH 6.0) and 15% (v/v) ethylene glycol. A single monochromatic data set was collected at the Advanced Light Source (ALS) beam line 8.2.1, at a wavelength of 1.069 Å using a two-dimensional ADSC Q210 detector. Data were processed using the HKL2000 package including DENZO and SCALEPACK³⁶ (Table 1).

The PGRP-SA structure was determined by molecular-replacement method with AMoRe³⁷ from the CCP4 suite³⁸ using the 2.0 Å crystal structure of PGRP-LB (PDB code 1OHT,¹⁸) as a starting model. A simulated annealing refinement step was first performed on the incomplete model, for which sequence differences were replaced by alanine, using CNS version 1.1.³⁹ The final model was obtained after several rounds of alternate manual building with O⁴⁰ and refinement with Refmac5⁴¹ of CCP4 using a maximum-likelihood target and data in the 45–1.56 Å resolution range. The water molecules were constructed automatically with ARP/wARP version 5.0.⁴² Finally, as negative electron density in the $F_o - F_c$ difference map was found for both disulfide bonds, they were refined as partially reduced. The final model consists of 184 residues including the entire C-terminal six-histidine tag, 389 water molecules, two sulfate ions from the crystallization solution and four ethylene glycol molecules from the cryo-protecting solution. The final refinement statistics are shown in Table 1.

MOLSCRIPT,⁴³ Bobscript⁴⁴ and Raster3D⁴⁵ were used to generate the structure Figures and ESPript2.2⁴⁶ for the sequence alignment Figure.

Protein Data Bank accession numbers

The structure factors and coordinates of *Drosophila* PGRP-SA are deposited in the Protein Data Bank (PDB) under the accession number 1SXR.

Acknowledgements

We thank Jean-Marc Reichhart of University of Strasbourg (France) for providing cDNA of PGRP-SA, the Advance Light Source (ALS) at Berkeley and in particular the beamline 8.2.1 staff for assistance. We are also indebted to R. Stanfield for her assistance on synchrotron trip. This work is supported, in part, by National Institutes of Health Grants CA58896 (I.A.W.), AI42266 (I.A.W.) and GM037696 (L.T.) and by the Skaggs Institute for Chemical Biology.

References

- Janeway, C. A., Jr & Medzhitov, R. (2002). Innate immune recognition. *Annu. Rev. Immunol.* **20**, 197–216.
- Schleifer, K. H. & Stackebrandt, E. (1983). Molecular systematics of prokaryotes. *Annu. Rev. Microbiol.* **37**, 143–187.
- Gupta, D., Kirkland, T. N., Viriyakosol, S. & Dziarski, R. (1996). CD14 is a cell-activating receptor for bacterial peptidoglycan. *J. Biol. Chem.* **271**, 23310–23316.
- Takeuchi, O. & Akira, S. (2001). Toll-like receptors; their physiological role and signal transduction system. *Int. Immunopharmacol.* **1**, 625–635.
- Inohara, N., Ogura, Y. & Nunez, G. (2002). Nods: a family of cytosolic proteins that regulate the host response to pathogens. *Curr. Opin. Microbiol.* **5**, 76–80.
- Ferrandon, D., Imler, J. L. & Hoffmann, J. A. (2004). Sensing infection in *Drosophila*: Toll and beyond. *Semin. Immunol.* **16**, 43–53.
- Dziarski, R. (2004). Peptidoglycan recognition proteins (PGRPs). *Mol. Immunol.* **40**, 877–886.
- Dziarski, R. (2003). Recognition of bacterial peptidoglycan by the innate immune system. *Cell. Mol. Life Sci.* **60**, 1793–1804.
- Yoshida, H., Kinoshita, K. & Ashida, M. (1996). Purification of a peptidoglycan recognition protein from hemolymph of the silkworm, *Bombyx mori*. *J. Biol. Chem.* **271**, 13854–13860.
- Werner, T., Liu, G., Kang, D., Ekengren, S., Steiner, H. & Hultmark, D. (2000). A family of peptidoglycan recognition proteins in the fruit fly *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA*, **97**, 13772–13777.
- Liu, C., Xu, Z., Gupta, D. & Dziarski, R. (2001). Peptidoglycan recognition proteins: a novel family of four human innate immunity pattern recognition molecules. *J. Biol. Chem.* **276**, 34686–34694.
- Kang, D., Liu, G., Lundstrom, A., Gelius, E. & Steiner, H. (1998). A peptidoglycan recognition protein in innate immunity conserved from insects to humans. *Proc. Natl Acad. Sci. USA*, **95**, 10078–10082.
- Liu, C., Gelius, E., Liu, G., Steiner, H. & Dziarski, R. (2000). Mammalian peptidoglycan recognition protein binds peptidoglycan with high affinity, is expressed in neutrophils, and inhibits bacterial growth. *J. Biol. Chem.* **275**, 24490–24499.
- Hoffmann, J. A. & Reichhart, J. M. (2002). *Drosophila* innate immunity: an evolutionary perspective. *Nature Immunol.* **3**, 121–126.
- Gottar, M., Gobert, V., Michel, T., Belvin, M., Duyk, G., Hoffmann, J. A. *et al.* (2002). The *Drosophila* immune response against Gram-negative bacteria is

- mediated by a peptidoglycan recognition protein. *Nature*, **416**, 640–644.
16. Takehana, A., Katsuyama, T., Yano, T., Oshima, Y., Takada, H., Aigaki, T. & Kurata, S. (2002). Over-expression of a pattern-recognition receptor, peptidoglycan-recognition protein-LE, activates imd/relish-mediated antibacterial defense and the prophenol-oxidase cascade in *Drosophila* larvae. *Proc. Natl Acad. Sci. USA*, **99**, 13705–13710.
 17. Michel, T., Reichhart, J. M., Hoffmann, J. A. & Royet, J. (2001). *Drosophila* Toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein. *Nature*, **414**, 756–759.
 18. Kim, M. S., Byun, M. & Oh, B. H. (2003). Crystal structure of peptidoglycan recognition protein LB from *Drosophila melanogaster*. *Nature Immunol.* **4**, 787–793.
 19. Cheng, X., Zhang, X., Pflugrath, J. W. & Studier, F. W. (1994). The structure of bacteriophage T7 lysozyme, a zinc amidase and an inhibitor of T7 RNA polymerase. *Proc. Natl Acad. Sci. USA*, **91**, 4034–4038.
 20. Wang, Z. M., Li, X., Cocklin, R. R., Wang, M., Fukase, K., Inamura, S. *et al.* (2003). Human peptidoglycan recognition protein-L is an *N*-acetylmuramoyl-L-alanine amidase. *J. Biol. Chem.* **278**, 49044–49052.
 21. Coleman, J. E. (1998). Zinc enzymes. *Curr. Opin. Chem. Biol.* **2**, 222–234.
 22. Gelius, E., Persson, C., Karlsson, J. & Steiner, H. (2003). A mammalian peptidoglycan recognition protein with *N*-acetylmuramoyl-L-alanine amidase activity. *Biochem. Biophys. Res. Commun.* **306**, 988–994.
 23. Mellroth, P., Karlsson, J. & Steiner, H. (2003). A scavenger function for a *Drosophila* peptidoglycan recognition protein. *J. Biol. Chem.* **278**, 7059–7064.
 24. Rawlings, N. D. & Barrett, A. J. (1994). Families of serine peptidases. *Methods Enzymol.* **244**, 19–61.
 25. Nardini, M. & Dijkstra, B. W. (1999). Alpha/beta hydrolase fold enzymes: the family keeps growing. *Curr. Opin. Struct. Biol.* **9**, 732–737.
 26. Dodson, G. & Wlodawer, A. (1998). Catalytic triads and their relatives. *Trends Biochem. Sci.* **23**, 347–352.
 27. Gaboriaud, C., Serre, L., Guy-Crotte, O., Forest, E. & Fontecilla-Camps, J. C. (1996). Crystal structure of human trypsin 1: unexpected phosphorylation of Tyr151. *J. Mol. Biol.* **259**, 995–1010.
 28. Fulop, V., Bocskei, Z. & Polgar, L. (1998). Prolyl oligopeptidase: an unusual beta-propeller domain regulates proteolysis. *Cell*, **94**, 161–170.
 29. Larsen, N. A., Turner, J. M., Stevens, J., Rosser, S. J., Basran, A., Lerner, R. A. *et al.* (2002). Crystal structure of a bacterial cocaine esterase. *Nature Struct. Biol.* **9**, 17–21.
 30. Wei, Y., Schottel, J. L., Derewenda, U., Swenson, L., Patkar, S. & Derewenda, Z. S. (1995). A novel variant of the catalytic triad in the *Streptomyces scabies* esterase. *Nature Struct. Biol.* **2**, 218–223.
 31. Tjalsma, H., Stover, A. G., Driks, A., Venema, G., Bron, S. & van Dijk, J. M. (2000). Conserved serine and histidine residues are critical for activity of the ER-type signal peptidase SipW of *Bacillus subtilis*. *J. Biol. Chem.* **275**, 25102–25108.
 32. Talarico, T. L., Dev, I. K., Bassford, P. J., Jr & Ray, P. H. (1991). Inter-molecular degradation of signal peptidase I *in vitro*. *Biochem. Biophys. Res Commun.* **181**, 650–656.
 33. Craik, C. S., Roczniak, S., Largman, C. & Rutter, W. J. (1987). The catalytic role of the active site aspartic acid in serine proteases. *Science*, **237**, 909–913.
 34. Gobert, V., Gottar, M., Matskevich, A. A., Rutschmann, S., Royet, J., Belvin, M. *et al.* (2003). Dual activation of the *Drosophila* toll pathway by two pattern recognition receptors. *Science*, **302**, 2126–2130.
 35. Pili-Floury, S., Leulier, F., Takahashi, K., Saigo, K., Samain, E., Ueda, R. & Lemaitre, B. (2004). *In vivo* RNA interference analysis reveals an unexpected role for GGBP1 in the defense against Gram-positive bacterial infection in *Drosophila* adults. *J. Biol. Chem.* **279**, 12848–12853.
 36. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
 37. Navaza, J. (2001). Implementation of molecular replacement in AMoRe. *Acta Crystallog. sect. D*, **57**, 1367–1372.
 38. Collaborative Computational Project Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
 39. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
 40. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110–119.
 41. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D*, **53**, 240–253.
 42. Lamzin, V. & Wilson, K. (1993). Automated refinement of protein model. *Acta Crystallog. sect. D*, **49**, 129–147.
 43. Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.
 44. Esnouf, R. M. (1997). An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J. Mol. Graph. Model.* **15**, 132–134. see also pp. 112–113.
 45. Merrit, E. A. & Bacon, D. J. (1997). Raster3D: photo-realistic molecular graphics. *Methods Enzymol.* **277**, 505–524.
 46. Gouet, P., Courcelle, E., Stuart, D. I. & Metz, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics*, **15**, 305–308.
 47. Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucl. Acids Res.* **16**, 10881–10890.

Edited by D. Rees

(Received 8 March 2004; received in revised form 13 April 2004; accepted 15 April 2004)

Chapitre III

**Reconnaissance et réponse
antigéniques spécifiques des
lymphocytes T**

Steric Hindrance and Fast Dissociation Explain the Lack of Immunogenicity of the Minor Histocompatibility HA-1^{Arg} Null Allele¹

Eric Spierings,^{2,*} Stéphanie Gras,^{2,3,‡} Jean-Baptiste Reiser,^{2,‡} Bregje Mommaas,^{*} Mathilde Almekinders,^{*} Michel G. D. Kester,[†] Anne Chouquet,[‡] Madalen Le Gorrec,[‡] Jan W. Drijfhout,^{*} Ferry Ossendorp,^{*} Dominique Housset,^{4,‡} and Els Goulmy^{4,*}

The di-allelic HLA-A2 restricted minor histocompatibility Ag HA-1 locus codes for the highly immunogenic HA-1^{His} and the nonimmunogenic HA-1^{Arg} nonapeptides, differing in one amino acid. The HA-1^{His} peptide is currently used for boosting the graft-vs-tumor responses after HLA matched HA-1 mismatched stem cell transplantation; usage of the HA-1^{Arg} peptide would significantly enlarge the applicability for this therapy. Our studies on mechanisms causing the HA-1 unidirectional immunogenicity revealed marginal differences in proteasomal digestion, TAP translocation, and binding affinity, whereas both dissociation rates and structural analyses clearly showed marked differences in the stability of these two HLA-A2 bound alleles. These data provide a rationale for the lack of HA-1^{Arg} peptide immunogenicity essential for the choice of tumor peptides for stem cell-based immunotherapeutic application. *The Journal of Immunology*, 2009, 182: 4809–4816.

Minor histocompatibility Ags play important roles in the outcome of stem cell transplantation (SCT)⁵ and organ transplantation. In HLA-identical SCT minor histocompatibility Ags are involved in the development of graft-versus-host disease and in the graft-versus-tumor reactivity (1). Minor histocompatibility Ags are peptides from polymorphic cellular proteins encoded by autosomal genes and by genes of the Y chromosome (2). Minor histocompatibility peptides are presented by MHC class I and class II molecules and are capable of inducing strong T cell immune responses between minor histocompatibility Ag disparate individuals. One important characteristic of minor histocompatibility Ags is their expression pattern; some minor histocompatibility peptides are ubiquitously expressed on all body

cells, whereas others show a restricted tissue distribution (3). Therefore, these two minor histocompatibility Ag types have discriminative roles in graft-versus-host disease and graft-versus-tumor reactivity: the broad minor histocompatibility peptides are involved in both disorders, whereas the restricted peptides are crucial for the graft-versus-tumor responses. The HA-1 minor histocompatibility Ag is hematopoietic system specific but shows additional expression on epithelial cancer cells while absent on the normal epithelial cell counterpart (4, 5). Consequently, this minor histocompatibility Ag can be applied as an immunotherapeutic tool in SCT such as in cellular adoptive immunotherapy with HA-1-specific CTLs or in boosting the graft-versus-tumor responses with HA-1 peptide (6). Indeed, the HA-1 peptide has recently entered a clinical phase I/II study for boosting the antileukemic and antisolid tumor responses after HLA matched minor histocompatibility Ag HA-1 mismatched SCT for hematologic malignancies and solid tumors, respectively (E. Goulmy and L. Hambach, unpublished observations).

The first identified and most studied HA-1 T cell epitope is an HLA-A2-restricted nonapeptide derived from a protein encoded by a polymorphic locus comprising two distinct alleles that differ by one amino acid (6). The allele with a histidine at position 3 (HA-1^{His}, VLHDDLLEA) functions as a T cell epitope in the context of HLA-A2, whereas the other allele with arginine at position 3 (HA-1^{Arg}, VLRDDLLEA) does not. To explain the absence of the HA-1^{Arg} immunogenicity, we earlier determined the IC₅₀ binding values, half maximum lysis values, and the number of copies per cell of both the HA-1^{His} and the HA-1^{Arg} peptides. A noteworthy difference between the two alleles was solely observed in the estimated number of copies/cell, i.e., 80 for HA-1^{His} and less than 5 for HA-1^{Arg} (6). Among the prevalent class I HLA molecules found in the Caucasian population studied, HLA-B60 also showed a strong and stable binding with both alleles of another HA-1 derived nonameric peptide (KECVL^H_RDDLLEA) (7). However, albeit the immunogenicity of these two HLA-B60-restricted peptide alleles remains to be precisely determined, the T cell response to HA-1^{His} is clearly immunodominant in HLA-A2⁺ individuals,

*Department of Immunohematology and Blood Transfusion, [†]Department of Hematology, Leiden University Medical Center, Leiden, The Netherlands; and [‡]Institut de Biologie Structurale Jean-Pierre Ebel, Unité Mixte de Recherche 5075 (CEA, CNRS, UJF, PSB), Grenoble, France

Received for publication December 2, 2008. Accepted for publication February 10, 2009.

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 grants of the Netherlands Organization for Scientific Research, the 6th Framework Programme "Allostem" of the European Commission, and by Grant ANR-05-MIIM-019 from the Agence Nationale de la Recherche.

² E.S., S.G., and J.-B.R. contributed equally to this work.

³ Current address: The Protein Crystallography Unit, Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria 3800, Australia.

⁴ Address correspondence and reprint requests to Dr. Els Goulmy, Department of Immunohematology and Blood Transfusion, Leiden University Medical Center, PO Box 9600, 2300RC Leiden, The Netherlands, E-mail address: e.a.j.m.goulmy@lumc.nl or to Dr. Dominique Housset, Institut de Biologie Structurale Jean-Pierre Ebel, Unité Mixte de Recherche 5075 (CEA, CNRS, UJF, PSB), 41 rue Jules Horowitz, F-38027 Grenoble, France, E-mail address: dominique.housset@ibs.fr

⁵ Abbreviations used in this paper: SCT, stem cell transplantation; ER, endoplasmic reticulum.

when compared with other HLA-A2-restricted minor histocompatibility Ags (8–10).

To elucidate the mechanisms underlying the differential cell surface presentation of the HA-1 allelic peptides, we investigated the impact of the HA-1^{His/Arg} polymorphism on molecular and cellular processes involved in the intracellular generation and stable cell surface presentation of HLA class I-bound peptides. Crucial processes for proper generation of MHC class I-restricted peptides comprise precursor protein ubiquitination, degradation by proteasomes into peptides and peptide translocation into the endoplasmic reticulum (ER), trimming by proteases in the ER as well as the peptide binding capacity to HLA molecules (11). In addition, the peptide immunogenicity has been shown to correlate with both its dissociation rate from and its affinity for the MHC molecule (12–15). Therefore, proteasome-mediated digestion experiments, TAP translocation analyses, and HLA-dissociation assays with HA-1^{His} and HA-1^{Arg} peptides were performed. Moreover, the crystal structures of HLA-A2 in complex with HA-1^{His}, HA-1^{Arg}, or a HA-1 variant with a citrulline residue at position 3 (HA-1^{Cit}) were determined to obtain atomic level insights into the conformation of the HLA-A2/HA-1 peptide complexes. Indeed, because the citrulline differs from the arginine only by the replacement of one of the arginine guanidino nitrogen atoms by an oxygen atom (see data in supplemental Fig. S1),⁶ it is sterically very similar to the arginine but uncharged. Our results exclude a role for Ag processing in preventing HA-1^{Arg} to be presented at the cell surface and clearly show by both the structural and HLA dissociation data that the lack of cell surface expression essentially results from an increased instability of the HA-1^{Arg} allele in the HLA-A2 peptide binding groove. The *in vivo* and *in vitro* stability demonstration of the HLA-A2/HA-1^{His} ligand in combination with its extensively proven immunogenicity strongly support its use as a therapeutic tool in SCT therapy for hematologic malignancies and solid tumors.

Materials and Methods

T cell clones

The HLA-A2-restricted HA-1-specific CD8⁺ CTL clones 3HA15 and 5W38 were both isolated from patients after HLA-identical SCT (8, 16). These CTLs were maintained and used in cytotoxicity and T cell epitope reconstitution assays as previously described (17).

Synthetic peptides

Peptides were synthesized on a Syro II multiple peptide synthesizer (MultiSyntech) using solid-phase Fmoc chemistry and Tentagel resins. Peptides were HPLC purified to >98% on a C18 reversed phase column using a water/acetonitrile/trifluoroacetic acid system. Identity of all synthetic peptides was confirmed using electrospray ionization mass spectrometry on an LCQ mass spectrometer.

Fluorescent peptides were synthesized by replacing leucine at position 7 with a fluorescence-labeled cysteine-derivative (C*). These replacements were shown not affect peptide binding to HLA-A*0201 in competition-based binding assays (data not shown). Fluorescent labeling was performed with 4-(iodoacetamido)-fluorescein (Fluka Chemie) at pH 7.5 (sodium phosphate in a water to acetonitrile ratio of 1:1). The fluorescent peptides were desalted over Sephadex G-10, purified by C18 reverse phase HPLC and analyzed by MALDI mass spectrometry.

Proteasomal digestion assays

20S proteasomes from HeLa, containing mostly constitutive proteasomes, and EBV-LCL ROF, containing mostly immunoproteasomes, were prepared and purified (purity >95%) as described (18). To determine the proteasome-mediated cleavage, 10 µg of 28 or 32 amino acid long, >98% pure HA-1^{His} and HA-1^{Arg} synthetic peptides were incubated with 1 µg of purified 20S (immuno)proteasomes in 100 µl of assay buffer (20 mM HEPES/KOH (pH 7.8), 2 mM MgAc₂, 5 mM DTT) at 37°C for different

time periods ranging from 15 min to 24 h. The proteolysis products were analyzed by electrospray ionization mass spectrometry on a hybrid quadrupole time of flight mass spectrometer (Micromass), equipped with an online nano-electrospray interface as described (19).

In parallel, 1- and 3-h fractions were HPLC fractionated. Aliquots of each fraction were preincubated with ⁵¹Cr-labeled HLA-A2-positive, HA-1^{Arg} EBV-LCL and tested for their ability to reconstitute epitope activity of the HA-1^{His}-specific T cell clone 5W38 at an E:T ratio of 10:1.

Streptolysin O peptide transport assay

In vitro assays of TAP-mediated peptide transport were performed as previously described (20). In short, serial dilutions of peptides of interest were tested for their ability to compete for TAP-dependent translocation of a ¹²⁵I-iodinated model peptide in streptolysin O-permeabilized EBV-LCLs. The results were expressed as IC₅₀ values.

Class I MHC-peptide binding and dissociation assay

The dissociation rates of the recombinant HLA-A2 molecules complexed with fluorescent peptides VLHDDL*EA, VLRDDL*EA, VLFDDL*EA, VLCitDDL*EA were determined using size exclusion HPLC. The fluorescent peptide FLPSDC*FPSV, known to display high affinity for HLA-A2, was included as a positive control. Fluorescent peptides (1 nM) were added to a freshly prepared mixture of recombinant HLA-A2 (20 nM), β₂-microglobulin (150 nM; Sigma-Aldrich), and protease inhibitors (Complete; Roche) in binding buffer (75 mM NaCl, 100 mM sodium phosphate (pH 7), 1 mM CHAPS). The HLA-peptide mixture was kept in the dark and mixed gently for 48 h at room temperature. To determine the level of binding at *t* = 0, 75 µl of each mix was injected on HPLC (column GCP100 SynChropack 250 × 4.0 mm ID; Eichrom Technologies) equipped with a fluorescence detector (em 491 nm, ex 512 nm). Immediately thereafter, the HLA binding of fluorescent peptides was inhibited by adding 10,000-fold nonfluorescent peptides. The level of fluorescent peptides bound to HLA was then determined on HPLC at various time points over a maximum time span of 40 days at room temperature. Additional dissociation experiments with HA-1^{His}, HA-1^{Arg}, HA-1^{Phc}, and HA-1^{Cit} were performed at 37°C. Regression analyses on the percentages HLA-A2-bound peptide at the different time points were executed using the two-phase exponential decay model in GraphPad Prism software (version 4.02).

HLA-A2 binding was tested using a quantitative assay based on the inhibition of binding of the fluorescent-labeled standard peptide FLPSDFC*FPSV as described before (21). In short, various doses of the tested peptides were incubated in the presence of 1 nM fluorescent standard peptide, HLA-A2, and 150 nM β₂-microglobulin for 1 day at room temperature in the dark in a volume of 100 µl, yielding ~40–60% HLA-bound fluorescent standard peptide. The percentage of HLA-bound fluorescence was determined after size exclusion chromatography, as described. The IC₅₀ concentrations were deduced for each peptide by using one-site competition nonlinear regression analysis with the GraphPad Prism software.

Production of recombinant HLA-A2-peptide complexes for crystallographic studies

Both HLA-A2 and β₂-microglobulin chains were produced as previously described by Garboczi et al. (22) and Bodinier et al. (23). In brief, the A245V mutant of the HLA-A2 H chain tagged in C-ter with a biotinylation sequence was cloned into pHN1 expression vector as described in Bodinier et al. (23). Recombinant proteins were produced as inclusion bodies in *Escherichia coli* XA90F'LaqQ1 strain. The inclusion bodies were resuspended in 8 M Urea, 50 mM MES (pH 6.5), 0.1 mM DTT, 0.1 mM EDTA, incubated overnight at 4°C and centrifuged 30 min at 100,000 × *g*. The supernatant was collected and frozen at –80°C. The HLA-peptide complex refolding step was done by flash dilution of a mix of 21 mg of HLA-A2, 10 mg of β₂-microglobulin and 10 mg of the desired synthetic peptide into 350 ml of 100 mM Tris (pH 8.0), 400 mM L-arginine HCl, 2 mM EDTA, 5 mM reduced glutathione, 0.5 mM oxidized glutathione, 2 Complete EDTA-free Cocktail Inhibitor tablets (Roche Diagnostics). The refolding solution was subsequently incubated for 4–5 days at 4°C and concentrated with a 10- or 30-kDa cutoff membrane (Vivascience; Vivascience). The MHC-peptide complex was purified on a MonoQ 5/50 (GE Healthcare BioScience) column with an FPLC system (Purifier 10; GE Healthcare BioScience) equilibrated in a 10 mM Tris (pH 8.0) buffer. The complex was eluted with 100–150 mM NaCl and concentrated on Amicon-10 or Amicon-30 to reach a final protein concentration of 2.5–3.5 mg/ml.

⁶ The online version of this article contains supplemental material.

Crystallization and data collection

The crystals were obtained using a vapor diffusion technique at 20°C and microseeding. The three HLA-A2/HA-1 complexes were crystallized in presence of 12–15% w/v PEG 6000, 0.1 M NaCl and 0.1 M tri-sodium citrate (pH 6.5). Before x-ray diffraction experiments, crystals were soaked into the crystallization solution to which PEG 6000 was added to reach 30% w/v concentration and flash cooled in liquid nitrogen. Data were collected at European Synchrotron Radiation Facilities (Grenoble, France) on either beamline ID29 or ID23. Complete data sets processing was performed using XDS package (24). Crystals belong to space group P2₁ and contain one HLA-A2/HA-1 complex per asymmetric unit (see data in supplemental Table S1).⁶

Structure determination and refinement

The structures were determined by molecular replacement using the program AMoRe (25) and with HLA-A2/MAGE-A4 structure as search model (Protein Data Bank no. 1I4F (26)). Structures of HLA-A2/HA-1 complexes were refined using iterative cycles of manual rebuilding with O (27) or Coot (28) and refinement with REFMAC5 (29) included in CCP4 package (30). Final and detailed refinement statistics are summarized (see supplemental data in Table S1).⁶ The coordinates have been deposited in the Protein Data Bank under entry codes 3FT3 (HLA-A2/HA-1^{His}), 3FT4 (HLA-A2/HA-1^{Arg}), and 3FT2 (HLA-A2/HA-1^{Cit}).

Results

Proteasome-mediated digestion results in equal amounts of HA-1^{His} and HA-1^{Arg} nonameric peptides

Proteasome-mediated digestion was analyzed for its impacts on the absence of the nonimmunogenic HA-1^{Arg} minor histocompatibility Ag at the surface of cells expressing HLA-A2 molecules. Indeed, degradation by proteasomes is the major pathway for the antigenic peptide generation in the cytosol (31). Because the HA-1^{His} minor histocompatibility Ag production is proteasome-dependent (see supplemental data in Fig. S2),⁶ proteasome-mediated cleavage experiments in combination with tandem mass spectrometry were conducted to investigate differential cleavage of the HA-1^{His} and HA-1^{Arg} alleles. The digested products of both the 28-mer HA-1^{His} and the HA-1^{Arg} peptide contained significant and similar material amounts exactly corresponding to the mass of both nonameric peptides using both constitutive (HeLa) and immunoproteasomes (EBV-LCL) (Table I) (also see supplemental Fig. S2).⁶ Indeed, 1% of detected relevant fragments is sufficient to be effectively processed for MHC class I presentation (19, 32, 33). Subsequent tandem mass spectrometry sequencing analysis on the digestion products confirmed the presence of the nonameric peptides HA-1^{His} and HA-1^{Arg} in the corresponding fractions, demonstrating that the HA-1^{His/Arg} polymorphism did not affect the proteasome-mediated generation of the correct sizes of HA-1^{His/Arg} peptides. Thus, proteasomal cleavage of the HA-1^{Arg} peptide cannot account for the absence of the HA-1^{Arg} at the surface of HLA-A2-positive cells.

TAP translocates the HA-1^{His} and HA-1^{Arg} peptides similarly

Antigenic peptide translocation into the ER by TAP molecules is important for adequate peptide cell surface presentation. As TAP can transport peptides with different length up to 15-mers (34), we tested the TAP translocation of 9- to 15-mer HA-1^{His/Arg} peptides. TAP translocated the nonameric HA-1^{His/Arg} peptides in equal quantities, as was the case with several longer HA-1^{His/Arg} variants. Only the 12-mer HA-1^{Arg} peptide (KECVLRDDLLEA) was transported less efficiently as compared with the corresponding HA-1^{His} peptide. However, because these 12-mer peptides were not generated by *in vitro* proteasome-mediated digestion of 28- or 32-mer HA-1^{His} and HA-1^{Arg} peptides, this difference is likely to be irrelevant. (Table II). Thus, the differential HA-1^{His} and HA-1^{Arg} cell surface presentation is not caused by differences in TAP translocation efficiency.

Table I. The HA-1^{His} and HA-1^{Arg} nonameric peptide are equally produced by the proteasome

	Percentage of Fragment ^c	
	HeLa	EBV-LCL
HA-1 ^{His} <i>a</i>		
GLEKLKECVL	0.00	1.65
GLEKLKECVLHD	0.83	1.26
DLLEARRPRAHE	1.27	1.36
DLLEARRPRA	1.20	0.58
LEKLKECVLHDDL	1.25	1.05
VLHDDLLEA	3.24	1.02
HA-1 ^{Arg} <i>b</i>		
GLEKLKECVL	1.11	1.73
RDDLLEARRPRAHE	0.82	0.53
RDDLLEARRPRAH	0.48	0.34
RDDLLEARRPRA	0.43	0.54
GLEKLKECVLRD	2.00	1.65
DLLEARRPRAHE	1.53	0.28
DLLEARRPRAH	0.48	0.34
DLLEARRPRA	1.71	0.71
DLLEARRPR	0.25	0.22
LEARRPRAHE	0.52	0.00
LEARRPRAH	0.54	0.00
VLRDDLLEA	2.58	1.05

^a Proteasome-mediated *in vitro* digestion of a 28 aa long HA-1^{His} peptide GLEKLKECVLHDDLLEARRPRAHECLGE.

^b Proteasome-mediated *in vitro* digestion of a 28 aa long HA-1^{Arg} peptide GLEKLKECVLRDDLLEARRPRAHECLGE.

^c 20S proteasomes, derived from HeLa cells (constitutive proteasomes) and by EBV-LCL cells (immunoproteasomes), was performed as described in *Materials and Methods*. Fragments listed indicate cleavage within the epitope or fragments that match the nonameric CTL epitope HA-1^{His} and its nonimmunogenic allelic counterpart HA-1^{Arg} (marked in bold). Other fragments that add up to 100%, but which are not related to epitope formation, are not shown. Percentage is calculated based upon the total amount of HA-1-related digestion products in each experiment. All fragments have been grouped per cleavage site.

HLA-A2/HA-1^{Arg} complexes dissociate faster than HLA-A2/HA-1^{His} complexes

The very low level of the HLA-A2/HA-1^{Arg} peptide complex on the cell surface may be caused by its dissociation rate. We therefore compared the HA-1^{His} and HA-1^{Arg} peptide dissociation rates in a competition-based cell-free HLA binding assay using recombinant HLA-A2 molecules and fluorescent analogs of the HA-1^{His} and HA-1^{Arg} peptides (Fig. 1). HLA-A2/HA-1^{His} complexes are relatively stable at room temperature ($t_{1/2} = 2593$ min). In contrast, the HLA-A2/HA-1^{Arg} complexes dissociated rapidly ($t_{1/2} = 29$ min) and no intact HLA-A2/HA-1^{Arg} complexes were found after 2 h. These relative differences in dissociation rates between

Table II. *In vitro* TAP translocation of HA-1^{His} and HA-1^{Arg} peptides^a

Sequence	Amino Acid Length	IC ₅₀ , μM
KECVLHDDLLEARRP	15	>100
KECVLRDDLLEARRP	15	>100
KECVLHDDLLEA	12	8
KECVLRDDLLEA	12	53
ECVLHDDLLEA	11	>100
ECVLRDDLLEA	11	45
CVLHDDLLEA	10	92
CVLRDDLLEA	10	87
VLHDDLLEA	9	17
VLRDDLLEA	9	10

^a Different lengths of HA-1^{His} and HA-1^{Arg} peptides were tested for their ability to compete for translocation of the radio-iodinated model peptide TVNKTERAY in streptolysin O-permeabilized EBV-LCLs. The IC₅₀ values represent the concentration of the test peptide that inhibits the transport of the model peptide with 50%. The HLA-A*0201-restricted nonameric HA-1^{His} CTL epitope is indicated in bold.

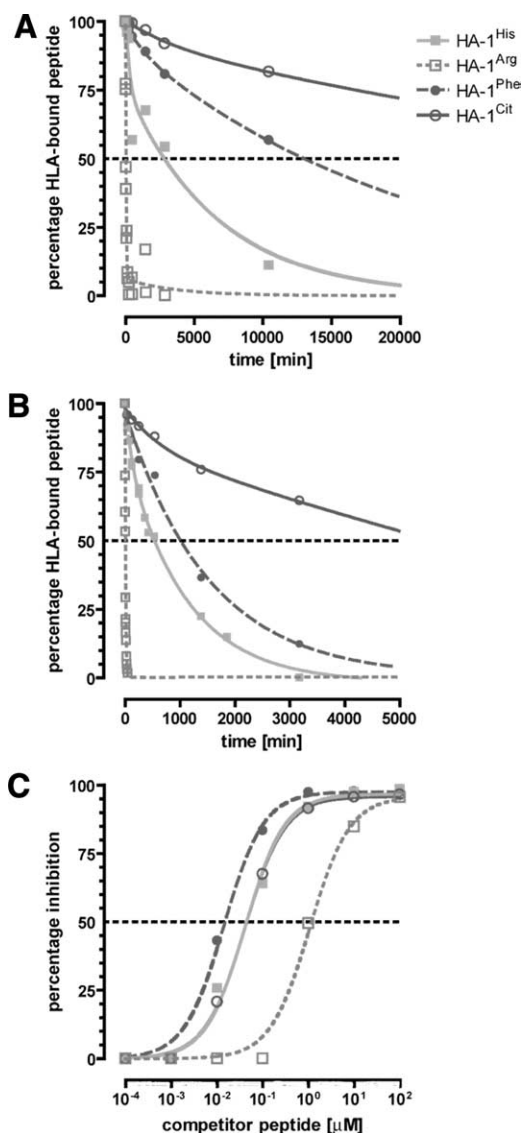


FIGURE 1. HA-1 peptide variants have different HLA-A2 affinity and kinetics. Peptides HA-1^{His} (■ continuous line), HA-1^{Arg} (□ dashed line), HA-1^{Phe} (● dotted line), HA-1^{Cit} (○ continuous line) are shown. *A*, Dissociation rates of HA-1^{His} and its p3 variants at room temperature. Data are expressed as the percentage of peptide in complex with HLA-A2. All samples were analyzed up to 60,000 min. *B*, Dissociation rates of HA-1^{His} and its p3 variants at 37°C. Data are presented as the percentage of fluorescent peptide in complex with HLA-A2. All samples were analyzed up to 9000 min. *C*, Competition based binding of the HA-1^{His} peptide and its p3 variants using a fluorescent-labeled reference peptide. Percentage indicates the binding inhibition of the fluorescent-labeled reference peptide. All experiments were performed as described in *Materials and Methods*.

the two HA-1 alleles were also observed at 37°C ($t_{1/2} = 536$ min for HA-1^{His} and $t_{1/2} = 8$ min for HA-1^{Arg}) (Table III). Thus, a high dissociation rate of the HA-1^{Arg} peptide, as compared with that of the HA-1^{His} peptide, may account for the absence of the HA-1^{Arg} peptide presented by HLA-A2 on the cell surface.

Binding characterization of the HA-1 peptide variants

We questioned whether the high HA-1^{Arg} peptide off-rate might be caused by either steric hindrance induced by the arginine side chain size or by the arginine guanidinium group positive charge at position 3. In this experiment, we substituted HA-1 peptide position 3 by a phenylalanine (HA-1^{Phe}) or by a citrulline (HA-1^{Cit})

Table III. Dissociation rates on cell surface between HA-1 alleles^a

Peptide	BIMAS Score	Half-life RT, min	Half-life 37°C, min	IC ₅₀ , nM
VLHDDLLEA	79.642	2593	536	44
VLRDDLLEA	6.053	29	8	1171
VLFDLLEA	294.675	11985	1011	15
VLCitDDLLEA	n.t.	47348	5622	44

^a Overview of the estimate of half time of disassociation as predicted by BIMAS (http://www-bimas.cit.nih.gov/molbio/hla_bind/), experimentally observed half-life values (in minutes) of the HLA-peptide complexes at room temperature (RT) and 37°C, and the IC₅₀ values (in nanomoles) derived from the competition-based HLA binding assay expressed. The half time of disassociation could not be estimated (n.t.) by BIMAS because the citrulline residue is not included in the algorithms.

residue. The results of the competition-based binding assay showed a slightly stronger interaction with the HLA-A2 molecules for HA-1^{Phe} as compared with HA-1^{His} (Fig. 1C), whereas HA-1^{Cit} was very similar to HA-1^{His} in this respect. Consistently with the dissociation data, the HA-1^{Arg} peptide showed a significant weaker interaction with the HLA-A2 molecules compared with HA-1^{His} peptide (Fig. 1C and Table III), in agreement with binding data obtained previously (6). The HA-1^{Arg} peptide IC₅₀ is clearly the weakest binder of the four variants tested. As shown in Fig. 1B, the HLA-A2-peptide complex half-life with HA-1^{Phe} and HA-1^{Cit} peptides is 4.6 and 18 times longer, respectively, than that with HA-1^{His} peptide, whereas the HA-1^{Arg} peptide exhibits the shortest half-life at both temperatures with two orders of magnitude difference compared with HA-1^{His} (Table III).

If all HA-1 variants are assumed to adopt the same side chain conformation at position 3, these binding data suggest that the side chain size could not directly account for the HA-1^{Arg} allele instability because both HA-1^{Cit} and HA-1^{Phe} peptides are more stable than the HA-1^{His} peptide. They also indicate that the positive charge removal by substitution of the arginine residue by a citrulline residue seems to enhance the stability of the HLA-A2/HA-1 complex dramatically. Structural data presented below show that the citrulline side chain adopts a different conformation and lead to a completely different interpretation of the binding data.

Crystal structure determination of the HLA-A2/HA-1^{His}, HLA-A2/HA-1^{Arg}, and HLA-A2/HA-1^{Cit} complexes

To elucidate the structural bases for the difference in the HA-1 peptide binding capacities, the structures of HLA-A2 in complex with HA-1^{His}, HA-1^{Arg}, and HA-1^{Cit} have been determined by x-ray crystallography at 1.95, 1.90, and 1.80 Å resolution, respectively. Overall, the peptide backbones are very well defined in the electron density (Fig. 2, A–C) and their conformations are similar for the three HA-1 variants with root mean square differences on C α positions ranging from 0.23 Å to 0.41 Å (the superposition is performed on the HLA-A2 $\alpha 1\alpha 2$ domain). This conformation is closest to the one adopted by the HIV p17 Gag peptide (SLYN-TVATL, Protein Data Bank entry 1T21 (35)) with a root mean square difference of 0.38 Å with the HA-1^{His} peptide. The peptide solvent accessible surfaces are 343 Å² (HA-1^{His}), 368 Å² (HA-1^{Arg}) and 381 Å² (HA-1^{Cit}) with the most solvent exposed residues found at positions 4, 5, 7, and 8. The five other positions form the main and secondary anchors. The HA-1^{His} histidine at position 3 is essentially buried in the D pocket (only 5 Å² of solvent accessible surface), as defined by Saper et al. (36). Its side chain contacts HLA-A2 residues Lys^{66H}, His^{70H}, Tyr^{99H}, Gln^{155H}, Leu^{156H}, and Tyr^{159H} (Fig. 2D). The histidine ring is almost exactly superimposed on the Tyr^{3P} aromatic ring of HLA-A2-p17 Gag complex. The HA-1^{Arg} arginine side chain is roughly superposed to that

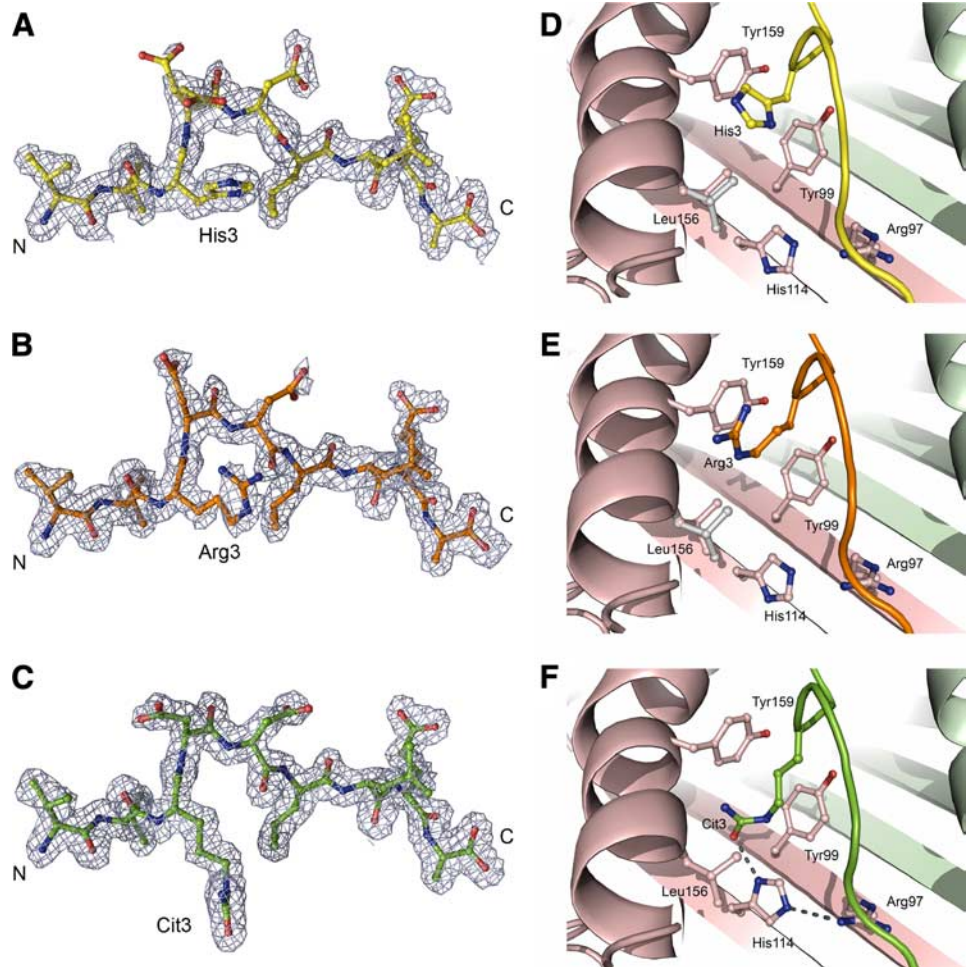


FIGURE 2. Crystal structure of HA-1^{His}, HA-1^{Arg}, and HA-1^{Cit} peptides in complex with HLA-A2. *A–C*, $F_{\text{obs}} - F_{\text{calc}}$ electron density map contoured at the 1σ level for the three HA-1 variants. HA-1^{His} is depicted in yellow (*A*), HA-1^{Arg} in orange (*B*), HA-1^{Cit} in green (*C*). *D–F*, View of the HA-1 peptide in the HLA-A2 peptide binding groove. The side chain of the peptide position 3 and the HLA-A2 residues that interact with it are shown as ball and stick. The HLA-A2 $\alpha 1$ and $\alpha 2$ domains are depicted in green and pink, respectively. The HA-1^{His} (*D*), HA-1^{Arg} (*E*), and HA-1^{Cit} (*F*) peptides are shown in yellow, orange, and green, respectively.

of the histidine in HA-1^{His}, with its guanidinium group pointing toward the surface (Fig. 2*E*). Up to the N ϵ atom, it superposes fairly well the Lys^{3P} of HIV-1 reverse transcriptase peptide (residues 309–317) bound to HLA-A2 (Protein Data Bank entry 1hhj (37)). However, its electron density is much weaker than that of the HA-1^{His} variant, indicating an increased mobility. Accordingly, the arginine side chain atomic occupancy had to be set to 0.5 for the HA-1^{Arg} variant refinement, whereas a full occupancy was used for the HA-1^{His} and HA-1^{Cit} variants. The arginine side chain contacts the same HLA-A2 amino acids and its solvent accessible surface is rather limited (25 Å² of solvent accessible surface). The citrulline side chain conformation differs unexpectedly from that adopted by the arginine. It is buried in the HLA-A2 peptide binding groove (only 4 Å² of solvent accessible surface) and points toward the bottom of the peptide binding groove, the Cit3^P O η 1 atom forming one hydrogen bond with His^{114H} N δ 1 and the Cit3^P N ϵ making a special hydrogen bond with the π -electron of Tyr^{99H} aromatic ring (Fig. 2*F*). Its electron density is very well defined all along the side chain, corroborating its stabilization through a dense network of interatomic interactions.

Overall, the structural impact of the histidine/arginine polymorphism on the peptide structure is limited to a small systematic shift of the peptide toward the $\alpha 1$ HLA-A2 helix. This shift is consistent with the larger size of the arginine side chain. However, the D pocket seems not optimal to welcome the arginine side chain: the charge of the guanidinium group is neither compensated by negatively charged HLA-A2 residues nor by a hydrogen bond network. Moreover, there is a moderate but significant steric hindrance generated by short contacts between the Arg^{3P} C δ and N ϵ atoms and

the Leu^{156H} C δ 2 (3.19 and 3.09 Å, respectively) and the Arg^{3P} N ϵ NH group pointing toward the hydrophobic Leu^{156H} (Fig. 2*E*). Last, despite their size and chemical similarities, the arginine and the citrulline do not interact with the same HLA-A2 amino acids and this difference will impact the analysis of the binding data mentioned.

The analysis of the thermal displacement parameters (B-factors) shows that comparatively to the HLA-A2 $\alpha 1\alpha 2$ domain, the HA-1^{Cit} variant is the most stable and the HA-1^{Arg} peptide the most

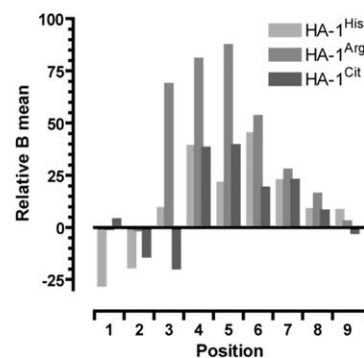


FIGURE 3. Histogram of relative mean B factors for the HA-1 peptide variants. For each peptide residue, $100 \times (\langle B \rangle_{\text{residue}} - \langle B \rangle_{\text{HLA-A2 } \alpha 1\alpha 2}) / (\langle B \rangle_{\text{HLA-A2 } \alpha 1\alpha 2})$ is calculated. A positive value represents a peptide residue more mobile than the overall HLA-A2 $\alpha 1\alpha 2$ domain. A negative value represents a peptide residue more stable than the overall HLA-A2 $\alpha 1\alpha 2$ domain.

mobile variant (Fig. 3), corroborating the observed electron density. The difference is the most contrasted at position 3, for which the citrulline is even slightly more stable than the $\alpha 1\alpha 2$ HLA domain, whereas the histidine is almost as stable as the $\alpha 1\alpha 2$ domain and the arginine much more mobile. These B-factors are very consistent with the number of contacts that side chains at position 3 establish with the HLA-A2 molecule: the one that forms the largest number of contacts with the HLA-A2 molecule and that is the less exposed to the solvent, i.e., the HA-1^{Cit} variant, is also the most stable.

Discussion

The observed unidirectional immunogenicity of the di-allelic minor histocompatibility Ag HA-1 locus restricts its usage for immunotherapeutic purposes to the one immunogenic HA-1^{His} allele. Therefore, we aimed at analyzing the cause of the HA-1 allele differential immunogenicity. Most di-allelic minor histocompatibility Ags display a unidirectional immune response (38) and the phenomenon of “one-allele immunogenicity” can generally be attributed to the lack of peptide presentation at cell surface. The latter can result from gene deletion (i.e., the minor histocompatibility Ag UGT2B17 (39)), from frame-shifts due to nucleotide insertion (i.e., the minor histocompatibility Ag LRH-1 (40)), from alternative splicing (i.e., the minor histocompatibility Ag ACC-6 (41)), or from the introduction of a translation termination codon (i.e., the minor histocompatibility Ag PANE1 (42)). Other mechanisms responsible for one-allele immunogenicity are due to cell biological processes such as differential Ag processing by the proteasome (i.e., the minor histocompatibility Ag HA-3 (43)) and differential Ag transport into the ER by TAP (i.e., minor histocompatibility Ag HA-8 (44)).

So far, the HA-1^{Arg} allele nonimmunogenicity has remained unexplained. Our results rule out a role for proteasome-mediated cleavage resulting in the absence of the HA-1^{Arg} peptide at the cell surface because no differential effect of the HA-1^{His/Arg} polymorphism on proteasome-mediated digestion was observed. Our results on cleavage sites are in line with previous findings that once the relevant peptides are generated by proteasome mediated cleavage, other cleavage sites found within the epitope weigh less heavily for functional peptides cell surface presentation (19). Moreover, as the exact lengths of HLA-A2 binding nonameric peptides were generated by proteasomes, it was not necessary to study the possibility of an impact of the HA-1^{His/Arg} polymorphism on N-terminal trimming by amino peptidases in the ER (45–48). Our data also clearly indicate that the loss of HA-1^{Arg} surface presentation is not caused by differential TAP translocation.

Low binding affinity to HLA-A2 molecules certainly influences the absence of detectable HA-1^{Arg} peptide amounts on the cell surface. The HA-1^{Arg} peptide binds with 27-fold lower affinity to recombinant HLA-A2 molecules as compared with the HA-1^{His} peptide. Although peptide binding to recombinant HLA molecules may not always be representative for intracellular peptide binding to TAP-associated HLA molecules in the ER, a 10-fold lower affinity to intracellular TAP-associated HLA-A2 has also been estimated for the HA-1^{Arg} peptide (data not shown). Because both assays detect the net result of association and dissociation, the lower affinity of HA-1^{Arg} for HLA-A2 is likely the result of its 90-fold higher dissociation rates as compared with the HLA-A2/HA-1^{His} complexes. Given the cell surface presentation of 80 HA-1^{His} peptides per cell (6), the affinity differences between the two HA-1 alleles is insufficient to explain the lack of the HA-1^{Arg} presentation because peptide density at the cell surface is not necessarily correlated with their MHC binding affinity (49). Moreover,

when compared with other HLA-A2 bound peptides, HA-1^{His} peptide can be considered as a high binder ($IC_{50} < 50$ nM) and HA-1^{Arg} as moderate binder (50 nM $< IC_{50} < 500$ nM (14)), both being in the affinity range usually observed for effective T cell epitopes. Even low affinity peptides can represent important T cell epitopes (50, 51), most probably slowly dissociating from the MHC, thereby having sufficient Ag on the cell surface. Therefore, the 90-fold increase in dissociation rate is clearly identified as the main reason for the lack of immunogenicity for HA-1^{Arg} allele, as it has been already observed for other peptide-MHC complexes (15). The magnitude of the increased dissociation rate for HA-1^{Arg} clearly accounts for the lack of HLA-A2/HA-1^{Arg} complexes at the cell surface. Most likely, the HLA-A2/HA-1^{Arg} complexes already dissociate intracellularly and never reach the cell surface.

The fast dissociation observed for the HA-1^{Arg} does not prevent the formation of HLA-A2/HA-1^{Arg} crystals at room temperature. This indicates that in the crystallization solution, the complex half-life should be at least as long as the time needed for the crystals to appear in the drops, estimated to last 1–3 days. It confirms that the HLA-A2/HA-1 complex half-life depends on the temperature, as exemplified by Table III and previous experiments that showed that the HA-1^{His} affinity at 20°C is comparable to that of HA-1^{Arg} at 4°C (7). It also suggests that the chemical composition of the protein solution may have an impact.

The structural data on the two naturally occurring HA-1 allelic peptides and the artificial HA-1^{Cit} variant provide a clear rationale for these binding data. All peptides bind HLA-A2 with a similar backbone conformation, albeit with different level of flexibility. More quantitatively, the B-factor analysis shows that the average B-factors for the HA-1^{Arg}, HA-1^{His}, and HA-1^{Cit} peptides are 37%, 16%, and 11% higher than that of the HLA-A2 $\alpha 1\alpha 2$ domain, respectively. The difference is more contrasted at position 3 (Fig. 3). The analysis of the interactions between the peptide and the HLA-A2 molecule indicates that the HA-1^{Cit} variant establishes the largest number of hydrogen bonds and van der Waals contacts (15 hydrogen bonds and 54 C-C contacts < 4 Å), followed by the HA-1^{His} peptide (14 hydrogen bonds and 47 C-C contacts < 4 Å), whereas the HA-1^{Arg} peptides is the less tightly bound (12 hydrogen bonds and 39 C-C contacts < 4 Å). This number of contacts correlate very well with the stability data; high peptide off-rates correspond to complexes in which less energy is involved in the dissociation, i.e., in which fewer interactions have to be broken to dissociate the peptide from the HLA-A2 molecule. Indeed, the lower number of contacts, the high mobility of the arginine in HA-1^{Arg} peptide suggested by the crystallographic data and the steric hindrance generated by the close contacts between Arg^{3P} and Leu^{156H} are corroborating an extremely rapid HLA dissociation rate for the HA-1^{Arg} peptide when compared with the HA-1^{His} peptide. Replacing histidine and arginine residues while maintaining or slightly increasing the size of the substituted amino acids by using HA-1^{Phe} and HA-1^{Cit} variants respectively, leads to a significant increase in $t_{1/2}$ values. For the HA-1^{Phe} variant bound to HLA-A2, the phenylalanine side chain conformation is likely to be very close to that adopted by the histidine in HA-1^{His} peptide. Indeed, among the peptide-HLA-A2 complexes deposited with the Protein Data Bank and whose peptide backbone conformation is similar to that adopted by HA-1^{His} peptide, aromatic residues (phenylalanine, or tyrosine in Protein Data Bank entries 2V2X (52) and 1T21 (35)) can be found at position 3. These residues also fill the D pocket with a conformation very similar to the one adopted by either the histidine or the arginine in HA-1. In this context, burying a slightly larger and hydrophobic aromatic ring in the HLA-A2 D pocket enhances the HA-1 binding as exemplified by the binding data. Because HA-1^{Arg} and HA-1^{Cit} variants adopt very different

side chain conformation for position 3, the difference observed in binding properties cannot be attributed to the sole difference in net charge. Due to its positive charge, the arginine cannot adopt the conformation observed for the citrulline side chain. The His^{114H}, which is hydrogen bonded to the citrulline side chain (Fig. 2F) is positively charged at the crystallization (pH 6.5) and the His^{114H} N δ 1 atom is certainly protonated because the His^{114H} N ϵ 2 atom is not as it is involved in a hydrogen bond with the protonated N η 2 of Arg^{97H}. As a result, the presence of an arginine in place of the observed citrulline would be electrostatically very unfavorable. Therefore, the increased HA-1^{Cit} variant half-life is certainly due to the larger number of interactions formed by citrulline side chain that plunges more deeply into the HLA-A2 D pocket. The arginine guanidinium group positive charge per se is not responsible for the increased the HA-1^{Arg} peptide instability because an immunogenic peptide bearing a lysine at position 3 bound in the D pocket has already been observed (Protein Data Bank entry 1HHJ (37)). It just prevents the arginine to adopt the very stable conformation adopted by the citrulline.

Serendipitously, binding and structural data on HA-A1^{Cit} clearly demonstrate that the citrulline in position 3 behaves as an additional anchor in the HLA-A2 binding groove. Its conformation is reminiscent of that of the arginine side chain in the HLA-B8-restricted EBV epitope (FLRGRAYGL), when bound to HLA-B8 (Protein Data Bank entry 1m05 (53)). In HLA-B8, a lysine or an arginine at position 3 constitutes a main anchor. Because the three first amino acid conformations are very conserved in nonapeptide-HLA-A2 complex structures determined so far, introducing a citrulline at position 3 may constitute a quite general approach to further stabilize a nonapeptide in the HLA-A2 peptide groove. The structural data presented here show that this change has almost no impact on the peptide backbone conformation, and suggest that T cells could thus be likely cross-reactive for HLA-A2-restricted nonapeptide citrulline variants at position 3.

We show for the first time functional peptide processing studies combined with the fully concordant HLA-A2/peptide crystallographic data that are explaining the lack of immunogenicity of one of two allelic peptides. Structural basis for this change of stability have been unveiled. It strongly suggests that the HLA-A2/HA-1^{His} and HLA-A2/HA-1^{Arg} complexes have a half-life, respectively, above and below a threshold defining the immunologic fate of a subtle different peptide variant. It emphasizes that the use of HLA-A2-restricted HA-1-based immunotherapy is restricted to HA-1^{His} patients and that other targets should be identified for the patients bearing the HA-1^{Arg} allele.

Finally, the present data demonstrate how combined functional and structural information on peptides before application in immunotherapeutic tumor therapy are essential to ensure optimal cellular immune responses.

Acknowledgments

We thank J. McCarthy (ID29), D. Flot (ID23-2), and A. McCarthy (ID23-1) for help with synchrotron data collections at the European Synchrotron Radiation Facilities (Grenoble, France), and X. Saulquin and M. Bonneville (INSERM, Nantes, France) for providing the pHN1-HLA-A2 plasmid and the *E. coli* XA90F/LaqQ1 strain.

Disclosures

The authors have no financial conflict of interest.

References

- Goulmy, E. 1997. Human minor histocompatibility antigens: new concepts for marrow transplantation and adoptive immunotherapy. *Immunol. Rev.* 157: 125–140.
- Goulmy, E. 1996. Human minor histocompatibility antigens. *Curr. Opin. Immunol.* 8: 75–81.
- de Bueger, M., A. Bakker, J. J. Van Rood, F. Van der Woude, and E. Goulmy. 1992. Tissue distribution of human minor histocompatibility antigens: ubiquitous versus restricted tissue distribution indicates heterogeneity among human cytotoxic T lymphocyte-defined non-MHC antigens. *J. Immunol.* 149: 1788–1794.
- Fujii, N., A. Hiraki, K. Ikeda, Y. Ohmura, I. Nozaki, K. Shinagawa, F. Ishimaru, K. Kiura, N. Shimizu, M. Tanimoto, and M. Harada. 2002. Expression of minor histocompatibility antigen, HA-1, in solid tumor cells. *Transplantation* 73: 1137–1141.
- Klein, C. A., M. Wilke, J. Pool, C. Vermeulen, E. Blokland, E. Burghart, S. Krostina, N. Wendler, B. Passlick, G. Riethmueller, and E. Goulmy. 2002. The hematopoietic system-specific minor histocompatibility antigen HA-1 shows aberrant expression in epithelial cancer cells. *J. Exp. Med.* 196: 359–368.
- den Haan, J. M., L. M. Meadows, W. Wang, J. Pool, E. Blokland, T. L. Bishop, C. Reinhardus, J. Shabanowitz, R. Offringa, D. F. Hunt, et al. 1998. The minor histocompatibility antigen HA-1: a diallelic gene with a single amino acid polymorphism. *Science* 279: 1054–1057.
- Mommaas, B., J. Kamp, J. W. Drijfhout, N. Beekman, F. Ossendorp, P. Van Veelen, J. Den Haan, E. Goulmy, and T. Mutis. 2002. Identification of a novel HLA-B60-restricted T cell epitope of the minor histocompatibility antigen HA-1 locus. *J. Immunol.* 169: 3131–3136.
- van Els, C. A., J. D'Amato, J. Pool, E. Blokland, A. Bakker, P. J. van Elsen, J. J. van Rood, and E. Goulmy. 1992. Immunogenetics of human minor histocompatibility antigens: their polymorphism and immunodominance. *Immunogenetics* 35: 161–165.
- Goulmy, E., R. Schipper, J. Pool, E. Blokland, J. H. Falkenburg, J. Vossen, A. Gratwohl, G. B. Vogelsang, H. C. van Houwelingen, and J. J. van Rood. 1996. Mismatches of minor histocompatibility antigens between HLA-identical donors and recipients and the development of graft-versus-host disease after bone marrow transplantation. *N. Engl. J. Med.* 334: 281–285.
- Rufer, N., E. Wolpert, C. Helg, J. M. Tiercy, A. Gratwohl, B. Chapuis, M. Jeannel, E. Goulmy, and E. Roosnek. 1998. HA-1 and the SMCY-derived peptide FIDSYICQV (H-Y) are immunodominant minor histocompatibility antigens after bone marrow transplantation. *Transplantation* 66: 910–916.
- Pamer, E., and P. Cresswell. 1998. Mechanisms of MHC class I-restricted antigen processing. *Annu. Rev. Immunol.* 16: 323–358.
- Chen, Y., J. Sidney, S. Southwood, A. L. Cox, K. Sakaguchi, R. A. Henderson, E. Appella, D. F. Hunt, A. Sette, and V. H. Engelhard. 1994. Naturally processed peptides longer than nine amino acid residues bind to the class I MHC molecule HLA-A2.1 with high affinity and in different conformations. *J. Immunol.* 152: 2874–2881.
- Ressing, M. E., A. Sette, R. M. Brandt, J. Ruppert, P. A. Wentworth, M. Hartman, C. Oseroff, H. M. Grey, C. J. Melief, and W. M. Kast. 1995. Human CTL epitopes encoded by human papillomavirus type 16 E6 and E7 identified through in vivo and in vitro immunogenicity studies of HLA-A*0201-binding peptides. *J. Immunol.* 154: 5934–5943.
- Sette, A., A. Vitiello, B. Rehman, P. Fowler, R. Nayarsina, W. M. Kast, C. J. Melief, C. Oseroff, L. Yuan, J. Ruppert, et al. 1994. The relationship between class I binding affinity and immunogenicity of potential cytotoxic T cell epitopes. *J. Immunol.* 153: 5586–5592.
- van der Burg, S. H., M. J. Visseren, R. M. Brandt, W. M. Kast, and C. J. Melief. 1996. Immunogenicity of peptides bound to MHC class I molecules depends on the MHC-peptide complex stability. *J. Immunol.* 156: 3308–3314.
- Goulmy, E., J. W. Gratama, E. Blokland, F. E. Zwaan, and J. J. van Rood. 1983. A minor transplantation antigen detected by MHC-restricted cytotoxic T lymphocytes during graft-versus-host disease. *Nature* 302: 159–161.
- den Haan, J. M., T. Mutis, E. Blokland, A. P. IJzerman, and E. Goulmy. 2002. General T-cell receptor antagonists to immunomodulate HLA-A2-restricted minor histocompatibility antigen HA-1-specific T-cell responses. *Blood* 99: 985–992.
- Sijts, A. J., T. Ruppert, B. Rehmann, M. Schmidt, U. Koszinowski, and P. M. Kloetzel. 2000. Efficient generation of a hepatitis B virus cytotoxic T lymphocyte epitope requires the structural features of immunoproteasomes. *J. Exp. Med.* 191: 503–514.
- Kessler, J. H., N. J. Beekman, S. A. Bres-Vloemans, P. Verdijk, P. A. van Veelen, A. M. Kloosterman-Joosten, D. C. Vissers, G. J. ten Bosch, M. G. Kester, A. Sijts, et al. 2001. Efficient identification of novel HLA-A(*)0201-presented cytotoxic T lymphocyte epitopes in the widely expressed tumor antigen PRAME by proteasome-mediated digestion analysis. *J. Exp. Med.* 193: 73–88.
- Neisig, A., J. Roelse, A. J. Sijts, F. Ossendorp, M. C. Feltkamp, W. M. Kast, C. J. Melief, and J. J. Neefjes. 1995. Major differences in transporter associated with antigen presentation (TAP)-dependent translocation of MHC class I-presentable peptides and the effect of flanking sequences. *J. Immunol.* 154: 1273–1279.
- Ottenhoff, T. H., A. Geluk, M. Toebes, W. E. Benckhuijsen, K. E. van Meijgaarden, and J. W. Drijfhout. 1997. A sensitive fluorometric assay for quantitatively measuring specific peptide binding to HLA class I and class II molecules. *J. Immunol. Methods* 200: 89–97.
- Garboczi, D. N., D. T. Hung, and D. C. Wiley. 1992. HLA-A2-peptide complexes: refolding and crystallization of molecules expressed in *Escherichia coli* and complexed with single antigenic peptides. *Proc. Natl. Acad. Sci. USA* 89: 3429–3433.
- Bodnier, M., M. A. Peyrat, C. Tournay, F. Davodeau, F. Romagne, M. Bonneville, and F. Lang. 2000. Efficient detection and immunomagnetic sorting of specific T cells using multimers of MHC class I and peptide with reduced CD8 binding. *Nat. Med.* 6: 707–710.

24. Kabsch, W. 1993. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* 26: 795–800.
25. Navaza, J. 1994. AMoRe: an automated package for molecular replacement. *Acta Crystallogr. A* 50: 157–163.
26. Hillig, R. C., P. G. Coulie, V. Stroobant, W. Saenger, A. Ziegler, and M. Hulsmeyer. 2001. High-resolution structure of HLA-A*0201 in complex with a tumour-specific antigenic peptide encoded by the MAGE-A4 gene. *J. Mol. Biol.* 310: 1167–1176.
27. Jones, T. A., J. Y. Zou, S. W. Cowan, and M. Kjeldgaard. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* 47(Pt. 2): 110–119.
28. Emsley, P., and K. Cowtan. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* 60: 2126–2132.
29. Murshudov, G. N., A. A. Vagin, and E. J. Dodson. 1997. Refinement of macromolecular structures by the Maximum-Likelihood Method. *Acta Crystallogr. D.* 53: 240–255.
30. Collaborative Computational Project Number 4. 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* 50: 760–763.
31. York, I. A., and K. L. Rock. 1996. Antigen processing and presentation by the class I major histocompatibility complex. *Annu. Rev. Immunol.* 14: 369–396.
32. Ossendorp, F., M. Eggers, A. Neisig, T. Ruppert, M. Groettrup, A. Sijts, E. Mengede, P. M. Kloetzel, J. Neeffjes, U. Koszinowski, and C. Melief. 1996. A single residue exchange within a viral CTL epitope alters proteasome-mediated degradation resulting in lack of antigen presentation. *Immunity* 5: 115–124.
33. Beekman, N. J., P. A. van Veelen, T. van Hall, A. Neisig, A. Sijts, M. Camps, P. M. Kloetzel, J. J. Neeffjes, C. J. Melief, and F. Ossendorp. 2000. Abrogation of CTL epitope processing by single amino acid substitution flanking the C-terminal proteasome cleavage site. *J. Immunol.* 164: 1898–1905.
34. Reits, E. A., A. C. Griekspoor, and J. Neeffjes. 2000. How does TAP pump peptides? insights from DNA repair and traffic ATPases. *Immunol. Today* 21: 598–600.
35. Martinez-Hackert, E., N. Anikeeva, S. A. Kalams, B. D. Walker, W. A. Hendrickson, and Y. Sykulev. 2006. Structural basis for degenerate recognition of natural HIV peptide variants by cytotoxic lymphocytes. *J. Biol. Chem.* 281: 20205–20212.
36. Saper, M. A., P. J. Bjorkman, and D. C. Wiley. 1991. Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 Å resolution. *J. Mol. Biol.* 219: 277–319.
37. Madden, D. R., D. N. Garboczi, and D. C. Wiley. 1993. The antigenic identity of peptide-MHC complexes: a comparison of the conformations of five viral peptides presented by HLA-A2. *Cell* 75: 693–708.
38. Spierings, E., and E. Goulmy. 2005. Expanding the immunotherapeutic potential of minor histocompatibility antigens. *J. Clin. Invest.* 115: 3397–3400.
39. Murata, M., E. H. Warren, and S. R. Riddell. 2003. A human minor histocompatibility antigen resulting from differential expression due to a gene deletion. *J. Exp. Med.* 197: 1279–1289.
40. de Rijke, B., A. van Horsen-Zoetbrood, J. M. Beekman, B. Otterud, F. Maas, R. Woestenenk, M. Kester, M. Leppert, A. V. Schattenberg, T. de Witte, et al. 2005. A frameshift polymorphism in P2X5 elicits an allogeneic cytotoxic T lymphocyte response associated with remission of chronic myeloid leukemia. *J. Clin. Invest.* 115: 3506–3516.
41. Kawase, T., Y. Akatsuka, H. Torikai, S. Morishima, A. Oka, A. Tsujimura, M. Miyazaki, K. Tsujimura, K. Miyamura, S. Ogawa, et al. 2007. Alternative splicing due to an intronic SNP in HMSD generates a novel minor histocompatibility antigen. *Blood* 110: 1055–1063.
42. Brickner, A. G., A. M. Evans, J. K. Mito, S. M. Xuereb, X. Feng, T. Nishida, L. Fairfull, R. E. Ferrell, K. A. Foon, D. F. Hunt, et al. 2006. The PANE1 gene encodes a novel human minor histocompatibility antigen that is selectively expressed in B-lymphoid cells and B-CLL. *Blood* 107: 3779–3786.
43. Spierings, E., A. G. Brickner, J. A. Caldwell, S. Zegveld, N. Tatsis, E. Blokland, J. Pool, R. A. Pierce, S. Mollah, J. Shabanowitz, et al. 2003. The minor histocompatibility antigen HA-3 arises from differential proteasome-mediated cleavage of the lymphoid blast crisis (Lbc) oncoprotein. *Blood* 102: 621–629.
44. Brickner, A. G., E. H. Warren, J. A. Caldwell, Y. Akatsuka, T. N. Golovina, A. L. Zarlino, J. Shabanowitz, L. C. Eisenlohr, D. F. Hunt, V. H. Engelhard, and S. R. Riddell. 2001. The immunogenicity of a new human minor histocompatibility antigen results from differential antigen processing. *J. Exp. Med.* 193: 195–206.
45. Saric, T., S. C. Chang, A. Hattori, I. A. York, S. Markant, K. L. Rock, M. Tsujimoto, and A. L. Goldberg. 2002. An IFN- γ -induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I-presented peptides. *Nat. Immunol.* 3: 1169–1176.
46. Serwold, T., F. Gonzalez, J. Kim, R. Jacob, and N. Shastri. 2002. ERAAP customizes peptides for MHC class I molecules in the endoplasmic reticulum. *Nature* 419: 480–483.
47. Snyder, H. L., J. W. Yewdell, and J. R. Bennink. 1994. Trimming of antigenic peptides in an early secretory compartment. *J. Exp. Med.* 180: 2389–2394.
48. York, I. A., S. C. Chang, T. Saric, J. A. Keys, J. M. Favreau, A. L. Goldberg, and K. L. Rock. 2002. The ER aminopeptidase ERAP1 enhances or limits antigen presentation by trimming epitopes to 8–9 residues. *Nat. Immunol.* 3: 1177–1184.
49. Engelhard, V. H., A. G. Brickner, and A. L. Zarlino. 2002. Insights into antigen processing gained by direct analysis of the naturally processed class I MHC associated peptide repertoire. *Mol. Immunol.* 39: 127–137.
50. Cox, A. L., J. Skipper, Y. Chen, R. A. Henderson, T. L. Darrow, J. Shabanowitz, V. H. Engelhard, D. F. Hunt, and C. L. Slingluff, Jr. 1994. Identification of a peptide recognized by five melanoma-specific human cytotoxic T cell lines. *Science* 264: 716–719.
51. Parkhurst, M. R., M. L. Salgaller, S. Southwood, P. F. Robbins, A. Sette, S. A. Rosenberg, and Y. Kawakami. 1996. Improved induction of melanoma-reactive CTL with peptides from the melanoma antigen gp100 modified at HLA-A*0201-binding residues. *J. Immunol.* 157: 2539–2548.
52. Lee, J. K., G. Stewart-Jones, T. Dong, K. Harlos, K. Di Gleria, L. Dorrell, D. C. Douek, P. A. van der Merwe, E. Y. Jones, and A. J. McMichael. 2004. T cell cross-reactivity and conformational changes during TCR engagement. *J. Exp. Med.* 200: 1455–1466.
53. Kjer-Nielsen, L., C. S. Clements, A. G. Brooks, A. W. Purcell, M. R. Fontes, J. McCluskey, and J. Rossjohn. 2002. The structure of HLA-B8 complexed to an immunodominant viral determinant: peptide-induced conformational changes and a mode of MHC class I dimerization. *J. Immunol.* 169: 5153–5160.

Structural Bases for the Affinity-Driven Selection of a Public TCR against a Dominant Human Cytomegalovirus Epitope¹

Stéphanie Gras,^{2,3*} Xavier Saulquin,^{2†} Jean-Baptiste Reiser,* Emilie Debeaupuis,[†] Klara Echasserieu,[†] Adrien Kissenpfennig,^{4‡} François Legoux,[†] Anne Chouquet,* Madalen Le Gorrec,* Paul Machillot,* Bérandère Neveu,[†] Nicole Thielens,* Bernard Malissen,[‡] Marc Bonneville,^{5†} and Dominique Housset^{5*}

Protective T cell responses elicited along chronic human CMV (HCMV) infections are sometimes dominated by CD8 T cell clones bearing highly related or identical public TCR in unrelated individuals. To understand the principles that guide emergence of these public T cell responses, we have performed structural, biophysical, and functional analyses of an immunodominant public TCR (RA14) directed against a major HLA-A*0201-restricted HCMV Ag (pp65_{495–503}) and selected in vivo from a diverse repertoire after chronic stimulations. Unlike the two immunodominant public TCRs crystallized so far, which focused on one peptide hotspot, the HCMV-specific RA14 TCR interacts with the full array of available peptide residues. The conservation of some peptide-MHC complex-contacting amino acids by lower-affinity TCRs suggests a shared TCR-peptide-MHC complex docking mode and supports an Ag-driven selection of optimal TCRs. Therefore, the emergence of a public TCR of an oligoclonal Ag-specific response after repeated viral stimulations is based on a receptor displaying a high structural complementarity with the entire peptide and focusing on three peptide hotspots. This highlights key parameters underlying the selection of a protective T cell response against HCMV infection, which remains a major health issue in patients undergoing bone marrow transplantation. *The Journal of Immunology*, 2009, 183: 430–437.

Human CMV (HCMV)⁶ is a ubiquitous β -herpesvirus that infects 60–90% of the population. After primary infection, HCMV persists in a latent stage and can undergo transient reactivations. Although HCMV infections or reactivations are usually kept in check by the immune system of immunocompetent individuals, they can cause life-threatening diseases in immunocompromised patients (1). CD8⁺ CTLs have a central

role in controlling HCMV reactivation (2–6). Although CD8⁺ T cells specific for the immediate early gene product IE-1 are found at substantial frequencies in some donors (7, 8), the predominant CTL response is directed against the viral tegument protein pp65 (5, 9). In individuals sharing the widespread HLA-A*0201 allele (referred to as A2), HCMV-specific CTLs recognize the same epitope pp65_{495–503} (NLVPMVATV), hereafter referred to as NLV (5, 10, 11). Albeit polyclonal, the NLV-specific CTL response in healthy donors is usually made of a limited number of peptide-specific CTL clones expressing a restricted set of TCR β variable gene segments that may differ from one individual to another (5, 11, 12). However, a further dramatic reduction of clonal diversity occurs during chronic inflammation (e.g., in rheumatoid arthritis patients) and immunodepression, resulting in the selection of a few dominant clones bearing high-affinity TCRs (13). Because immunodepressed patients are prone to HCMV reactivation, this clonal focusing presumably reflects favored expansion of the best-fit clonotypes along recurrent antigenic stimulations. NLV-specific TCRs that predominate in the latter patients carry several public features (i.e., shared by clonotypes from different individuals), such as restricted V α or V β usage and conserved motifs within the CDR3 loops (13). These public TCRs can be expressed by up to 15% of peripheral blood CD8 T cells of patients undergoing HCMV reactivation, highlighting their major contribution to the NLV-specific response. Recent structural analysis of public TCRs in complex with peptide-MHC (pMHC) revealed unique structural features of the selecting pMHC complex, such as limited solvent accessibility or marked bulging of one peptide residue, that accounts for selection of an homogeneous repertoire (14, 15). However, it remains unclear whether similar rules apply to public T cell responses selected after repeated stimulation, as observed in HCMV infection. To gain insight into the mechanisms underlying clonal focusing along chronic Ag stimulation, we have determined the crystal structure of the immunodominant NLV peptide bound

*Institut de Biologie Structurale Jean-Pierre Ebel, Unité mixte de recherche 5075 (CEA, CNRS, UJF, PSB), Grenoble, France; †INSERM, Unité 892, Institut de Recherche Thérapeutique, Université de Nantes, Nantes, France; and ‡Centre d'Immunologie de Marseille-Luminy, Université de la Méditerranée, INSERM, Unité 631, CNRS, Marseille, France

Received for publication February 18, 2009. Accepted for publication April 29, 2009.

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 study was supported in part by the Agence Nationale de la Recherche (Grant ANR-05-MIIM-019). B.M. and M.B. were also supported by the EPI-PEPVAC European Union Grant. The authors have declared that no competing interests exist.

² S.G. and X.S. equally contributed to the work.

³ Current address: The Protein Crystallography Unit, Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria 3800, Australia.

⁴ Current address: Infection and Immunity Division, Centre for Cancer Research and Cell Biology, School of Biomedical Sciences, Queens University, Belfast, Northern Ireland.

⁵ Address correspondence and reprint requests to Dr. Marc Bonneville, INSERM, Unité 892, Institut de Recherche Thérapeutique, 9 quai Moncousu, Université de Nantes, F-44035 Nantes, France. E-mail address: bonnevil@nantes.inserm.fr, or Dominique Housset, Institut de Biologie Structurale Jean-Pierre Ebel, Unité mixte de recherche 5075 (CEA, CNRS, UJF, PSB), 41 rue Jules Horowitz, F-38027 Grenoble, France. E-mail address: dominique.housset@ibs.fr

⁶ Abbreviations used in this paper: HCMV, human CMV; A2, HLA-A*0201 allele; FLR, FLRGRAYGL; NLV, NLVPMVATV; PEG, polyethylene glycol; pMHC, peptide-MHC; pMHCI, pMHC class I; RA14-NLV-A2, RA14 TCR-NLV-A2 complex; SPR, surface plasmon resonance; TAP, transporter associated with Ag processing.

Copyright © 2009 by The American Association of Immunologists, Inc. 0022-1767/09/\$2.00

to A2 in isolation or in complex with a dominant public TCR (RA14) derived from an immunodepressed patient (13). These structural data, combined with biophysical and functional experiments, allowed us to determine a new and quite extensive peptide readout mode with three peptide hot spots sensed by the RA14 TCR. It also provided a possible mechanism underlying preferred usage of some TRAV, TRBV, CDR3 α , and CDR3 β for the recognition of this immunodominant peptide in an immunodepressed context associated with HCMV reactivation, as is the case for patients with oncohematological diseases undergoing bone marrow transplantation.

Materials and Methods

Plasmid construction

The cDNA encoding the full-length α - and β -chains were cloned into pGEM-T Easy (Promega). The sequence encoding the V and C ectodomain of α - and β -chain were then amplified by PCR using the following primers: 5' sense primers (5'-ggaattccaggaggaatttaaatgatactgaacctggaacaaagctcc-3') or (5'-ggaattccaggaggaatttaaatgggtctcaactcagaccctcaattcc) containing an *EcoRI* restriction site (underlined) and a strong consensus Shine Dalgarno sequence (bold) around the initiation codon to promote optimal translation of the α -chain and β -chain respectively, and 5' anti-sense primers (5'-gctctagattacaaccaccaccctctcttctggctggggaagaaggtctcttctgg-3') or (5'-gctctagagcctagtcacaaccaccaccctctctctgctctcaccagcctcg-3') containing an *XbaI* restriction site and encoding for C α - and β -terminal constant domain extension (NDGGCK) and (QDRGGGCD*) (14). PCR products were cloned into the pLM1 vector in frame with a C-terminal biotin acceptor peptide sequence tag for the α -chain and into pET22b (Novagen) for the β -chain.

Protein expression and purification of R14 TCR

Both RA14 α - and β -chains were produced separately as inclusion bodies in BL21(DE3)RIL *Escherichia coli* strain (Stratagene), transformed by either pLM1- α or pET22b- β plasmids. Inclusion bodies were resuspended in 8 M urea, 50 mM MES (pH 6.5), 0.1 mM DTT, and 0.1 mM EDTA. Equal quantity of both chains were then mixed in 6 M guanidine HCl, 50 mM MES (pH 6.5), 10 mM EDTA, 2 mM DTT, and 1 mM sodium acetate. RA14 TCR was refolded by flash dilution in a solution containing 3 M urea, 200 mM arginine HCl, 150 mM Tris (pH 8), 1.5 mM reduced glutathione, and 0.15 mM oxidized glutathione at 4°C. After an incubation of 72 h at 4°C, the folding solution was dialyzed against 10 mM Tris (pH 8.5) and 50 mM NaCl for 24 h and against 10 mM Tris (pH 8.5) for 48 h. The resulting protein solution was then concentrated using a 10 kDa membrane (Millipore) and purified on a MonoQ 5/50 (GE Healthcare) column with a fast protein liquid chromatography system (Purifier 10, GE Healthcare). The elution was performed with a 0–0.5 M NaCl gradient.

Protein expression and purification of NLV-HLA-A2

Both HLA-A*0201 heavy chains and β 2-microglobulin were produced separately as previously described (16, 17). In brief, the A245V mutant of the HLA-A*0201 H chain tagged at the C-ter with a biotinylation sequence was cloned into pHN1 expression vector as described in (17). Recombinant proteins were produced as inclusion bodies in XA90F⁺LaqQ1 *Escherichia coli* strain. The inclusion bodies were resuspended in 8 M urea, 50 mM MES (pH 6.5), 0.1 mM DTT, and 0.1 mM EDTA, incubated overnight at 4°C, and centrifuged 30 min at 100,000 g. The supernatant was collected and frozen at –80°C. The pMHC complex refolding step was done by flash dilution of a mix of 21 mg HLA-A*0201, 10 mg β 2-microglobulin, and 10 mg of the desired synthetic peptide into 350 ml of 100 mM Tris (pH 8.0), 400 mM L-Arginine HCl, 2 mM EDTA, 5 mM reduced glutathione, 0.5 mM oxidized glutathione, and two Complete EDTA-free Cocktail Inhibitor Tablets (Roche). The refolding solution was then incubated for 4 to 5 days at 4°C and concentrated with a 10 or 30 kDa cutoff membrane (Vivacell System; Vivascience). The pMHC complex was purified on a MonoQ 5/50 column with a fast protein liquid chromatography system equilibrated in a 10 mM Tris (pH 8.0) buffer. It was eluted with 100 to 150 mM NaCl and concentrated with Amicon-10 or Amicon-30 devices to reach a final protein concentration of 2.5 to 3.5 mg/ml.

Surface plasmon resonance (SPR) experiments

Binding of HLA-A*0201 loaded with the different NLV peptide variants to TCR RA14 was analyzed by SPR. All experiments were performed using a Biacore 3000 instrument (GE Healthcare) at 25°C in HBS-EP buffer (10

Table I. Affinity of RA14 TCR for NLV variants and functional impact on cytotoxic response of the RA14 T cell clone^a

	Sequence	K _D (μM)	EC ₅₀ (nM)
NLV	NLVPMVATV	27.7 ± 2.3	5.10 ⁻² ± 0.01
P4G	NLVGMVATV	69.2 ± 1.5	50 ± 3
P4A	NLVAMVATV	56.5 ± 4.0	5.10 ⁻¹ ± 0.1
M5S	NLVPSVATV	n.d.	5.10 ² ± 20
M5V	NLVPSVATV	60.2 ± 2.6	5.10 ² ± 50
M5T	NLVPTVATV	58.2 ± 1.7	10 ³ ± 30
M5Q	NLVQVATV	57.5 ± 1.3	10 ² ± 20
T8A	NLVPMVAAV	n.d.	>10 ⁴
T8S	NLVPMVASV	42.7 ± 8.4	10 ⁻¹ ± 0.05
T8V	NLVPMVAVV	n.d.	>10 ⁴

^a Dissociation constants (K_D) were determined by steady-state SPR experiments for the interaction between RA14 TCR and HLA-A*0201 loaded with either the native viral NLV peptide or one of the nine different peptide variants (the mutated amino acids are underlined). The ability of a T cell clone bearing the RA14 TCR to recognize HLA-A*0201-positive target cells loaded with each peptide variant was evaluated by determining the peptide concentration necessary to achieve half maximal target-cell lysis (EC₅₀ at an effector to target ratio of 10:1). n.d., No binding detected.

mM HEPES (pH 7.4), 150 mM NaCl, 3 mM EDTA, 0.005% P20). RA14 TCR was diluted to 20 μg/ml in 10 mM acetate buffer (pH 5.0) and coupled at 3300–6000 resonance units onto three adjacent cells of a CM5 sensor chip (GE Healthcare) by classical amine coupling at a flow rate of 5 μl/min. The blank cell was made by ethanolamine deactivation of the activated surface. Before SPR experiments, HLA-A*0201-peptide complexes were dialyzed against 10 mM Tris (pH 8) containing 150 mM NaCl and a final step of purification by gel-filtration on a Superdex-200 column (GE Healthcare) was performed to eliminate aggregates. HLA-A*0201-peptide complexes were injected onto the four flow cells simultaneously at various concentrations ranging from 0.31 to 80 μM. SPR data were analyzed using the steady state affinity model of the BIAevaluation 3.1 software. Analysis of the steady-state data was then performed by averaging the equilibrium response over the final 10 to 25 s of each injection, plotting the equilibrium response against concentration, and fitting using a 1:1 binding model (supplemental Fig. S1).⁷ Final affinity constants were obtained by averaging values fitted over the triplicate data sets and all errors reported are SD from the mean (Table I).

T cell clones and functional assays

T cell clones directed against NLV-HLA-A*0201 were obtained and cultured as previously described (13). The NLV and NLV variant synthetic peptides were dissolved in DMSO, diluted at 2 mg/ml in 10 mM acetic acid, and then diluted to the required concentration in RPMI 1640 10% FCS. Cytotoxicity was measured in a standard 4 h ⁵¹CR assay (18). Briefly, transporter associated with Ag processing (TAP)-deficient HLA-A*0201-positive T2 cells were labeled with 100 μCi of Na₂⁵¹CRO₄ for 1 h at 37°C, washed three times in RPMI 1640 10% FCS, and pulsed with various concentrations of NLV peptides for 1 h at 37°C. After two washes, target cells were incubated with the different T cell clones at an effector:target ratio of 10:1 in 100 μl of RPMI 1640 10% FCS for 4 h at 37°C. Supernatant (25 μl) from each well was removed and counted in a gamma scintillation counter. Peptide concentrations necessary to achieve half-maximal target-cell lysis (EC50) were calculated from three independent experiments and errors reported correspond to SD from the mean EC50 values (Table I).

Crystallographic data collection

Crystals were grown by vapor diffusion with the hanging drop technique at 293K. The crystals of the NLV-A2 and the NLV_{variants}-A2 complexes were obtained by mixing 2 μl of 3–5 mg/ml protein solution and 2 μl of the reservoir solution (9–20% polyethylene glycol (PEG) 6000, 0.1 M tri-Na Citrate (pH 6.5), 0–0.1 M NaCl). Before crystallization, the RA14-TCR-NLV-A2 complex was formed by mixing TCR and pMHC with a molar ratio of 1:1 and at a final concentration of 2 mg/ml for the complex. The crystals were grown in 18–24% PEG 3350, 0.3 M LiSO₄. The crystals were cryoprotected with 30% of either glycerol or PEG added to the mother liquor before flash freezing. All data were collected at 100 K. The

⁷ The online version of this article contains supplemental material.

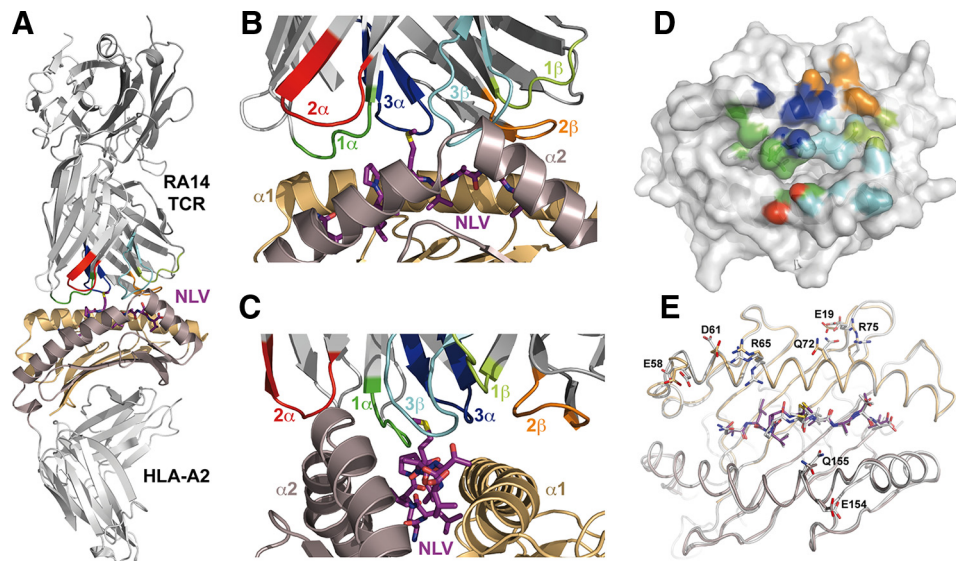


FIGURE 1. Overall view of RA14-NLV-A2 complex structure. *A*, Ribbon representation of the RA14-NLV-A2 ternary complex. TCR α - and β -chains are depicted in light gray and medium gray, respectively. CDR1 α , CDR2 α , CDR3 α , CDR1 β , CDR2 β , and CDR3 β are shown in green, red, blue, chartreuse green, orange, and cyan, respectively. HLA-A2 is depicted in gold and silver mauve for the α 1 and α 2 domains, respectively, and gray for the α 3 and β 2-microglobulin domains. The peptide is shown in purple balls and sticks. *B* and *C*, Two perpendicular enlarged views of the RA14-NLV-A2 interface. *D*, Footprint of the RA14 TCR on the molecular surface of the NLV-A2 complex. NLV-A2 molecular surface buried by the RA14 CDRs has been colored as in *A*. *E*, Superposition of the NLV-A2 α 1 α 2 domain derived from both NLV-A2 (depicted in gray) and RA14-NLV-A2 (depicted as in *A*) complex structures.

data sets for the NLV-A2 and NLV_{variants}-A2 crystals were collected on beamlines ID14-eh2 and ID23-eh2 of the European Synchrotron Radiation Facility (ESRF) at a wavelength of 0.933 Å using an ADSC Q4 CCD detector. The data set of the RA14-NLV-HLA-A2 crystal was collected on beamline ID29 of the ESRF at a wavelength of 0.976 Å using an ADSC Q315 CCD detector. Data processing was performed using XDS (19) and summarized in Table S1.

Structure determination and refinement

All structures were solved by molecular replacement with AMoRe (20). The NLV-A2 structure was solved using the GVDYDGREHTV-A2 structure (Protein Data Bank (PDB) entry 1I4F, (21)) as the initial model. For the NLV_{variants}-A2 structures, the NLV-A2 structure was used as starting model. The RA14 TCR-NLV-A2 complex (hereafter denoted RA14-NLV-A2) structure was solved using the NLV-A2 binary complex structure and PDB entry 2BNU (22) as pMHC and TCR initial models. All structures include one complex per asymmetric unit and the refinement protocol used included several cycles of refinement with REFMAC (23) followed by manual model rebuilding with O (24) and Coot (25), until no interpretable electron density could be identified in the residual map. The electron density is well defined for the whole NLV-A2 and NLV_{variants}-A2 binary complexes. However, some peptide side chains showed significant mobility. For instance, the Met5^P side chain has a weak electron density beyond its C β atom, and Met5^P, Val6^P, and Thr8^P side chains have been modeled with two discrete conformations, indicating a significant mobility for the most solvent-exposed side chains as well as for the Val6^P in the A2 C pocket (Fig. S2). The whole RA14-NLV-A2 complex structure is well defined in the electron density map (Fig. S3) except for two stretches of the TCR C α domain corresponding to residues 143–150 and 160–169. Final refinement statistics and PDB entry are summarized in Table S1. Coordinates for RA14-NLV-A2, NLV-A2, NLV_{M5S}-A2, NLV_{M5V}-A2, NLV_{M5T}-A2, NLV_{M5Q}-A2, NLV_{T8A}-A2, and NLV_{T8V}-A2 structures were deposited with the PDB (<http://www.rcsb.org/pdb>) under 3GSN, 3GSO, 3GSQ, 3GSR, 3GSU, 3GSV, 3GSW, and 3GSX accession numbers, respectively.

Results

Structural overview of the NLV-A2 and TCR-NLV-A2 complexes

The crystal structure of the NLV-A2 binary complex has been determined at 1.6 Å resolution. NLV carries the canonical A2 binding motif (Leu in P2, Val in P6, and Leu in P9, (26)) and the most solvent-exposed residues are found at positions 4, 5, and 8.

The crystal structure of RA14-NLV-A2 has been refined to 2.8 Å resolution (Fig. 1A). The RA14 TCR is encoded by rearranged TRAV24 and TRAJ49 gene segments for the α -chain and TRBV6-5, TRBD1, and TRBJ1-2 gene segments for the β -chain (27). The RA14 TCR docking orientation on NLV-A2 (35°) falls within the range of orientations already observed for other TCRs in complex with pMHC class I (pMHCI) (28), and reflects the generally adopted diagonal docking mode. As illustrated by the TCR footprint on the pMHC surface (Fig. 1D), the CDR1 α and CDR3 α loops interact with the N-terminal half of the peptide and the α 1 MHC helix, whereas the CDR1 β and CDR3 β loops primarily contact the C-terminal end of the peptide and the α 2 A2 helix. CDR2 α and CDR2 β exclusively contact the α 2 and α 1 A2 helices, respectively. When docking onto the NLV-A2 complex, RA14 buries 91% (324 Å²) of the peptide solvent accessible surface while in other TCR-pMHCI complex structures known to date, this percentage varies from 60 to 89%. Only 32 Å² of the peptide surface is left exposed to the solvent. Therefore, while the RA14-NLV-A2 complex adopts the canonical TCR diagonal docking mode, it manages to form the most extensive peptide covering among the TCR-pMHCI structures known to date.

The TCR-A2 interface highlights a new TRAV6-5-A2 binding mode

The RA14-A2 interaction is mediated by six hydrogen bonds and 43 van der Waals contacts (Table II) and involves all CDRs except CDR2 α . RA14 sits on A2 α 1 helix through a hydrophobic cluster between the CDR2 β (Val50 ^{β} , Ile54 ^{β}) on one side and the aliphatic chain of Gln72^H, Arg75^H and Val76^H, on the other, as well as through two hydrogen bonds between the framework residues Tyr48 ^{β} hydroxyl group and Asp56 ^{β} O δ 1 atom and the Gln72^H N ϵ 2 atom (Fig. 2A). This anchor is complemented with two hydrogen bonds between the tip of RA14 CDR3 α (Asn96 ^{α} amine group) and both Gln72^H O ϵ 1 atom and Ala69^H carbonyl group (Fig. 2B). Gln72^H is therefore truly identified as a tight anchor for

Table II. TCR-peptide and TCR-MHC interactions^a

TCR RA14: HLA-A2 van der Waals Contacts (C-C < 4.5 Å)					
CDR	Gene segment	TCR contact residues		HLA contact residues	No. of contacts
CDR1 α	TRAV24	Asn29		Lys66	2
CDR1 α	TRAV24	Tyr31		Gln155	5
CDR2 α	TRAV24	Thr51		Gln155	2
CDR2 α	TRAV24	Leu52		Ala158	1
CDR3 α	TRAJ49	Thr94		Arg65	3
CDR3 α	TRAJ49	Gly95		Ala69	2
CDR3 α	TRAJ49	Asn96		Ala69	4
FW	TRBV6-5	Tyr48		Gln72	2
CDR2 β	TRBV6-5	Val50		Gln72	1
CDR2 β	TRBV6-5	Val50		Val76	2
CDR2 β	TRBV6-5	Ile54		Arg75	7
CDR2 β	TRBV6-5	Ile54		Val76	1
CDR3 β	N-TRBD1-N	Val96		Lys146	4
CDR3 β	N-TRBD1-N	Ile100		Gln155	3
CDR3 β	TRBJ1-2	Tyr101		Ala149	1
CDR3 β	TRBJ1-2	Tyr101		Ala150	3

TCR RA14: HLA-A2 Hydrogen Bonds						
CDR	Segment	TCR contact atom		HLA contact atom		Length (Å)
CDR3 α	TRAJ49	Asn96	ND2	Ala69	O	2.83
CDR3 α	TRAJ49	Asn96	ND2	Gln72	OE1	2.70
CDR1 β	TRBV6-5	Glu30	OE2	Lys146	NZ	3.23
FW	TRBV6-5	Tyr48	OH	Gln72	NE2	2.67
FW	TRBV6-5	Asp56	OD1	Gln72	NE2	2.85
CDR3 β	TRBJ1-2	Tyr101	OH	Ala149	O	3.11

TCR RA14: NLV van der Waals Contacts (C-C < 4.5 Å)					
CDR	Segment	TCR contact residues		Peptide contact residues	No. of contacts
CDR1 α	TRAJ49	Asn29		Pro4	3
CDR1 α	TRAJ49	Phe30		Pro4	1
CDR1 α	TRAJ49	Tyr31		Val3	1
CDR1 α	TRAJ49	Tyr31		Pro4	11
CDR1 α	TRAJ49	Tyr31		Met5	14
CDR3 α	TRAJ49	Gly95		Met5	1
CDR1 β	TRBV6-5	Glu30		Thr8	1
CDR3 β	N-TRBD1-N	Thr97		Ala7	1
CDR3 β	N-TRBD1-N	Thr97		Thr8	2
CDR3 β	N-TRBD1-N	Gly98		Met5	3
CDR3 β	N-TRBD1-N	Gly98		Val6	2
CDR3 β	N-TRBD1-N	Gly98		Ala7	4
CDR3 β	N-TRBD1-N	Gly99		Met5	2
CDR3 β	N-TRBD1-N	Gly99		Ala7	1
CDR3 β	N-TRBD1-N	Ile100		Met5	1

TCR RA14: NLV Hydrogen Bonds						
CDR	Segment	TCR contact atom		Peptide contact atom		Length (Å)
CDR3 α	TRAJ49	Asn96	NH	Met5	SD	3.61
CDR1 β	TRBV6-5	Glu30	OE1	Thr8	OH	2.74
CDR3 β	N-TRBD1	Thr97	O	Thr8	N	3.09
CDR3 β	N-TRBD1	Thr97	O	Thr8	OH	2.92
CDR3 β	N-TRBD1	Thr97	NH	Thr8	OH	3.44

^a FW, Frame work residues.

RA14, thanks to TRAJ49-TRBV6-5 combination. The particular focusing of Tyr48 ^{β} and Asp56 ^{β} on the Gln72^H is unique to RA14-NLV-A2 and differs from that highlighted in several other TCR-pMHC structures in which these two well conserved residues preferentially interact with Arg65^H (29). RA14 contacts A2 α 2 helix mainly through its CDR1 α , CDR1 β and CDR3 β . Tyr31 α side chain is sandwiched by A2 Gln155^H and peptide Pro4^P and Met5^P, while its main chain carbonyl group further stabilizes the Gln155^H side chain through an elongated hydrogen bond (3.86 Å,

Fig. 2C). Such an interaction is reminiscent of that observed for Tyr31 α in 1G4 TCR-NY-ESO-1-A2 complex (22), but differs from the one identified by Marrack et al. (29) as a commonly used V α /MHC contact and involving a tyrosine at position 32 α (numbered as Tyr31 α (29)) and position 155^H on the α 2 MHC helix. CDR1 β and CDR3 β are involved through an ionic interaction between Glu30 ^{β} and Lys146^H side chains (Fig. 2D) and one hydrogen bond between Tyr101 ^{β} hydroxyl group and Ala149^H carbonyl oxygen atom and a few van der Waals contacts. Remarkably,

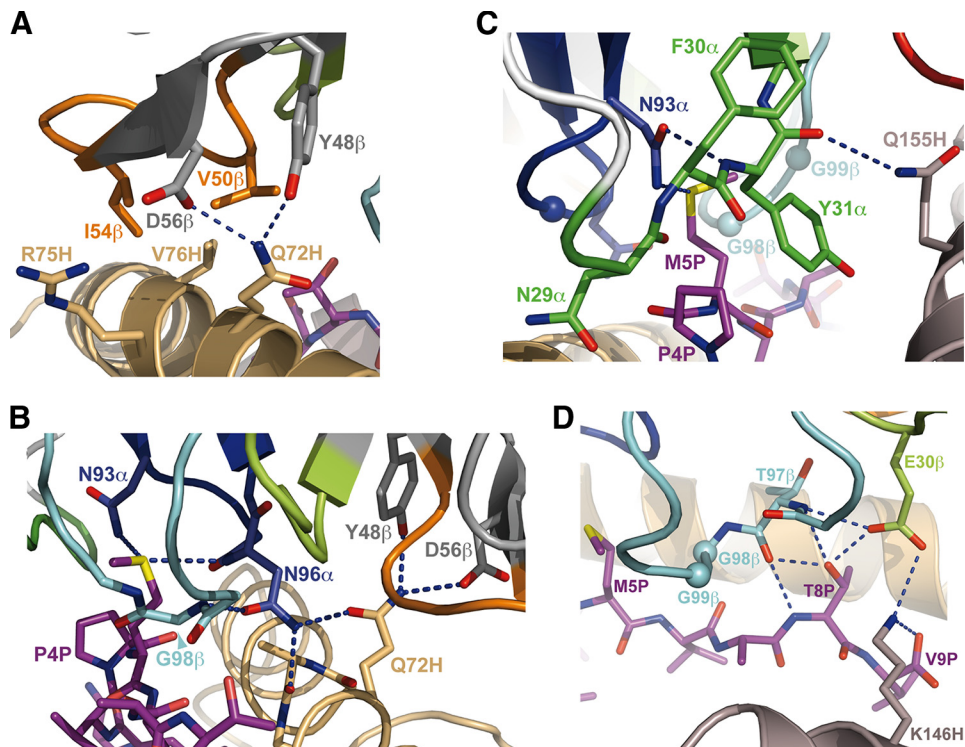


FIGURE 2. Detailed views of the RA14-NLV-A2 interaction hot spots. *A*, Interaction of the RA14 CDR2 β with the A2 α 1 helix. *B*, Interaction of the RA14 CDR1 α (Asn96 α) with the NLV-A2. *C*, Interaction of the CDR1 α (Tyr31 α) with the NLV-A2 complex. *D*, Interaction of the CDR1 β with NLV-A2.

although RA14 TCR shares with A6, B7, and 1G4 TCRs the same TRBV6-5 segment, its binding mode to A2 differs and constitutes a new structural codon for the TRBV6-5-A2 interaction as defined by Feng et al. (30) likely favored by the presence of the TRAJ49 encoded Asn96 α . This reinforces the concept that a V β (or V α) domain encoded by given TRBV (or TRAV) segment can adopt several different binding modes with the same MHC molecule, depending on the other segments used to encode the whole TCR.

A TCR-pMHC complex with a nearly maximal peptide readout

Overall, the NLV peptide establishes five hydrogen bonds and 48 van der Waals contacts with RA14 (Table II). RA14 contacts the peptide at nearly all solvent-exposed positions, namely 3, 4, 5, 6, 7, and 8, but \sim 80% of the van der Waals contacts focus on positions 4 and 5, while the hydrogen bonds essentially engage Thr8 P . The bulging Pro4 P and Met5 P fit a cavity formed by CDR1 α , CDR3 α , and CDR3 β (Fig. 2C). Pro4 P is sandwiched between Asn29 α and Tyr31 α side chains and establishes hydrophobic contacts with them (Fig. 2C). The tips of the CDR3 α (Asn93 α , Gly95 α , Asn96 α) and CDR3 β (Gly98 β , Gly99 β) loops and the Tyr31 α aromatic ring form a molecular cage for the Met5 P side chain and thus fully stabilize it. The Met5 P sulfur atom is further stabilized by a 3.61 Å long hydrogen bond with the Asn96 α NH group (Table II, Fig. 2B). Contrasting with the hydrophobic nature of the TCR-peptide interactions involving positions 4 and 5, a dense network of hydrogen bonds connects Thr8 P with the side chain of Glu30 $^\beta$ and the main chain of Thr97 $^\beta$ (Table II). The Glu30 $^\beta$ side chain is further stabilized through two hydrogen bonds with Thr97 $^\beta$ N and Lys146 H N ζ (Fig. 2D). Importantly, the CDR3 β contribution to the TCR-peptide interface is essentially mediated through its main chain. The only CDR3 β sequence requirement seems to be the glycine at position 98 β , consistent with the van der Waals contacts established between the C α atom and position 5, 6, and 7 of the peptide. Overall, although all the significantly solvent-exposed peptide residues contribute to the interface with the TCR, the present structural data identify Pro4 P , Met5 P , and Thr8 P as main peptide recognition spots.

Minor pMHC structural changes induced upon TCR binding

The structure of the NLV-A2 alone or in complex with RA14 are similar, the root mean square difference on C α atom pairs being 0.51 Å for the α 1 α 2 A2 domain and 0.53 Å for the peptide (Fig. 1E). Upon TCR binding, a 1 Å shift of Pro4 P and Met5 P backbone toward the α 1 A2 helix is observed and correlates with both the insertion of Tyr31 α between the peptide and the A2 α 2 helix and the clamping of Met5 P by CDR1 α , CDR3 α , and CDR3 β . Moreover, TCR docking onto pMHC clearly affects the mobility of Met5 P and Thr8 P , as shown by the electron density, as the side chains of both residues adopt a unique and well-defined conformation and interact with RA14 CDRs (Fig. S3). Among the 24 A2 residues implicated in the RA14-NLV-A2 interface, eight of them have their side chain conformation altered by the docking of RA14. These changes were mainly caused by steric hindrance generated by the TCR (Glu58 H , Arg65 H , Arg75 H , Gln155 H), by favorable interactions with the TCR (Gln72 H , Gln155 H), or by concerted movement with one of the latter residues (Glu19 H , Asp61 H , Glu154 H).

RA14-NLV-A2 affinity measurements identify three peptide hot spots

To assess the contribution of RA14-NLV contact spots in the stabilization of the RA14-NLV-A2 complex, the affinity of the RA14 TCR for A2 bound to the native NLV peptide and nine NLV variants (P4G, P4A, M5S, M5T, M5V, M5Q, T8A, T8S, and T8V) was measured by SPR. Both kinetic and steady state binding experiments were performed. For the native NLV peptide and all the variants, binding of A2 onto immobilized TCR was characterized by fast association and dissociation. Indeed, due to the speed of the association/dissociation process, none of the association (K_{on}) or dissociation (K_{off}) constants could be reliably evaluated, except for the k_{off} corresponding to wild-type NLV peptide (data not shown). Affinities (K_D) were thus derived from steady state experiments (Table I and Fig. S1). The wild-type NLV-A2 has the highest affinity for RA14 with a K_D of 27.7 μ M, followed by the T8S NLV

Table III. Overrepresentation of motifs involved in key interactions between RA14 TCR and NLV-A2 among TCR dominating the NLV-specific repertoire in patients prone to HCMV reactivation^a

Sequence	TCR	No. Binding Motifs	Healthy Donors		IS Patients	
			Rec.	f(%)	Rec.	f(%)
Public 1	AV24-AJ49 BV6-5	4	3/5	<45	4/12	>91
RA15	AV24-AJ49 BV6-1	3			1/12	100
M1	AV24-AJ49 BV27	4	1/5	<6		
Public 2	AV24-AJ49 BV27	4	1/5	14	1/12	100
RA17R4	AV8-3-AJ49 BV6-5	3			1/12	43
M2	AV8-3-AJ49 BV27	3	1/5	<6		
GR2	AV8-3-AJ20 BV6-5	2			1/12	97
HD33.5	AV3-AJ35 BV28	2	1/5	nd		
GR3	AV3-AJ26 BV28	1			1/12	98
RA5PBL	AV26-2-AJ53 BV27	2			1/12	100
R22	AV26-2-AJ43 BV30	1			1/12	33
HD55.2	AV26-2-AJ36 BV4-3	0	1/5	nd		
HD55.23	AV5-AJ36 BV13	0	1/5	nd		
Public 3	AV5-AJ44 BV30	0	2/5	<63	1/12	99
HD44.1	AV5-AJ6 BV24-1	0	1/5	nd		
	AV12-3-AJ49 BV12-3	1	1/5	nd		
Public 4	AV35-AJ50 BV12-3	1	2/5	>67		
	AV41-AJ48 BV27	2			1/12	10
	AV21-AJ44 BV20-1	0	1/5	2		

^a Each line corresponds to highly related or identical TCR sequences whose TRAV-TRAJ-TRBV gene composition and number of key NLV-A2 binding motifs shared with the RA14 TCR (Asn29 and Tyr31 in CDR1 α , Asn96 in CDR3 α , Glu30 in CDR1 β , and Tyr48 and Asp56 in CDR2 β) are indicated in the second and third columns, respectively. When a given TCR sequence is found in several individuals, the maximal (<*x*) or minimal (>*x*) percentage of cells expressing the corresponding TCR are indicated. Note that most NLV-specific TCR carrying three or more tightly involved motifs are expressed by the vast majority (>90%) of NLV-specific T cells in six of 12 immunosuppressed patients, but by a small fraction only of NLV-specific T cells in healthy donors. By contrast, TCR carrying two or less tightly involved motifs are preferentially found within NLV-specific T cell lines from healthy donors.

IS, Immunosuppressed; nd, not determined; Rec., the interindividual recurrence of the corresponding TCR sequence among NLV-specific T cell lines derived from five healthy donors and 12 immunosuppressed patients; f(%), the percentage of cells expressing the corresponding TCR within each NLV-specific T cell line (numbers are in bold when this percentage is above 66.67%).

variants with 42.7 μ M. Five variants (P4G, P4A, M5T, M5V, and M5Q) have an affinity between 58 and 69 μ M, and the last three (M5S, T8A, and T8V) do not interact detectably with the RA14 TCR. The K_D obtained for the NLV peptide differs significantly from that recently determined by Gakamsky et al. (6.3 μ M) (31), although the estimated k_{off} is quite similar (0.675 s^{-1} vs 0.44 s^{-1} in the present and previous study, respectively). The different experimental setup (see *Materials and Methods*) accounts for such a difference because the binding of RA14 to the immobilized NLV-A2 through the C-ter end of the A2 H chain provided affinity values closer to that observed by Gakamsky et al. (31 and data not shown) and presumably more directly comparable to other TCR-pMHC K_D values published previously. As the absolute K_D value remains to be precisely cross-checked with other techniques, such as isothermal titration calorimetry, the present setup is optimal for the affinity comparison of the different NLV variants because all NLV-A2 complexes were tested on the same RA14-loaded surface. Binding data drawn from NLV variants clearly establish the significant contribution of peptide positions 4, 5, and 8 to the stability of the RA14-NLV-A2 complex, fully supporting structural data described above. It truly identifies Pro4, Met5, and Thr8 as peptide hot spots.

RA14 T cell clone activation by NLV variants confirms the three peptide hot spots

The functional impact of each of the above mutations on RA14 clone activation was assessed in cytotoxicity assays against TAP-deficient A2-positive target cells loaded with graded doses of peptide. The highest EC_{50} value was obtained with the wild-type peptide, followed by T8S and P4A variants (Table I). P4G and all the M5 variants were still able to activate RA14, though at significantly higher concentrations than the wild-type peptide, whereas

T8V and T8A mutants were no longer recognized by RA14. Because all these mutants showed similar or even higher affinity for A2, as suggested by their ability to stabilize surface-A2 expression on T2 TAP-deficient cells (Fig. S4), decreased recognition of peptide variants is likely accounted for by decreased functional avidity of RA14 clone for the corresponding pMHC complexes. Thus, the EC_{50} values of most mutants correlate well with the affinity of the corresponding peptide-A2 complexes for RA14 TCR estimated by SPR, except for M5S-A2. It thus confirms the role of interaction network formed between peptide position 4, 5, and 8 and the TCR in the recognition of NLV-A2 by RA14 TCR.

Discussion

The mechanisms underlying the emergence of public TCRs remain controversial and probably involve multiple parameters, such as recombination biases during TCR gene rearrangements, intrathymic selection, peripheral homeostatic processes, and Ag-driven selection (32). The latter process is most likely the main parameter contributing to predominant usage of public TCRs by NLV-A2-specific CD8 T cells derived from immunodepressed patients undergoing HCMV reactivation. Indeed, the NLV-A2-specific repertoire in immunosuppressed patients, who are prone to HCMV reactivation, is more limited than that of healthy donors. This repertoire focusing parallels enrichment for high affinity public TCRs, with recurrent CDR1 and CDR2 motifs and conserved CDR3 length and sequence (13). Among public TCRs, those carrying TRAV24-TRAJ49 and TRBV6-5 chains were the most frequent and represented the vast majority of NLV-A2-specific T cells in four of 12 immunosuppressed patients. Those carrying at least three of the four tightly involved motifs identified in RA14 represented more than 90% of the NLV-A2-specific T cells in six of 12 immunosuppressed patients (Table III). This suggests that

TRAV24-TRBV6-5 TCRs carry optimal CDR combinations to productively interact with their cognate pMHC, and are thus favored during chronic Ag stimulation. Structural analysis of the RA14 TCR, which is representative of the most frequent public high avidity and dominant NLV-A2-specific TCRs, supports this assumption and provides new insights into the structural basis of public TCR selection.

Several structures of human TCRs in complex with pMHC derived from T cell responses with reported biased usage of V(D)J gene segments have been reported. They include JM22 TCR in complex with an influenza-derived peptide (MP₅₈₋₆₆) presented by A2 (14), LC13 TCR in complex with an EBV nuclear Ag 3-derived peptide (FLRGRAYGL, hereafter referred to as FLR) presented by HLA-B8 (15), SB27 TCR in complex with EBV BZLF1₅₂₋₆₄ peptide presented by HLA-B*3508 (33), ELS4 TCR in complex with BZLF1₅₄₋₆₄ peptide presented by HLA-B*3501 (34), and HA1.7 TCR in complex with an influenza hemagglutinin-derived peptide (HA₃₀₆₋₃₁₈) presented by HLA-DR1 (35). However, the most relevant structural comparison is certainly with the two public TCRs in complex with immunodominant pMHC I that have been crystallized so far, JM22-MP₅₈₋₆₆-A2 and LC13-FLR-HLA-B8.

The JM22-MP₅₈₋₆₆-A2 complex structure showed involvement of the three β -chain CDRs in peptide interactions but limited contribution of the V α domain to the TCR-pMHC interface, in agreement with a much less stringent selection of the latter chain (14). Moreover, the JM22 Arg98 ^{β} is inserted into a notch between the peptide and the A2 α 2 helix while the peptide is essentially featureless, in that it has a minimal fraction of its molecular surface bulging out of the peptide binding groove (247 Å² of peptide accessible surface exposed). Therefore, the JM22 CDR3 β motif (Arg98-Ser99) associated with TRBV19*01 CDR1 β and CDR2 β provides a unique solution for the MP₅₈₋₆₆-A2 recognition and explains the highly constrained TRBV repertoire and CDR3 β sequences of MP₅₈₋₆₆-specific TCR (36, 37).

In the case of the public T cell response against the HLA-B8-restricted FLR epitope, the repertoire comprises clones expressing almost identical TCR α - and β -chains and rests primarily on the TRAV26-TRAJ52 and TRBV7-8-TRBD1-TRBJ2-7 gene segment combination (38), although public TCRs with different V(D)J usage may cope with the same pMHC complex (39). The crystal structure of one such TCR (LC13) showed a significant involvement of five CDRs in the interaction with the pMHC (15), in full agreement with selection constraints on both chains. The CDR1 and CDR3 tightly contact through residues unique to TRAV26-2*01, TRBD1/2 and TRBJ2-7*01, the bulging tyrosine at peptide position 7, the latter contributing the most to the peptide solvent accessible surface (95 of 250 Å² for the whole peptide). For these two public TCRs, recognition essentially relies on optimal interactions with one specific feature of the pMHC surface. Presumably, the number of TCRs able to productively interact with such a confined part of the pMHC is very limited and consistent with the quasi-monoclonal repertoire observed for these two cases. Indeed, depletion of TRBV19*01⁺ T cells abrogates the MP₅₈₋₆₆-A2-specific response in most donors (36). For FLR-B8, the T cell clone type that dominates the response is deleted in HLA-B44*01⁺ individuals through a self-tolerance mechanism. In such individuals, the oligoclonal response that includes public clones targets a different epitope involving the N- rather than the C-terminal region of the FLR peptide (39).

RA14-NLV-A2 departs from these two aforementioned examples. Because the antigenic peptide is significantly more exposed to the solvent in NLV-A2 (355 Å²), it cannot be considered as featureless as MP₅₈₋₆₆. RA14 does not focus on one particular

bulging residue, as observed for LC13, nor does it conform to the peg-notch type of recognition adopted in the two previous examples. Instead, four of its CDRs (CDR1 α , CDR1 β , CDR3 α , and CDR3 β) bury most of the peptide surface and interact with three peptide hot spots through a network of hydrophobic and polar interactions involving the moderately bulging Pro4^P, Met5^P, and Thr8^P. The significance of these three hot spots in stabilizing the TCR-pMHC complex is demonstrated by binding and T cell activation data on NLV variants. The wild-type NLV peptide displays the highest affinity for RA14 and triggers the most robust cytotoxic responses when used to stimulate the RA14 CTL clone (Table I), whereas a decrease of the affinity is observed for amino acid substitution at positions 4, 5, and 8. Modification of Met5^P has the most pronounced effect, M5A being the only variant that remains able to activate RA14 CTL clones, albeit much less efficiently (data not shown). Conservative change of Thr8^P to a serine induces a moderate decrease in the affinity and activation efficacy, whereas a change to either a valine or alanine impairs both RA14 interaction and RA14 CTL clone activation and highlights the role of the hydrogen bonds involving Thr8^P side chain. As confirmed by NLV variants-A2 binary structures, no peptide conformational change induced by replacement of Met5^P or Thr8^P with other amino acids which could have accounted for the change in TCR binding was observed for these variants (Fig. S5 and Table S1). Overall, these results confirm the substantial role of three different peptide hot spots and corroborate structural data showing that these three positions are actually the most contacted by the RA14 TCR.

As previously described for other TCRs (40, 41), several key residues, located on CDR1 α , CDR3 α , CDR1 β , and CDR2 β , provide a structural rationale for biased usage of particular V(D)J segments within the NLV-specific repertoire. For instance, Asn29 ^{α} and Tyr31 ^{α} belong to a key motif that is exclusively found in TRAV24, whereas Tyr31 ^{α} alone is found in the TRAV21 segment, which is also used by other NLV-specific TCRs. Asn96 ^{α} in RA14 (TRAJ49) is also shared by the TRAJ43 gene segment found in several NLV-specific clones. Glu30 ^{β} , which is quite scarce among TRBV genes, is shared by three TRBV segments (TRBV6-5, -27, and -28) frequently used by NLV-specific TCR. The hydrophobic character of the CDR2 β observed in TRBV6-5 and TRBV27 and the conservation of Tyr48 ^{β} and Asp56 ^{β} in TRBV6-1, -7-6, -12-3, -27, and -30 suggests a conserved mode of contact with A2 for this CDR. The occurrence of each of the above key motifs is much higher in NLV-A2-specific TCR sequence than in their respective V or J segments. Therefore, their selection either alone or in combination during chronic CMV reactivation is likely due to their ability to form specific contacts with NLV-A2.

Our results further exemplify that a specific TRBV segment (6-5) can bind the same MHC molecule in different ways. Thanks to a combination with TRAJ49, the positioning of TRBV6-5 seems driven by a dense hydrogen bond network with Gln72^H highlighting the latter residue as a fourth interacting spot.

In conclusion, our study highlights the structural characteristics that could explain the immunodominance of the RA14 TCR in response to NLV-A2 in an immunodepressed context associated with HCMV reactivation. The mechanism contrasts from that observed for the immunodominant MP₅₈₋₆₆-A2 and FLR-B8 pMHC because the quasi-unique TCRs used for their recognition essentially interact with a unique feature on the pMHC surface (14, 15). Instead, RA14 buries most of the peptide and forms tight contact with three peptide residues and one A2 amino acid. The combination of TRAV, TRAJ, and TRBV appears optimal to establish this four-spot recognition. Sequence comparison of other NLV-A2-specific TCRs shows that they share most of the key contacting residues and suggests that most of the high avidity TCRs could

adopt a similar recognition mode. Confirming this hypothesis would require supplemental structural and interaction data on different private and public NLV-specific TCRs. Nevertheless, the present data indicate that forming optimal interaction with the four identified spots appears to be the structural solution for the affinity-driven emergence of an optimal public TCR among an oligoclonal Ag-specific response after repeated antigenic stimulations. This solution is based on a TCR with a marked structural complementarity with the full array of the available peptide residues.

Acknowledgments

We thank J. McCarthy (ID29), D. Flot (ID23-2), and S. McSweeney and D. Hall (ID14-2) for help with synchrotron data collections at the ESRF (Grenoble, France), the staff of the EMBL HTX laboratory (Grenoble, France) for the use of the PSB crystallization platform, E. Forest and B. Dublet for the use of the mass spectroscopy PSB platform and D. Hart (EMBL Grenoble, France) and J. Rossjohn (Monash University, Clayton, Australia) for careful reading of the manuscript.

Disclosures

The authors have no financial conflict of interest.

References

- Sissons, J. G., M. Bain, and M. R. Wills. 2002. Latency and reactivation of human cytomegalovirus. *J. Infect.* 44: 73–77.
- Borysiewicz, L. K., J. K. Hickling, S. Graham, J. Sinclair, M. P. Cranage, G. L. Smith, and J. G. Sissons. 1988. Human cytomegalovirus-specific cytotoxic T cells: relative frequency of stage-specific CTL recognizing the 72-kD immediate early protein and glycoprotein B expressed by recombinant vaccinia viruses. *J. Exp. Med.* 168: 919–931.
- Quinnan, G. V., Jr., W. H. Burns, N. Kirmani, A. H. Rook, J. Manischewitz, L. Jackson, G. W. Santos, and R. Saral. 1984. HLA-restricted cytotoxic T lymphocytes are an early immune response and important defense mechanism in cytomegalovirus infections. *Rev. Infect. Dis.* 6: 156–163.
- McLaughlin-Taylor, E., H. Pande, S. J. Forman, B. Tanamachi, C. R. Li, J. A. Zaia, P. D. Greenberg, and S. R. Riddell. 1994. Identification of the major late human cytomegalovirus matrix protein pp65 as a target antigen for CD8+ virus-specific cytotoxic T lymphocytes. *J. Med. Virol.* 43: 103–110.
- Wills, M. R., A. J. Carmichael, K. Mynard, X. Jin, M. P. Weekes, B. Plachter, and J. G. Sissons. 1996. The human cytotoxic T-lymphocyte (CTL) response to cytomegalovirus is dominated by structural protein pp65: frequency, specificity, and T-cell receptor usage of pp65-specific CTL. *J. Virol.* 70: 7569–7579.
- Saulquin, X., C. Ibsch, M. A. Peyrat, E. Scotet, M. Hourmant, H. Vie, M. Bonneville, and E. Houssaint. 2000. A global appraisal of immunodominant CD8 T cell responses to Epstein-Barr virus and cytomegalovirus by bulk screening. *Eur. J. Immunol.* 30: 2531–2539.
- Kern, F., I. P. Surel, N. Faulhaber, C. Frommel, J. Schneider-Mergener, C. Schonemann, P. Reinke, and H. D. Volk. 1999. Target structures of the CD8(+)-T-cell response to human cytomegalovirus: the 72-kilodalton major immediate-early protein revisited. *J. Virol.* 73: 8179–8184.
- Khan, N., M. Cobbold, R. Keenan, and P. A. Moss. 2002. Comparative analysis of CD8+ T cell responses against human cytomegalovirus proteins pp65 and immediate early 1 shows similarities in precursor frequency, oligoclonality, and phenotype. *J. Infect. Dis.* 185: 1025–1034.
- Engstrand, M., C. Tournay, M. A. Peyrat, B. M. Eriksson, J. Wadstrom, B. Z. Wirtgart, F. Romagne, M. Bonneville, T. H. Totterman, and O. Korsgren. 2000. Characterization of CMVpp65-specific CD8+ T lymphocytes using MHC tetramers in kidney transplant patients and healthy participants. *Transplantation* 69: 2243–2250.
- Diamond, D. J., J. York, J. Y. Sun, C. L. Wright, and S. J. Forman. 1997. Development of a candidate HLA A*0201 restricted peptide-based vaccine against human cytomegalovirus infection. *Blood* 90: 1751–1767.
- Peggs, K., S. Verfuert, A. Pizzey, J. Ainsworth, P. Moss, and S. Mackinnon. 2002. Characterization of human cytomegalovirus peptide-specific CD8(+) T-cell repertoire diversity following in vitro restimulation by antigen-pulsed dendritic cells. *Blood* 99: 213–223.
- Weekes, M. P., M. R. Wills, K. Mynard, A. J. Carmichael, and J. G. Sissons. 1999. The memory cytotoxic T-lymphocyte (CTL) response to human cytomegalovirus infection contains individual peptide-specific CTL clones that have undergone extensive expansion in vivo. *J. Virol.* 73: 2099–2108.
- Trautmann, L., M. Rimbart, K. Echasserieu, X. Saulquin, B. Neveu, J. Dechanet, V. Cerundolo, and M. Bonneville. 2005. Selection of T cell clones expressing high-affinity public TCRs within human cytomegalovirus-specific CD8 T cell responses. *J. Immunol.* 175: 6123–6132.
- Stewart-Jones, G. B., A. J. McMichael, J. I. Bell, D. I. Stuart, and E. Y. Jones. 2003. A structural basis for immunodominant human T cell receptor recognition. *Nat. Immunol.* 4: 657–663.
- Kjer-Nielsen, L., C. S. Clements, A. W. Purcell, A. G. Brooks, J. C. Whisstock, S. R. Burrows, J. McCluskey, and J. Rossjohn. 2003. A structural basis for the selection of dominant $\alpha\beta$ T cell receptors in antiviral immunity. *Immunity* 18: 53–64.
- Garboczi, D. N., D. T. Hung, and D. C. Wiley. 1992. HLA-A2-peptide complexes: refolding and crystallization of molecules expressed in *Escherichia coli* and complexed with single antigenic peptides. *Proc. Natl. Acad. Sci. USA* 89: 3429–3433.
- Bodinier, M., M. A. Peyrat, C. Tournay, F. Davodeau, F. Romagne, M. Bonneville, and F. Lang. 2000. Efficient detection and immunomagnetic sorting of specific T cells using multimers of MHC class I and peptide with reduced CD8 binding. *Nat. Med.* 6: 707–710.
- Neveu, B., K. Echasserieu, T. Hill, K. Kuus-Reichel, E. Houssaint, M. Bonneville, and X. Saulquin. 2006. Impact of CD8-MHC class I interaction in detection and sorting efficiencies of antigen-specific T cells using MHC class I peptide multimers: contribution of pMHC valency. *Int. Immunol.* 18: 1139–1145.
- Kabsch, W. 1993. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* 26: 795–800.
- Navaza, J. 1994. AMoRe: an automated package for molecular replacement. *Acta Crystallogr. A* 50: 157–163.
- Hillig, R. C., P. G. Coulie, V. Stroobant, W. Saenger, A. Ziegler, and M. Hulsmeijer. 2001. High-resolution structure of HLA-A*0201 in complex with a tumour-specific antigenic peptide encoded by the MAGE-A4 gene. *J. Mol. Biol.* 310: 1167–1176.
- Chen, J. L., G. Stewart-Jones, G. Bossi, N. M. Lissin, L. Wooldridge, E. M. Choi, G. Held, P. R. Dunbar, R. M. Esnouf, M. Sami, et al. 2005. Structural and kinetic basis for heightened immunogenicity of T cell vaccines. *J. Exp. Med.* 201: 1243–1255.
- Collaborative Computational Project Number 4. 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* 50: 760–763.
- Jones, T. A., J. Y. Zou, S. W. Cowan, and Kjeldgaard. 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.
- Emsley, P., and K. Cowtan. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* 60: 2126–2132.
- Rammensee, H., J. Bachmann, N. P. Emmerich, O. A. Bachor, and S. Stevanovic. 1999. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics* 50: 213–219.
- Lefranc, M. P. 2001. IMGT, the international ImMunoGeneTics database. *Nucleic Acids Res.* 29: 207–209.
- Rudolph, M. G., R. L. Stanfield, and I. A. Wilson. 2006. How TCRs bind MHCs, peptides, and coreceptors. *Annu. Rev. Immunol.* 24: 419–466.
- Marrack, P., J. P. Scott-Browne, S. Dai, L. Gapin, and J. W. Kappler. 2008. Evolutionarily conserved amino acids that control TCR-MHC interaction. *Annu. Rev. Immunol.* 26: 171–203.
- Feng, D., C. J. Bond, L. K. Ely, J. Maynard, and K. J. Garcia. 2007. Structural evidence for a germline-encoded T cell receptor-major histocompatibility complex interaction “codon.” *Nat. Immunol.* 8: 975–983.
- Gakamsky, D. M., E. Lewitzki, E. Grell, X. Saulquin, B. Malissen, F. Montero-Julian, M. Bonneville, and I. Pecht. 2007. Kinetic evidence for a ligand-binding-induced conformational transition in the T cell receptor. *Proc. Natl. Acad. Sci. USA* 104: 16639–16644.
- Venturi, V., D. A. Price, D. C. Douek, and M. P. Davenport. 2008. The molecular basis for public T-cell responses? *Nat. Rev. Immunol.* 8: 231–238.
- Tynan, F. E., S. R. Burrows, A. M. Buckle, C. S. Clements, N. A. Borg, J. J. Miles, T. Beddoe, J. C. Whisstock, M. C. Wilce, S. L. Silins, et al. 2005. T cell receptor recognition of a “super-bulged” major histocompatibility complex class I-bound peptide. *Nat. Immunol.* 6: 1114–1122.
- Tynan, F. E., H. H. Reid, L. Kjer-Nielsen, J. J. Miles, M. C. Wilce, L. Kostenko, N. A. Borg, N. A. Williamson, T. Beddoe, A. W. Purcell, et al. 2007. A T cell receptor flattens a bulged antigenic peptide presented by a major histocompatibility complex class I molecule. *Nat. Immunol.* 8: 268–276.
- Hennecke, J., A. Carfi, and D. C. Wiley. 2000. Structure of a covalently stabilized complex of a human $\alpha\beta$ T-cell receptor, influenza HA peptide and MHC class II molecule, HLA-DR1. *EMBO J.* 19: 5611–5624.
- Lehner, P. J., E. C. Wang, P. A. Moss, S. Williams, K. Platt, S. M. Friedman, J. I. Bell, and L. K. Borysiewicz. 1995. Human HLA-A0201-restricted cytotoxic T lymphocyte recognition of influenza A is dominated by T cells bearing the V β 17 gene segment. *J. Exp. Med.* 181: 79–91.
- Moss, P. A., R. J. Moots, W. M. Rosenberg, S. J. Rowland-Jones, H. C. Bodmer, A. J. McMichael, and J. I. Bell. 1991. Extensive conservation of α and β chains of the human T-cell antigen receptor recognizing HLA-A2 and influenza A matrix peptide. *Proc. Natl. Acad. Sci. USA* 88: 8987–8990.
- Argaet, V. P., C. W. Schmidt, S. R. Burrows, S. L. Silins, M. G. Kurilla, D. L. Doolan, A. Suhrbier, D. J. Moss, E. Kieff, T. B. Sculley, and I. S. Misko. 1994. Dominant selection of an invariant T cell antigen receptor in response to persistent infection by Epstein-Barr virus. *J. Exp. Med.* 180: 2335–2340.
- Burrows, S. R., S. L. Silins, D. J. Moss, R. Khanna, I. S. Misko, and V. P. Argaet. 1995. T cell receptor repertoire for a viral epitope in humans is diversified by tolerance to a background major histocompatibility complex antigen. *J. Exp. Med.* 182: 1703–1715.
- Turner, S. J., P. C. Doherty, J. McCluskey, and J. Rossjohn. 2006. Structural determinants of T-cell receptor bias in immunity. *Nat. Rev. Immunol.* 6: 883–894.
- Gras, S., L. Kjer-Nielsen, S. R. Burrows, J. McCluskey, and J. Rossjohn. 2008. T-cell receptor bias and immunity. *Curr. Opin. Immunol.* 20: 119–125.

Jean-Baptiste Reiser,^a François Legoux,^{b,c} Paul Machillot,^a Emilie Debeaupuis,^{b,c} Béatrice Le Moullac-Vaydie,^{b,c} Anne Chouquet,^a Xavier Saulquin,^{b,c} Marc Bonneville^{b,c,*} and Dominique Housset^{a*}

^aInstitut de Biologie Structurale Jean-Pierre Ebel, UMR 5075 (CEA, CNRS, UJF, PSB), 41 Rue Jules Horowitz, F-38027 Grenoble, France,

^bINSERM, UMR 892, Institut de Recherche Thérapeutique de l'Université de Nantes, 8 Quai Moncoussu, BP 70721 F-44007 Nantes, France, and ^cUniversité de Nantes, Nantes, France

Correspondence e-mail: bonneville@nantes.inserm.fr, dominique.housset@ibs.fr

Received 27 July 2009

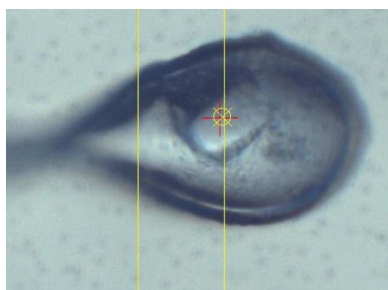
Accepted 18 September 2009

Crystallization and preliminary X-ray crystallographic characterization of a public CMV-specific TCR in complex with its cognate antigen

The T-cell response to human cytomegalovirus is characterized by a dramatic reduction of clonal diversity in patients undergoing chronic inflammation or immunodepression. In order to check whether all the selected high-avidity T-cell clones recognize the immunodominant pp65 peptide antigen pp65_{495–503} (NLVPMVATV) presented by the major histocompatibility complex (MHC) molecule HLA-A2 in a similar manner, several public high-affinity T-cell receptors (TCRs) specific for the pp65_{495–503}–HLA-A2 complex have been investigated. Expression, purification and crystallization were performed and preliminary crystallographic data were collected to 4.7 Å resolution for the RA15 TCR in complex with the pp65_{495–503}–HLA-A2 complex. Comparison of the RA15–pp65_{495–503}–HLA-A2 complex molecular-replacement solution with the structure of another high-affinity pp65_{495–503}–HLA-A2-specific TCR, RA14, shows a shared docking mode, indicating that the clonal focusing could be accompanied by the selection of a most favoured peptide-readout mode. However, the position of the RA15 V β domain is significantly shifted, suggesting a different interatomic interaction network.

1. Introduction

Human cytomegalovirus (HCMV) is a ubiquitous β -herpesvirus that infects 60–90% of the population. HCMV persists in a latent stage after the primary infection and can undergo transient reactivations. Although HCMV infections are kept in check by the immune system of healthy individuals, they can cause life-threatening diseases in immunodeficient patients (Sissons *et al.*, 2002). Cytotoxic T cells (CTLs) play a central role in controlling HCMV reactivation (Borysiewicz *et al.*, 1988; Quinnan *et al.*, 1984; McLaughlin-Taylor *et al.*, 1994; Wills *et al.*, 1996; Saulquin *et al.*, 2000) and the predominant CTL response is directed against the viral tegument protein pp65 (Wills *et al.*, 1996; Engstrand *et al.*, 2000). In individuals sharing the widespread HLA-A*0201 allele (referred to as A2), HCMV-specific CTLs recognize the same epitope pp65_{495–503} (NLVPMVATV), hereafter referred to as NLV (Wills *et al.*, 1996; Diamond *et al.*, 1997; Peggs *et al.*, 2002). During chronic inflammation (*e.g.* in rheumatoid arthritis patients) and immunodepression, a dramatic reduction in clonal diversity occurs, resulting in the selection of a few dominant clones bearing high-affinity TCRs (Trautmann *et al.*, 2005). We have recently determined the crystal structures of the immunodominant NLV peptide bound to A2 in isolation or in complex with a high-avidity public TCR expressed by a predominant CTL clone (RA14) derived from a rheumatoid arthritis (RA) patient (Trautmann *et al.*, 2005) and have gained insight into the mechanisms underlying clonal focusing upon chronic antigen stimulation (Gras, Saulquin *et al.*, 2009). Here, we present preliminary X-ray diffraction data for another high-avidity NLV–A2 specific TCR, RA15, in complex with NLV–A2. This TCR shares the same V α domain (TRAV24–TRAJ49, according to IMGT nomenclature; Lefranc, 2001) with RA14 but has a different V β domain (TRBV6-1 instead of TRBV6-5) and lacks some of the



© 2009 International Union of Crystallography
All rights reserved

essential NLV-contacting residues identified in RA14. The structure of RA15 in complex with NLV-A2 would therefore be extremely useful to check whether the extraordinary clonal focusing that occurs upon chronic reactivation is accompanied by the selection of a common NLV-A2 readout mode. The molecular-replacement solution indicates at least a convergent mode of docking.

2. Experimental methods

2.1. RA15 TCR cloning

Total RNA was isolated with TRIzol (Invitrogen) from 5×10^6 RA15 T-cell clones. cDNAs encoding the V and C ectodomain of the α -chain and β -chain were obtained by RT-PCR using the following primers: 5'-CCATATGAATGCTGGTGTCACTCAGACCCC-3' and 5'-GGAATTCTTAATCACAACCACCACCACGATCTTGGTCTGCTTACCCCAGGCCTCGGCGCTGAC-3' for the β -chain and 5'-CCATATGATACTGAACGTGGAACAAGTCC-3' and 5'-ACTCGAGTTTACAACCACCACCATCGTTTTCTGGGCTGGGG-AAGAAGGTGTCTTCTGG-3' for the α -chain. The PCR products were cloned into the pET-22b vector (Novagen). The resulting constructs were checked by sequencing.

However, attempts to obtain soluble RA15 TCR using this construct inexorably led to aggregates of TCR and an unsuitable yield of soluble protein for protein crystallization. In 2003, Boutler and coworkers demonstrated the efficacy of introducing non-native interchain disulfide bonds between the two constant domains in order to produce soluble TCRs without altering either the pMHC-binding capacity or the ternary complex crystallization (Boutler *et al.*, 2003). Cysteine codons were then introduced into pET22b-RA15 α and pET22b-RA15 β by PCR mutagenesis. Complementary primers were designed in order to mutate into cysteines (bold codons) Thr48 in TRAC (5'-GATGTGTATATCACAGACAAAT**GTGT**GCTAGACATGAGGTCTATG-3' and 5'-CATAGACCTCATGTCTAGCACACATTGTCTGTGATATACACATC-3) and Ser57 in TRBC (5'-GGTGCACAGTGGGGT**TGT**TACAGACCCGACGCC-3' and 5'-GGGCTGCGGGTCTGT**AC**AGACCCCACTGTGCACC-3'). These primers were then used with the QuikChange II kit (Stratagene, La Jolla, California, USA) and the pET22b constructs previously obtained as PCR templates. The site-directed mutagenesis results were confirmed by DNA sequencing.

2.2. Protein expression and purification of RA15 TCR

Cysteine-mutated RA15 α -chain and β -chain were produced separately as inclusion bodies in *Escherichia coli* strain BL21 (DE3) RIPL (Stratagene, La Jolla, California, USA) by the induction of a 3 l culture with 1 mM IPTG when the optical density reached 0.6 at a wavelength of 600 nm followed by 3 h incubation at 310 K. Inclusion bodies were isolated by cell lysis performed with a French press followed by successive washing and centrifugation steps with 15 ml of a solution containing 50 mM Tris pH 8, 100 mM NaCl, 0.5% Triton X-100, 1 mM DTT. The last wash step was made with a solution containing the same constituents apart from Triton X-100. Finally, inclusion bodies were solubilized in 8 M urea, 50 mM MES pH 6.5, 0.1 mM DTT and 0.1 mM EDTA and stored at 193 K before use.

Prior to refolding, 30 mg of both chains were mixed in 6 M guanidine-HCl, 50 mM MES pH 6.5, 10 mM EDTA, 2 mM DTT and 1 mM sodium acetate. RA15 TCR was then refolded by flash dilution into 500 ml of a cold (277 K) solution containing 3 M urea, 200 mM arginine-HCl, 150 mM Tris pH 8, 1.5 mM reduced glutathione and 0.15 mM oxidized glutathione. After incubation for 72 h at 277 K, the folding solution was dialyzed against 10 mM Tris pH 8.5 and 50 mM

Table 1

Crystallographic data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Data-collection statistics	
Space group	$P4_22_2$
Unit-cell parameters (Å)	$a = b = 99.14, c = 229.06$
Resolution (Å)	50.0–4.7 (4.95–4.7)
$R_{\text{merge}}^{\dagger}$ (%)	8.3 (40.7)
Completeness (%)	91.0 (92.7)
$I/\sigma(I)$	10.4 (2.99)
No. of reflections	21804 (3122)
No. of unique reflections	5856 (817)
Molecular-replacement statistics	
Resolution (Å)	15.0–4.7
Two-body rigid-body refinement	
Correlation coefficient (%)	68.2
R factor ‡ (%)	52.3
Seven-body rigid-body refinement	
Correlation coefficient (%)	69.8
R factor ‡ (%)	50.9

$\dagger R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$. $\ddagger R = \frac{\sum_{hkl} ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum_{hkl} |F_{\text{obs}}|}$. The R factor is calculated on all reflections.

NaCl for 24 h and against 10 mM Tris pH 8.5 for 48 h at 277 K. The resulting protein solution was then concentrated using a 10 kDa membrane (Vivacell, Vivascience AG, Germany) and purified on a size-exclusion column (HiLoad 16/60 Superdex 200) equilibrated in 10 mM Tris pH 8.5 and 20 mM NaCl and connected to an FPLC system (Purifier 10, GE Healthcare Bio-Sciences AB, Uppsala, Sweden). The fractions containing soluble TCR were analyzed by Coomassie-stained SDS-PAGE electrophoresis under reducing and non-reducing conditions. Fractions containing equal quantities of both α -chain and β -chain were pooled, concentrated on Amicon with a cutoff of 30 kDa to a concentration of about 2 mg ml⁻¹ and frozen at 193 K before further use. An aliquot was analyzed by mass spectrometry (MALDI-TOF) and confirmed the presence of RA15 TCR α/β dimers.

Prior to crystallization, samples of TCR were thawed and cleaned from aggregates and remaining single chains on a size-exclusion column (HiLoad 16/60 Superdex 75) equilibrated as previously and finally concentrated to a final concentration of 2–3 mg ml⁻¹.

2.3. Protein expression and purification of NLV-HLA-A2

HLA-A*0201 heavy chain and β_2 -microglobulin were produced separately as described previously (Garboczi *et al.*, 1992; Bodinier *et al.*, 2000; Gras, Saulquin *et al.*, 2009). In brief, the A245V mutant of the HLA-A*0201 heavy chain tagged at the C-terminus with a biotinylation sequence was cloned into pHN1 expression vector as described in Bodinier *et al.* (2000). Recombinant proteins were produced as inclusion bodies in *E. coli* strain XA90F/LaqQ1. The inclusion bodies were resuspended in 8 M urea, 50 mM MES pH 6.5, 0.1 mM DTT, 0.1 mM EDTA, incubated overnight at 277 K and centrifuged for 30 min at 100 000g. The supernatant was collected and frozen at 193 K. The pMHC complex refolding step was performed by flash dilution of a mixture of 21 mg HLA-A*0201, 10 mg β_2 -microglobulin and 10 mg of the NLV synthetic peptide into 350 ml 100 mM Tris pH 8.0, 400 mM L-arginine-HCl, 2 mM EDTA, 5 mM reduced glutathione, 0.5 mM oxidized glutathione with two Complete EDTA-free Cocktail protease-inhibitor tablets (Roche Diagnostics). The refolding solution was then incubated for 4–5 d at 277 K and concentrated with a 30 kDa cutoff membrane (Vivacell system). The pMHC complex was purified on a MonoQ 5/50 column with an FPLC system equilibrated in 10 mM Tris pH 8.0 buffer. It was eluted with

100–150 mM NaCl and concentrated with Amicon-10 or Amicon-30 devices to a final protein concentration of 2.5–3.5 mg ml⁻¹.

2.4. Crystallization

Prior to crystallization, RA15 TCR and HLA-A2-NLV were mixed in an equal molar ratio and at a final complex concentration of 3 mg ml⁻¹. Crystallization conditions were initially screened at 293 and 277 K using a Cartesian nanodrop crystallization robot available at the PSB HTX laboratory (Grenoble, France) and using JBScreen Classic screens (Jena Biosciences, Jena, Germany) and home-made screens with PEG (3350, 4000 and 6000), buffers at several pH values (NaCH₃COO, MES, HEPES and Tris) and salts (LiCl, LiSO₄, CaCl₂, MgCl₂ and MgSO₄).

Initial tiny crystals of the RA15-NLV-A2 complex were obtained after several weeks from these screens at 277 K and in conditions containing PEG 4000, Tris pH 8.0–8.5 and MgSO₄. These conditions were then refined using the hanging-drop technique by mixing 2 µl protein solution and 2 µl reservoir solution. The conditions were optimized by varying the concentration of PEG 4000 and MgSO₄ and the pH in the range 8.0–8.5 for the 100 mM Tris buffer.

The best crystals were obtained after several weeks at 277 K in conditions containing 17–18% PEG 4000, 100 mM Tris pH 8.5, 100–35 mM MgSO₄.

2.5. Diffraction data collection, processing and molecular replacement solution

Prior to being flash-frozen in liquid nitrogen for diffraction experiments, crystals were soaked in a cryoprotectant solution (mother liquor with the PEG 6000 concentration increased to 30%). Several crystals were tested for diffraction on beamline ID23-eh2 of the European Synchrotron Radiation Facility (ESRF, Grenoble, France) using an ADSC Q4 CCD detector at a wavelength of 0.873 Å. For most of the crystals, very weak diffraction was observed to 6 Å resolution. Only one crystal gave reasonable diffraction to 4.7 Å resolution (Fig. 1). A full data set was then collected at 100 K within 90° total angular range and a 1° oscillation step per image.

Data processing was performed using *XDS* (Kabsch, 1993) and is summarized in Table 1. The RA15-NLV-A2 complex crystals belonged to the tetragonal space group *P*4₂2₁2, with unit-cell parameters *a* = *b* = 99.14, *c* = 229.06 Å.

Structure determination of RA15-NLV-A2 was performed by molecular replacement with *AMoRe* (Navaza, 1994) using the RA14 and NVL-A2 crystal structures (Gras, Saulquin *et al.*, 2009; PDB code 3gsn) as initial and independent models. Clear contrasted rotation-

function solutions were found for both the TCR and the pMHC independently and the translation function provided a unique solution for one TCR-pMHC ternary complex per asymmetric unit. Extra rigid-body refinement cycles were performed with *AMoRe* by defining seven rigid bodies (NLV MHCα1α2 and MHCα3 domains, β₂-microglobulin, TCR Vα, Cα, Vβ and Cβ domains) with a final *R* factor of 50.9% at 4.7 Å resolution (Table 1).

3. Results and discussion

We have previously solved the structure of the public RA14 TCR in complex with its cognate NLV-A2 epitope (Gras, Saulquin *et al.*, 2009). This structure highlighted the structural characteristics that explain the immunodominance of this particular TCR in response to NLV-A2 associated with HCMV reactivation. Particularly, the emergence of an optimal public solution from an oligoclonal antigen-specific repertoire after repeated HCMV stimulations seems to be based on a TCR with a very favourable combination of TRAV, TRAJ and TRBV translating into an optimal structural complementarity between the TCR and the pMHC surface. Moreover, a significant number of RA14 TCR-contacting amino acids appear to be conserved by lower affinity TCRs, suggesting a shared TCR-pMHC docking mode and an antigen-driven selection of the best-fitted TCR. In order to validate this hypothesis, we focused our attention on another high-avidity public TCR directed against the immunodominant HCMV epitope, RA15 (TRAV24-TRAJ49-TRBV6-1). Interestingly, RA15 has less avidity for NLV-A2 than RA14 (Trautmann *et al.*, 2005) and shares only three of the four motifs of RA14 that contact the NLV-A2 surface (Gras, Saulquin *et al.*, 2009).

As our attempts to produce soluble RA15 TCR heterodimers with the strategy used for RA14 was unsuccessful, we looked for methods to stabilize the α/β heterodimer. Indeed, several strategies have been developed to achieve this goal, some of which work for several TCRs: expression of single-chain variable domain (scFv; Gregoire *et al.*, 1996; Housset *et al.*, 1997), fusion with a coil-coiled heterodimeriza-

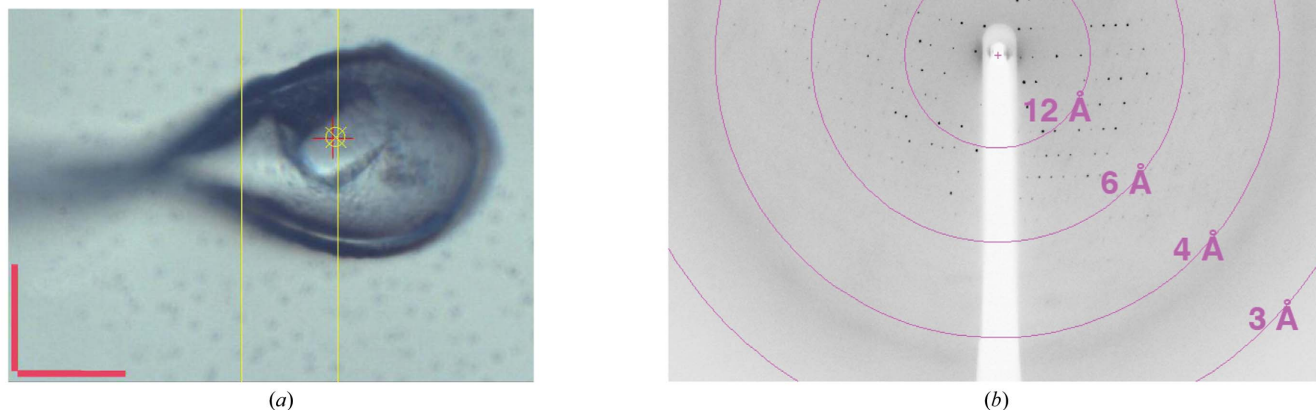


Figure 1

(a) View of the frozen RA15-NLV-A2 complex crystal in the loop. Scale bars represent 50 µm. (b) Diffraction pattern of the first 1° oscillation. Resolution circles and corresponding resolution are shown in pink.

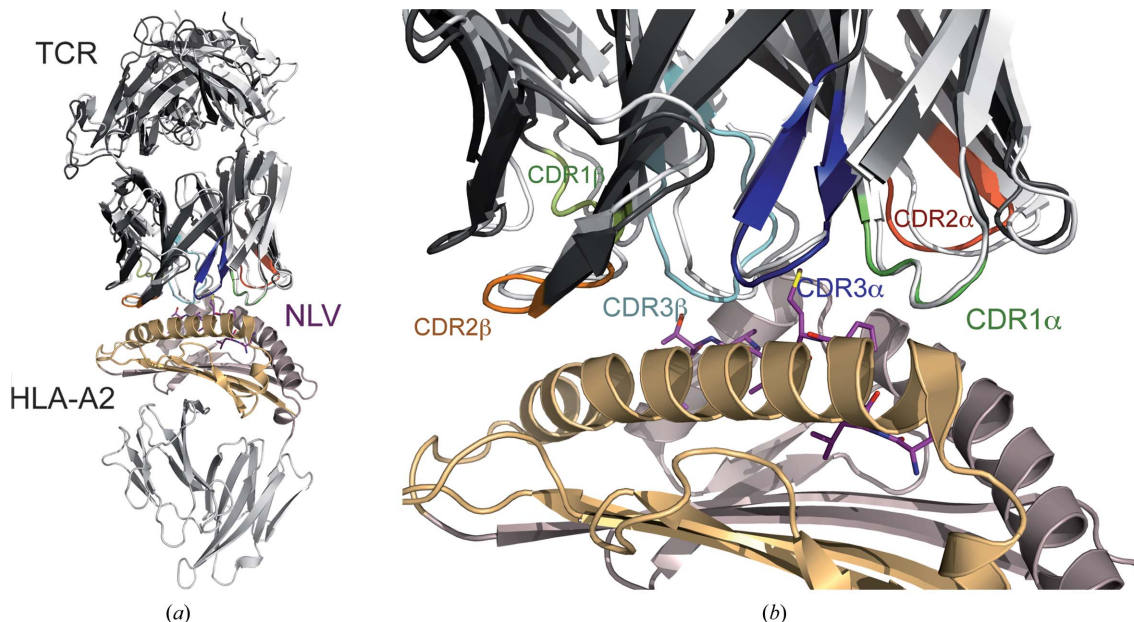


Figure 2 Superimposition of the molecular-replacement solution of the RA15–NLV–A2 complex (coloured) on the RA14–NLV–A2 complex structure (white). Superposition was performed on the $\alpha 1\alpha 2$ HLA–A2 domain. (a) Overall view of the superimposition, (b) closer view of the TCR–NLV–A2 interface. CDRs are coloured as follows: CDR1 α , green; CDR2 α , red; CDR3 α , dark blue; CDR1 β , pale green; CDR2 β , orange; CDR3 β , light blue. The NLV peptide is represented by violet sticks, the HLA–A2 $\alpha 1$ helix in gold and the $\alpha 2$ helix in silver.

tion motif (Willcox *et al.*, 1999), a membrane-proximal disulfide bridge (Stewart-Jones *et al.*, 2003) and a non-native interchain disulfide bridge (Boulter *et al.*, 2003). As the latter was successfully applied to characterize TCR–pMHC interactions (Boulter *et al.*, 2003) and ternary complex structure determination (Sami *et al.*, 2007; Archbold *et al.*, 2009; Dunn *et al.*, 2006; Tynan *et al.*, 2007; Gras, Burrows *et al.*, 2009), we decided to test it on RA15. By creating an artificial disulfide bridge between residue 48 of the C α domain and residue 57 of the C β domain, we were able to produce several milligrams of cysteine mutated RA15 TCR in a soluble and homogeneous form and to crystallize it in complex with its cognate pMHC.

The crystallization strategy was based on the usage of commercial screening kits and home-made screens designed from the conditions in which other TCR–pMHC complexes were previously crystallized. This strategy appears to be suitable to find optimized conditions for a particular complex. Indeed, although TCR–pMHC ternary complexes are prone to crystallize in the presence of PEG, the buffer and pH, as well as the presence and type of secondary salts and additives, are critical to obtain well diffracting crystals. It is therefore necessary to test large screens of buffers and additives. To date, the best crystals of RA15–NLV–A2 appeared in the presence of 17–18% PEG 4000, 100 mM Tris pH 8.0–8.5 and MgSO₄ and the crystals tested diffracted to 4.7 Å resolution.

Using the single and complete data set obtained using synchrotron radiation, molecular replacement was employed to solve the low-resolution structure of the RA15–NLV–A2 complex as it crystallized in a different space group to that observed for RA14–NLV–A2 (Gras, Saulquin *et al.*, 2009). Rigid-body refinement was then performed at 4.7 Å resolution.

Although at a very low resolution, this new structure of a public TCR directed against the immunogenic NLV–HLA–A2 antigen provides some interesting structural features of the public nature of CMV epitope recognition. Indeed, the structure shows that RA15 docks on NLV–A2 in quite a similar way compared with the public RA14 TCR for which we have previously solved the structure. When

the $\alpha 1\alpha 2$ A2 domains are superimposed, the RA15 variable domain has to be rotated by 8.5° with a pivotal point close to the tip of the V α CDR1 loop and shifted by about 0.5 Å to be fitted onto the RA14 variable domain (Fig. 2). At the TCR–pMHC interface, the best superposed CDR is the CDR1 of the V α domain with a 0.3 Å translational shift, while the positions of the five other CDRs vary from 1.1 to 2.7 Å for CDR2 of the V β domain. Quite logically, the most pronounced positional difference is observed for the domain that differs most in sequence. A small difference in the V α –V β domain pairing of RA15 is also observed, with a rotation of 4.4° for the V β domain relative to the orientation of the V β domain in RA14. Despite these observed differences, the NLV–A2 surface buried by the RA15 TCR is likely to be very similar to that buried by RA14. However, owing to the resolution of the present structure of the RA15–NLV–A2 complex, it is not possible to obtain a detailed view of the interactions formed by this RA15 TCR to recognize its epitope. Higher resolution crystallographic data will be required. These may possibly be obtained by further crystallization screenings, either by using finer steps in screening around the already determined crystallization conditions or by looking for additives that may help to obtain crystals of better quality.

We thank the ID23-eh2 team for help with synchrotron data collection at ESRF (Grenoble, France), the staff of the EMBL HTX laboratory (Grenoble, France) for the use of the PSB crystallization platform and E. Forest (IBS, Grenoble, France) for the use of mass spectroscopy. This study was supported in part by the Agence Nationale de la Recherche (grant ANR-05-MIIM-019). MB was also supported by the EPI-PEPVAC European Union grant.

References

Archbold, J. K., Macdonald, W. A., Gras, S., Ely, L. K., Miles, J. J., Bell, M. J., Brennan, R. M., Beddoe, T., Wilce, M. C., Clements, C. S., Purcell, A. W.,

- McCluskey, J., Burrows, S. R. & Rossjohn, J. (2009). *J. Exp. Med.* **206**, 209–219.
- Bodiniér, M., Peyrat, M. A., Tournay, C., Davodeau, F., Romagne, F., Bonneville, M. & Lang, F. (2000). *Nature Med.* **6**, 707–710.
- Borysiewicz, L. K., Hickling, J. K., Graham, S., Sinclair, J., Cranage, M. P., Smith, G. L. & Sissons, J. G. (1988). *J. Exp. Med.* **168**, 919–931.
- Boulter, J. M., Glick, M., Todorov, P. T., Baston, E., Sami, M., Rizkallah, P. & Jakobsen, B. K. (2003). *Protein Eng.* **16**, 707–711.
- Diamond, D. J., York, J., Sun, J. Y., Wright, C. L. & Forman, S. J. (1997). *Blood*, **90**, 1751–1767.
- Dunn, S. M., Rizkallah, P. J., Baston, E., Mahon, T., Cameron, B., Moysey, R., Gao, F., Sami, M., Boulter, J., Li, Y. & Jakobsen, B. K. (2006). *Protein Sci.* **15**, 710–721.
- Engstrand, M., Tournay, C., Peyrat, M. A., Eriksson, B. M., Wadstrom, J., Wirgart, B. Z., Romagne, F., Bonneville, M., Totterman, T. H. & Korsgren, O. (2000). *Transplantation*, **69**, 2243–2250.
- Garboczi, D. N., Hung, D. T. & Wiley, D. C. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 3429–3433.
- Gras, S., Burrows, S. R., Kjer-Nielsen, L., Clements, C. S., Liu, Y. C., Sullivan, L. C., Bell, M. J., Brooks, A. G., Purcell, A. W., McCluskey, J. & Rossjohn, J. (2009). *Immunity*, **30**, 193–203.
- Gras, S., Saulquin, X., Reiser, J.-B., Debeauvais, E., Echasserieu, K., Kissenpfennig, A., Legoux, F., Chouquet, A., Le Gorrec, M., Machillot, P., Neveu, B., Thielens, N., Malissen, B., Bonneville, M. & Housset, D. (2009). *J. Immunol.* **183**, 430–437.
- Gregoire, C., Malissen, B. & Mazza, G. (1996). *Eur. J. Immunol.* **26**, 2410–2416.
- Housset, D., Mazza, G., Gregoire, C., Piras, C., Malissen, B. & Fontecilla-Camps, J. C. (1997). *EMBO J.* **16**, 4205–4216.
- Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
- Lefranc, M. P. (2001). *Nucleic Acids Res.* **29**, 207–209.
- McLaughlin-Taylor, E., Pande, H., Forman, S. J., Tanamachi, B., Li, C. R., Zaia, J. A., Greenberg, P. D. & Riddell, S. R. (1994). *J. Med. Virol.* **43**, 103–110.
- Navaza, J. (1994). *Acta Cryst. A* **50**, 157–163.
- Peggs, K., Verfuërth, S., Pizzey, A., Ainsworth, J., Moss, P. & Mackinnon, S. (2002). *Blood*, **99**, 213–223.
- Quinnan, G. V. Jr, Burns, W. H., Kirmani, N., Rook, A. H., Manischewitz, J., Jackson, L., Santos, G. W. & Saral, R. (1984). *Rev. Infect. Dis.* **6**, 156–163.
- Sami, M., Rizkallah, P. J., Dunn, S., Molloy, P., Moysey, R., Vuidepot, A., Baston, E., Todorov, P., Li, Y., Gao, F., Boulter, J. M. & Jakobsen, B. K. (2007). *Protein Eng. Des. Sel.* **20**, 397–403.
- Saulquin, X., Ibsch, C., Peyrat, M. A., Scotet, E., Hourmant, M., Vie, H., Bonneville, M. & Houssaint, E. (2000). *Eur. J. Immunol.* **30**, 2531–2539.
- Sissons, J. G., Bain, M. & Wills, M. R. (2002). *J. Infect.* **44**, 73–77.
- Stewart-Jones, G. B., McMichael, A. J., Bell, J. I., Stuart, D. I. & Jones, E. Y. (2003). *Nature Immunol.* **4**, 657–663.
- Trautmann, L., Rimbart, M., Echasserieu, K., Saulquin, X., Neveu, B., Dechanet, J., Cerundolo, V. & Bonneville, M. (2005). *J. Immunol.* **175**, 6123–6132.
- Tynan, F. E., Reid, H. H., Kjer-Nielsen, L., Miles, J. J., Wilce, M. C., Kostenko, L., Borg, N. A., Williamson, N. A., Beddoe, T., Purcell, A. W., Burrows, S. R., McCluskey, J. & Rossjohn, J. (2007). *Nature Immunol.* **8**, 268–276.
- Willcox, B. E., Gao, G. F., Wyer, J. R., O’Callaghan, C. A., Boulter, J. M., Jones, E. Y., van der Merwe, P. A., Bell, J. I. & Jakobsen, B. K. (1999). *Protein Sci.* **8**, 2418–2423.
- Wills, M. R., Carmichael, A. J., Mynard, K., Jin, X., Weekes, M. P., Plachter, B. & Sissons, J. G. (1996). *J. Virol.* **70**, 7569–7579.

Chapitre IV

**Une nouvelle voie de signalisation de la
phagocytose**



Research paper

The C-terminal polyproline-containing region of ELMO contributes to an increase in the life-time of the ELMO-DOCK complex

Marion Sévajol^{a,1}, Jean-Baptiste Reiser^{a,1}, Anne Chouquet^a, Julien Pérard^{a,b}, Isabel Ayala^c, Pierre Gans^c, Jean-Philippe Kleman^{a,*}, Dominique Housset^{a,**}

^a Immune response to pathogens and altered-self group, Institut de Biologie Structurale Jean-Pierre Ebel, Commissariat à l'Energie Atomique et aux Energies Alternatives, Centre National de la Recherche Scientifique, Université Joseph Fourier-Grenoble 1, 41, rue Jules Horowitz, F-38027 Grenoble, France

^b Unit of Virus Host Cell Interactions, Unité Mixte Internationale 3265, Centre National de la Recherche Scientifique, European Molecular Biology Laboratory, Université Joseph Fourier-Grenoble 1, 6 rue Jules Horowitz, F-38042 Grenoble, France

^c Biomolecular NMR spectroscopy group, Institut de Biologie Structurale Jean-Pierre Ebel, Commissariat à l'Energie Atomique et aux Energies Alternatives, Centre National de la Recherche Scientifique, Université Joseph Fourier-Grenoble 1, 41, rue Jules Horowitz, F-38027 Grenoble, France

ARTICLE INFO

Article history:

Received 10 May 2011

Accepted 29 November 2011

Available online 6 December 2011

Keywords:

DOCK180

ELMO

Phagocytosis

Cell motility

NMR

SPR

ABSTRACT

The eukaryotic Engulfment and Cell Motility (ELMO) proteins form an evolutionary conserved family of key regulators which play a central role in Rho-dependent biological processes such as engulfment and cell motility/migration. ELMO proteins interact with a subset of Downstream of Crk (DOCK) family members, a new type of guanine exchange factors (GEF) for Rac and cdc42 GTPases. The physiological function of DOCK is to facilitate actin remodeling, a process which occurs only in presence of ELMO. Several studies have determined that the last 200 C-terminal residues of ELMO1 and the first 180 N-terminal residues of DOCK180 are responsible for the ELMO-DOCK interaction. However, the precise role of the different domains and motifs identified in these regions has remained elusive. Divergent functional, biochemical and structural data have been reported regarding the contribution of the C-terminal end of ELMO, comprising its polyproline motif, and of the DOCK SH3 domain. In the present study, we have investigated the contribution of the C-terminal end of ELMO1 to the interaction between ELMO1 and the SH3 domain of DOCK180 using nuclear magnetic resonance spectroscopy and surface plasmon resonance. Our data presented here demonstrate the ability of the SH3 domain of DOCK180 to interact with ELMO1, regardless of the presence of the polyproline-containing C-terminal end. However, the presence of the polyproline region leads to a significant increase in the half-life of the ELMO1-DOCK180 complex, along with a moderate increase on the affinity.

© 2011 Elsevier Masson SAS. All rights reserved.

1. Introduction

The highly conserved eukaryotic ELMO proteins cooperate with certain members of the DOCK family and are involved in several cellular processes such as engulfment and cell motility [1]. The ELMO family protein from *C. elegans*, CED12, was originally identified in a screen for genes that control necrotic and apoptotic cell phagocytosis [2]. CED12 and its mammalian orthologs ELMO1 and ELMO2 have been shown to cooperate with CRKII and DOCK180 upstream of the activation of the small GTPase Rac1 [3]. DOCK family members are multidomain proteins with a Guanine

Exchange Factor (GEF) function for Rho-family GTPases, conveyed by a DOCK Homology Region (DHR-2). DOCK proteins regulate several Rac1- and cdc42-dependent processes that rely on actin cytoskeleton remodeling including phagocytosis and cell migration. Although the DHR-2 domain of DOCK180 has been shown to be sufficient to act as a GEF for Rac1 both *in vitro* and *in cellulo* [4,5], DOCK180 needs to cooperate with ELMO for proper Rac1 signaling *in vivo* [3,6–9]. Indeed, ELMO has been shown to help DOCK180 stabilize the nucleotide free form of Rac1 [7], to relieve the auto-inhibition of DOCK [10], and to facilitate the recruitment of DOCK to the membrane through a RhoG-activated ELMO interaction [11–13].

The N-terminal sequence of ELMO contains a first domain called RBD (residues 1–115) which interacts with GTP-loaded RhoG [11,14]; a second domain (EID, residues 115–280) which is known to interact with ERM proteins [14] and which is related to the Diaphanous Inhibitory Domain [14,15]; and the ELMO domain

* Corresponding author. Tel.: +33 (0)4 76 20 94 48.

** Corresponding author. Tel.: +33 (0)4 76 20 94 69.

E-mail addresses: jean-philippe.kleman@ibs.fr (J.-P. Kleman), dominique.housset@ibs.fr (D. Housset).¹ These authors contributed equally to the work.

(residues 318–491), which is distantly related to ELMOD2 Arl2 GAP proteins [16]. The C-terminal end of ELMO contains a Pleckstrin homology (PH) domain (residues 555–675) flanked by two α -helices (N α h: 532–558 and C α h: 680–699) and a polyproline motif containing flexible fragment (PxxP: 700–727, including proline at positions 707, 710–712, 714, 717).

A direct interaction involving the C-terminal region of ELMO1 and the N-terminal region of DOCK180 was strongly suggested by co-precipitation and engulfment rescuing experiments [3,8,9] and was mapped to the last 200 residues of ELMO1 and the first 200 residues of DOCK180 [5]. More specifically, co-precipitation experiments initially suggested that the SH3 domain of DOCK could interact with the PxxP motif of ELMO [3]. Subsequently, different functional assays have demonstrated the role of the ELMO PxxP motif and/or the SH3 domain of DOCK in engulfment [9] or cell migration [5,10]. The involvement of the ELMO C α h helix has also been demonstrated. Extensive mutagenesis of residue in C α h (L689A, M691S, E692D, R696K, L697A and L698A) was shown to cause a complete loss of cell migration [12]. It was also demonstrated, rather surprisingly, that a DOCK180 construct lacking the SH3 domain was still able to interact with ELMO and promote cell migration. Côté and colleagues subsequently demonstrated using pull-down assays that both the ELMO PxxP motif and the DOCK180 SH3 domains were dispensable for a tight ELMO-DOCK interaction [5]. From these data a minimal fragment of ELMO capable of interacting with DOCK was defined. This region includes the PH related homology region and its two flanking α -helix extensions (532–707). Hydrophobic residues on the N α h α -helix appear to be critical for DOCK180 binding. On the DOCK side, the region spanning residues 1–187 has been demonstrated to harbor the ELMO binding site, with a preponderant role for the 69–187 stretch [5].

The coordinates of a three-dimensional (3D) structure of the ELMO1 C-terminal region including the PH domain and the PxxP motif in complex with the first 167 residues of DOCK2 were recently released by the protein data bank (PDB; entry code 3A98). The 3D structure confirms that both of the α -helices that flank the PH domain of ELMO and the 69–167 α -helical bundle of DOCK are involved in the previously proposed binding interface [5]. However, an extensive interaction between the SH3 of DOCK and the PxxP motif of ELMO was also revealed. This interaction is consistent with functional data that suggest that the PxxP motif of ELMO and the SH3 domain of DOCK contribute to enhance actin cytoskeleton remodeling. It further confirms the ability of the ELMO polyproline motif to canonically interact with SH3 domains, as previously observed for the Hck SH3 domain [17].

In order to assess how the ELMO1 polyproline motif and the SH3 domain of DOCK could be implicated in the ELMO-DOCK complex formation, we investigated the interaction between ELMO1 and the SH3 domain of DOCK180 by surface plasmon resonance (SPR) and nuclear magnetic resonance (NMR) spectroscopy. Our data clearly show that the SH3 domain of DOCK180 interacts significantly with ELMO1. Furthermore, the C-terminal end of ELMO, containing the last α -helix C α h and the PxxP motif, is shown to contribute to the affinity of ELMO1 for the SH3 domain of DOCK180 by a factor 2.5 to 5 and leads to a significant decrease of the dissociation rate (by 3–5 times), indicating that the PxxP-SH3 interaction may increase the life-time of the ELMO-DOCK complex *in vivo*, supporting its functional significance.

2. Materials and methods

2.1. Cloning

The cDNA encoding the wild type murine ELMO1 was kindly provided by Dr Yoshinori Fukui. Full length and truncated forms of ELMO1 (Fig. 1) were generated by PCR and cloned into the pPROEXHtb vector using ligation-independent cloning [18] in frame with a 6His tag.

The C-terminal ELMO1 polypeptide ELMO_C (residues 678–727) was sub-cloned in fusion with Ubiquitin and a 6His tag into pHub-H SacII EcoRI (see supplementary material).

The cDNA coding for the human DOCK1-SH3 domain (9–72) was optimized for bacterial expression, synthesized (Geneart AG) and cloned into the NcoI and EcoRI sites of pET-28a.

2.2. Expression and purification

The protein constructs were overexpressed in *E. coli* BL21-DE3 cells using conventional IPTG induction. For ELMO1 constructs, cells were lysed using a Microfluidizer[®] (Microfluidics) in Phosphate-Buffered Saline, supplemented with DNase (25 μ g/ml), 1 mM MgCl₂ and Complete[®] cocktail (Roche Diagnostics). The lysate was clarified by centrifugation and loaded on a Ni-NTA column (Qiagen). The protein was eluted by imidazole and further purified by gel filtration. DOCK180 SH3 is expressed as inclusion bodies. Cells were lysed (Fastprep[®], MP biochemicals) in 20 mM Tris pH 8.0, 100 mM NaCl. Inclusion bodies were pelleted by centrifugation and washed successively with 2 M NaCl, 2% Triton, and lysis buffer. A final wash was performed before the inclusion bodies were solubilized in 8 M urea. Unfolded SH3 was further

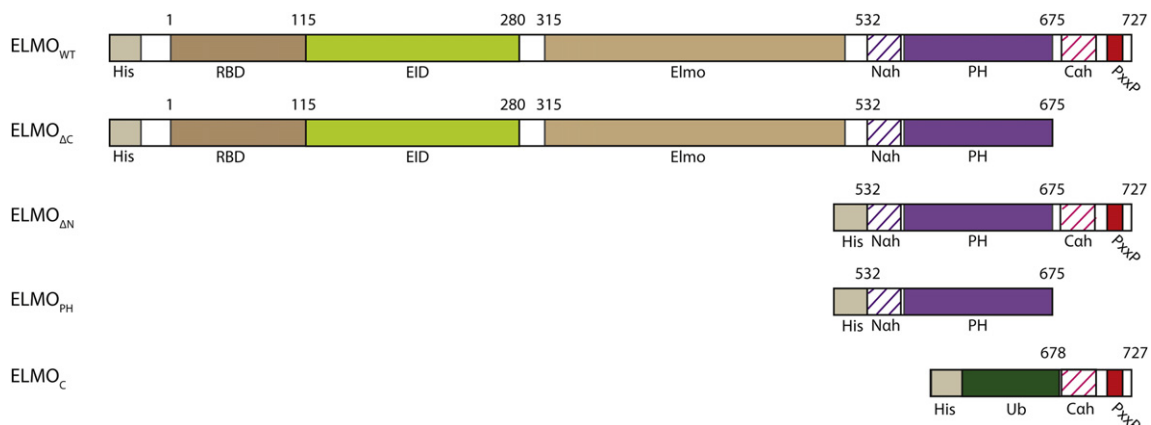


Fig. 1. Sketch of the ELMO constructs. Full length ELMO_{WT} (top) and ELMO_{ΔC} 1–675, ELMO_{PH} 532–675 and ELMO_{ΔN} 532–727 truncated forms of ELMO1 were expressed fused to an N-terminal 6His tag. A polypeptide comprising the C-terminus of ELMO (ELMO_C; residues 678–727) was expressed as a fusion with His-Ubiquitin (see supplementary material for details).

purified by size exclusion chromatography in 6 M Urea. Purified SH3 was refolded by dialysis against Tris-NaCl buffer and stored in 5% glycerol.

ELMO_C purification was achieved by nickel affinity chromatography. NiNTA resin was equilibrated in 20 mM Tris pH 7.5, 100 mM NaCl, Complete[®], 5 mM 2-mercaptoethanol and the bound protein was eluted with imidazole before being further purified by size exclusion chromatography. Where specified, Ubiquitin was removed by Ubiquitin-Hydrolase as published [19]. The quality of the purified polypeptide was assessed by mass spectrometry.

2.3. SPR experiments

The interaction between ELMO1 and the SH3 domain of DOCK180 was analyzed by SPR using a BIAcore 3000 (GE Healthcare) at 25 °C. Purified SH3 domain was diluted to 20 µg/ml in 10 mM sodium acetate, pH 4.1 and was immobilized by classical amine coupling onto two adjacent cells of a CM5 sensor-chip at a flow rate of 5 µl/min using 25 mM sodium phosphate pH 8.0, 100 mM NaCl and 0.005% surfactant P20 as running buffer. The blank cell was generated by ethanolamine deactivation of the activated surface. Preliminary experiments carried out with different ELMO constructs unambiguously showed a specific binding to the sensor-chip coupled SH3 domain (data not shown). However, as most common chip regeneration procedures failed to remove the bound ELMO without damaging the coupled SH3 domain, we decided to use the single cycle kinetic titration approach [20]. Each cycle consisted of five successive injections of increasing protein concentrations, at a flow rate of 30 µl/min using a 200 s contact time and a 400 s dissociation time. Prior to each protein binding cycle, the running buffer was injected for a five-step cycle. This buffer response was used as a blank for double-reference.

Kinetics parameters were initially evaluated within the context of a Langmuir 1:1 model using the BIAevaluation3.1 software. However, the fit obtained was not optimal, with χ^2 values ranging from 2.5 to 4 for data obtained with the higher level of SH3 immobilization. Consequently, the freely available general purpose fitting software, Fityk [21], was used to analyze our data. Fityk scripts defining integrated rate equations for the 1:1 and double ligand models were used to perform non-linear global fitting on the two curves corresponding to the two SH3 immobilization levels simultaneously. We based our analysis on the 1:1 model derived kinetic constants.

2.4. NMR spectroscopy

ELMO1 peptides 702–716 and 700–727 were chemically synthesized (NeoMPS, Strasbourg, France). Recombinant uniformly-labeled SH3 was overexpressed in *E. coli* bacteria grown in M9 minimal medium, as previously described [22], and purified as described above.

Uniformly [¹⁵N]-labeled samples of the SH3 domain of DOCK180 were prepared to a concentration of 0.2 mM in 18 mM Tris-base buffer, pH 7.2, 10% D₂O and [¹³C ¹⁵N]-labeled SH3 in 25 mM Na₂PO₄ pH 7.5, 0.1 M NaCl, 10% D₂O, 5% glycerol. Interactions between ELMO1 peptides and the SH3 domain were investigated by the addition of aliquots of a concentrated solution of unlabeled chemically synthesized peptide (typically 12 mM ELMO_{702–716}, or 8 mM ELMO_{700–727}) directly to isotope-labeled SH3 domain to a final molar ratio of 1:7. An equimolar sample of ELMO_C and SH3 were prepared by adding 500 µL of a 60 µM sample of cleaved ELMO_C to the same volume of a 60 µM SH3 before concentration of the mixture to a final volume of 450 µL. Complexes between the SH3 domain and ELMO_{WT}, ELMO_{ΔN} or peptides were prepared by addition of aliquots of a concentrated solution of the unlabeled

sample (typically 800 µM) directly to an 80 µM sample of isotope-labeled SH3 domain to achieve a final molar ratio of 1:1.1. NMR spectra were recorded at 25 °C on either a Varian Inova 800 MHz or DirectDrive 600 MHz spectrometer equipped with a triple-resonance (¹H, ¹³C, ¹⁵N) cryoprobe and shielded z-gradients. All NMR experiments were acquired using standard pulse sequences supplied in the Varian Protein Pack (available at www.varianinc.com) with an experimental duration between 42 and 44 min. Data were processed and analyzed with the nmrPipe/nmrDraw suite of programs [23].

3. Results and discussion

In order to determine the precise contribution of the ELMO C-terminal region to the ELMO-DOCK complex formation, we performed a series of different binding experiments by NMR spectroscopy and SPR techniques. We tested the interaction of a short construct of DOCK180 SH3 domain (residues 9–72, hereafter named SH3) with several ELMO1 constructs (Fig. 1), including: the full length ELMO1 (ELMO_{WT}); the ELMO1-1-675 N-terminal fragment (ELMO_{ΔC}); the ELMO1 fragment lacking the first 531 residues (ELMO_{ΔN}); the ELMO1 PH domain including the N α h helix (residues 532–675, ELMO_{PH}); and C-terminal peptides containing the PxxP motif (702–716, 700–727, and ELMO_C 678–727).

In contrast with published experiments [5], our preliminary NMR experiments clearly show that the DOCK180 SH3 domain can interact with ELMO_{WT} in solution. Addition of a stoichiometric amount of 727 residue ELMO_{WT} protein to [¹⁵N]-labeled SH3 domain leads to complete disappearance of the SH3 NMR signal, confirming the binding of the SH3 domain to this large protein (data not shown). Addition of the ELMO1 C-terminal region (ELMO_{ΔN}), leads to a similar result (Fig. 2), clearly indicating that the SH3 binding zone is located in this C-terminal part. Fig. 2 also shows that the impact on SH3 NMR spectra increases with the amount of ELMO_{ΔN}, suggesting a dose dependent effect.

However, previously published results suggested a very limited role, if any, of the ELMO polyproline motif in the complex formation with DOCK180 [5,12]. In order to evaluate this assumption, we studied the interactions between the ¹⁵N labeled SH3 domain and the ELMO_{702–716} or ELMO_{700–727} peptides. Very small chemical shift variations were observed for a small number of NH correlations and only with a large excess of peptide. Chemical shifts also increase with the amount of peptide added to the solution, further supporting a dose dependent phenomenon (Fig. 3). Taking into account the experimental conditions, these data suggests that the affinity is very weak.

These preliminary results induced a more quantitative investigation using SPR. Binding experiments of both ELMO_{WT} and ELMO_{ΔC} on the sensor-chip-immobilized SH3 domain were performed and clearly demonstrated a specific binding of both ELMO constructs. The use of the single cycle kinetics titration approach [20] allowed us to obtain for the first time, estimates of the kinetic parameters characterizing the ELMO1-SH3 interaction. The 1:1 Langmuir model was used and yielded a reasonable fit between measured and theoretical curves (Fig. 4). Interestingly, we observed that the ELMO_{ΔN} and ELMO_{PH} constructs also interacted specifically with the immobilized SH3 domain (data not shown), although no kinetic parameter could be reliably estimated. The comparison of the equilibrium dissociation constant (K_D) obtained by SPR showed up to a 5-fold higher affinity for wild type ELMO1 than for the ELMO_{ΔC} construct. However, this difference suggests a small energetic contribution (about 1 kcal/mol, as calculated by the difference in K_D) of the PxxP containing C-terminal end of ELMO1 to the overall stabilization of the complex, in agreement with NMR data described above. Differences in k_{off} are the most significant and

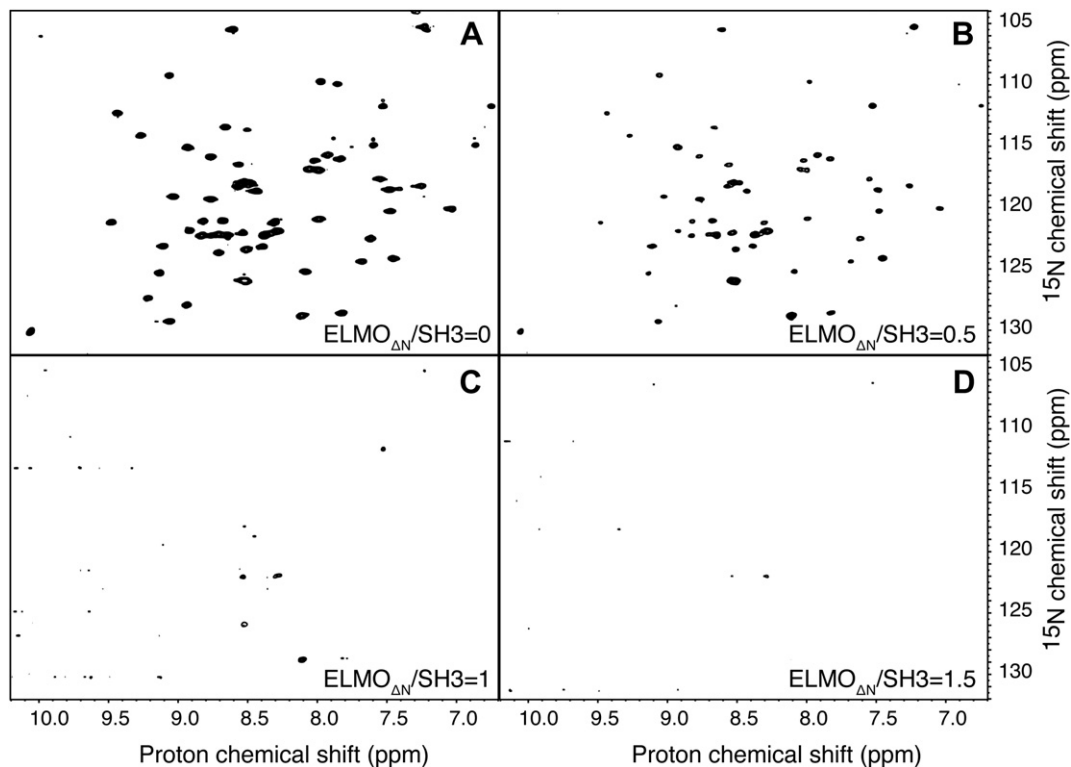


Fig. 2. 2D ^{15}N -HSQC spectra of Ubiquitin-tagged [^{15}N]-labeled SH3 (Ub- ^{15}N) SH3) domain in the presence of (A) 0; (B) 0.5; (C) 1.0; and (D) 1.5 equivalents of ELMO $_{\Delta\text{N}}$. NMR data was collected using an 800 MHz spectrometer at 25 °C.

indicate a 5 times slower dissociation rate for the full length ELMO1 in comparison to its truncated form.

The structure of ELMO1 C-terminus (530–727) in complex with the N-terminal region of DOCK2 (1–177, PDB entry 3A98)

reveals the nature of the interface between the DOCK SH3 domain (1–69) and ELMO. The most surprising observation is the extent of the interaction between the PxxP containing C-terminal end of ELMO (700–727) and the DOCK SH3 domain. This interaction,

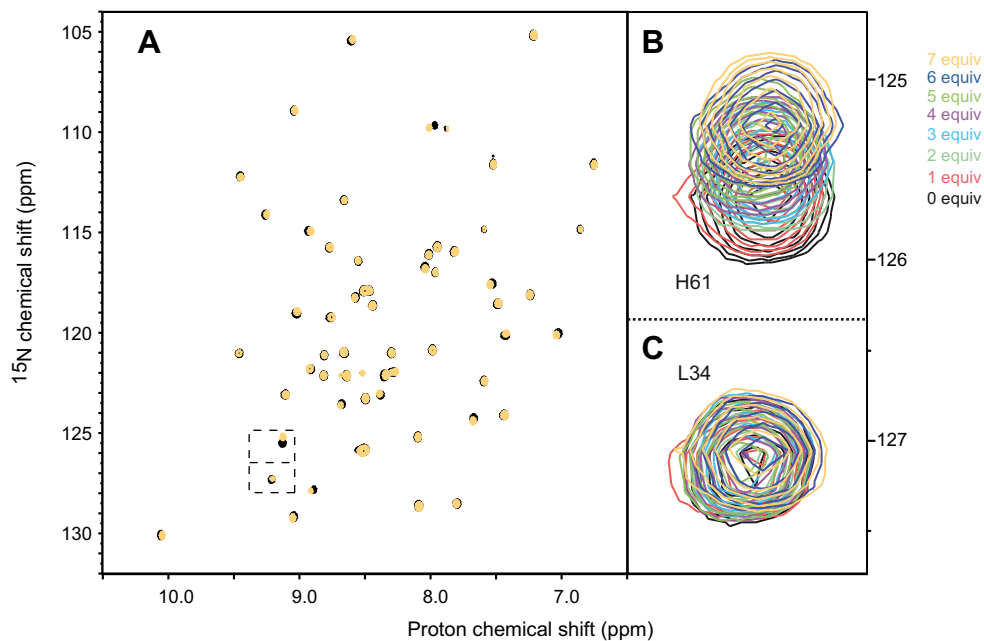


Fig. 3. 2D ^{15}N -HSQC spectra of 200 μM Ub- ^{15}N SH3 acquired using a 600 MHz spectrometer at 25 °C. (A) In the left panel, Ub- ^{15}N SH3 alone (black) and in presence of 7 equivalent amounts of the 702–716 peptide (yellow). In the right panel, the changes in (^1H , ^{15}N) resonance frequencies of (B) H61 and (C) L34 in response to the addition of increasing concentrations of 702–716 peptide. Each superposed spectrum is colored as described on the figure. The final 703–726 peptide concentration was 1.4 mM. The large excess of peptide required to induce variations in chemical shift indicates that, under these conditions, the affinity of the SH3 domain for the 702–716 peptide is in the millimolar range. Peak assignments were previously made on a U- ^{13}C , ^{15}N SH3 using conventional 3D NMR experiments (data not shown). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

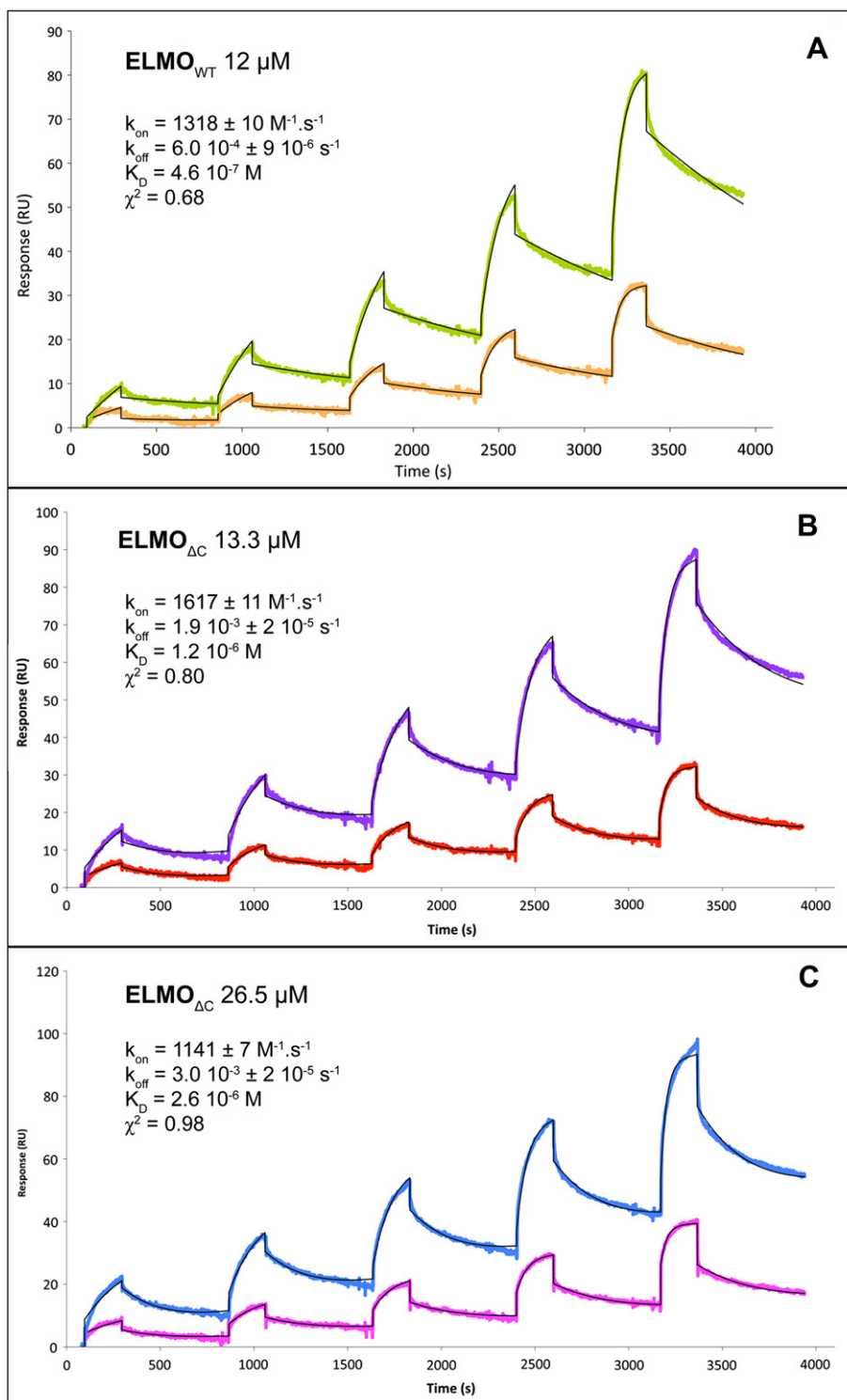


Fig. 4. SPR sensorgrams of ELMO_{WT} (A) and ELMO_{ΔC} (B, C) binding to immobilized SH3. The single cycle kinetic method was used with 5 two-fold dilutions series. Green and orange curves (panel A) correspond to the measured response for ELMO_{WT} with two different levels of SH3 immobilization (730 and 1700 RU, respectively). ELMO_{WT} was injected at 5 increasing concentrations: 0.75 μM, 1.5 μM, 3 μM, 6 μM, 12 μM. Panel B and C correspond to the measured response for ELMO_{ΔC} with the same two levels of SH3 immobilization. ELMO_{ΔC} was injected at 5 increasing concentrations: 0.83 μM, 1.66 μM, 3.3 μM, 6.6 μM, 13.3 μM (purple and red curves, panel B) and 1.66 μM, 3.3 μM, 6.6 μM, 13.3 μM, 26.5 μM (blue and pink curves, panel C). Fitted curves were calculated using a 1:1 Langmuir model with the software, Fityk [21] using non-linear optimization of integrated rate equations and are depicted with black solid lines. Kinetic constants derived from these curves are provided. The fit quality was reasonable, with χ^2 values of 0.68 (1.03 for the green curve and 0.33 for the orange curve) for ELMO_{WT}, 0.80 (1.27 for the purple curve and 0.30 for the red curve) for ELMO_{ΔC} at 13.25 μM concentration and 0.98 (1.53 for the blue curve and 0.42 for the pink curve) for ELMO_{ΔC} at 26.5 μM concentration. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

which buries approximately 1200 Å² for each fragment, represents about 52% of the whole ELMO-DOCK interface area (Figure S1). Another 3D structure of a chimeric protein in which the polyproline stretch of ELMO1 was fused to DOCK2-SH3,

clearly confirms the extent of this binding (PDB entry 2RQR). This chimeric structure also shows that the polyproline of ELMO1 interacts in an essentially canonical way with the SH3 domain of DOCK2. Since the N-terminal 167 residues of DOCK180 shares 60%

identity and 87% similarity with DOCK2, ELMO1 binding to DOCK180 is very likely to be similar. This large interface seems at odds with co-precipitation experiments in which both the ELMO polyproline and the DOCK SH3 domains appeared to be dispensable for a full ELMO-DOCK interaction [5,12]. A significant part of the SH3 surface (340 Å²) is buried by the two α -helices (N α h: 210 Å² and C α h: 130 Å²) that flank the ELMO1 PH domain (Figure S1). Our binding data suggest that most of the binding free energy is contributed by the SH3-ELMO1-N α h interface. A possible explanation for low energetic contribution of ELMO1_{700–727} may be found in its intrinsic flexibility. This region is predicted to be unstructured and the conformation adopted in both 3A98 and 2RQR structures does not contain any secondary structure element. This region is thus likely to be fully disordered in the unbound form of ELMO1. Accordingly, the entropic cost to stabilize this stretch in the conformation observed in the complex with DOCK2_{1–177} may compensate the enthalpic benefit of forming this large interface. Our data also suggest that the ELMO1 C α h helix, absent in our ELMO Δ C construct does not greatly contribute to the interaction with the SH3 domain. This result contradicts previous observations where mutations in the C α h helix were shown to impair the promotion of cell migration and co-precipitation with DOCK180 [12]. In solution, NMR spectroscopic studies suggest that the ELMO_C region binds weakly to SH3 (data not shown), further confirming our observations. However, it is not possible to exclude a slight change of the SH3 docking mode in the absence of this α -helix, where the loss of interactions with the ELMO N α h helix could be balanced by new ones within the C α h helix. Yet, the limited change in affinity reported here argues against a drastically different binding mode.

Our data may help to provide an explanation for the apparent contradiction between previously published functional and interaction data that sought to describe the role of both the ELMO PxxP motif and the DOCK SH3 domain in the ELMO-DOCK interaction. Indeed, on one hand, several functional assays indicated that the loss of the ELMO1 PxxP motif had a negative impact on engulfment rescuing capacity [9] or on cell migration [5]. On the other hand, the absence of either the SH3 domain or the ELMO1 PxxP had no effect on the strength of the ELMO-DOCK interaction, as analyzed by pull-down assays [5]. In order to resolve this apparent contradiction, Komander et al. suggested that the PxxP motif and the SH3 could interact with functionally significant, yet unidentified biological partners. The small impact of the presence of the PxxP motif on the affinity is consistent with co-precipitation experiments. The fact that the formation of a complex between the DOCK180 SH3 domain and ELMO1 was not observed [5,12] may also be explained by the relatively low affinity (0.5 μ M) between these two entities. Remarkably, we observed a five-fold difference in k_{off} between wild type ELMO1 and its Δ C fragment. As a consequence, the half-life of the ELMO-SH3 complex is five times longer (20 min) than that of the ELMO Δ C-SH3 complex (4 min). Thus, we propose that the functional role of the PxxP motif – SH3 interaction relies on its ability to increase both the life-time of the ELMO-DOCK complex and, albeit modestly, the overall stability of the complex.

Acknowledgments

We thank Isabelle Bally and Nicole Thielens, from the IBS platform of the Partnership for Structural Biology and the Institut de Biologie Structurale in Grenoble (PSB/IBS), for assistance and access to the SPR facility, Ombeline Pessey and Madalen Le Gorrec for their assistance in the protein production step and Michael Plevin for the careful reading of the manuscript. Part of this work was financially supported by the Cluster 10 from Région Rhône-Alpes.

Appendix. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.biochi.2011.11.014.

References

- [1] M. Lu, K.S. Ravichandran, Dock180-ELMO cooperation in Rac activation, *Meth Enzymol.* 406 (2006) 388–402.
- [2] S. Chung, T.L. Gumienny, M.O. Hengartner, M. Driscoll, A common set of engulfment genes mediates removal of both apoptotic and necrotic cell corpses in *C. elegans*, *Nat. Cell Biol.* 2 (2000) 931–937.
- [3] T. Gumienny, E. Brugnera, A. Tosello-Trampont, J. Kinchen, L. Haney, K. Nishiwaki, S.F. Walk, M. Nemerout, I. Macara, R. Francis, T. Schedl, Y. Qin, L. Van Aelst, M. Hengartner, K. Ravichandran, CED-12/ELMO, a novel member of the Crkl/Dock180/Rac pathway, is required for phagocytosis and cell migration, *Cell* 107 (2001) 27–41.
- [4] J.-F. Côté, K. Vuori, Identification of an evolutionarily conserved superfamily of DOCK180-related proteins with guanine nucleotide exchange activity, *J. Cell Sci.* 115 (2002) 4901–4913.
- [5] D. Komander, M. Patel, M. Laurin, N. Fradet, A. Pelletier, D. Barford, J. Cote, An alpha-helical extension of the ELMO1 pleckstrin homology domain mediates direct interaction to DOCK180 and is critical in Rac signaling, *Mol. Biol. Cell* 19 (2008) 4837–4851.
- [6] E. Brugnera, L. Haney, C.M. Grimsley, M. Lu, S.F. Walk, A. Tosello-Trampont, I. Macara, H. Madhani, G. Fink, K. Ravichandran, Unconventional Rac-GEF activity is mediated through the Dock180-ELMO complex, *Nat. Cell Biol.* 4 (2002) 574–582.
- [7] M. Lu, J. Kinchen, K. Rossman, C.M. Grimsley, C. deBakker, E. Brugnera, A. Tosello-Trampont, L. Haney, D. Klingele, J. Sondek, M. Hengartner, K. Ravichandran, PH domain of ELMO functions in trans to regulate Rac activation via Dock180, *Nat. Struct. Mol. Biol.* 11 (2004) 756–762.
- [8] Y. Wu, M. Tsai, L. Cheng, C. Chou, N. Weng, *C. elegans* CED-12 acts in the conserved crkl/Dock180/Rac pathway to control cell migration and cell corpse engulfment, *Dev. Cell* 1 (2001) 491–502.
- [9] Z. Zhou, E. Caron, E. Hartwig, A. Hall, H. Horvitz, The *C. elegans* PH domain protein CED-12 regulates cytoskeletal reorganization via a Rho/Rac GTPase signaling pathway, *Dev. Cell* 1 (2001) 477–489.
- [10] M. Lu, J. Kinchen, K. Rossman, C.M. Grimsley, M. Hall, J. Sondek, M. Hengartner, V. Yajnik, K. Ravichandran, A steric-inhibition model for regulation of nucleotide exchange via the Dock180 family of GEFs, *Curr. Biol.* 15 (2005) 371–377.
- [11] C. deBakker, L. Haney, J. Kinchen, C.M. Grimsley, M. Lu, D. Klingele, P. Hsu, B. Chou, L. Cheng, A. Blangy, J. Sondek, M. Hengartner, Y. Wu, K. Ravichandran, Phagocytosis of apoptotic cells is regulated by a UNC-73/TRIO-MIG-2/RhoG signaling module and armadillo repeats of CED-12/ELMO, *Curr. Biol.* 14 (2004) 2208–2216.
- [12] C.M. Grimsley, J. Kinchen, A. Tosello-Trampont, E. Brugnera, L. Haney, M. Lu, Q. Chen, D. Klingele, M. Hengartner, K. Ravichandran, Dock180 and ELMO1 proteins cooperate to promote evolutionarily conserved Rac-dependent cell migration, *J. Biol. Chem.* 279 (2004) 6087–6097.
- [13] H. Katoh, M. Negishi, RhoG activates Rac1 by direct interaction with the Dock180-binding protein Elmo, *Nature* 424 (2003) 461–464.
- [14] C.M. Grimsley, M. Lu, L.B. Haney, J.M. Kinchen, K.S. Ravichandran, Characterization of a novel interaction between ELMO1 and ERM proteins, *J. Biol. Chem.* 281 (2006) 5928–5937.
- [15] M. Patel, Y. Margaron, N. Fradet, Q. Yang, B. Wilkes, M. Bouvier, K. Hofmann, J.-F. Côté, An evolutionarily conserved autoinhibitory molecular switch in ELMO proteins regulates Rac signaling, *Curr. Biol.* (2010) 1–7.
- [16] J. Bowzard, D. Cheng, J. Peng, R. Kahn, ELMOD2 is an Arl2 GTPase-activating protein that also acts on Arfs, *J. Biol. Chem.* 282 (2007) 17568–17580.
- [17] M.P. Scott, F. Zappacosta, E.Y. Kim, R.S. Annan, W.T. Miller, Identification of novel SH3 domain ligands for the Src family kinase Hck. Wiskott-Aldrich syndrome protein (WASP), WASP-interacting protein (WIP), and ELMO1, *J. Biol. Chem.* 277 (2002) 28238–28246.
- [18] C. Aslanidis, P. de Jong, Ligation-independent cloning of PCR products (LIC-PCR), *Nucleic Acids Res.* 18 (1990) 6069–6074.
- [19] T. Kohno, H. Kusunoki, K. Sato, K. Wakamatsu, A new general method for the biosynthesis of stable isotope-enriched peptides using a decahistidine-tagged ubiquitin fusion system: an application to the production of mastoparan-X uniformly enriched with ¹⁵N and ¹⁵N/¹³C, *J. Biomol. NMR* 12 (1998) 109–121.
- [20] R. Karlsson, P.S. Katsamba, H. Nordin, E. Pol, D.G. Myszka, Analyzing a kinetic titration series using affinity biosensors, *Anal. Biochem.* 349 (2006) 136–147.
- [21] M. Wojdyr, Fityk: a general-purpose peak fitting program, *J. Appl. Cryst.* 43 (2010) 1126–1128.
- [22] D. Bouvier, C. Vanhaverbeke, J. Simorre, G.J. Arlaud, I. Bally, V. Forge, R.L. Margolis, P. Gans, J.-P. Kleman, Unusual Ca(2+)-calmodulin binding interactions of the microtubule-associated protein F-STOP, *Biochemistry* 42 (2003) 11484–11493.
- [23] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, Nmrpipe – a multidimensional spectral processing system based on unix pipes, *J. Biomol. Nmr* 6 (1995) 277–293.

Chapitre V

**Recherche de nouveaux agents
anticancéreux et antibactériens**

New potent dual inhibitors of CK2 and Pim kinases: discovery and structural insights

Miriam López-Ramos,^{*,†,‡,§,1} Renaud Prudent,^{||,¶,1} Virginie Moucadel,^{||,¶,1}
Céline F. Sautel,^{||,¶,1} Caroline Barette,^{**,††} Laurence Lafanechère,^{**,††}
Liliane Mouawad,^{‡,§} David Grierson,^{*,†} Frédéric Schmidt,^{*,†} Jean-Claude Florent,^{*,†}
Panagis Filippakopoulos,^{‡‡} Alex N. Bullock,^{‡‡} Stefan Knapp,^{‡‡,§§}
Jean-Baptiste Reiser,^{||,¶,1,2} and Claude Cochet^{||,¶,1,2}

*Centre de Recherche, Institut Curie, Paris, France; †Centre National pour la Recherche Scientifique (CNRS), Unité Mixte de Recherche (UMR) 176, Paris, France; ‡Centre de Recherche, Institut Curie, Université Paris-Sud, Orsay, France; §Institut National de la Santé et de la Recherche Médicale (INSERM) U759, Université Paris-Sud, Orsay, France; ||INSERM, U873, Grenoble, France; ¶Commissariat à l'Énergie atomique (CEA), Institut de Recherche en Technologies et Sciences pour le Vivant (iRTSV)/Laboratoire Transduction du Signal, Grenoble, France; #Université Joseph Fourier (UJF), Grenoble, France; **CEA, Direction des Sciences du Vivant (DSV), iRTSV/Centre de Criblage pour Molécules Bio-Actives (CBMA), Grenoble, France; ††CNRS, UMR 5168; CEA, DSV, iRTSV/CMBA, Grenoble, France; ‡‡Nuffield Department of Clinical Medicine and §§Department of Clinical Pharmacology, Oxford University, Oxford, UK; |||Institut de Biologie Structurale Jean-Pierre Ebel, CEA-CNRS-UJF, Grenoble, France; and ¶¶Partnership for Structural Biology, Grenoble, France

ABSTRACT Protein kinase casein kinase 2 (CK2) is a serine/threonine kinase with evidence of implication in growth dysregulation and apoptosis resistance, making it a relevant target for cancer therapy. Several CK2 inhibitors have been developed showing variable efficiency, emphasizing the need to expand the chemical diversity of those inhibitors. We report the identification and characterization of 2,8-difurandicarboxylic acid derivatives as a new class of nanomolar ATP-competitive inhibitors. Selectivity profiling pointed out proviral insertion Moloney virus kinases (Pim kinases) as the only other kinases that are significantly inhibited. By combining structure-activity relationship analysis with structural determination, we were able to determine the binding mode of these inhibitors for both kinases and to explain their strong inhibitory potency. Essential chemical features necessary for activity on both kinases were then identified. The described compounds are not cell permeable: however, they could provide a lead for developing novel inhibitors usable also *in vivo*. Given the similar but not redundant pathophysiological functions of CK2 and Pim family members, such inhibitors would provide new attractive leads for targeted cancer therapy. This work highlights that 2 functionally related kinases from different kinase branches display exquisite sensitivity to a common inhibitor.—López-Ramos, M., Prudent, R., Moucadel, V., Sautel, C. F., Barette, C., Lafanechère, L., Mouawad, L., Grierson, D., Schmidt, F., Florent, J.-C., Filippakopoulos, P., Bullock, A. N., Knapp, S., Reiser, J.-B., Cochet, C. New potent dual inhibitors of CK2 and Pim kinases: discovery and structural insights. *FASEB J.* 24, 3171–3185 (2010). www.fasebj.org

Key Words: X-ray crystallography • binding mode • comparative structural study • high-throughput screening • selective inhibition

PROTEIN KINASE CASEIN KINASE 2 (CK2) and proviral insertion Moloney virus kinases (Pim kinases) are serine/threonine kinases that participate in a wide variety of cellular processes, including cell differentiation, proliferation, and survival. Their key role in various stages of the cell cycle and in the regulation of apoptosis has been demonstrated (1–4). Both kinases act on similar but nonredundant pathways involved in tumorigenesis, making simultaneous inhibition of these kinases relevant for cancer therapy, in particular for treatment of leukemia, prostate cancer, and lymphomas (5–7).

Of particular interest are reports suggesting that abnormal CK2 expression and function are associated with a number of pathologies, including inflammation, infections, and neoplastic growth. In several experimental models, the CK2 catalytic subunits (α , α') behave as oncogenes and their overexpression may contribute to the enhancement of the tumor phenotype, notably by promoting cell survival (3). Notably, the recent demonstration that CK2 overexpression is an unfavorable prognostic marker in prostate cancer (8), in lung cancer (9), and in acute myeloid leukemia (10) suggests that CK2 may represent a promising therapeutic

¹ These authors contributed equally to this work.

² Correspondence: C.C., INSERM, U873, Grenoble, F-38054, France. E-mail: ccochet1@cea.fr; J.B.R., Institut de Biologie Structurale Jean-Pierre Ebel, CEA-CNRS-UJF, 41 rue Jules Horowitz, F-38027 Grenoble. E-mail: jean-baptiste.reiser@ibs.fr
doi: 10.1096/fj.09-143743

target. The feasibility of targeting CK2 for cancer therapy (5) has been shown by experiments wherein disruption of CK2 activity *in vitro* with antisense RNA resulted in induction of cell apoptosis. *In vivo*, the concept of inhibiting CK2 for cancer treatment has been validated in a prostate cancer xenograft (11). In addition, many viral proteins are CK2 substrates, suggesting that selective inhibitors may find applications for the treatment of viral infection. For these reasons, compounds targeting CK2 could behave both as anti-neoplastic and anti-infectious agents (12).

Similarly to CK2, Pim kinases display constitutive activity (13). Diverging from the other members of the kinase, CK2 and Pim kinases display unique characteristics explaining this constitutive activity. Unlike most other eukaryotic Ser/Thr kinases, the absence of a phosphorylation site in their activation loop appears as one of the crucial determinants of their constitutive activation. In the case of CK2, an unusual tryptophan that replaces the central Phe residue in the canonical amino acid triplets DFG (DWG in CK2) allows an additional hydrogen bond that favors an active conformation (14). Furthermore, a structural chloride ion together with a conserved water cluster supports a tight contact between the N-terminal domain and the activation loop in its open and consequently active conformation (15, 16). Similarly, among the Pim kinases, phosphorylation of Pim-1 is not necessary for its activity. The Pim-1 structure reveals an extensive network of hydrophobic interactions and hydrogen bond between the activation and the catalytic loops that are critical for opening the ATP pocket and maintaining its active conformation (13).

In addition to structural analogies, a striking functional similarity has been observed between CK2 and Pim-1 kinases. Like CK2, members of the Pim kinase family are overexpressed in prostatic adenocarcinoma (6, 17) where they promote chemoresistance of cancer cells by increasing hypoxia-induced drug resistance (18) and by inactivating BAD (7, 19). Pim-1, like CK2, interacts with Myc and is required for Myc-dependent cell transformation. The mechanistic basis of CK2 and Pim kinase dysregulation is currently unclear. However, their overexpression in cancer cells might not only be related to the increased proliferative capacity of dedifferentiated tumor cells but also to their marked resistance to apoptotic signals. CK2 and Pim activity has been implicated in transcriptional regulation, differentiation, and development, DNA damage signaling, and cell survival (3, 20–25). This broad range of cellular response is orchestrated through the regulation of survival and hypoxia pathways. For example, both kinases participate in the NF- κ B activation observed in transformed cells. Increased CK2 activity induces aberrant activation of NF- κ B, which, in turn, modulates the survival and the transformed phenotype of breast cancers (26, 27). Similarly, lymphocyte transformation by Pim-2 is dependent on NF- κ B activation (28) and expression of Pim-1 was shown to be a critical component of a NF- κ B-dependent survival pathway activated by docetaxel in prostate cancer cells (24, 29). Phosphorylation of the Bcl-2 family member Bad on specific residues by CK2 and Pim kinases inhibits its proapoptotic function in lym-

phoma and prostate cancer promoting cell proliferation and preventing cell death (10, 19, 30). HIF-1 and VEGF are mechanistically linked to intratumor hypoxia by activating pathways that contribute to cancer cell survival and by playing important role in angiogenesis and tumor progression (31). Indeed, it has been observed that CK2 activity is increased in hypoxic conditions and participates to the up-regulation of HIF-1 transactivation activity (32, 33). Similarly, Pim-1 is induced by hypoxia and plays a pivotal role in hypoxia-induced chemoresistance (18, 34). Thus, overexpression of CK2 and Pim kinases in cancer could participate in the molecular circuits that propagate the cancer phenotype through the HIF-1/VEGF interplay. Altogether, a high expression level of CK2 and Pim kinases in cancer cells may generate a favorable cellular environment to the progression of a tumor toward an aggressive phenotype.

Therefore, the development of potent and selective CK2 and Pim inhibitors has a strong therapeutic interest and will also provide a powerful chemical tool for dissecting the specific biological functions of these pleiotropic protein kinases (20). Extensive efforts have been focused on developing CK2 inhibitors, and several active, and in certain cases selective, ATP site-directed compounds have been identified (35). These include condensed polyphenolic compounds such as emodin and derivatives of hydroxycoumarins (3-carboxy-4(1H)-quinolone), the indoloquinazoline derivative (IQA), tetrabromocinnamic acid (TBCA), 4,5,6,7-tetrabromo-1-benzotriazole (TBB), and pyrazolo[1,5-a](1,3,5)triazine derivatives (36). Recently, CX-4945 (Cylene Pharmaceuticals, San Diego, CA, USA) has been reported to be a very efficient CK2 inhibitor with good *in vivo* potency (37). Also, several groups (38–41) have reported the crystal structure of Pim-1 in complex with a number of ATP mimetic kinase inhibitors.

In the present study, we report the identification of three 2,8-difurandicarboxylic acid derivatives as new potent ATP-competitive CK2 inhibitors 2,8-dicarboxynaphtho[2,1-b:7,6-b']difuran 1 (CPA 1), 2,8-dicarboxyanthra[2,1-b:7,6-b']difuran 2 (CPB 2), and AMR 14, which are structurally unrelated to previously described inhibitors. Compounds 1 and 2 were discovered by automated screening of a proprietary small-molecule library using *in vitro* kinase assays, and compound 14 was obtained from subsequent hit expansion. The evaluation of the kinase selectivity of CPA 1 against a selected panel of 61 kinases involved in oncogenic signaling pathways identified Pim kinases as the only other kinase family targeted by this compound.

To gain structural insights into the potency of this new class of ATP competitors, we solved the crystal structure of CK2 in complex with CPA and AMR, as well as that of Pim-1 with CPA. Structural information obtained from these crystallographic structures and from the docking of all compounds in CK2 together with the structure-activity relationship (SAR) analysis allowed us to clearly identify key features that underlie inhibitory potency. This will be useful for further optimization of these leads.

MATERIALS AND METHODS

Human recombinant CK2 α subunit was expressed in *Escherichia coli* and purified to homogeneity as described previously (42). Human recombinant N-terminal GST-tagged Pim-1 was obtained from Millipore (cat. no. 14-573; Bedford, MA, USA).

Cell culture

Prostate cancer cell lines were purchased from American Type Culture Collection (ATCC; Manassas, VA, USA). LNCaP, PC3, and DU145 were maintained in RPMI 1640 supplemented with 10% fetal bovine serum. Western blot analysis was performed using polyclonal anti-CK2 α/α' (43) at a dilution at 1:2000, monoclonal anti-Pim-1 (12-H8, 1:500; Santa Cruz Biotechnology, Santa Cruz, CA, USA), and polyclonal anti-HSP90 (4874, 1:1000; Cell Signaling, Boston, MA, USA).

Automated screening

Automated screening of the Institut Curie/CNRS small-molecule library was performed at 1.5 and 15 μ M in 96-well plates using a TECAN Genesis robot of the CMBA facilities. CK2 activity was assayed for 65 min at room temperature in a final volume of 30 μ l containing 10 μ l of compounds or DMSO or TBB (10 μ M) controls, 10 μ l of CK2 α (50 ng) in 50 mM Tris, 150 mM NaCl, and 10 μ l of a mixture containing 100 μ M peptide substrate (RRRE-DEESDDEE), 1 μ M ATP, and 10 mM MgCl₂. The final DMSO concentration in the assay was 1%. At the end of the reaction, kinase activity was determined by the luminescence-based Kinase-Glo Assay (Promega, Madison, WI, USA) according to the manufacturer's recommendations. This assay lies on the use of ATP depletion as a readout for kinase activity (44).

CK2 phosphorylation assays

As a second screen, selected compounds were tested in a radioactive assay performed in a final volume of 18 μ l containing 1 μ l of compound, 3 μ l of CK2 α (36 ng), and a mixture containing 100 μ M of peptide substrate (RRREDEESDDEE), 10 mM MgCl₂, and 100 μ M [γ -³²P]ATP. Assays were performed at room temperature for 5 min before termination by the addition of 60 μ l of 4% TCA. ³²P-incorporation in the peptide substrate was determined as described previously (45).

Kinase selectivity profiling

Kinase selectivity was performed on a panel of 54 recombinant protein kinases. The assays were performed at 10 μ M ATP in the presence of 1 μ M inhibitor using the Upstate Kinase Profiler Panel Service (Millipore). Inhibition, expressed as the percentage of activity determined in the absence of inhibitor, was calculated from the residual activity measured in the presence of 1 μ M inhibitor.

Thermal stability measurements

Thermal melting experiments were carried out using a real-time PCR Mx3005p machine (Stratagene, La Jolla, CA, USA) according to the protocol described by Vedadi *et al.* (46). Recombinant protein kinases were subjected to gradually increasing temperature. The temperature shift between the melting temperature in the presence or absence of inhibitors was measured.

Pim-1 kinase assay

Pim-1 kinase assays were performed according to the manufacturer's recommendations (Upstate, Lake Placid, NY, USA). Briefly, the assay was performed in a final volume of 25 μ l containing 5 μ l of concentrated reaction buffer (40 mM MOPS, pH 7.0, and 1 mM EDTA), 2.5 μ l of 100 μ M peptide substrate (KKRNRTLTV), 5 μ l of compounds at the indicated concentrations, 2.5 μ l of Pim-1 (12.5 ng), and 10 μ l of a mixture of 100 μ M [γ -³²P]ATP and 10 mM MgCl₂ (final assay concentration). Assays were performed at 30°C for 5 min before termination by the addition of 5 μ l of 3% phosphoric acid. Twenty microliters of assay medium was then transferred on P30 filtermat. Filters were then washed in 75 mM phosphoric acid, and ³²P-incorporation in the peptide substrate was determined by scintillation counting.

Chemical synthesis and characterization of compounds

Starting materials and solvents were purchased from Acros Organics (Geel, Belgium), Sigma-Aldrich (St. Louis, MO, USA), SDS (Vitry sur Seine, France), or Maybridge (Geel, Belgium). Dry THF and Et₂O were prepared by distillation from benzophenone/sodium. Dry CH₂Cl₂ was prepared by distillation on P₂O₅. Reactions were monitored by analytical thin-layer chromatography on neutral Kieselgel G Merck (60F254) plates (Merck, Darmstadt, Germany), and products were visualized by exposure to ultraviolet light. Merck silica gel (230–400 Mesh ASTM) was used for flash chromatography. All reported yields are unoptimized. Compound purity was determined by HPLC analysis (Waters W Delta 600, Photodiode Array Detector 996; Waters Corp., Milford, MA, USA) using one of the following isocratic methods: 1) Waters X-Terra MS C18 reverse-phase column, 5 μ m, 4.6 \times 150 mm; eluent 80% CH₃CN and 20% H₂O; *t*_R (DMSO) = 1.72 min; 2) Merck Nucleosil 100 Å Phenyl column, 5 μ m, 4 \times 250 mm; eluent 80% CH₃CN and 20% H₂O; *t*_R (DMSO) = 4.00 min; 3) Interchrom Spherisorb S5PH-25F column (Interchim, Montluçon, France), 5 μ m, 4.6 \times 250 mm; eluent 80% CH₃CN and 20% H₂O; *t*_R (DMSO) = 3.47 min; or 4) Interchrom Spherisorb S5PH-25F column, 5 μ m, 4.6 \times 250 mm; eluent 80% CH₃CN and 20% sodium phosphate buffer (pH 3, 20 mM); *t*_R (DMSO) = 3.57 min. The samples were dissolved in DMSO and diluted to 80 μ l/ml in elution solvents. The injection volumes were typically 20 μ l, and a constant flow rate of 1 ml/min was maintained. Melting points were measured on a Köfeler hot stage apparatus and are uncorrected. The ¹H (300 MHz) and ¹³C (75 MHz) NMR spectra were recorded on a Bruker AC 300 spectrometer (Bruker Optics, Ettlingen, Germany). Chemical shifts are expressed as parts per million downfield from tetramethylsilane. Splitting patterns have been designated as follows: s, singlet; d, doublet; dd, doublet of doublet; t, triplet; q, quadruplet; m, multiplet; and br, broad signal. Coupling constants (*J* values) are listed in hertz. Peak assignment was performed using COSY, HMQC, and HMBC techniques. Mass spectra were obtained with a Nermag-Ribermag R10-10C spectrometer (Nermag, Rueil-Malmaison, France) applying a desorption chemical ionization (CI) technique using ammonia as the reagent gas or by electrospray (ES/MS) with a Waters ZQ 2000 apparatus.

Diethyl 1,2,9,10-tetrahydronaphtho[2,1-b:7,8-b']difuran-2,9-dicarboxylate (compound 11)

K₂CO₃ (966 mg, 7.2 mmol, 5.6 eq) and ethyl 2,3-dibromopropanoate (0.4 ml, 2.8 mmol, 2.2 eq) were added to a solution of 2,7-dihydroxynaphthalene (203.4 mg, 1.3 mmol) in dry acetone (5 ml), and the mixture was refluxed for 15 h, then cooled to room temperature, and filtered. The filtrate was concentrated, partitioned between EtOAc and water, and extracted. The combined organic phases were dried and concentrated, and the residue obtained was purified by silica flash chromatogra-

phy using the following gradient (EtOAc/cyclohexane; 2:98, 4:98, 6:98, and finally 10:90). Compound 11 was obtained as a white solid (299 mg, 65%): HPLC₍₁₎ *t*_R 2.72 min (98.5%); mp: 155°C; ¹H NMR (CDCl₃) δ 7.59 (d, *J*=8.8 Hz, 2H, H-5, H-6), 6.98 (d, *J*=8.8 Hz, 2H, H-4, H-7), 5.27 (dd, *J*_{2,1a}=*J*_{9,10a}=10.5 Hz, *J*_{2,1b}=*J*_{9,10b}=7.1 Hz, 2H, H-2, H-9), 4.27 (q, *J*=7.1 Hz, 4H, 2×CH₂), 3.85 (dd, *J*_{1a,2}=*J*_{10a,9}=10.5 Hz, *J*_{gem}=13.0 Hz, 2H, H-1a, H-10a), 3.72 (dd, *J*_{1b,2}=*J*_{10b,9}=7.1 Hz, *J*_{gem}=13.0 Hz, 2H, H-1b, H-10b), 1.31 (t, *J*=7.1 Hz, 6H, 2×CH₃). ¹³C NMR (CDCl₃) δ 171.2 (2×COOEt), 157.7 (C-3a, C-7a), 130.5 (C-5, C-6), 128.3 (C-10b), 126.2 (C-5a), 113.9 (C-10a, C-10c), 109.5 (C-4, C-7), 79.5 (C-2, C-9), 61.8 (2×CH₂), 34.1 (C-1, C-10), 14.2 (2×CH₃). MS (ES⁺) *m/z* 380 [M+Na]⁺; MS (ES⁻) *m/z* 356 [M-H]⁻.

2,9-Dicarboxy-1,2,9,10-tetrahydronaphtho[2,1-b:7,8-b']difuran (compound 12)

An aqueous solution of sodium hydroxide (0.5 M; 0.4 ml, 0.20 mmol, 2 eq) was added to a solution of compound 11 (34.3 mg, 0.10 mmol) in THF-MeOH-H₂O (2-2-1; 5 ml), and the mixture was stirred for 1 h at room temperature. It was then concentrated, and the residue was partitioned between Et₂O and water. The aqueous phase was acidified to pH ~1 with 1 M hydrochloric acid. Acidification of the aqueous phase did not lead to precipitation of the product, so the acid aqueous layer was extracted with CH₂Cl₂. The resulting organic phase was dried over Mg₂SO₄ and concentrated, and the resultant residue was purified by flash chromatography (CH₂Cl₂/MeOH 1:99, 3:97 then 10:90). The fractions containing the expected product were precipitated from Et₂O, which yielded compound 12 as white crystals (8.3 mg, 28%): HPLC₍₃₎ *t*_R 3.09 min (100.0%); mp: 181°C; ¹H NMR (CDCl₃) δ 7.64 (d, *J*=8.7 Hz, 2H, H-5, H-6), 7.02 (d, *J*=8.7 Hz, 2H, H-4, H-7), 5.36 (dd, *J*_{9,10a}=*J*_{2,1a}=10.6 Hz, *J*_{9,10b}=*J*_{2,1b}=6.9 Hz, 2H, H-2, H-9), 3.96 (dd, *J*_{10a,9}=*J*_{1a,2}=10.6 Hz, *J*_{gem}=15.6 Hz, 2H, H-1a, H-10a), 3.84 (dd, *J*_{10b,9}=*J*_{1b,2}=6.9 Hz, *J*_{gem}=15.6 Hz, 2H, H-1b, H-10b). ¹³C-NMR (CDCl₃) δ 171.7 (2×COOH), 157.8 (C-3a, C-7a), 130.6 (C-5, C-6), 128.3 (C-10b), 125.9 (C-5a), 113.9 (C-10a, C-10c), 109.7 (C-4, C-7), 79.4 (C-2, C-9), 34.3 (C-1, C-10). MS (ES⁺) *m/z* 341 [M+K+H]⁺, 381 [M+2K+H]⁺.

Diethyl naphtho[2,1-b:7,8-b']difuran-2,9-dicarboxylate (compound 13)

2,3-Dicyano-5,6-dichloro-parabenzoquinone (DDQ; 9.053 g, 26 mmol, 3.1 eq) was added to a solution of compound 11 (4.520 g, 12.7 mmol) in toluene (125 ml), and the mixture was refluxed for 5 h and then neutralized by addition of saturated aqueous NaHCO₃ and extracted with EtOAc using large volumes of solvent. The combined organic phases were washed with water, dried over Mg₂SO₄, and concentrated to give a solid residue that was separated by silica flash chromatography (CH₂Cl₂/cyclohexane 1:1 then 7:3). Compound 13 was obtained as a white solid (3.069 g, 69%): HPLC₍₁₎ *t*_R 3.86 min (100.0%); mp: 217°C; ¹H NMR (CDCl₃) δ 8.21 (s, 2H, H-1, H-10), 8.01 (d, *J*=9.0 Hz, 2H, H-5, H-6), 7.77 (d, *J*=9.0 Hz, 2H, H-4, H-7), 4.53 (q, *J*=7.1 Hz, 4H, 2×CH₂), 1.51 (t, *J*=7.1 Hz, 6H, 2×CH₃). ¹³C NMR (CDCl₃) δ 199.7 (2×COOEt), 159.4 (C-2, C-9), 154.9 (C-3a, C-7a), 145.7 (C-10a, C-10c), 129.9 (C-5, C-6), 127.5 (C-5a), 121.3 (C-10b), 113.5 (C-1, C-10), 111.8 (C-4, C-7), 61.7 (2×CH₂), 14.4 (2×CH₃). MS (ES⁺) *m/z* 375 [M+Na]⁺.

2,9-Dicarboxynaphtho[2,1-b:7,8-b']difuran (compound 14; AMR)

Compound 14 was obtained from compound 13 (507.0 mg, 1.4 mmol) as described for compound 12. The reaction

mixture was sonicated for 24 h. After work-up and acidification of the aqueous phase, the expected product compound 14 was collected as a white precipitate (423.7 mg, 100%). Compound 14 was recrystallized from acetic acid: HPLC₍₂₎ *t*_R 1.64 min (100.0%); mp > 260°C; ¹H NMR (DMSO-*d*₆) δ 8.69 (s, 2H, H-1, H-10), 8.17 (d, *J*=9 Hz, 2H, H-5, H-6), 7.91 (d, *J*=9 Hz, 2H, H-4, H-7). ¹³C NMR (DMSO-*d*₆) δ 160.5 (2×COOH), 154.6 (C-3a, C-7a), 146.7 (C-2, C-9), 130.3 (C-5, C-6), 127.3 (C-5a), 122.8 (C-10b), 121.3 (C-10a, C-10c), 114.4 (C-1, C-10), 112.1 (C-4, C-7). MS (ES⁻) *m/z* 295 [M-H]⁻.

Diethyl 1,2-dihydronaphtho[2,1-b:7,8-b']difuran-2,9-dicarboxylate (compound 15)

Compound 15 was obtained from compound 11 (48.9 mg, 0.14 mmol) as described for compound 13 but using only 1.05 equivalents of DDQ (32.7 mg, 0.15 mmol). Purification by flash chromatography (acetone/toluene 1:99) gave compound 15 as a white solid (47 mg, 95%): HPLC₍₁₎ *t*_R 3.34 min (96.9%); mp: 113°C; ¹H NMR (CDCl₃) δ 7.79 (s, 1H, H-10), 7.77 (d, *J*=8.9 Hz, 2H, H-5, H-6), 7.49 (d, *J*=8.9 Hz, 1H, H-7), 7.19 (d, *J*=8.9 Hz, 1H, H-4), 5.44 (dd, *J*_{2,1a}=10.9 Hz, *J*_{2,1b}=7.0 Hz, 1H, H-2), 4.46 (q, *J*=7.1 Hz, 2H, CH₂CH₃-9), 4.28 (q, *J*=7.1 Hz, 2H, CH₂CH₃-2), 3.99 (dd, *J*_{1a,2}=10.9 Hz, *J*_{gem}=15.6 Hz, 1H, H-1a), 3.83 (dd, *J*_{1b,2}=7.0 Hz, *J*_{gem}=15.6 Hz, 1H, H-1b), 1.46 (t, *J*=7.1 Hz, 3H, CH₂CH₃-9), 1.35 (t, *J*=7.1 Hz, 3H, CH₂CH₃-2). ¹³C NMR (CDCl₃) δ 171.0 (COOEt-2), 159.4 (COOEt-9), 158.4 (C-3a), 154.9 (C-7a), 144.7 (C-9), 130.0 (C-5, C-6), 126.4 (C-5a), 125.4 (C-10b), 120.2 (C-10a), 116.4 (C-10c), 114.1 (C-10), 110.5 (C-4), 110.1 (C-7a), 79.8 (C-2), 61.9 (CH₂CH₃-2), 61.5 (CH₂CH₃-9), 34.4 (C-1), 14.4 (CH₂CH₃-9), 14.2 (CH₂CH₃-2). MS (ES⁺) *m/z* 355 [M+H]⁺.

2,9-Dicarboxy-1,2-dihydronaphtho[2,1-b:7,8-b']difuran (compound 16)

Compound 16 was obtained from compound 15 (48.2 mg, 0.14 mmol) and purified as described for compound 12, which yielded compound 16 as a pale yellow solid (27.7 mg, 66%): HPLC₍₄₎ *t*_R 2.39 min (100.0%); mp > 260°C; ¹H NMR (DMSO-*d*₆) δ 7.97 (d, 2H, H-5, H-6), 7.92 (s, *J*=8.7 Hz, 1H, H-10), 7.65 (d, *J*=8.7 Hz, 1H, H-7), 7.23 (d, *J*=8.7 Hz, 1H, H-4), 5.49 (dd, *J*_{2,1a}=10.5 Hz, *J*_{2,1b}=6.1 Hz, 1H, H-2), 4.11 (dd, *J*_{1a,2}=10.5 Hz, *J*_{gem}=16.3 Hz, 1H, H-1a), 3.78 (dd, *J*_{1b,2}=6.1 Hz, *J*_{gem}=16.3 Hz, 1H, H-1b). ¹³C NMR (DMSO-*d*₆) δ 172.7 (COOH-2), 160.4 (COOH-9), 158.8 (C-3a), 154.6 (C-7a), 145.9 (C-9), 130.9 (C-5), 130.4 (C-6), 126.3 (C-5a), 125.5 (C-10b), 120.3 (C-10a), 117.5 (C-10c), 114.3 (C-10), 110.6 (C-4), 110.4 (C-7), 80.0 (C-2), 34.1 (C-1). MS (ES⁺) *m/z* 321 [M+Na]⁺; MS (ES⁻) *m/z* 297 [M-H]⁻.

Diethyl 1,2,6,7-tetrahydronaphtho[2,1-b:6,5-b']difuran-2,7-dicarboxylate (compound 17)

Compound 17 was obtained from 2,6-dihydroxynaphthalene (205.2 mg, 1.3 mmol) as described for compound 11, as a white solid (201.7 mg, 44%): HPLC₍₁₎ *t*_R 2.67 min (100.0%); mp: 179°C; ¹H NMR (CDCl₃) δ 7.45 (d, *J*=8.8 Hz, 2H, H-5, H-10), 7.22 (d, *J*=8.8 Hz, 2H, H-4, H-9), 5.35 (dd, *J*_{2,1a}=*J*_{7,6a}=10.8 Hz, *J*_{2,1a}=*J*_{7,6a}=7 Hz, 2H, H-2, H-7), 4.29 (q, *J*=7 Hz, 4H, 2×CH₂), 3.82 (dd, *J*_{1a,2}=*J*_{6a,7}=10.8 Hz, *J*_{gem}=15.8 Hz, 2H, H-1a, H-6a), 3.64 (dd, *J*_{1b,2}=*J*_{6b,7}=7 Hz, *J*_{gem}=15.8 Hz, 2H, H-1b, H-6b), 1.33 (t, *J*=7 Hz, 6H, 2×CH₃). ¹³C NMR (CDCl₃) δ 171.3 (2×COOEt), 155.0 (C-3a, C-8a), 126.5 (C-5b, C-10b), 123.4 (C-5, C-10), 118.0 (C-5a, C-10a0), 113.0 (C-4, C-9), 79.4 (C-2, C-7), 61.7 (2×CH₂), 33.4 (C-1, C-6), 14.2 (2×CH₃). MS (ES⁺) *m/z* 380 [M+Na]⁺; MS (ES⁻) *m/z* 356 [M-H]⁻.

2,7-Dicarboxy-1,2,6,7-tetrahydronaphtho[2,1-b:6,5-b']difuran
(compound 18)

Compound 18 was obtained from compound 17 (31.3 mg, 0.09 mmol) as described for compound 12 (white solid, 1.3 mg, 5%): HPLC₍₄₎ t_R 2.50 min (100.0%); mp > 260°C; ¹H NMR (DMSO-*d*₆) δ 7.50 (d, J = 8.8 Hz, 2H, H-5, H-10), 7.18 (d, J = 8.8 Hz, 2H, H-4, H-9), 5.34 (dd, $J_{2,1a}$ = $J_{7,6a}$ = 10.7 Hz, $J_{2,1b}$ = $J_{7,6b}$ = 6.3 Hz, 2H, H-2, H-7), 3.78 (dd, $J_{1a,2}$ = $J_{6a,7}$ = 10.7 Hz, J_{gem} = 16 Hz, 2H, H-1a, H-6a), 3.48 (dd, $J_{1b,2}$ = $J_{6b,7}$ = 6.3 Hz, J_{gem} = 16 Hz, 2H, H-1b, H-6b). ¹³C NMR (DMSO-*d*₆) δ 173.2 (2×COOH), 155.0 (C-3a, C-8a), 126.4 (C-5b, C-10b), 124.4 (C-5, C-10), 118.7 (C-5a, C-10a), 113.1 (C-4, C-9), 79.3 (C-2, C-7), 33.2 (C-1, C-6). MS (ES⁺) m/z 301 [M+H]⁺, 323 [M+Na]⁺; MS (ES⁻) m/z 299 [M-H]⁻.

Diethyl naphtho[2,1-b:6,5-b']difuran-2,7-dicarboxylate
(compound 19)

Compound 19 was obtained from compound 17 (55.2 mg, 0.16 mmol) as described for compound 15 (white solid, 56.7 mg, 100%): HPLC₍₁₎ t_R 4.50 min (100.0%); mp: 240°C; ¹H NMR (CDCl₃) δ 8.18 (d, J = 9.0 Hz, 2H, H-5, H-10), 8.02 (s, 2H, H-1, H-6), 7.86 (d, J = 9.0 Hz, 2H, H-4, H-9), 4.49 (q, J = 7.1 Hz, 4H, 2×CH₂), 1.47 (t, J = 7.1 Hz, 6H, 2×CH₃). ¹³C NMR (CDCl₃) δ 159.4 (2×COOEt), 153.5 (C-3a, C-8a), 145.6 (C-2, C-7), 124.8 (C-5b, C-10b), 124.0 (C-5a, C-10a), 123.9 (C-5, C-10), 113.5 (C-4, C-9), 112.5 (C-1, C-6), 61.6 (2×CH₂), 14.4 (2×CH₃). MS (ES⁺) m/z 353 [M+H]⁺, 375 [M+Na]⁺.

2,7-Dicarboxynaphtho[2,1-b:6,5-b']difuran (compound 20)

Compound 20 was obtained from compound 19 (65.8 mg, 0.19 mmol) as described for compound 12 (white solid, 3.8 mg, 7%): HPLC₍₄₎ t_R 2.44 min (98.3%); mp > 260°C; ¹H NMR (DMSO-*d*₆) δ 8.55 (d, J = 9.0 Hz, 2H, H-5, H-10), 8.43 (s, 2H, H-1, H-6), 8.02 (d, J = 9.0 Hz, 2H, H-4, H-9). ¹³C NMR (DMSO-*d*₆) δ 160.4 (2×COOH), 153.2 (C-3a, C-8a), 146.4 (C-2, C-7), 124.9 (C-5, C-10, C-5b, C-10b), 124.2 (C-5a, C-10a), 113.9 (C-4, C-9), 113.5 (C-1, C-6). MS (ES⁻) m/z 295 [M-H]⁻.

CK2 protein purification and crystallization

The catalytic α -subunit of human CK2 was expressed in *E. coli* and purified as described previously (47). Crystals of rhCK2 α^{AC} complex were first grown by using hanging drop vapor-diffusion technique at 293 K by mixing 2 μ l of rhCK2 α^{AC} -magnesium-substrate peptide mix (4.6 mg/ml rhCK2 α^{AC} , 114 μ M peptide substrate RRREDEESDDEE, 680 μ M MgCl₂, and 340 μ M AMPNP) with 2 μ l of precipitant solution (36% polyethylene glycol 5000 monomethyl ether, 150 mM ammonium sulfate, and 100 mM Tris-HCl pH 7.5). Crystals of the CPA-rhCK2 α^{AC} and AMR-rhCK2 α^{AC} complexes were then obtained by soaking rhCK2 α^{AC} native crystals for 1 wk in a 4 μ l drop of same composition as crystallization drop except that AMPNP was replaced by 300 μ M of CPA or AMR.

CK2 complex structure determination and refinement

For X-ray diffraction data collection, inhibitor-rhCK2 α^{AC} crystals were directly frozen in liquid nitrogen without adding any extra cryoprotectant. Data were collected at 100 K on beamline ID23-eh1 (CPA-rhCK2 α^{AC}) or ID14-eh2 (AMR-rhCK2 α^{AC}) of the European Synchrotron Radiation Facility at a wavelength of 1.0722Å (CPA-rhCK2 α^{AC}) or 0.933Å (AMR-rhCK2 α^{AC}). Data processing was performed using an XDS package (48) and is summarized in Supplemental Table 4.

The inhibitor-rhCK2 α^{AC} complex structures were determined by molecular replacement method with AMoRe (49) from the CCP4 suite (50) using the rhCK2 α^{AC} structure as a search model [Protein Data Bank (PDB) code: 1PJJK; ref. 47]. The final models were obtained after several rounds of alternate manual building with COOT (51) and maximum likelihood refinement with Refmac5 of CCP4 (52) and using data in the 20- to 1.75-Å resolution range for CPA-rhCK2 α^{AC} or 36–1.65 Å for AMR-rhCK2 α^{AC} . The final refinement statistics are summarized in Supplemental Table 4. Pymol (DeLano Scientific, San Francisco, CA, USA) was used to generate structure figures. The structure factors and coordinates are deposited in the PDB under the accession number 3MB6 for CPA-rhCK2 α^{AC} and 3MB7 for AMR-rhCK2 α^{AC} .

Pim-1 purification, crystallization, and structure determination

Pim-1 has been purified as described previously (53). Crystallization was carried out using the sitting drop vapor diffusion method at 4°C. The complex with CPA was grown by mixing 150 nl of the protein (10 mg/ml) with an equal volume of reservoir solution containing 20% isopropanol and 0.1 M Tris buffer pH 8.5. Crystals grew to diffracting quality within a few days. Crystals were cryoprotected using the well solution supplemented with 25% ethylene glycol and were flash frozen in liquid nitrogen. Data were collected at the Swiss Light Source (SLS) beam-line SA10. Indexing and integration were carried out using MOSFLM (54), and scaling was performed with SCALA (55). Initial phases were calculated by molecular replacement with PHASER (56) using 2C3I as a starting model. The model was completed manually in COOT (51) and was refined with REFMAC5 (52). Thermal motions were analyzed using TLSMD (57), and hydrogen atoms were included in late refinement cycles. Data collection and refinement statistics can be found in Supplemental Table 4. The model and structure factors have been deposited with PDB accession code: 3MA3.

Docking

Docking of CPA, CPB, and AMR was performed on both CK2 and Pim-1 using Gold (58), FlexX (59), and DOCK (60, 61) with standard parameters. The resulting structures were further refined by energy minimizations using CHARMM (62). CPA, CPB, and AMR were docked in the ATP-binding site of the catalytic subunit CK2 α of human CK2. Starting with CPA, standard docking techniques were tried using the programs Gold and DOCK. However, since the difference between the scores of the resulting poses of the ligand in the active site was not significant enough to allow the selection of a preferred position, further refinement was needed. Thus, for each of the poses obtained, the energy of the complex CK2-CPA was minimized with the program CHARMM. After rescoring the resulting protein-ligand complexes with the scoring function Goldscore and using other criteria as accessible surface, hydrophobic contacts, and interaction energy, 1 position could be retained as the most likely one. Extending this technique for CPB and AMR in CK2 using Gold and CHARMM, 1 good pose was found for each compound. For Pim-1, FlexX was used for CPA and CPB and Gold for AMR. We describe hereafter each of the steps of this workflow: protein and ligand preparation, standard docking, energy minimization of the protein-ligand complexes, and finally analysis of the resulting structures to select the most likely position of the ligand in the active site.

Protein setup

The structure of human CK2 cocrystallized with AMPNP was taken from the PDB (1JWH). Only 1 CK2 α subunit was considered with all the crystal waters in its vicinity. For the Pim-1 catalytic subunit (PDB code IYXT), all the crystal

waters were considered, except molecule 39, which interacts with the natural ligand (AMPPNP). AMPPNP was removed from both proteins. All Asp, Glu, Lys, and Arg side chains were considered ionized, and His was neutral.

For docking with Gold or DOCK and minimization with CHARMM, hydrogen atoms were added to the protein structure and water molecules using the Hbuild procedure in CHARMM. For docking with FlexX, the protein was prepared using the Protein Preparation Tool of the Biopolymer module of Sybyl 7.1 (63). For docking with DOCK, all crystal water molecules were removed from the protein.

Ligand setup

CPA, CPB, and AMR were converted to their 3-D optimized structure with Gaussian 03W (64), using a Hartee-Fock method with the basis set 6-31g(d). Electrostatic potential charges and normal modes of the optimized geometry were calculated with the same basis. These features were used to determine a large part of the ligand parameters that should be included in CHARMM (charges, equilibrium distances, and angles). To determine some force constants, we compared our ligands to amino acids, whose parameters are already known. Special care was taken when defining the junction of the cycles for which no parameters were available for comparison. In this case, the missing parameters were adjusted to better reproduce with CHARMM the normal modes that were calculated by Gaussian. The parameters for the carboxylates were deduced from the rotational energy barriers obtained with Gaussian. The sets of parameters thus defined were subsequently tested with CHARMM by minimizing the ligands energy and carrying out molecular dynamics in order to prevent the ligands from being distorted.

Docking protocol for CK2

Docking of all the compounds of the difuran series in CK2 was carried out with Gold version 3.2. The binding pocket was defined using the following residues: Arg43, Leu45, Gly46, Arg47, Ser-51, Glu52, Val53, Val66, Val67, Lys68, Glu81, Ile84, Leu85, Leu88, Arg89, Ile95, Thr96, Leu97, Leu111, Val112, Phe113, Glu114, His115, Val116, Asn117, Asn118, Asp156, Lys158, His160, Asn161, Met163, Ile174, Asp175, and Trp176. The crystallographic water molecules situated in the active site were enabled to spin. The docking procedure was repeated 20 times for each compound, using the default parameters and the scoring function Goldscore. No early termination was allowed.

Compound CPA was also docked in CK2 with version 6.1 of DOCK. All the files needed to describe the binding site were prepared using the programs present in the DOCK distribution (DMS, SPHGEN, SHOWBOX, and GRID). Among the spheres used by DOCK to describe the active site of the protein and to place the ligand, the ones situated in the vicinity of Lys68 and Asp175 were designated as critical. A first docking run was performed to orient the ligand using these defined critical points and a bump filter. All the clustered conformations obtained were prepared to be used with AMBER (65) force field with the script provided and scored with AMBER score after minimization with the whole system considered as movable.

Docking protocol for Pim-1

CPA and CPB were docked in Pim-1 with FlexX as implemented in Sybyl. Default values were kept for all settings. Serial CScore calculations with scoring functions ChemScore, D_Score, G_Score, and PMF_Score were requested. AMR was docked with Gold as described for CK2. The

active site was defined by the following list of residues: Phe49, Leu44, Val52, Ala65, Leu120, Pro123, Ile104, Val126, Leu174, Ile185, Phe187, Asp186, Glu89, Glu121, Lys67, Arg122, Leu93, Leu118, Asp128, Leu184, Ser-46, Gly50, Gly48, Asn172, Gly45, and Glu171.

Energy minimizations

The complexes formed by the crystal structure of the protein and the active ligands in the orientations found by DOCK (CPA in CK2), Gold (CPA, CPB and AMR in CK2 and AMR in Pim-1), and FlexX (CPA and CPB in Pim-1) were subject to energy minimizations by the conjugate gradient method using the all-H CHARMM22 force field. Crystal waters were all taken into account. The electrostatic and van der Waals energies were truncated using the switching function between 11 and 15 Å. In the first 600 steps, harmonic constraints were applied to heavy atoms to allow smooth minimization without abrupt deviation from the initial structure. The force constant was decreased every 100 steps: it took the values 250, 100, 50, 25, 10, and 5 kcal/mol/Å² and then followed a free minimization of 1000 steps.

Analysis of the protein-ligand complexes obtained after docking and energy minimization

All protein-ligand complexes obtained after minimization with CHARMM were rescored using the scoring function Goldscore of Gold. Since some parameters representative of the quality of the binding of the ligand to the protein are not taken into account in Goldscore, some characteristic data of the protein-ligand complex were calculated with CHARMM: 1) the solvent-accessible surface of the hydrophobic core of the ligand; 2) the number of hydrophobic contacts between the ligand and the protein, *i.e.*, the number of contacts between the atoms of the polycyclic core of the ligand on the one hand and the carbon atoms of the protein on the other hand; and 3) the protein-ligand interaction energy, taking into account all van der Waals and electrostatic terms.

It is known that the best positions for the ligand are those in which it is well buried in the hydrophobic pocket of the protein. This situation corresponds to a pose with a low solvent-accessible surface and a high number of hydrophobic contacts (data 1 and 2 in the paragraph above). In addition, the interaction energy (data 3 from above) must be as favorable as possible. Analysis of these data combined with the Goldscore values after energy minimization allowed us to select reasonable binding poses for the ligands and to retain in the end 1 binding mode for each compound in each protein (see Supplemental Fig. 4 for CPB in CK2 and AMR and CPB in Pim-1).

Docking: methods used to dock compounds 11 to 20 in CK2

Docking of compounds 11 to 20 in CK2 was performed with Gold using the default parameters. Since for these compounds the scoring with Goldscore discriminated well between the different poses, no further refinement (energy minimization) was performed.

RESULTS

High-throughput screening and biochemical characterization of the hits

A screen of 6,560 compounds from the Institut-Curie/CNRS small-molecule library against the rhCK2 α catalytic

subunit was performed using an automated luminescence-based *in vitro* kinase assay at a 15 μM concentration. This led to the identification of 158 compounds belonging to several distinct structural families that showed inhibition >75%. Further screening of these initial hits at a concentration of 1.5 μM reduced to 19 the number of compounds that still inhibited CK2 activity with an efficacy of > 50% (Supplemental Chart 1). The activity of the identified hits was confirmed using a secondary radiometric CK2 kinase assay. The 2 most active compounds, inhibiting the enzyme >80%, were CPA (66) and CPB (67). Moreover, difurandicarboxylic acid derivatives that inhibit the CK2 α catalytic subunit were also fully active on the CK2 holoenzyme (Supplemental Table 1). Given that these active compounds are hydrophobic, planar, and rigid with peripheral polar groups, it is possible that they could behave as “promiscuous” inhibitors (68), acting *via* an aggregation mechanism (69). However, in all experiments described in this work, BSA was added to stabilize the kinases. Since BSA is known to attenuate aggregation-based inhibition, our assays were less prone to identify promiscuous inhibitors (70). However, it was decided to further investigate this point. Since aggregate-forming inhibitors often display steep dose-response curves and high Hill coefficients (70, 71), we performed dose-response curves for CK2 and Pim-1 inhibition on our screening hits. It was observed that the inhibitors displayed standard dose-response curves with Hill slopes ranging between 0.5 to 1.0, indicating that the compounds likely bind their target enzymes through a classical 1:1 inhibition mechanism (Supplemental Fig. 1A, B). Another feature of aggregation-prone inhibitory mechanisms is sensitivity of enzyme inhibition to detergents (72). By disrupting enzyme/aggregate interactions, detergents allow us to distinguish between promiscuous and classical inhibitors. The effect of such inhibitors that act based on their aggregation is dramatically reduced by the addition of detergents at concentrations below the critical micellar concentrations. Therefore, we explored the inhibition profile of CPA on CK2 in the absence or presence of detergents (Triton X-100 or Tween-20; Supplemental Fig. 1C). It was observed that the presence of these detergents did not induce any change in the IC_{50} for CK2, ruling out the possibility that CPA could act as a promis-

cuous aggregator. To further investigate the mechanism of CK2 inhibition by CPA, the kinase activity of a CK2-CPA mix was assayed either undiluted or diluted 10-fold to displace the equilibrium from CPA-CK2 complex to unbound CK2 (Supplemental Fig. 1D). On dilution, a significant increase in activity was observed, indicating that CPA inhibition was reversible in our assay conditions.

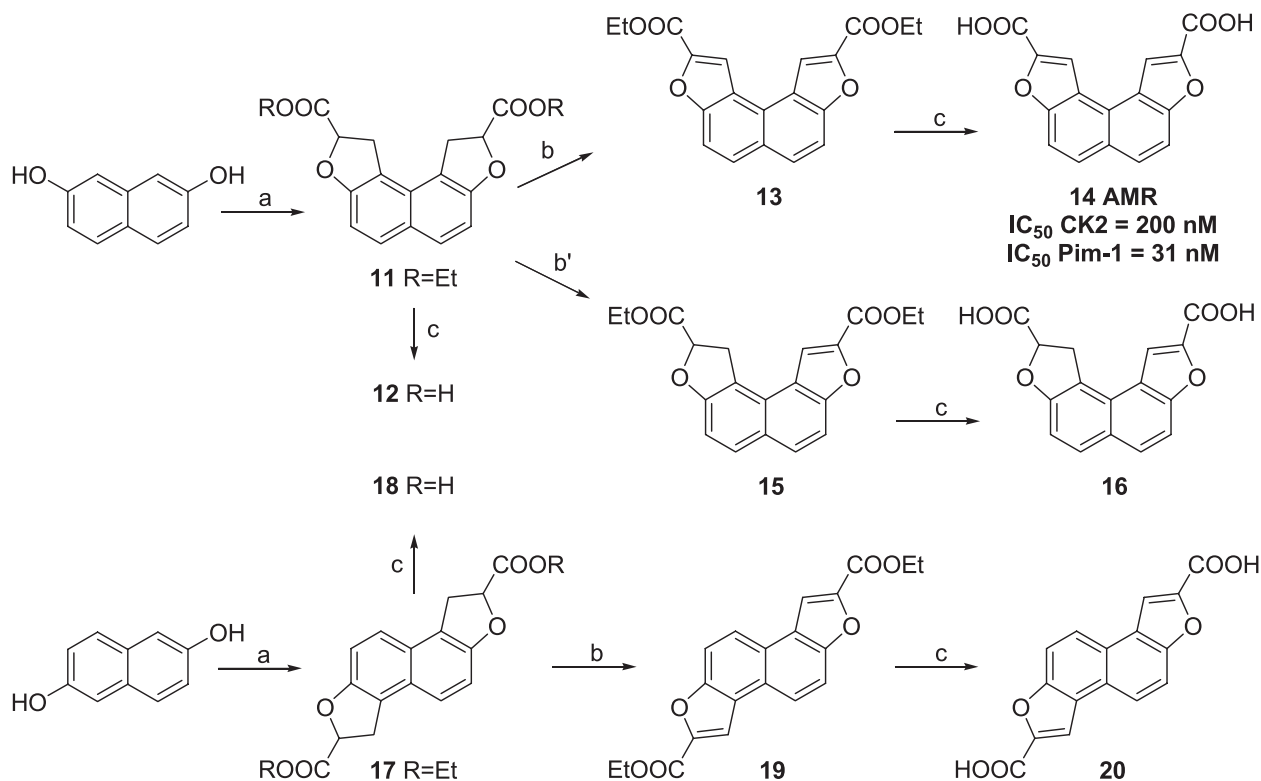
Both CPA and CPB were identified as ATP-competitive inhibitors (Supplemental Fig. 2A and data not shown) with an IC_{50} (and K_i) of 100 nM ($K_i=47$ nM) and 63 nM ($K_i=13$ nM), respectively (Table 1). Interestingly, a substructure search indicated that 72 other compounds also containing a benzofuran moiety substituted by a carboxylic acid function were present in the screened library. However, most of these compounds, and in particular compounds 3–5 and 7–10, which are closely related to CPA and CPB, displayed no significant activity. To explore the influence of the carboxylic-acid-functionalized furan rings on inhibitory activity, a series of CPA analogues were prepared in which the relative positioning of these moieties, as well as the degree of unsaturation of the furan rings were modified (Scheme 1). CPA 1 was chosen for this study, as it is more water soluble and therefore more suitable for optimization than CPB. Partly saturated derivatives of this molecule were prepared in order to move away from the geometrically flat shapes of compounds 1 and 2.

Chemistry

Adapting a route developed by Arrault *et al.* (66) for the construction of dihydro and benzofuran derivatives from a monophenol, 2,7-dihydroxynaphthalene was reacted with ethyl dibromopropanoate to give the dihydrofuran intermediate compound 11 as the unique product in 65% yield (Scheme 1). Treatment of this compound with excess DDQ gave the bisfuran compound 13, whereas its reaction with a single equivalent of DDQ gave the mixed furan-dihydrofuran product compound 15. The ester functions in compounds 11, 13, and 15 were then saponified to give the desired carboxylic acid compounds 12, 14, and 16. In a similar way, 2,6-dihydroxynaphthalene was converted *via* the

TABLE 1. CK2 and Pim-1 percentage of residual activity in the presence of 2,8-difurandicarboxylic acid derivatives

Compound	R1	R2	CK2 activity at 1.5 μM (%)	Pim-1 activity at 1.0 μM (%)
CPA derivatives				
1 (CPA)	COOH	COOH	19 ($\text{IC}_{50}=100$ nM)	9 ($\text{IC}_{50}=60$ nM)
3	H	H	89	94
4	H	COOH	73	92
5	H	COOEt	100	89
6	COOH	COOEt	59	7
7	COOEt	COOEt	100	84
CPB derivatives				
2 (CPB)	COOH	COOH	2 ($\text{IC}_{50}=63$ nM)	4 ($\text{IC}_{50}=20$ nM)
8	COOEt	COOEt	100	86
9	COOEt	COOH	33 ($\text{IC}_{50}=500$ nM)	81
10	H	H	98	71



Scheme 1. Synthesis of compounds 11 to 20 (all these compounds are inactive except AMR, compound 14). Reagents and conditions: a) $\text{BrCH}_2\text{CHBrCOOEt}$, K_2CO_3 , acetone, reflux, 15 h, 65% (compound 11) and 44% (compound 17); b) DDQ, toluene, reflux, 5 h, 69% (compound 13) and 100% (compound 19); b') default DDQ, toluene, reflux, 5 h, 95% (compound 15); c) NaOH , MeOH , THF , room temperature, 1 h, 28% (compound 12), 100% (compound 14), 66% (compound 16), 5% (compound 18) and 7% (compound 20).

intermediate compound 17 to the bisfuran dicarboxylic ester derivative compound 19. Hydrolysis of the ester functions in compounds 17 and 19 gave the corresponding free acid compounds 18 and 20, which have the same arrangement of the furan rings as their high ring homologue CPB.

In vitro evaluation for CK2 inhibition activity of the 10 new compounds revealed that none was active except compound 14 (AMR; IC_{50} =200 nM) with the COOH -substituted furan motifs on adjacent positions on each ring of the bicyclic naphthalene core. Contrary to expectation and interestingly, compound 20 was inactive.

Selectivity

Selectivity is a major issue in the development of ATP-competitive kinase inhibitors for a given therapeutic application. It has been reported that several commercially available CK2 inhibitors like TBB, 2-dimethylamino-4,5,6,7-tetrabromo-1H-benzimidazole (DMAT), and TBI tend to inhibit other kinases such as Pim1, HIPK2, and DYRK1a (73). In this context, CPA 1 was selected for evaluation against a panel of 112 protein kinases using 2 independent selectivity profiling methods. This selectivity profiling is representative of nearly all the branches of the kinome. The results presented in Fig. 1 (74) and Supplemental Table 2 show that the only other protein kinases strongly inhibited by CPA are

the serine/threonine kinases Pim-1 (IC_{50} =60 nM and K_i =31 nM), Pim-2, and Pim-3 and to a lower extent DRAK2 and PRAK. Since Pim kinases are highly homologous and are likely to be involved in the same biological pathways, CPA can be considered to possess a very narrow selectivity profile. Notably, kinases that tend to be inhibited by other CK2 inhibitors (HIPK2, DYRK1a, and PKD) or close evolutionary homologs in the GMGC and CAMK subfamilies are insensitive to difurandicarboxylic acid derivatives. Except for compound 4, all the hits of that series showed similar activities on Pim-1 and CK2. Both CPB 2 and AMR 14 were subsequently found to be also potent ATP-competitive inhibitors of Pim-1 (Table 1, Scheme 1, and Supplemental Fig. 2B), with IC_{50} values of 20 and 31 nM, respectively, which ranks them among the most potent Pim-1 and CK2 inhibitors described to date. Indeed, all 3 compounds proved to be slightly more potent inhibitors of Pim-1 relative to CK2.

Both CK2 and Pim1 kinases are overexpressed in prostate cancer cell lines

We examined the expression of both kinases in several well-established prostate cancer cell lines, including LNCaP, PC3, and DU145 cells, by Western blot and kinase assays. The data illustrated in Supplemental Fig.

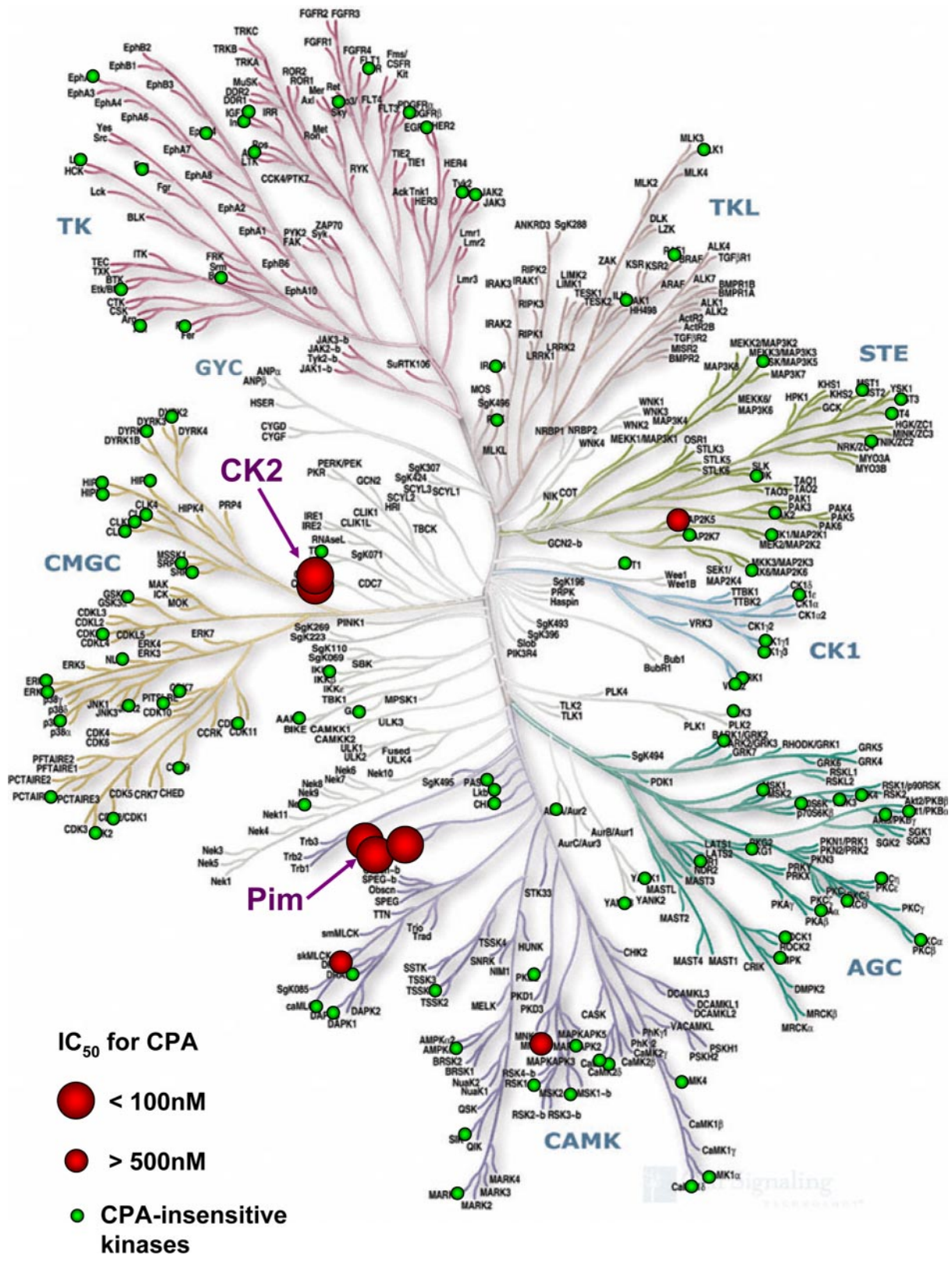


Figure 1. Kinase selectivity profile of CPA on a panel of 112 different protein kinases. Each kinase represented in the assay panel is marked with green or red circles. Circle size is proportional to inhibitory potency (IC_{50}). The kinome illustration was adapted from Manning *et al.* (74) and is reproduced courtesy of Cell Signaling Technology (<http://www.cell-signal.com>).

3 show that both the expression level of Pim1 and CK2 α' and the CK2 activity are correlated with the known tumorigenicity of these prostate cancer cells.

The observed concomitant dysregulation of CK2 and Pim1 in prostate cancer cell lines suggests a potential role of these kinases in prostate cancer progression.

Cross-inhibition of CK2 and Pim kinases

To check whether the dual potency of difurans for CK2 and Pim-1 is specific to these new compounds, different known CK2 and Pim-1 inhibitors (38–41, 75–77), for which the inhibitor-kinase complex crystal structures have been solved, were assayed *in vitro* using both kinases. The results (Supplemental Table 3) highlight that all tested CK2 inhibitors are potent Pim-1 inhibitors, whereas known Pim-1 inhibitors did not inhibit CK2. Unfortunately, when tested in cellular assays for both CK2 and Pim-1 kinase activities, CPA, CPB, and AMR displayed weak or no activity (data not shown). However, we observed that endogenous CK2 present in a cell extract was as sensitive as recombinant CK2 to CPA inhibition (Supplemental Fig. 4). These data show that CPA is biologically active, being able to efficiently hit its CK2 target in complex protein mixtures.

Binding mode of CPA and AMR in CK2 α

To gain insights into the structural basis of the mode of inhibition of these dicarboxydifuran hits, we determined the crystal structure of a truncated form of rhCK2 α (78) in complex with CPA and AMR by X-ray crystallography at 1.75

Å (CPA-CK2) and 1.65 Å (AMR-CK2) resolution (Fig. 2). The determined CK2 α structures superimpose closely with structures that have been reported previously for the human and maize enzyme (79, 80): the root mean square deviation of C α backbone ranged from 0.41Å (compared with 2PVR; ref. 81) to 1.22 Å (compared with 3BW5; ref. 82) when the entire CK2 α molecules were superimposed. Surprisingly, the region comprising residues 118 to 123 (within the E0 linker; see Supplemental Fig. 5 for important structural features) was found in 2 different conformations in the complex structures: although partially disordered in CPA-CK2, the linker was built using a conformation similar to the one found in the CK2 complex structures 2RKP, 1OM1, 2ZOH, 1M2R, and 1F0Q. In contrast, in the AMR-CK2 complex, this linker was found in a more canonical conformation as in 1NA7 or 1PJK. (see Supplemental Fig. 5 for the important structural elements of CK2 α catalytic subunit).

Electron densities for the ligands CPA and AMR were well defined and confirm binding within the ATP-binding site. The quality of the electron density difference maps led to the unambiguous assignment of the orientations and binding modes of the inhibitors. CPA and AMR bound to the ATP-binding site are located between the N-terminal and C-terminal

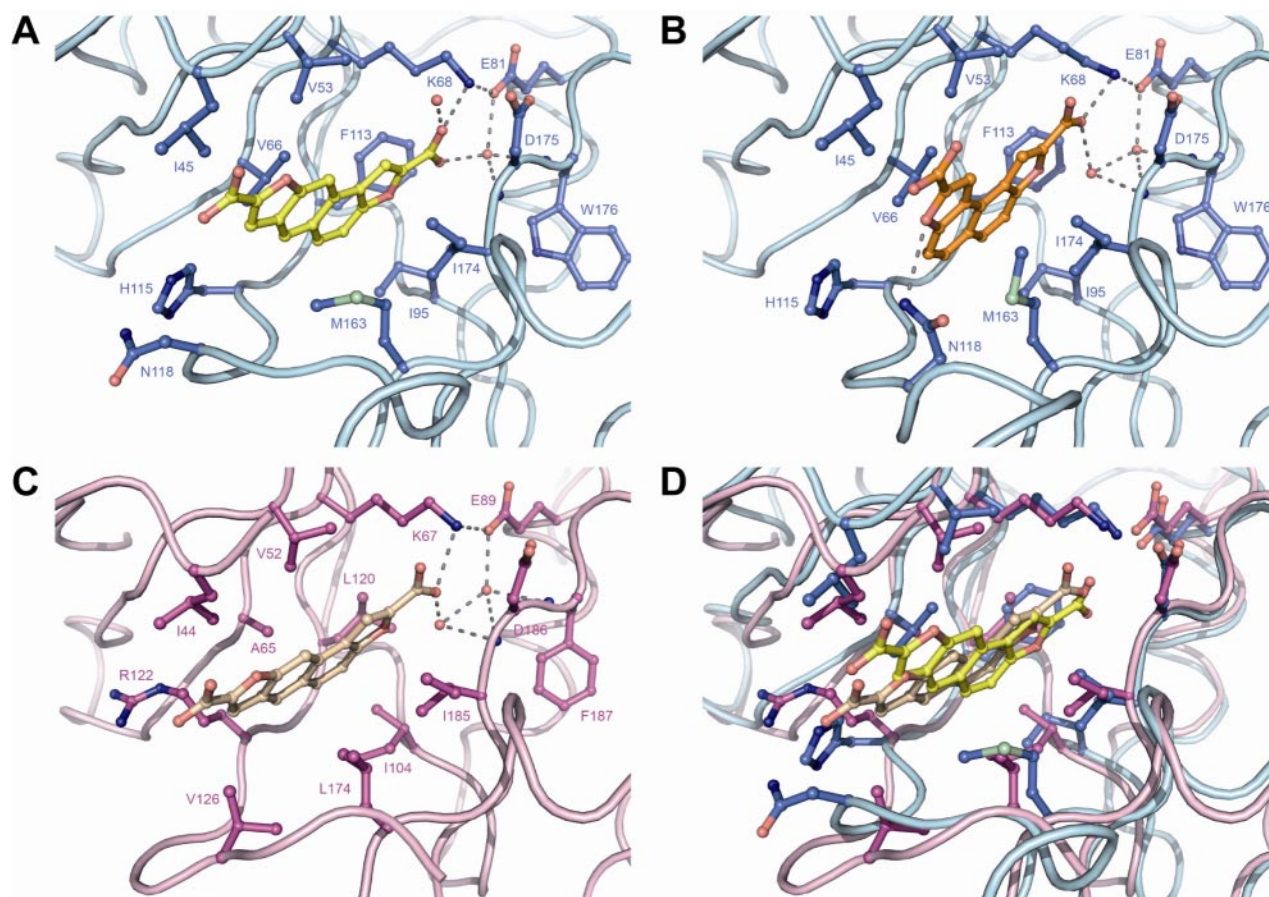


Figure 2. Views of the ATP-binding site of the CK2 and Pim-1 in complex with difuranedicarboxyl acid derivatives. CK2 is represented as light blue coils; Pim-1 is represented as light pink coils. Residues most involved in ligand interactions are the corresponding dark colored balls and sticks. Hydrogen bonds are dark dashed lines; water molecules involved in ligand stabilization are red balls. *A*) CK2 site in complex with CPA 1 (yellow ball and sticks). *B*) CK2 site in complex with AMR (orange). *C*) Pim-1 site in complex with CPA 1 (wheat). *D*) Superimposition of Pim-1 and CK2 site in complex with CPA 1.

lobes of CK2 α and occupy a position similar to other inhibitors for which the complex structures are known (Fig. 2A, B): the plane formed by the naphthalene core fits the A pocket (83) naturally occupied by the adenine moiety of ATP or analogues like AMPPNP, while one of the carboxyfuran groups partially occupies the ribose (R) binding pocket and the other one the inner catalytic lysine (K) pocket (see Supplemental Fig. 5 for important structural features). The CK2-inhibitor interactions are mainly due to hydrophobic contacts that involve mainly the G-loop residues Leu45, Val53, Val66, the hydrophobic residue preceding the DWG motif (the phenylalanine of the DFG-motif is a tryptophan in CK2), the gatekeeper Phe113, Ile174, and the Met163 close to the catalytic loop (Fig. 2A, B). The inner carboxylic acid group, which occupies the K pocket, forms a crucial hydrogen bond with the conserved catalytic lysine 68. However, the orientations of CPA and AMR are slightly different and AMR is tilted $\sim 10^\circ$ compared with CPA. As a consequence, the interaction network between the inner carboxylic acid and Lys68 is slightly different for both ligands. In addition, CPA and AMR are further stabilized by the mean of the carboxyfuran moiety that is oriented toward the solvent-exposed region. Indeed, this moiety interacts with a mainly hydrophilic region comprising the N-terminal part of the sheet $\beta 1$ and the entrance regions E0 and E1, which adopt 2 different conformations for CPA and AMR (see above). This stabilization is particularly obvious for AMR, for which the furan oxygen establishes a direct hydrogen bond with the side chain of Asn118 (in E0) and a tight network of water molecules is made between the outer carboxylate oxygens and Asn118 (twice), Asp120 (in E0), His160 (catalytic loop), and the backbone of Leu45 (G loop). By contrast, the carboxylic acid of CPA makes weak interactions with this CK2 region, which appears more flexible. CPA is only sandwiched between carbonyl groups of Leu45 and Asn118 and no water molecules could be found.

Binding mode of CPA in Pim-1

To investigate the binding modes of the dicarboxyfuran hits in Pim-1, we also determined the crystal structure of Pim-1 in complex with CPA at 2.3 Å resolution. In a similar way to the CK2 complex, CPA occupied the Pim-1 ATP site in an ATP-competitive manner. In analogy with the CK2 complexes, the inhibitor establishes numerous hydrophobic contacts with the equivalent Pim-1 residues, which include the P-loop residues Leu44, Val52, and Ala65, the DFG-motif-preceding Ile185, and the residues Ile104 and Leu174. In addition, a few van der Waals contacts are made with the gatekeeper Leu120 and the alkyl chain of the hinge residue Arg122 (Fig. 2C). However, when the CK2 and Pim-1 ATP sites are superimposed, it appears that in Pim-1 the CPA orientation and position resemble more those of AMR in CK2, rather than those of CPA itself (see Fig. 2D). As a consequence, the interaction network between the inner CPA carboxylate and Pim-1 active site is different from the CK2 site and is similar to

AMR/CK2 (direct hydrogen bond with the catalytic Lys67 and interactions mediated through water molecules with the αC glutamate Glu89 and the backbones of the DFG-motif residues Asp186 and Phe187). The second carboxyl group, which is oriented toward the solvent, does not significantly interact with Pim-1 and forms only long range polar interactions with carbonyl backbone atoms.

Docking of dicarboxyfuran hits in CK2 and Pim-1

For the protein-ligand complexes for which no crystal structure was available (AMR in complex with Pim-1 and CPB with both CK2 and Pim-1), the analysis of the mode of interaction was investigated by docking. All 3 hits (CPA, CPB, and AMR) were therefore docked in both CK2 and Pim-1 and the good consistency obtained with the crystal structures of CPA/CK2, AMR/CK2, and CPA/Pim-1 complexes validated the docking method (data not shown).

In the docked position obtained for CPB in CK2, the part of the ligand that is common to CPA is oriented toward the inside of the active site and perfectly superimposed to CPA in the crystal structure (Supplemental Fig. 6). In the models obtained for Pim-1, the inner furancarboxylic moieties of AMR and CPB are perfectly superimposed but, compared with the crystal structure of CPA, they are slightly tilted toward the back of the active site so as to interact directly with the backbone of Asp186. Like in the crystal structures, all models present a dense network of hydrogen bonds and water molecules that anchor the inner carboxylate group to the catalytic lysine and the surrounding residues. The outer furancarboxylate moiety is exposed to the solvent and surrounded by polar residues, while the aromatic cores of AMR and CPB occupy the same region of the active site and lie in the same plane as that of CPA.

DISCUSSION

CK2 and Pim-1 kinases have been reported to be overexpressed in a number of prostate cancers and in leukemia. Moreover, the oncogenic potential of these kinases is well demonstrated by targeted overexpression in lymphoid lineages. Therefore, these 2 constitutively active kinases are attractive targets for drug development. Given their synergistic roles in oncogenic pathways, the development of dual kinase inhibitors is of particular interest (5, 6, 8, 10, 12, 17, 18, 84).

The CK2 and Pim-1 crystal structures in complex with CPA and AMR and the models obtained for AMR in Pim-1 and CPB in both proteins provide a rationale for the structural basis underlying the potency of the 2 studied compounds. Indeed, all 3 compounds can be accommodated not only in the ATP-binding site by numerous hydrophobic contacts between the polyaromatic core of the compounds and the hydrophobic residues of the cavity but also by a direct and strong interaction with the catalytic lysine and the participa-

tion in the surrounding tight network of water molecules. Previous structures of CK2 in complex with ATP-competitive inhibitors that contain carboxylic acid moieties such as IQA (PDB code: 1OM1), 3-carboxy-4(1H)-quinolone, and TBCA also revealed the crucial role of the carboxylate-Lys68 interaction for the inhibitory potency of these compounds and displayed a very similar geometry in this catalytic region when compared with CPA and AMR. The importance of the interaction with the conserved active site lysine has also been highlighted for the CK2 inhibitors TBB and pyrazolo[1,5-a](1,3,5)triazine derivatives (85). For both CK2 and Pim-1, CPA, CPB, and AMR are examples of this novel, noncanonical kinase binding mode based exclusively on the interaction with the catalytic lysine, leaving the hinge unbound, which is also shared by some of the reported CK2 inhibitors. Some Pim-1 inhibitors have also been described to bind in the same way (PDB code: 2C3I).

With the structures for the CK2 and Pim-1 inhibitor complexes in hand, the differences in potencies observed for CPA and CPB derivatives could be explained on a structural basis. The global shape of the active site makes that a different arrangement of the aromatic ring systems, as in the inactive analogues of CPA and AMR compounds 17, 18, 19, and 20, may result in steric clashes with CK2 residues or may expose the hydrophobic core to the solvent (no plausible position could be found by docking for those compounds). The compounds with the same shape as the hits but lacking the carboxyl group that interacts with the active site lysine (compounds 3 and 10) or in which the carboxyl group has been replaced by an ester (compounds 5, 7, 8, 11, 13, and 15) are devoid of activity for both CK2 and Pim-1 (Table 1 and Scheme 1). This highlights the crucial importance of the interaction with the conserved lysine-glutamate dyad, which cannot be formed for those compounds.

Interactions with the E0-E1 region may also possess a relevant role in the difurane inhibitory potency. Indeed, the mono-acid compound 4 derived from CPA lacks the outer carboxylate and then lost most of its inhibitory activity. The principal hypothesis for this lack of activity may be the loss of the long-range interactions between the outer carboxylate and the E0-E1 regions. In contrast, partial activity can be kept by adding an external ester group, which might favorably be stabilized as exemplified by the CPB-derived compound 9. Moreover, solvation may play a role through the outer carboxylate, which is exposed to the solvent, as observed for AMR and its water-molecule-mediated interaction with CK2. However, the different configurations observed for this structural element and the poorly sequence conservation within the CK2 family make the prediction of interactions with this region very difficult.

One exception to the general observed trend in this series of compounds is compound 6, which possesses an ester in place of the crucial carboxylic acid but still retains partial inhibitory activity on both CK2 and Pim-1 (Table 1). This might be explained by a different binding mode of that particular compound. Although only one orienta-

tion can be observed in the crystal structures, an alternative position of the ligand could be found when docking CPA in both proteins in which the molecule could be flipped so as to exchange the 2 groups R1 and R2 and maintain the interaction with the catalytic lysine. However, this flipped orientation appeared to be less favorable than the one observed in the crystal structures.

Partial saturation of one and both of the furan rings in the counterparts of AMR compounds 12 and 16 is correlated with a complete loss of activity. Structurally, partial saturation of a furan ring causes the loss of conjugation between the carboxylate and the aromatic core and leads to a geometry in which the carboxylate is no longer in the same plane as the rings.

Docking of compound 12 could only yield unfavorable orientations of the carboxylic acid that enables the formation of polar interactions with the active site lysine (data not shown), which explains its lack of activity. However, compound 16 may still bind to the catalytic lysine in the same way as AMR, since it still has a carboxylate in plane. However, docking showed that the additional hydrogen atoms and the unplanar 5-member ring could cause a steric hindrance with CK2, which could explain the loss of activity observed.

One interesting observation is the high selectivity of CPA for CK2 and Pim-1. To explain the origin of this selectivity, we considered CK2 and Pim-1 as CPA-sensitive kinases, Mnk2 and CDK2 as moderately CPA-inhibited kinases, and GSK3 β and Chk1, which are not inhibited by CPA. Docking of CPA in these proteins showed that the ligand could fit in all of them (data not shown). However, superposition of the crystal structures of those 6 kinases shows that, although the polar residues of the lysine pocket are conserved for all proteins, some differences are observed in the hydrophobic residues that assure the anchoring of the aromatic core of CPA in the active site. The most significant difference between CK2 and Pim-1 and the other kinases that are not (or less) inhibited by CPA are the CK2 residues Ile174 and Ile95, which are also present in Pim-1 (185 and 104, respectively). These bulky hydrophobic residues are replaced by either cysteine, serine, or alanine for Ile174 and by valine for Ile95, which is all smaller in size. It can be noticed that when these replacements are accompanied by a change of the hinge Val116 to a bulkier residue, CPA exerts a moderate inhibitory activity on the kinase (this is the case of CDK2 and Mnk2), which is completely lost otherwise (the case of GSK3 β and Chk1). In addition, a reduced sensitivity for emodin, TBB and IQA in the CK2 α Ile174Ala mutant has been observed, highlighting the importance of the involvement of this residue in the inhibitor binding. These considerations suggest that the highlighted bulky residues delimit a narrow ATP-binding site that is crucial for an efficient hydrophobic interaction with CPA. Therefore, proteins with a bigger cavity due to less bulky residues would establish weaker interactions, which could explain the lack of activity of CPA. The smaller size of CK2 active site as compared with the other kinases had already been pointed out


and was reported to account for the selectivity of other CK2 inhibitors such as TBB (86).

However, such selectivity might result from other factors than just the size of the ATP-binding site. Indeed, the mode in which the compound enters the pocket, the balance between favorable and unfavorable mutations in the binding site, some thermodynamic parameters (*e.g.*, stabilization of the E0-E1 loop and the hinge region, water exclusion, and structural changes), and the mode of kinase activation are relevant points that need to be addressed. The role of the gatekeeper nature in the conformation of the regulatory spine conformation and the binding ability of inhibitor has been pointed out recently. The results of Daley's group indeed highlight the relevant role of the hydrophobic nature of the gatekeeper residue (87, 88) in turning the regulatory spine in an active conformation. Interestingly, in CK2 and Pim-1, CPA interacts with the hydrophobic residues found at the position of the gatekeeper residue (Phe113 in CK2 and Leu120 in Pim-1) and one of the crucial isoleucines (Leu95 in CK2 and Leu104 in Pim-1) makes hydrophobic contacts with this gatekeeper and the first residue of the regulatory spine (Leu97/CK2 and Leu106/Pim-1). All together, this may participate in the stabilization of the regulatory spine in a constitutively active conformation, conferring then a restrictive selectivity of those kinases for CPA and derivative compounds.

CONCLUSIONS

Automated screening followed by SAR led to the identification of 3 new difurandicarboxylic acid-type dual CK2/Pim kinase inhibitors, which are among the most potent ATP-competitive inhibitors known to date against each of these therapeutically important kinases. Furthermore, the identified inhibitors are chemically unrelated to compounds that have been previously described. Crystal structures of the complexes of these inhibitors with CK2 and Pim-1 highlight the structural features that make those compounds potent ATP-competitive inhibitors: they have a planar ring arrangement in order to fit the conserved hydrophobic ATP-binding site of CK2 and Pim-1 and at least a carboxylate group to reach their catalytic site.

The dual inhibition of 2 oncogenic kinases like CK2 and Pim-1 by difuran compounds is of particular interest, since in prostate cancer dysregulated expression of both kinases correlates significantly with the clinical outcome of the disease (8, 84).

Optimization toward cell-potent inhibitors is ongoing. Hopefully, this will allow the discovery of new cell-potent dual inhibitors targeting pathological signal transduction pathways that control apoptosis resistance and growth dysregulation. 

This work was supported by INSERM, CNRS, CEA, the Institut Curie, the Ligue Nationale Contre le Cancer (équipe labélisée 2007), the Institut National du Cancer (grant 57). R.P. is the recipient of a fellowship from the Fondation pour

la Recherche Medicale. The authors thank Sylvie Dubruille for assessing the purity of all synthesized compounds by HPLC. The authors also thank the ChemAxon company (<http://www.chemaxon.com>) that has allowed the academic TAMIS software team (especially S. Wieczorek, S. Aci, C. Charavay, and S. Roy) to freely use the MarvinView package, a helpful graphic user interface for viewing structure files and query results. The authors are also indebted to Dr. J. Schwaller and S. Ehret for cellular Pim-1 kinase assays and to Dr. Michael Lilly (Loma Linda University School of Medicine, Loma Linda, CA, USA) for providing us with BaF3/Pim-1 cells. The authors thank Juan-Carlos Fontecilla-Camps and his group (Laboratoire de Cristallogénèse et Cristallographie des Protéines, Institut de Biologie Structurale Jean-Pierre Ebel, Grenoble, France) for laboratory technical support. The authors thank the SLS and ESRF MX team and local contacts for support during diffraction data collections on beamlines. The Structural Genomics Consortium is a registered charity (No. 1097737) that receives funds from the Canadian Institutes for Health Research, the Canadian Foundation for Innovation, Genome Canada through the Ontario Genomics Institute, GlaxoSmithKline, Karolinska Institutet, the Knut and Alice Wallenberg Foundation, the Ontario Innovation Trust, the Ontario Ministry for Research and Innovation, Merck & Co Inc., the Novartis Research Foundation, the Swedish Agency for Innovation Systems, the Swedish Foundation for Strategic Research, and the Wellcome Trust.

REFERENCES

1. Ahmed, K., Gerber, D. A., and Cochet, C. (2002) Joining the cell survival squad: an emerging role for protein kinase CK2. *Trends Cell Biol.* **12**, 226–230
2. Bibby, A. C., and Litchfield, D. W. (2005) The multiple personalities of the regulatory subunit of protein kinase CK2: CK2 dependent and CK2 independent roles reveal a secret identity for CK2beta. *Int. J. Biol. Sci.* **1**, 67–79
3. Litchfield, D. W. (2003) Protein kinase CK2: structure, regulation and role in cellular decisions of life and death. *Biochem. J.* **369**, 1–15
4. Meggio, F., and Pinna, L. A. (2003) One-thousand-and-one substrates of protein kinase CK2? *FASEB J.* **17**, 349–368
5. Ahmad, K. A., Wang, G., Slaton, J., Unger, G., and Ahmed, K. (2005) Targeting CK2 for cancer therapy. *Anticancer Drugs* **16**, 1037–1043
6. Valdman, A., Fang, X., Pang, S. T., Ekman, P., and Egevad, L. (2004) Pim-1 expression in prostatic intraepithelial neoplasia and human prostate cancer. *Prostate* **60**, 367–371
7. Xie, Y., Xu, K., Dai, B., Guo, Z., Jiang, T., Chen, H., and Qiu, Y. (2006) The 44 kDa Pim-1 kinase directly interacts with tyrosine kinase Etk/BMX and protects human prostate cancer cells from apoptosis induced by chemotherapeutic drugs. *Oncogene* **25**, 70–78
8. Laramas, M., Pasquier, D., Filhol, O., Ringeisen, F., Descotes, J. L., and Cochet, C. (2007) Nuclear localization of protein kinase CK2 catalytic subunit (CK2alpha) is associated with poor prognostic factors in human prostate cancer. *Eur. J. Cancer* **43**, 928–934
9. O-charoenrat, P., Rusch, V., Talbot, S. G., Sarkaria, I., Viale, A., Socci, N., Ngai, I., Rao, P., and Singh, B. (2004) Casein kinase II alpha subunit and C1-inhibitor are independent predictors of outcome in patients with squamous cell carcinoma of the lung. *Clin. Cancer Res.* **10**, 5792–5803
10. Kim, J. S., Eom, J. I., Cheong, J. W., Choi, A. J., Lee, J. K., Yang, W. I., and Min, Y. H. (2007) Protein kinase CK2alpha as an unfavorable prognostic marker and novel therapeutic target in acute myeloid leukemia. *Clin. Cancer Res.* **13**, 1019–1028
11. Slaton, J. W., Unger, G. M., Sloper, D. T., Davis, A. T., and Ahmed, K. (2004) Induction of apoptosis by antisense CK2 in human prostate cancer xenograft model. *Mol. Cancer Res.* **2**, 712–721

12. Pagano, M. A., Cesaro, L., Meggio, F., and Pinna, L. A. (2006) Protein kinase CK2: a newcomer in the "druggable kinome." *Biochem. Soc. Trans.* **34**, 1303–1306
13. Qian, K. C., Wang, L., Hickey, E. R., Studts, J., Barringer, K., Peng, C., Kronkaitis, A., Li, J., White, A., Mische, S., and Farmer, B. (2005) Structural basis of constitutive activity and a unique nucleotide binding mode of human Pim-1 kinase. *J. Biol. Chem.* **280**, 6130–6137
14. Niefind, K., Guerra, B., Pinna, L. A., Issinger, O. G., and Schomburg, D. (1998) Crystal structure of the catalytic subunit of protein kinase CK2 from *Zea mays* at 2.1 Å resolution. *EMBO J.* **17**, 2451–2462
15. Sarno, S., Ghisellini, P., and Pinna, L. A. (2002) Unique activation mechanism of protein kinase CK2. The N-terminal segment is essential for constitutive activity of the catalytic subunit but not of the holoenzyme. *J. Biol. Chem.* **277**, 22509–22514
16. Raaf, J., Issinger, O. G., and Niefind, K. (2008) Insights from soft X-rays: the chlorine and sulfur sub-structures of a CK2 α /DRB complex. *Mol. Cell. Biochem.* **316**, 15–23
17. Cibull, T. L., Jones, T. D., Li, L., Eble, J. N., Ann Baldrige, L., Malott, S. R., Luo, Y., and Cheng, L. (2006) Overexpression of Pim-1 during progression of prostatic adenocarcinoma. *J. Clin. Pathol.* **59**, 285–288
18. Chen, J., Kobayashi, M., Darmanin, S., Qiao, Y., Gully, C., Zhao, R., Yeung, S. C., and Lee, M. H. (2009) Pim-1 plays a pivotal role in hypoxia-induced chemoresistance. *Oncogene* **28**, 2581–2592
19. Yan, B., Zenskova, M., Holder, S., Chin, V., Kraft, A., Koskinen, P. J., and Lilly, M. (2003) The PIM-2 kinase phosphorylates BAD on serine 112 and reverses BAD-induced cell death. *J. Biol. Chem.* **278**, 45358–45367
20. Duncan, J. S., and Litchfield, D. W. (2008) Too much of a good thing: The role of protein kinase CK2 in tumorigenesis and prospects for therapeutic inhibition of CK2. *Biochim. Biophys. Acta* **1784**, 33–47
21. Ruzzene, M., and Pinna, L. A. (2009) Addiction to protein kinase CK2: A common denominator of diverse cancer cells? *Biochim. Biophys. Acta* **1804**, 499–504
22. Morishita, D., Katayama, R., Sekimizu, K., Tsuruo, T., and Fujita, N. (2008) Pim kinases promote cell cycle progression by phosphorylating and down-regulating p27Kip1 at the transcriptional and posttranscriptional levels. *Cancer Res.* **68**, 5076–5085
23. Aho, T. L., Sandholm, J., Peltola, K. J., Ito, Y., and Koskinen, P. J. (2006) Pim-1 kinase phosphorylates RUNX family transcription factors and enhances their activity. *BMC Cell Biol.* **7**, 21
24. Zenskova, M., Sahakian, E., Bashkirova, S., and Lilly, M. (2008) The PIM1 kinase is a critical component of a survival pathway activated by docetaxel and promotes survival of docetaxel-treated prostate cancer cells. *J. Biol. Chem.* **283**, 20635–20644
25. Lilly, M., Sandholm, J., Cooper, J. J., Koskinen, P. J., and Kraft, A. (1999) The PIM-1 serine kinase prolongs survival and inhibits apoptosis-related mitochondrial dysfunction in part through a bcl-2-dependent pathway. *Oncogene* **18**, 4022–4031
26. Romieu-Mourez, R., Landesman-Bollag, E., Seldin, D. C., Traish, A. M., Mercurio, F., and Sonenshein, G. E. (2001) Roles of IKK kinases and protein kinase CK2 in activation of nuclear factor-kappaB in breast cancer. *Cancer Res.* **61**, 3810–3818
27. Landesman-Bollag, E., Song, D. H., Romieu-Mourez, R., Sussman, D. J., Cardiff, R. D., Sonenshein, G. E., and Seldin, D. C. (2001) Protein kinase CK2: signaling and tumorigenesis in the mammary gland. *Mol. Cell. Biochem.* **227**, 153–165
28. Hammerman, P. S., Fox, C. J., Cinalli, R. M., Xu, A., Wagner, J. D., Lindsten, T., and Thompson, C. B. (2004) Lymphocyte transformation by Pim-2 is dependent on nuclear factor-kappaB activation. *Cancer Res.* **64**, 8341–8348
29. Ayala, G. E., Dai, H., Ittmann, M., Li, R., Powell, M., Frolow, A., Wheeler, T. M., Thompson, T. C., and Rowley, D. (2004) Growth and survival mechanisms associated with perineural invasion in prostate cancer. *Cancer Res.* **64**, 6082–6090
30. Klumpp, S., Maurer, A., Zhu, Y., Aichele, D., Pinna, L. A., and Kriegelstein, J. (2004) Protein kinase CK2 phosphorylates BAD at threonine-117. *Neurochem. Int.* **45**, 747–752
31. Kallergi, G., Markomanolaki, H., Giannoukaraki, V., Papadaki, M. A., Strati, A., Lianidou, E. S., Georgoulas, V., Mavroudis, D., and Agelaki, S. (2009) Hypoxia-inducible factor-1 α and vascular endothelial growth factor expression in circulating tumor cells of breast cancer patients. *Breast Cancer Res.* **11**, R84
32. Mottet, D., Ruys, S. P., Demazy, C., Raes, M., and Michiels, C. (2005) Role for casein kinase 2 in the regulation of HIF-1 activity. *Int. J. Cancer* **117**, 764–774
33. Hubert, A., Paris, S., Piret, J. P., Ninane, N., Raes, M., and Michiels, C. (2006) Casein kinase 2 inhibition decreases hypoxia-inducible factor-1 activity under hypoxia through elevated p53 protein level. *J. Cell Sci.* **119**, 3351–3362
34. Reiser-Erkan, C., Erkan, M., Pan, Z., Bekasi, S., Giese, N. A., Streit, S., Michalski, C. W., Friess, H., and Kleeff, J. (2008) Hypoxia-inducible proto-oncogene Pim-1 is a prognostic marker in pancreatic ductal adenocarcinoma. *Cancer Biol. Ther.* **7**, 1352–1359
35. Prudent, R., Moucadel, V., Lopez-Ramos, M., Aci, S., Laudet, B., Mouawad, L., Barette, C., Einhorn, J., Einhorn, C., Denis, J. N., Bisson, G., Schmidt, F., Roy, S., Lafanechere, L., Florent, J. C., and Cochet, C. (2008) Expanding the chemical diversity of CK2 inhibitors. *Mol. Cell. Biochem.* **316**, 71–85
36. Cozza, G., Bortolato, A., and Moro, S. (2009) How druggable is protein kinase CK2? [E-pub ahead of print] *Med. Res. Rev.* doi: 10.1002/med.20164
37. Anderes, K., Siddiqui-Jain, A., Streiner, P., Chua, F., Pierre, F., Omori, M., Darjania, L., Stansfield, R., Phung, J., Bliesath, J., Ho, C., Drygin, D., O'Brien, S., and Rice, W. (2009) Discovery and characterization of CX-4945, a selective orally bioavailable small molecule inhibitor of protein kinase CK2: phase I initiated. Poster presentation, American Association for Cancer Research (AACR) Annual Meeting, April 18–22, 2009, Denver, CO, USA
38. Cheney, I. W., Yan, S., Appleby, T., Walker, H., Vo, T., Yao, N., Hamatake, R., Hong, Z., and Wu, J. Z. (2007) Identification and structure-activity relationships of substituted pyridones as inhibitors of Pim-1 kinase. *Bioorg. Med. Chem. Lett.* **17**, 1679–1683
39. Debreczeni, J. E., Bullock, A. N., Atilla, G. E., Williams, D. S., Bregman, H., Knapp, S., and Meggers, E. (2006) Ruthenium half-sandwich complexes bound to protein kinase Pim-1. *Angew. Chem. Int. Ed. Engl.* **45**, 1580–1585
40. Pierce, A. C., Jacobs, M., and Stuver-Moody, C. (2008) Docking study yields four novel inhibitors of the protooncogene Pim-1 kinase. *J. Med. Chem.* **51**, 1972–1975
41. Pogacic, V., Bullock, A. N., Fedorov, O., Filippakopoulos, P., Gasser, C., Biondi, A., Meyer-Monard, S., Knapp, S., and Schwaller, J. (2007) Structural analysis identifies imidazo[1,2-b]pyridazines as PIM kinase inhibitors with in vitro antileukemic activity. *Cancer Res.* **67**, 6916–6924
42. Heriche, J. K., Lebrin, F., Rabilloud, T., Leroy, D., Chambaz, E. M., and Goldberg, Y. (1997) Regulation of protein phosphatase 2A by direct interaction with casein kinase 2 α . *Science* **276**, 952–955
43. Gerber, D. A., Souquere-Besse, S., Puvion, F., Dubois, M. F., Bensaude, O., and Cochet, C. (2000) Heat-induced relocalization of protein kinase CK2. Implication of CK2 in the context of cellular stress. *J. Biol. Chem.* **275**, 23919–23926
44. Prudent, R., Lopez-Ramos, M., Moucadel, V., Barette, C., Grierson, D., Mouawad, L., Florent, J. C., Lafanechere, L., Schmidt, F., and Cochet, C. (2008) Salicylaldehyde derivatives as new protein kinase CK2 inhibitors. *Biochim. Biophys. Acta* **1780**, 1412–1420
45. Filhol, O., Cochet, C., Wedegaertner, P., Gill, G. N., and Chambaz, E. M. (1991) Coexpression of both alpha and beta subunits is required for assembly of regulated casein kinase II. *Biochemistry* **30**, 11133–11140
46. Vedadi, M., Niesen, F. H., Allali-Hassani, A., Fedorov, O. Y., Finerty, P. J., Jr., Wasney, G. A., Yeung, R., Arrowsmith, C., Ball, L. J., Berglund, H., Hui, R., Marsden, B. D., Nordlund, P., Sundstrom, M., Weigelt, J., and Edwards, A. M. (2006) Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 15835–15840
47. Ermakova, I., Boldyreff, B., Issinger, O. G., and Niefind, K. (2003) Crystal structure of a C-terminal deletion mutant of human protein kinase CK2 catalytic subunit. *J. Mol. Biol.* **330**, 925–934
48. Kabsch, W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* **26**, 795–800
49. Navaza, J. (1994) AMoRe: an automated package for molecular replacement *Acta Crystallogr. A* **50**, 157–163

50. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763
51. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132
52. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255
53. Bullock, A. N., Debreczeni, J., Amos, A. L., Knapp, S., and Turk, B. E. (2005) Structure and substrate specificity of the Pim-1 kinase. *J. Biol. Chem.* **280**, 41675–41682
54. Leslie, A. G. W., and Powell, H. (2007) *MOSFLM*, MRC Laboratory of Molecular Biology, Cambridge, UK
55. Evans, P. (2007) *SCALA—Scale Together Multiple Observations of Reflections*, MRC Laboratory of Molecular Biology, Cambridge, UK
56. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., and Read, R. J. (2005) Likelihood-enhanced fast translation functions. *Acta Crystallogr. D Biol. Crystallogr.* **61**, 458–464
57. Painter, J., and Merritt, E. A. (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 439–450
58. Rarey, M., Kramer, B., Lengauer, T., and Klebe, G. (1996) A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **261**, 470–489
59. Augustin, S., ed. (2005) *FlexX*, BioSolveIT, Darmstadt, Germany
60. Kuntz, I. D., Blaney, J. M., Oatley, S. J., Langridge, R., and Ferrin, T. E. (1982) A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.* **161**, 269–288
61. Moustakas, D. T., Lang, P. T., Pegg, S., Pettersen, E., Kuntz, I. D., Brooijmans, N., and Rizzo, R. C. (2006) Development and validation of a modular, extensible docking program: DOCK 5. *J. Comput. Aided Mol. Des.* **20**, 601–619
62. MacKerell, A., Jr., Bashford, D., Bellott, M., Dunbrack, R. L., Jr., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, III, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiórkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616
63. SYBYL. Tripos Inc., 1699 South Hanley Road, St. Louis, MO, 63144
64. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Montgomery, J. A., Jr., Vreven, T., Kudin, K. N., Burant, J. C., Millam, J. M., Iyengar, S. S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G. A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J. E., Hratchian, H. P., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Ayala, P. Y., Morokuma, K., Voth, G. A., Salvador, P., Dannenberg, J. J., Zakrzewski, V. G., Dapprich, S., Daniels, A. D., Strain, M. C., Farkas, O., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Ortiz, J. V., Cui, Q., Baboul, A. G., Clifford, S., Cioslowski, J., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Challacombe, M., Gill, P. M. W., Johnson, B., Chen, W., Wong, M. W., Gonzalez, C., and Pople, J. A. (2004) *Gaussian 03W*, Gaussian, Inc., Wallingford, CT, USA
65. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004) Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174
66. Arrault, A., Touzeau, F., Guillaumet, G., and Mèrou, J. Y. (1999) A straightforward synthesis of 1,2-dihydronaphtho[2,1-d]furans from 2-naphthols. *Synthesis* **7**, 1241–1245
67. Bilger, C., Demerseman, P., Buisson, J.-P., Royer, R., Gayral, P., and Fourniat, J. (1987) Sur quelques modifications d'activité biologique consécutives à l'accolement d'un second cycle nitrofuranne à l'homocycle de nitroarenofurannes. *Eur. J. Med. Chem.* **22**, 213–219
68. Seidler, J., McGovern, S. L., Doman, T. N., and Shoichet, B. K. (2003) Identification and prediction of promiscuous aggregating inhibitors among known drugs. *J. Med. Chem.* **46**, 4477–4486
69. McGovern, S. L., Caselli, E., Grigorieff, N., and Shoichet, B. K. (2002) A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. *J. Med. Chem.* **45**, 1712–1722
70. Feng, B. Y., Simeonov, A., Jadhav, A., Babaoglu, K., Inglese, J., Shoichet, B. K., and Austin, C. P. (2007) A high-throughput screen for aggregation-based inhibition in a large compound library. *J. Med. Chem.* **50**, 2385–2390
71. Shoichet, B. K. (2006) Interpreting steep dose-response curves in early inhibitor discovery. *J. Med. Chem.* **49**, 7274–7277
72. Ryan, A. J., Gray, N. M., Lowe, P. N., and Chung, C. W. (2003) Effect of detergent on “promiscuous” inhibitors. *J. Med. Chem.* **46**, 3448–3451
73. Pagano, M. A., Bain, J., Kazimierczuk, Z., Sarno, S., Ruzzene, M., Di Maira, G., Elliott, M., Orzeszko, A., Cozza, G., Meggio, F., and Pinna, L. A. (2008) The selectivity of inhibitors of protein kinase CK2: an update. *Biochem. J.* **415**, 353–365
74. Manning, G., Whyte, D. B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002) The protein kinase complement of human genome. *Science* **298**, 1912–1934
75. Battistutta, R., De Moliner, E., Sarno, S., Zanotti, G., and Pinna, L. A. (2001) Structural features underlying selective inhibition of protein kinase CK2 by ATP site-directed tetrabromo-2-benzotriazole. *Protein. Sci.* **10**, 2200–2206
76. Battistutta, R., Mazzorana, M., Sarno, S., Kazimierczuk, Z., Zanotti, G., and Pinna, L. A. (2005) Inspecting the structure-activity relationship of protein kinase CK2 inhibitors derived from tetrabromo-benzimidazole. *Chem. Biol.* **12**, 1211–1219
77. Battistutta, R., Sarno, S., De Moliner, E., Papinutto, E., Zanotti, G., and Pinna, L. A. (2000) The replacement of ATP by the competitive inhibitor emodin induces conformational modifications in the catalytic site of protein kinase CK2. *J. Biol. Chem.* **275**, 29618–29622
78. Niefind, K., Guerra, B., Ermakowa, I., and Issinger, O. G. (2001) Crystal structure of human protein kinase CK2: insights into basic properties of the CK2 holoenzyme. *EMBO J.* **20**, 5320–5331
79. Battistutta, R. (2009) Protein kinase CK2 in health and disease: Structural bases of protein kinase CK2 inhibition. *Cell. Mol. Life Sci.* **66**, 1868–1889
80. Niefind, K., Raaf, J., and Issinger, O. G. (2009) Protein kinase CK2 in health and disease: Protein kinase CK2: from structures to insights. *Cell. Mol. Life Sci.* **66**, 1800–1816
81. Niefind, K., Yde, C. W., Ermakowa, I., and Issinger, O. G. (2007) Evolved to be active: sulfate ions define substrate recognition sites of CK2 α and emphasise its exceptional role within the CMGC family of eukaryotic protein kinases. *J. Mol. Biol.* **370**, 427–438
82. Raaf, J., Brunstein, E., Issinger, O. G., and Niefind, K. (2008) The CK2 α /CK2 β interface of human protein kinase CK2 harbors a binding pocket for small molecules. *Chem. Biol.* **15**, 111–117
83. Liao, J. J. (2007) Molecular recognition of protein kinase binding pockets for design of potent and selective kinase inhibitors. *J. Med. Chem.* **50**, 409–424
84. Xu, Y., Zhang, T., Tang, H., Zhang, S., Liu, M., Ren, D., and Niu, Y. (2005) Overexpression of PIM-1 is a potential biomarker in prostate carcinoma. *J. Surg. Oncol.* **92**, 326–330
85. Mazzorana, M., Pinna, L. A., and Battistutta, R. (2008) A structural insight into CK2 inhibition. *Mol. Cell. Biochem.* **316**, 57–62
86. Sarno, S., Ruzzene, M., Frascella, P., Pagano, M. A., Meggio, F., Zambon, A., Mazzorana, M., Di Maira, G., Lucchini, V., and Pinna, L. A. (2005) Development and exploitation of CK2 inhibitors. *Mol. Cell. Biochem.* **274**, 69–76
87. Azam, M., Powers, J. T., Einhorn, W., Huang, W. S., Shakespeare, W. C., Zhu, X., Dalgarno, D., Clackson, T., Sawyer, T. K., and Daley, G. Q. AP24163 inhibits the gatekeeper mutant of BCR-ABL and suppresses in vitro resistance. *Chem. Biol. Drug Des.* **75**, 223–227
88. Azam, M., Seeliger, M. A., Gray, N. S., Kuriyan, J., and Daley, G. Q. (2008) Activation of tyrosine kinases by mutation of the gatekeeper threonine. *Nat. Struct. Mol. Biol.* **15**, 1109–1118

Received for publication November 12, 2009.

Accepted for publication March 25, 2010.

Antitumor Activity of Pyridocarbazole and Benzopyridoindole Derivatives that Inhibit Protein Kinase CK2

Renaud Prudent^{1,2,3}, Virginie Mouchadel^{1,2,3}, Chi-Hung Nguyen⁴, Caroline Barette⁵, Frédéric Schmidt^{6,7}, Jean-Claude Florent^{6,7}, Laurence Lafanechère⁵, Céline F Sautel^{1,2,3}, Eve Duchemin-Pelletier^{1,2,3}, Elodie Spreux^{1,2,3}, Odile Filhol^{1,2,3}, Jean-Baptiste Reiser⁸, and Claude Cochet^{1,2,3}

Abstract

The alkylid compound ellipticine derived from the berrywood tree is a topoisomerase II poison that is used in ovarian and breast cancer treatment. In this study, we report the identification of ellipticine derivatives and their tetracyclic angular benzopyridoindole analogues as novel ATP-competitive inhibitors of the protein kinase CK2. *In vitro* and *in vivo* assays showed that these compounds have a good pharmacologic profile, causing a marked inhibition of CK2 activity associated with cell cycle arrest and apoptosis in human cancer cells. Further, *in vivo* assays demonstrate antitumor activity in a mouse xenograft model of human glioblastoma. Finally, crystal structures of CK2–inhibitor complex provide structural insights on the molecular basis of CK2 inhibition. Our work lays the foundation for development of clinically useful CK2 inhibitors derived from a well-studied scaffold with suitable pharmacokinetics parameters. *Cancer Res*; 70(23); 9865–74. ©2010 AACR.

Introduction

Protein kinases represent major therapeutic targets and several kinase inhibitors have demonstrated powerful clinical activity in pathologies in which the target kinase is dysregulated.

Protein kinase CK2 plays critical roles in cell growth, cell differentiation, apoptosis, and oncogenic transformation (1, 2). CK2 has been found dysregulated in many cancers and its dual function in promotion of cell growth and in suppression of apoptosis may be particularly relevant to its oncogenic potential (3). The recent association of aberrant CK2 expression with unfavorable prognostic markers in prostate cancer (4) and in acute myeloid leukemia (5) supports the direct implication of CK2 in tumor formation and recurrence. Indeed, CK2 is now considered as a promising therapeutic target, supporting the development of chemical inhibitors. Consequently, several

classes of CK2 inhibitors have been reported showing various efficacy and specificity. Quercetin, emodin, 4,5,6,7-tetrabromo-1-benzotriazole (TBB), and indoloquinazolin derivative (IQA) are representative of main classes of ATP-competitive CK2 inhibitors (6, 7). Other families were also described, such as anthraquinone-related compounds, condensed polyphenolic, 7-substitued indoloquinazoline derivatives, and 2,8-difurandicarboxylic acid derivatives (8–11). However, so far, only a single report described a CK2 inhibitor (CX-4945; Cylene Pharmaceuticals) disclosing *in vivo* efficacy (http://www.cylenepharma.com/cylene/pr_1230934977). Thus, broadening the availability and diversity of CK2 inhibitors is warranted.

We have identified ellipticine and benzopyridoindole derivatives as new CK2 inhibitors. Ellipticine and its derivatives exhibit antitumoral properties and are established chemotherapeutics (12), acting either by distinct or common mechanisms of action. First, their intercalation in double-stranded DNA interferes with DNA replication or DNA transcription (13). Second, ellipticine induces topoisomerase II–mediated transient DNA double-strand breaks, which have been identified as the primary cellular target of the drug (14, 15). A third mechanism involves DNA adduct formation after P450 (CYP)–mediated activation of ellipticine in target tissues (16). Several studies have also reported apoptosis induction by ellipticine involving the p53 tumor suppressor protein (17–19). Finally, protein kinase inhibition may also represent a novel mechanism involved into ellipticine effect. Cdk2-dependent p53 phosphorylation is selectively inhibited by ellipticine and 9-hydroxyellipticine (20). Ellipticine derivatives have also inhibitory activity on c-Kit (21).

Here, we report that ellipticine and its benzopyridoindole derivatives are *in vitro* and *in vivo* ATP-competitive CK2

Authors' Affiliations: ¹INSERM, U873; ²CEA, iRTSV/LTS; ³Université Joseph Fourier, Grenoble, France; ⁴UMR 176 CNRS-Institut Curie, Institut Curie Bât 110, Orsay, France; ⁵CNRS, UMR 5168, CEA, iRTSV/CMBA, Grenoble, France; ⁶UMR 176 CNRS-Institut Curie, Centre de Recherche; ⁷CNRS, UMR 176, Paris, France; ⁸Institut de Biologie Structurale Jean-Pierre Ebel, CEA-CNRS-UJF-PSB, Grenoble, France

Note: Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org/>).

Corresponding Author: Claude Cochet, INSERM, U873, CEA, iRTSV/LTS, and Université Joseph Fourier, F-38054 Grenoble, France. Phone: 33-4-38-78-42-04; Fax: 33-4-38-78-50-58; E-mail: claudc.cochet@cea.fr or Jean-Baptiste Reiser, Institut de Biologie Structurale Jean-Pierre Ebel, CEACNRS-UJF-PSB, 41 rue Jules Horowitz, F-38027 Grenoble, France. Phone: 33-4-76-20-94-49; Fax: 33-4-76-20-94-80; E-mail: jean-baptiste.reiser@ibs.fr.

doi: 10.1158/0008-5472.CAN-10-0917

©2010 American Association for Cancer Research.

inhibitors displaying antitumoral activity. We provide a rationale basis for the mode of action of these compounds on the basis of the crystal structure of CK2 α -inhibitor complexes. Knowing that ellipticine displays good pharmacokinetics criteria, this new class of inhibitors is a source of attractive leads for further optimization.

Material and Methods

Chemical libraries

The Institut Curie-CNRS library (6,560 compounds) and The Prestwick Chemical Library (1,120 compounds) were screened for CK2 inhibitors. Active compounds in the primary screen were retested from freshly made solutions.

Cell culture

HeLa (human cervical adenocarcinoma), U373 (human glioblastoma), MCF-10A (human mammary epithelial cells), MESSA, and MESSA Dx5 (myosarcoma) cell lines were obtained from the American Type Culture Collection (ATCC) during 2005. Cells were cultured and stored according to the supplier's instructions and used at passages 5 to 20. Briefly, cells were routinely cultivated in Dulbecco's medium (Invitrogen Life Technologies, Inc.) supplemented with 10% (v/v) fetal calf serum (BioWest), except for MCF-10A cells that were maintained in a 1:1 mixture of Dulbecco's modified Eagle's medium and Ham's F-12 medium supplemented with 5% horse serum, L-glutamine, 25 ng/mL human recombinant EGF, 100 ng/mL cholera toxin, 10 μ g/mL insulin, and 0.5 μ g/mL hydrocortisone. Cells were last tested in July 2009 through cell morphology monitoring, growth curve analysis, and contamination checks.

Automated screening

Automated screening of chemical libraries was performed in 96-well plates using the automated platform of the CMBA (centre de Criblage pour Molécules Bio-Actives, Grenoble) screening facilities as described in the study of Prudent et al. (22).

CK2 phosphorylation assays

Compounds were tested in a radiometric CK2 assay performed in a final volume of 18 μ L containing 1 μ L of compound or equivalent amount of dimethyl sulfoxide (DMSO) as control, 3 μ L of CK2 α (36 ng), and a mixture containing 100 μ mol/L of peptide substrate (RRREDEESDDEE), 10 mmol/L MgCl₂, and 100 μ mol/L [γ -³²P]-ATP (6,000 Ci/mmol). Assays were performed at room temperature for 5 minutes before termination by the addition of 60 μ L of 4% trichloroacetic acid. Incorporation of ³²P in the peptide substrate was determined as previously described (23).

Protein purification and crystallization

A truncated active form of the catalytic α subunit of human CK2 (CK2 α^{AC}) was expressed in *Escherichia coli* and purified as described in the study of Ermakova et al. (24). CK2 α^{AC} was first crystallized using hanging drop vapor-diffusion technique. Crystallization drops were prepared by mixing 2 μ L of CK2 α^{AC} -Mg²⁺-substrate peptide mix [4.6 mgmL⁻¹ CK2 α^{AC} ,

114 μ mol/L peptide substrate RRREDEESDDEE, 680 μ mol/L MgCl₂, 340 μ mol/L adenosine 5'-(β , γ -imido)triphosphate] with 2 μ L of precipitant solution (36% polyethylene glycol 5000 monomethyl ether, 150 mmol/L ammonium sulfate, and 100 mmol/L Tris-HCl, pH 7.5). Crystals of ellipticine and benzopyridoindole derivative-CK2 α^{AC} complexes were then obtained by soaking techniques: crystals were soaked for 1 week in a 4- μ L drop of same composition as crystallization drop except that adenosine 5'-(β , γ -imido)triphosphate was replaced by 300 μ mol/L of these derivatives.

Structure determination and refinement

For X-ray diffraction data collection, CK2 α^{AC} complex crystals were directly flash frozen in liquid nitrogen without adding any extra cryoprotectant. All data were collected at -173.15°C on beamlines of either ID23-eh1 or ID23-eh2 of the European Synchrotron Radiation Facility (ESRF) at a wavelength of 1.072 or 0.834 Å using an ADSC Quantum 315R CCD detector. Data processing was performed using XDS (25) and summarized in Supplementary Table S1. The CK2 α^{AC} complex structures were determined by molecular replacement method with AMoRe (26) from the CCP4 suite (27) using the CK2 α^{AC} structure as search model [pdb code: 1PJK, (24)]. The final models were obtained after several rounds of alternate manual building with coot (28) and maximum likelihood refinement with Refmac5 of CCP4 (29), and using data in all available resolution range. The final refinement statistics are summarized in Supplementary Table S1. Pymol was used to generate structure figures. The structure factors and coordinates are deposited in the Protein Data Bank (PDB) under the accession numbers: 3OWJ, 3OWK, and 3OWL.

Kinase selectivity profiling

Kinase selectivity of compound **18** was assessed using a panel of 57 recombinant protein kinases. The assays were performed at 10 μ mol/L ATP in the presence of 10 μ mol/L inhibitor using the Kinase profiler panel service (Millipore). Inhibition, expressed as the percent of activity determined in the absence of inhibitor, was calculated from the residual activity measured in the presence of 10 μ mol/L inhibitor.

Cellular CK2 activity assay

Cellular CK2 activity was assayed as previously described (22, 30). Briefly, HeLa cells were transfected with a chimeric yellow fluorescent protein-based CK2 activity reporter (pEYFPc1-S β S). Phosphorylation status of the reporter was measured using the monoclonal antibody anti-GFP/yellow fluorescent protein (Roche, ref. 1814460) at 1/1,000 and the goat anti-mouse HRP as secondary antibody (Sigma, ref. A4416).

Murine xenograft tumor assay

All experimental procedures adhered to our local ethical committee (Comité régional d'éthique pour l'Expérimentation animale CREEA, Rhône Alpes, protocol no. 286). Female Harlan athymic nude mice (6–8 weeks) were inoculated subcutaneously into the right flank with 7.5×10^5 U373 cells. When tumors reached ± 50 mm³ (volume = length \times width \times height), animals were treated intraperitoneally

3 times weekly (every 2 days) for 2 weeks, with compound **19** (0.14 mg/100 μ L per injection), or equivalent amount of DMSO (10% final) dissolved in PEG3350 22.5%, bovine serum albumin 0.45%. Body weight and tumor volume were determined twice a week. The experiment was terminated when tumor volume reached about 1,000 mm³. Results are presented as mean \pm SEM.

Results

Discovery of new CK2 inhibitors by automated screening

So far, only few CK2 inhibitors display suitable pharmacokinetics criteria for *in vivo* investigations. To find new CK2 inhibitor scaffolds appropriate for *in vivo* uses, we screened the Prestwick Chemical Library. This library contains small molecules that were selected for their known bioavailability and safety in humans. The expectation was to identify hits that exhibit correct bioavailability and toxicity. These hits could provide good starting points for optimization program.

The library was screened in an automated luminescence-based kinase assay using human recombinant CK2 α catalytic subunit (Supplementary Fig. S1). Screening data were normalized and analyzed using an in-house software, TAMIS (Tool to Analysis and Manage Informations of Screening, Version 2.0; CEA, Grenoble, France). CK2 kinase inhibitory activity was determined at compound concentration of 10 μ mol/L, using TBB and DMSO as positive and negative controls, respectively.

Up to 24 compounds belonging to several distinct structural families showed inhibition greater than 50%. Among those hits, ellipticine was the only drug with therapeutic indication in oncology. To identify more potent and selective CK2 inhibitors we performed a focused screen on ellipticine derivatives present in the Institut Curie-CNRS small molecule library (Supplementary Fig. S1).

This led to the identification of benzopyridoindole derivatives (31, 32). Radiometric CK2 kinase assay showed that **1**, **8**, **16**, **18**, **19** efficiently inhibited kinase activity in a dose-dependent manner (Table 1) with IC₅₀ ranging from 0.31 to 1.5 μ mol/L. This represents a 10- to 50-fold improvement in inhibitor potency when compared with our primary hit, ellipticine (in these conditions, IC₅₀ for ellipticine is 15.4 μ mol/L).

Chemical features of pyridocarbazole derivatives conferring CK2 inhibitory potency

To define the mode of CK2 inhibition by compound **1**, we performed steady-state kinetic analysis (Supplementary Fig. S2). Lineweaver–Burk inhibition plots showed that CK2 inhibition by compound **1** is ATP competitive, with $K_i = 0.86$ μ mol/L. Similar inhibitory potencies were found using CK2 holoenzyme (data not shown).

To gain insights into the features conferring inhibitory potency to compound **1**, we tested a set of derivatives (Table 1). Position 11 is relatively tolerant because substitution of methyl by hydrogen also leads to an active compound (compound **8**) albeit less potent. In contrast, the free hydroxyl

substituent at position 9 (compounds **1** and **8**) is necessary because analogues bearing substituents other than hydroxyl (e.g., ether, ester) are less active or inactive even when assayed at 10 μ mol/L (compounds **2**, **4**, **6**, **7**, **9**, **10**, **11**, **12**, **13**, **14**, **15**, and **17**).

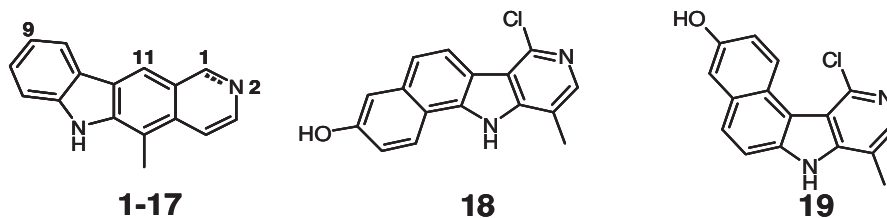
Similarly, methylation of N2 is also not tolerated because compound **3** is inactive. In line with this observation, compound **5**, which is fully substituted at positions 2, 9, and 11, is also inactive. Position 1 also seems to be relatively tolerant because replacement of lactam function (compound **8**) by chloro group (compound **16**) leads to active compound. Whether compounds **11**, **12**, and **13** are inactive because position 9 bears methoxy group or position 1 bears too bulky/hydrophobic substituents remains unknown.

Interestingly, although compounds **14**, **15**, and **17** bear methoxy groups at position 9, they display residual inhibitory potency. In the same trend, ellipticine is active (albeit at higher concentration). This observation indicates that unfavorable substitution at this important position can be counterbalanced by a network of favorable substitutions at other positions. We next investigated the features of the scaffold required for activity. Available pyridoindole (or γ -carboline) derivatives bearing 3 fused rings exhibit poor or no activity (data not shown), meaning that 4 rings are necessary for optimal activity. We tested several available angular tetracyclic compounds related to intopicine, which corresponds to antitumor agent inhibiting both topoisomerase I and II activities (33). Surprisingly, 2 compounds belonging to benzo[*g*]- and benzo[*e*]-pyridoindole (or BgPI and BePI) (compounds **18** and **19**, respectively) are potent CK2 inhibitors (Table 2). This indicates some tolerance into the scaffold of these new CK2 inhibitors.

In summary, regarding linear tetracyclic ellipticine derivatives, it appears that positions 9 and N2 are crucial for activity whereas positions 1 and 11 are relatively tolerant. Moreover, some variations into the scaffold of the molecule are allowed. Indeed, the angular tetracyclic analogues BgPI and BePI, which are structurally related to ellipticine, also exhibit significant CK2 inhibitory potencies.

Pyridocarbazole derivatives are multitarget compounds

Given the multiple targets of ellipticine and derivatives, we compared our structure-activity relationship data with published data (Supplementary Table S2), and on c-Kit (21) and topoisomerase II (15). This analysis showed differences in structure-activity relationship according to the targets. For instance, whereas C1 substitution does not have a strong influence on topoisomerase II and CK2 α activities, it is detrimental on c-Kit inhibitory potency when substituted by bulky groups. In contrast, N2 substitution is detrimental for CK2 α inhibitory potency, but has no effect on topoisomerase II and c-Kit inhibition. N6 substitution has no influence on topoisomerase II, but is detrimental for c-Kit and CK2 α inhibition. A common feature of the 3 targets is the presence of hydroxyl on position 9 in active compounds. Taken together, these observations indicate that distinct chemical requirements are necessary to selectively inhibit each of the 3 targets.

Table 1. Biochemical characterization of pyridocarbazole and B[g/e]PI in aminophenyl-4-methanesulfonamide

Compound	1-2 Bond	Substituents at position				CK2 residual activity	c-Kit residual activity	Topoisomerasell α residual activity	DNA intercalation
		1	2	9	11				
1	Single	C=O	H	OH	Me	4.5	79.9	97.8	65.1
2	Single	C=O	H	OMe	Me	99.5	131.2	94.2	52.0
3	Single	C=O	Me	OH	H	99.0	96.5	100.0	72.9
4	Single	C=O	H	OMe	H	100.1	81.4	99.2	66.8
5	Single	C=O	Bz	OBz	Me	100.1	183.3	133.6	71.2
6	Single	C=O	H	OBz	Me	100.1	103.5	112.2	46.9
7	Single	C=O	H	OAc	H	100.1	110.3	88.3	81.3
8	Single	C=O	H	OH	H	9.7 (IC ₅₀ = 1.5 μ mol/L)	88.8	65.3	79.4
9	Single	C=O	H	OBn	H	99.0	110.5	69.7	93.4
10	Single	C=O	H	OBn	Me	99.9	95.2	39.2	52.2
11	Double	A	H	OMe	Me	100.0	88.5	73.7	52.9
12	Double	B	H	OMe	H	99.8	90.9	37.3	69.4
13	Double	HNPPh	Me	OMe	H	100.1	142.7	81.5	88.3
14	Double	CN	/	OMe	H	62 (IC ₅₀ > 15 μ mol/L)	65.7	103.4	58.9
15	Double	Cl	/	OMe	Me	56.7 (IC ₅₀ > 15 μ mol/L)	93.3	74.5	65.5
16	Double	Cl	/	OH	H	4 (IC ₅₀ = 1.4 μ mol/L)	70.3	115.0	68.7
17	Double	H	Tosyl	OMe	Me	52.3 (IC ₅₀ > 15 μ mol/L)	61.4	82.8	91.7
18						1.4 (IC ₅₀ = 1.5 μ mol/L)	60.4	106.8	62.0
19						3.5 (IC ₅₀ = 0.67 μ mol/L)	16.9	77.0	66.4

CK2 α and c-Kit kinase assays and topoisomerase II α assay were performed at 10 μ mol/L compound concentration. DNA intercalation assay was performed at 2.5 μ mol/L compound concentration.

Table 2. Biochemical characterization of pyridocarbazole and B[g/e]PI in Aminophenyl-2-methoxy-4-methanesulfonamide

Compound	7-8 Bond	Substituents at position		CK2 residual activity	Compound	10-11 Bond	Substituents at position		CK2 residual activity
		3	7				3	11	
18	Double	OH	Cl	1.4 (IC ₅₀ = 1.5 μ mol/L)	19	Double	OH	Cl	3.5 (IC ₅₀ = 0.67 μ mol/L)
20	Double	OMe	Cl	94.0	22	Double	OMe	Cl	73.0
21	Single	OMe	C=O	80.0	23	Single	OMe	C=O	82.0

CK2 α assays were performed at 10 μ mol/L compound concentration. All results are expressed as percentage of activity without inhibitor.

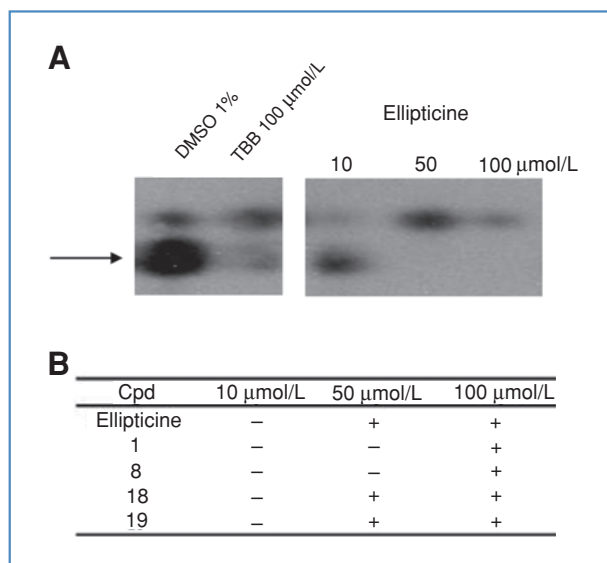


Figure 1. A, ellipticine inhibition of CK2 activity in living cells. HeLa cells transfected with a plasmid expressing a chimeric CK2 activity GFP-based reporter were incubated with the CK2 inhibitor TBB, ellipticine, or DMSO (as control) for 24 hours. Cell extracts were then analyzed by native electrophoresis and GFP was revealed by immunoblotting. The arrow indicates the CK2-phosphorylated form of the reporter. B, selected derivatives that were active *in vitro* were submitted to the cell-based CK2 activity assay as in A. +, a significant decrease in the phosphorylated form of the CK2 activity reporter (lower band in A and in Supplementary Fig. S3).

Pyridocarbazole derivatives are cell-potent CK2 inhibitors

We assayed the inhibitory potencies of pyridocarbazole derivatives on living cells using a cell-based CK2 kinase assay (Fig. 1) (22, 34). In control HeLa cells, the CK2 activity reporter could be detected as a phosphorylated form (lower band in Fig. 1A). When cells were treated with the known CK2 inhibitor TBB, the phosphorylated CK2 reporter form disappeared. Interestingly, 50 $\mu\text{mol/L}$ ellipticine triggered a similar strong inhibition (Fig. 1A). Other derivatives selected in the present work were found to be active in cells at similar or even lower concentrations (compounds **1**, **8**, **18**, **19**) (Fig. 1B; Supplementary Fig. S3).

Ellipticine and B[g/e]PI derivatives inhibit cell viability through cell cycle arrest and apoptosis

Because both ellipticine and CK2 inhibitors are known to affect cell survival, we measured cell viability of p53^{wt} HeLa cells after 48-hour treatment with TBB, ellipticine, or angular compounds **18** and **19** (Fig. 2A). B[g/e]PI derivatives reduced cell viability more efficiently than TBB (IC₅₀ of 2–10 $\mu\text{mol/L}$ and 20 $\mu\text{mol/L}$, respectively). Interestingly, in p53^{mut} U373 glioblastoma cells, 5 $\mu\text{mol/L}$ of pyridocarbazole derivatives was sufficient to decrease cell viability by 50% whereas in these conditions, TBB was not efficient (Fig. 2B). U373 cells harbor a TP53 mutant isoform and are particularly resistant to drug-induced apoptosis. Other cancer cell lines such as wild-type or resistant myosarcoma cells (MESSA and MESSA Dx5,

respectively) were highly sensitive to ellipticine and to a lower extent to compounds **18** and **19**. Of note, the cell viability of the nontransformed MCF10A mammary epithelial cells was also strongly affected by ellipticine, whereas compounds **18** and **19** had a much weaker effect. As cell viability inhibition induced by ellipticine was reported to be correlated with cell cycle arrest (35), we measured the effect of ellipticine and B[g/e]PI derivatives on cell cycle distribution of HeLa and U373 cells (Fig. 3A). Compared with vehicle-treated cells, **18**- or **19**-treated U373 cells exhibited a cell cycle arrest in G₂/M phase. The treatment also caused an appreciable accumulation of the cells in sub-G₁ phase. We next assessed the effects of the compounds on the induction of apoptosis in U373 cells by (terminal deoxynucleotidyl transferase-mediated dUTP nick end labeling (TUNEL). The results showed that **18** or **19** treatment results in the accumulation of TUNEL-positive cells. This effect was more pronounced with Ellipticine or staurosporine treatments (Fig. 3B). Moreover, the same treatments induced an increase of caspase-dependent apoptosis attested by the activation of caspase 3/7 activity (Fig. 3C).

Although these compounds are cell-potent CK2 inhibitors, we noticed that except for ellipticine, the active concentrations able to induce marked CK2 inhibition were higher than those required to induce cytotoxicity. Ellipticine is known as a multitarget agent, thus our selected active benzopyridoindole derivatives could, beside CK2 α , affect other targets involved in cell viability. Indeed, when kinase selectivity of compound **18** was profiled, several kinases involved in cancer-related pathways such as Pim-1, Mnk2, and DARK1 were significantly affected (Supplementary Table S3).

Benzopyridoindole derivatives inhibit colony formation and tumor growth

To ascertain the potential therapeutic interest of compounds **18** and **19**, we first tested their ability to inhibit colony formation using soft agar assay as an *in vitro* surrogate of tumorigenesis. U373 cells can grow without anchorage and form colonies in soft agar culture, reflecting their malignant properties. Both compounds **18** and **19** at a concentration of 5 $\mu\text{mol/L}$ inhibited drastically colony formation (Fig. 4A). In comparison, a weak inhibition of colony formation (10%) was observed in response to 50 $\mu\text{mol/L}$ of TBB (data not shown).

Next, murine xenograft tumor assays were carried out to assess both toxicity and efficiency of compound **18**, which showed the best inhibitory activities toward cell viability and colony formation. In mice treated with compound **18**, we observed a significant reduction of tumor incidence, whatever the time point. Three weeks after first compound injection, mice injected with compound **18** had tumors that were 5 to 6 times smaller than those injected with DMSO (Fig. 4B). Furthermore, mice were healthy after repetitive injections of the compound. They grew normally without significant weight loss. In addition, no obvious toxic effects were observed when the experiment was terminated and the tumors removed, indicating that compound **18** was relatively well tolerated at doses sufficient to induce antitumoral effect.

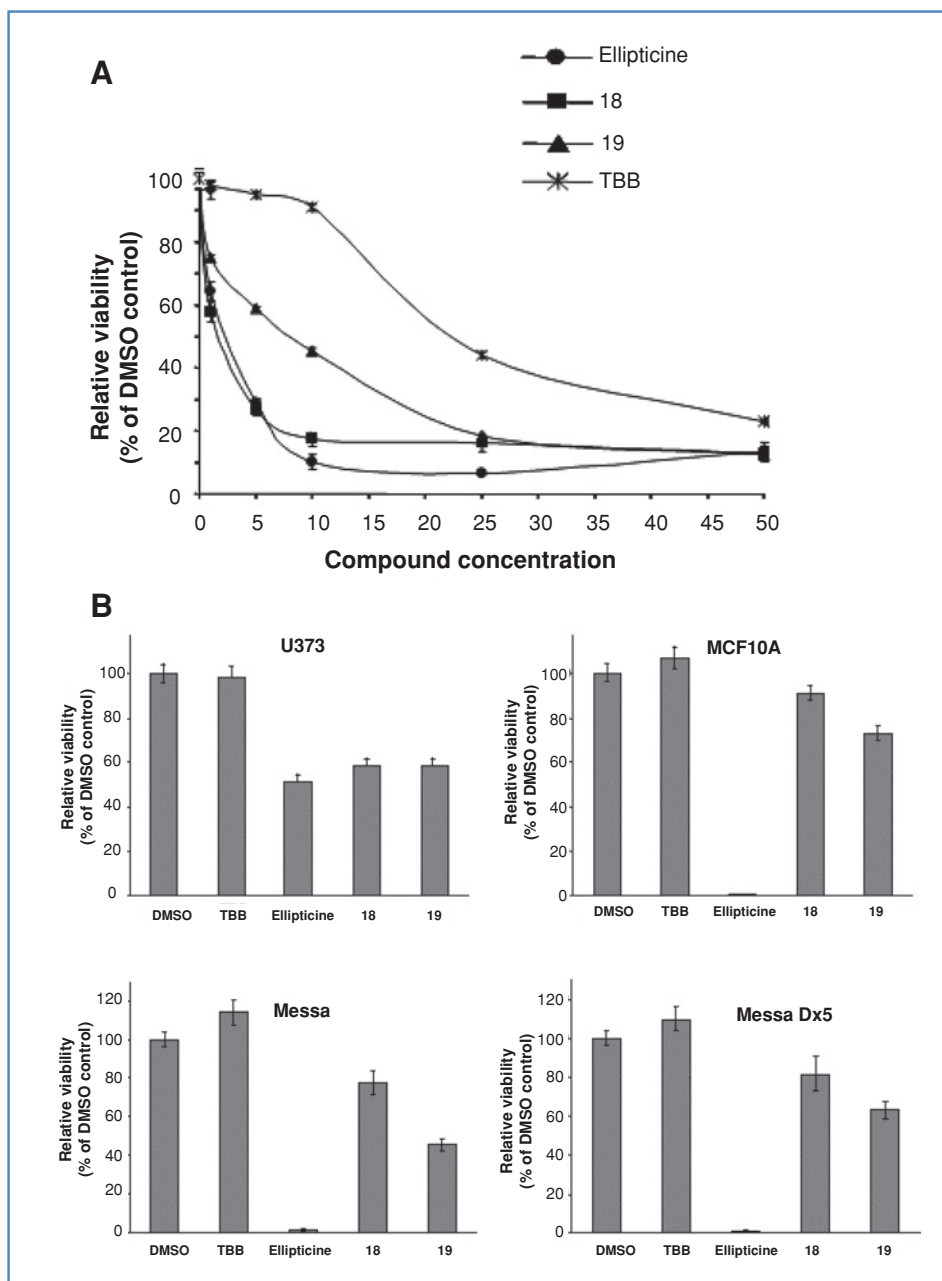


Figure 2. Effect of ellipticine and B[g/e]PI derivatives on cell viability. A, HeLa cells were incubated with increasing amount of compounds for 48 hours and cell viability was assayed by the Cell Titer Glo Luminescent Cell Viability Assay (Promega). Results are given relative to the luminescence recorded for DMSO. Experiment was done in triplicate and repeated twice. B, different cell lines as indicated were incubated with 5 $\mu\text{mol/L}$ compounds or equivalent amount of DMSO as control. Two days later, living cells were counted as in A. Results are given relative to the luminescence recorded for DMSO. Data are representative of 2 experiments performed in triplicate.

X-ray crystallographic structures of CK2-inhibitor complexes

To gain insights at the atomic level into the mode of action of these new inhibitors, we solved the structure of the truncated active form of CK2 α in complex with pyridocarbazole and B[g/e]PI derivatives by X-ray crystallography: compound **1**/CK2 α at 1.85 Å resolution, compound **18**/CK2 α at 1.80 Å resolution, and compound **19**/CK2 α at 2.1 Å resolution (Fig. 5). Electron densities of compounds **1**, **18**, and **19** are well defined and lead to no ambiguity regarding their orientations and positions in the ATP

binding site. Commonly, the interaction between the 3 compounds and CK2 are mainly hydrophobic and involve mainly residues 53, 66, and 174. They use the plane formed by 2 adjacent cycles to fit perfectly into the "purine plane" naturally occupied by the ATP adenine moiety (24, 36). However, they differ in their orientation to conserve specific interactions between the common atoms of their scaffold and of the ATP groove. For instance, the common hydroxyl group (at position 9) of the 3 derivatives forms a direct hydrogen bond with the main chain nitrogen of Val 116 that mimics the classical ATP/CK2 interaction. But, the

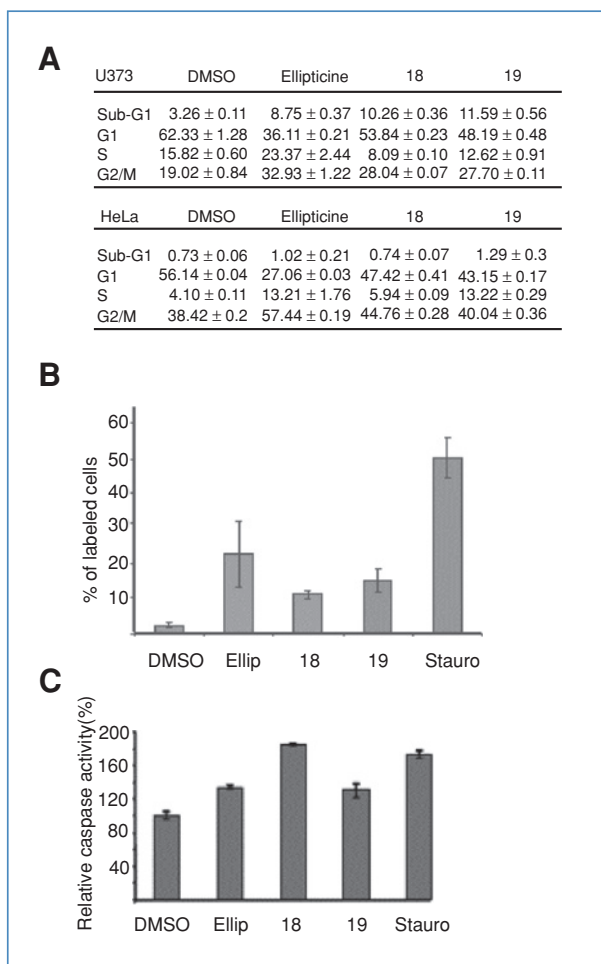


Figure 3. Ellipticine and B[g/e]PI derivatives promote cell cycle arrest and apoptosis. **A**, nearly confluent HeLa or U373 cells were treated with 5 $\mu\text{mol/L}$ compounds or equivalent amount of DMSO. After 24 hours, cells were harvested and labeled with propidium iodide. DNA content was analyzed with FACScalibur and Cell Quest software. Percentage of HeLa or U373 cells in sub-G₁, G₁, S, and G₂/M phases are summarized. Experiment was repeated twice. **B**, U373 cells were treated with 5 $\mu\text{mol/L}$ ellipticine (Ellip) and compounds **18** and **19** for 48 hours or with staurosporine (Stauro) for 4 hours. Then, TUNEL staining was measured with the DNA Fragmentation Kit (Calbiochem) according to the manufacturer recommendations. Around 200 cells were counted in randomly selected fields, for each point. **C**, U373 cells were treated with 5 $\mu\text{mol/L}$ compounds for 24 hours. Then, caspase 3/7 activation was measured with the Caspase-Glo 3/7 Assay (Promega) according to the manufacturer recommendations. Data are representative of 2 experiments performed in triplicate.

differences in geometry and orientation lead them to interact differently with the bottom of the binding pocket. Compound **1** fits into the ribose pocket and partially the α -phosphate site and interacts with the catalytic site via water molecules that connect N6 of compound **1** to the main chain of Asp 174 and the catalytic dyad (Lys68/Glu81). In addition, compound **1** is further stabilized by the external side of the ATP pocket by the mean of a π -interaction between its cycle D and the Histidine at position 160 (Supplementary Table S2). By contrast, compounds **18**

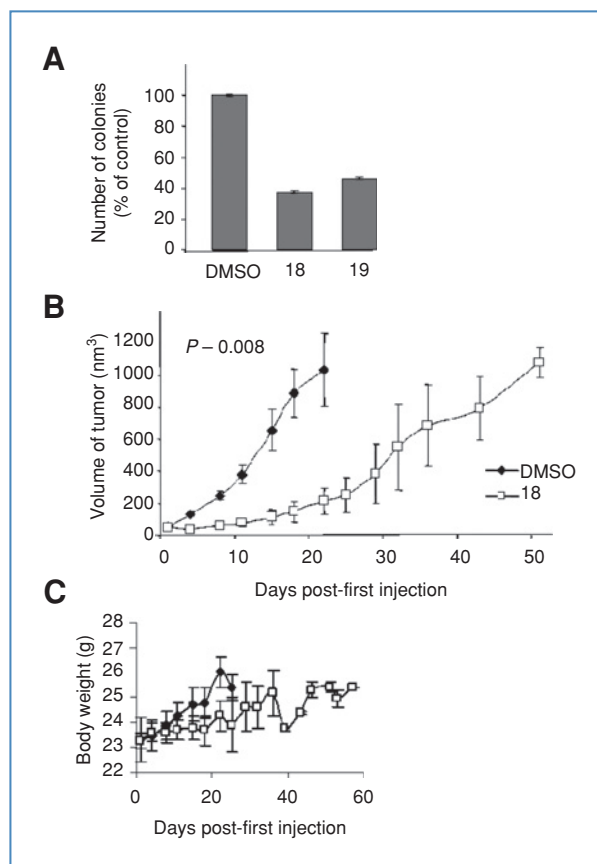


Figure 4. Compound **18** reduces colony formation *in vitro* and tumor growth *in vivo*. **A**, soft agar assay for colony formation. Experiment was done in duplicate at least 3 times. **B**, murine xenograft tumor growth assay. Athymic nude mice were inoculated subcutaneously into the right flank with 7.5×10^5 U373 cells. When tumors reached 50 mm³, mice were injected intraperitoneally every 2 days for 2 weeks with compound **18** or equivalent amount of DMSO. Body weight and tumor volume were determined 2 times weekly. Results are presented as mean \pm SEM.

and **19** are rotated by 180-degree rotation and binds deeply into the α -phosphate site to directly make a hydrogen bond between their nitrogen atom of chloropyridine group and Lys68 Nz atom.

Discussion

Despite many published CK2 inhibitors with various modes of inhibition (37, 38), only few were successfully validated *in vivo*. In this work, first we report that pyridocarbazole and its angular analogues B[g/e]PI are *in vitro* potent submicromolar ATP-competitive CK2 α inhibitors. Substitution requirements in pyridocarbazole derivatives are different for efficient inhibition of CK2 than substitution required for other known ellipticine targets (c-kit, topoisomerase II α , DNA intercalation). This can give guidelines for compound optimization. Second, these compounds are also active on cellular CK2 kinase activity and induce cell cycle arrest of a glioblastoma cell line.

Proteins involved in the cells' suicide program such as ARC, Bid, and p53 (2) are some of the many cellular substrates of

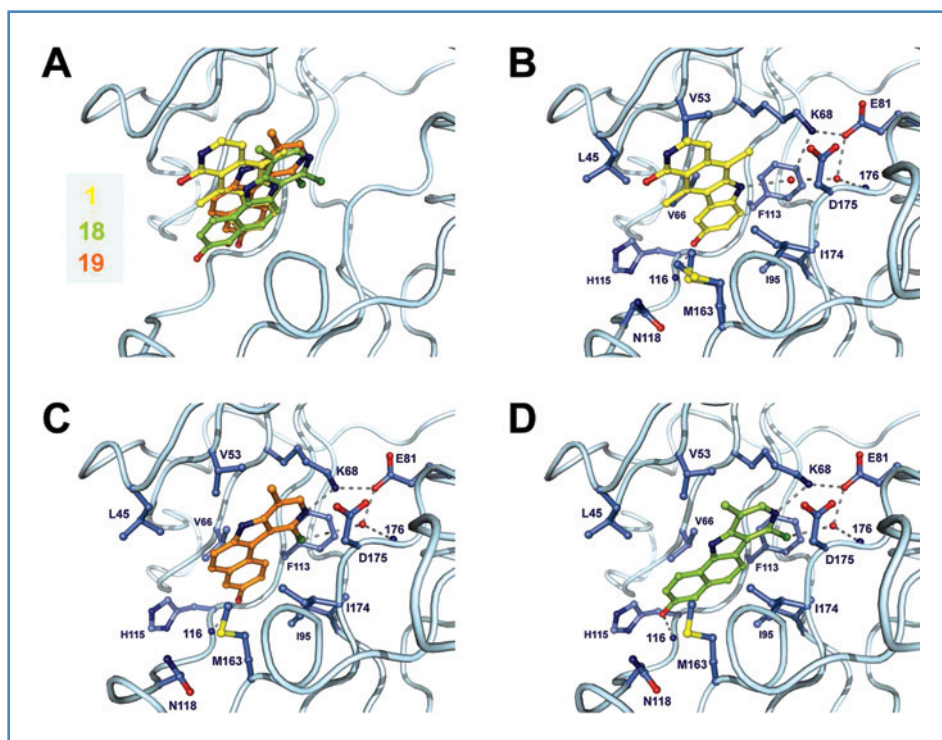


Figure 5. Close view of the structure of ellipticine and angular analogue derivatives in complex with CK2. A, superimposition of 3 CK2 α -inhibitor complexes. B, CK2 α -compound **1** complex structure. C, CK2 α -compound **19** complex structure. D, CK2 α -compound **18** complex structure. For all panels, CK2 α is represented in light blue coils and CK2 α residues involved in direct interaction with the compounds in darker blue ball-and-sticks. Hydrogen bonds between CK2 α atoms and these derivatives are highlighted in gray dashed lines. Compounds are represented in ball and sticks: **1** in yellow, **18** in green, and **19** in orange. Atoms are colored according to their nature: red, O; blue, N; light yellow, S; dark green, Cl.

CK2 (39). Indeed, we observed that in response to pyridocarbazole derivatives, HeLa cells showed caspase-dependent apoptosis. Cytotoxic effect may be partially due to CK2 inhibition because cytotoxicity appeared at concentrations lower than those required to achieve complete CK2 α inhibition. Most of cancer cells rely, however, on high CK2 activity as judged from their susceptibility to CK2 inhibitors or silencing of CK2 catalytic subunits (40). This undue reliance of tumor cells on CK2 activity may explain their high sensitivity to the pyridocarbazole derivatives. Another explanation is that these compounds target other cellular proteins such as prosurvival kinases. Indeed, kinase selectivity profiling of compound **18** showed that several kinases were significantly affected: Pim-1 and Mnk2, which are involved in cancer-related pathways (41), and DRAK1, which is related to death-associated protein kinases that trigger apoptosis (42). It is now accepted that high specificity is not a requisite for potent drugs (43) and that polypharmacology is an underlying mechanism essential in drug actions (43). Indeed, hitting more than 1 kinase at once can prevent cells to bypass the primary target by activating redundant signaling pathways (44). Interestingly, when assayed in a murine xenograft tumor growth assay, 1 of our compounds (**18**) exhibited antitumoral activity without apparent toxic effects. We conclude that this chemical class of leads seems to be safe enough to represent a good starting point for future optimization.

We have solved 3 CK2 α -inhibitor complex structures. Crystal structure of CK2-compound **1** complex provides a clear explanation for the differential potency of various pyridocarbazole derivatives. The intolerance to any substitution except hydroxyl group at position 9 is fully explained by the crucial hydrogen bond that compound **1** makes with CK2 α

backbone, and that mimics interactions between ATP and CK2. However, the disruption of this crucial hydrogen bond can be partially compensated by extra interactions involving other residues as in compounds **14**, **15**, and **17**, which bear polar groups either at position 2 or at position 11 and have residual inhibition capacity. Those groups would be there in favorable position to interact with CK2 α residues Leu45 and His160. This is exemplified by the substitution in compound **16**, which bears a chlorine group as compound **15** and a hydroxyl group as compound **1**, and retains full inhibition potency. However, in contrast to CK2 α -compound **1** structure, the structures of CK2 α in complex with compounds **18** and **19** revealed that their orientations are flipped by 180 degree, which lead their chlorine group to fit into the ATP cavity. Therefore, the position of such polar group cannot easily be predicted. Concerning position 11, reduction of this group volume by the substitution of a methyl group (in compound **1**) by an atom of hydrogen (in compound **3**) is correlated with a decreased inhibition (higher IC₅₀). This can be easily explained by the disruption of hydrophobic contacts between the methyl group and Met163 and Leu45 of CK2 α . However, as the number of those contacts is in minority as compared with the overall network between compound **1** and CK2 α , this disruption has a limited impact on IC₅₀ suggesting that position 11 would be tolerant to bulkier substituents. The structures of CK2 α in complex with compounds **18** and **19** illustrate the ability of the CK2 ATP-binding groove to bind benzopyridoindole derivatives with "cycle" configurations different from the original ellipticine, as long as they remain perfectly plane to fit into the purine plane. Furthermore, it appears crucial to keep the OH group at position 9 to mimic ATP/CK2 interaction as for compound **1**.

To conclude, this class of pyridocarbazole and benzopyridoindole derivatives has a well-mastered chemistry and displays new biological activities and good pharmacokinetic profile. Its high potency, the well-explored structure-activity relationship and the solved structure of these compounds in complex with CK2 will provide a good starting point for further optimization processes.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were declared.

Acknowledgments

The authors thank the ChemAxon company (<http://www.chemaxon.com>), which has allowed the academic TAMIS software team (especially C. Charavay

and S. Roy) to freely use the MarvinView package, a helpful Graphic User Interface for viewing structure files and query results. The authors also thank Drs. Marie-Paule Teulade-Fichou and David Monchaud for providing reagents and protocol for DNA intercalation assays. We thank H. Pointu and technicians for their efficiency for animals care. Special thanks to Juan-Carlos Fontecilla-Camps for the use of his lab facilities for crystallization, and MX group and local contacts of ESRF for their help in data collection on beamlines.

Grant Support

This work was supported by the Institut National de la Santé et de la Recherche Médicale (INSERM), the Centre National pour la Recherche Scientifique (CNRS), the Commissariat à l'Énergie atomique (CEA), the Institut Curie, the Ligue Nationale Contre le Cancer (équipe labellisée 2003–2012), the Institut National du Cancer (grant number 57). R. Prudent received a fellowship for the Fondation pour la Recherche Médicale.

Received 03/17/2010; revised 09/28/2010; accepted 09/30/2010; published OnlineFirst 11/30/2010.

References

- Ahmed K, Gerber DA, Cochet C. Joining the cell survival squad: an emerging role for protein kinase CK2. *Trends Cell Biol* 2002;12:226–30.
- Litchfield DW. Protein kinase CK2: structure, regulation and role in cellular decisions of life and death. *Biochem J* 2003;369:1–15.
- Tawfic S, Yu S, Wang H, Faust R, Davis A, Ahmed K. Protein kinase CK2 signal in neoplasia. *Histol Histopathol* 2001;16:573–82.
- Laramas M, Pasquier D, Filhol O, Ringeisen F, Descotes JL, Cochet C. Nuclear localization of protein kinase CK2 catalytic subunit (CK2 α -ph) is associated with poor prognostic factors in human prostate cancer. *Eur J Cancer* 2007;43:928–34.
- Kim JS, Eom JI, Cheong JW, Choi AJ, Lee JK, Yang WI, et al. Protein kinase CK2 α as an unfavorable prognostic marker and novel therapeutic target in acute myeloid leukemia. *Clin Cancer Res* 2007;13:1019–28.
- Pagano MA, Cesaro L, Meggio F, Pinna LA. Protein kinase CK2: a newcomer in the 'druggable kinome'. *Biochem Soc Trans* 2006;34:1303–6.
- Duncan JS, Gyenis L, Lenehan J, Bretner M, Graves LM, Haystead TA, et al. An unbiased evaluation of CK2 inhibitors by chemo-proteomics: characterization of inhibitor effects on CK2 and identification of novel inhibitor targets. *Mol Cell Proteom* 2008;7:1077–88.
- De Moliner E BN, Johnson LN. Alternative binding modes of an inhibitor to two different kinases. *Eur J Biochem* 2003;270:3174–81.
- Meggio F, Pagano MA, Moro S, Zagotto G, Ruzzene M, Sarno S, et al. Inhibition of protein kinase CK2 by condensed polyphenolic derivatives. An *in vitro* and *in vivo* study. *Biochemistry* 2004;43:12931–6.
- Loizou JI, El-Khamisy SF, Zlatanou A, Moore DJ, Chan DW, Qin J, et al. The protein kinase CK2 facilitates repair of chromosomal DNA single-strand breaks. *Cell* 2004;117:17–28.
- Lopez-Ramos M, Prudent R, Moucadel V, Sautel CF, Barette C, Lafanechère L, et al. New potent dual inhibitors of CK2 and Pim kinases: discovery and structural insights. *FASEB J*. 2010;24:3171–85.
- Auclair C. Multimodal action of antitumor agents on DNA: the ellipticine series. *Arch Biochem Biophys* 1987;259:1–14.
- Kohn KW, Waring MJ, Glaubiger D, Friedman CA. Intercalative binding of ellipticine to DNA. *Cancer Res* 1975;35:71–6.
- Tewey KM, Chen GL, Nelson EM, Liu LF. Intercalative antitumor drugs interfere with the breakage-reunion reaction of mammalian DNA topoisomerase II. *J Biol Chem* 1984;259:9182–7.
- Fosse P, Rene B, Le Bret M, Paoletti C, Saucier JM. Sequence requirements for mammalian topoisomerase II mediated DNA cleavage stimulated by an ellipticine derivative. *Nucleic Acids Res* 1991;19:2861–8.
- Stiborova M, Breuer A, Aimova D, Stiborova-Rupertova M, Wiessler M, Frei E. DNA adduct formation by the anticancer drug ellipticine in rats determined by ³²P postlabeling. *Int J Cancer* 2003;107:885–90.
- Peng Y, Li C, Chen L, Sebt S, Chen J. Rescue of mutant p53 transcription function by ellipticine. *Oncogene* 2003;22:4478–87.
- Sugikawa E, Hosoi T, Yazaki N, Gamanuma M, Nakanishi N, Ohashi M. Mutant p53 mediated induction of cell cycle arrest and apoptosis at G1 phase by 9-hydroxyellipticine. *Anticancer Res* 1999;19:3099–108.
- Kuo YC, Kuo PL, Hsu YL, Cho CY, Lin CC. Ellipticine induces apoptosis through p53-dependent pathway in human hepatocellular carcinoma HepG2 cells. *Life Sci* 2006;78:2550–7.
- Ohashi M, Sugikawa E, Nakanishi N. Inhibition of p53 protein phosphorylation by 9-hydroxyellipticine: a possible anticancer mechanism. *Jpn J Cancer Res* 1995;86:819–27.
- Vendome J, Letard S, Martin F, Svinarchuk F, Dubreuil P, Auclair C, et al. Molecular modeling of wild-type and D816V c-Kit inhibition based on ATP-competitive binding of ellipticine derivatives to tyrosine kinases. *J Med Chem* 2005;48:6194–201.
- Prudent R, Moucadel V, Lopez-Ramos M, Aci S, Laudet B, Mouawad L, et al. Expanding the chemical diversity of CK2 inhibitors. *Mol Cell Biochem* 2008;316:71–85.
- Filhol O, Cochet C, Delagoutte T, Chambaz EM. Polyamine binding activity of casein kinase II. *Biochem Biophys Res Commun* 1991;180:945–52.
- Ermakova I, Boldyreff B, Issinger OG, Niefind K. Crystal structure of a C-terminal deletion mutant of human protein kinase CK2 catalytic subunit. *J Mol Biol* 2003;330:925–34.
- Kabsch W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J Appl Cryst* 1993;26:795–800.
- Navaza J. AMoRe: an automated package for molecular replacement. *Acta Cryst* 1994;A50:157–63.
- Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 1994;50:760–3.
- Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 2004;60:2126–32.
- Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 1997;53:240–55.
- Prudent R, Lopez-Ramos M, Moucadel V, Barette C, Grierson D, Mouawad L, et al. Salicylaldehyde derivatives as new protein kinase CK2 inhibitors. *Biochim Biophys Acta* 2008;1780:1412–20.
- Nguyen CH, Lhoste JM, Lavelle F, Bissery MC, Bisagni E. Synthesis and antitumor activity of 1-[[[dialkylamino]alkyl]amino]-4-methyl-5H-pyrido[4,3-b]benzo[e]- and -benzo[g]indoles. A new class of antineoplastic agents. *J Med Chem* 1990;33:1519–28.
- Vinogradov S, Roig V, Sergueeva Z, Nguyen CH, Arimondo P, Thuong NT, et al. Synthesis and binding properties of oligo-2'-deoxyribonucleotides conjugated with triple-helix-specific interca-

- lators: benzo[e] and benzo[g] pyridoindoles. *Bioconjug Chem* 2003;14:120–35.
33. Riou JF, Fosse P, Nguyen CH, Larsen AK, Bissery MC, Grondard L, et al. Intopicine (RP 60475) and its derivatives, a new class of antitumor agents inhibiting both topoisomerase I and II activities. *Cancer Res* 1993;53:5987–93.
 34. Prudent R, Lopez-Ramos M, Moucadel V, Barette C, Grierson D, Mouawad L, et al. Salicylaldehyde derivatives as new protein kinase CK2 inhibitors. *Biochim Biophys Acta* 2008;1780:1412–20.
 35. Kuo PL, Hsu YL, Chang CH, Lin CC. The mechanism of ellipticine-induced apoptosis and cell cycle arrest in human breast MCF-7 cancer cells. *Cancer Lett* 2005;223:293–301.
 36. Yde CW, Ermakova I, Issinger OG, Niefind K. Inclining the purine base binding plane in protein kinase CK2 by exchanging the flanking side-chains generates a preference for ATP as a cosubstrate. *J Mol Biol* 2005;347:399–414.
 37. Prudent R, Cochet C. New protein kinase CK2 inhibitors: jumping out of the catalytic box. *Chem Biol* 2009;16:112–20.
 38. Prudent R, Moucadel V, Laudet B, Barette C, Lafanechère L, Hasenknopf B, et al. Identification of polyoxometalates as nanomolar noncompetitive inhibitors of protein kinase CK2. *Chem Biol* 2008;15: 683–92.
 39. Meggio F, Pinna LA. One-thousand-and-one substrates of protein kinase CK2? *FASEB J* 2003;17:349–68.
 40. Ruzzene M, Pinna LA. Addiction to protein kinase CK2: A common denominator of diverse cancer cells? *Biochim Biophys Acta* 2009;1804:499–504.
 41. Cibull TL, Jones TD, Li L, Eble JN, Ann Baldrige L, Malott SR, et al. Overexpression of Pim-1 during progression of prostatic adenocarcinoma. *J Clin Pathol* 2006;59:285–8.
 42. Sanjo H, Kawai T, Akira S. DRAKs, novel serine/threonine kinases related to death-associated protein kinase that trigger apoptosis. *J Biol Chem* 1998;273:29066–71.
 43. Frantz S. Drug discovery: playing dirty. *Nature* 2005;437:942–3.
 44. Sawyers CL. Opportunities and challenges in the development of kinase inhibitor therapy for cancer. *Genes Dev* 2003;17:2998–3010.

Potassium acts as a GTPase-activating element on each nucleotide-binding domain of the essential *Bacillus subtilis* EngA

Anne-Emmanuelle Foucher^{1,2,3}, Jean-Baptiste Reiser^{1,2,3}, Christine Ebel^{1,2,3}, Dominique Housset^{1,2,3} & Jean-Michel Jault^{1,2,3†}

¹Institut de Biologie Structurale, Université Joseph Fourier Grenoble 1, ²UMR 5075 CNRS, ³CEA, 41 rue Jules Horowitz 38027 Grenoble cedex 1, France.

†To whom correspondence should be addressed. Fax : 33 4 38 78 54 94 ; Phone : 33 4 38 78 31 19 ; E-mail : jean-michel.jault@ibs.fr

Abstract

EngA proteins form a unique family of bacterial GTPases with two GTP-binding domains in tandem, namely GD1 and GD2, followed by a KH (K-homology) domain. They have been shown to interact with the bacterial ribosome and to be involved in its biogenesis. Most prokaryotic EngA possess a high GTPase activity in contrast to eukaryotic GTPases that act mainly as molecular switches. Here, we have purified and characterized the GTPase activity of the *Bacillus subtilis* EngA and two shortened EngA variants that only contain GD1 or GD2-KH. Interestingly, the GTPase activity of GD1 alone is similar to that of the whole EngA, whereas GD2-KH has a 150-fold lower GTPase activity. At physiological concentration, potassium strongly stimulates the GTPase activity of each protein construct. Interestingly, it affects neither the affinities for nucleotides nor the monomeric status of EngA or the GD1 domain. Thus, potassium likely acts as a chemical GTPase-activating element as proposed for another bacterial GTPase like MnmE. However, unlike MnmE, potassium does not promote dimerization of EngA. In addition, we solved two crystal structures of full-length EngA. One of them contained for the first time a GTP-like analogue bound to GD2 while GD1 was free. Surprisingly, its overall fold was similar to a previously solved structure with GDP bound to both sites. Our data indicate that a significant structural change must occur upon K⁺ binding to GD2, and a comparison with *T. maritima* EngA and MnmE structures allowed us to propose a model explaining the chemical basis for the different GTPase activities of GD1 and GD2.

Keywords: EngA, GTPase, potassium activation, GMPPCP, 3D structure

† This work was supported by the ‘Agence Nationale de la Recherche’, grant N°: ANR-08-Blan-0143 to J.-M.J.

Introduction

1 GTPases form one of the most widely spread protein families in all kingdoms of life
2 and they play vital functions in many cellular processes, including translation, cellular
3 trafficking and cell division [1,2]. In eukaryotes, these enzymes have usually a very low
4 intrinsic GTPase activity which can be strongly stimulated, up to several orders of magnitude,
5 by additional protein partners known as GAP ('GTPase Activating Protein') [3]. Another
6 class of GTPases has been described in which the catalytic activity is promoted upon a
7 nucleotide-dependent dimerization [4]. In both cases, an optimal positioning of catalytic
8 residue(s) is thereby afforded, albeit using different ways. Some residues are brought in *trans*
9 by a protein partner or by the second subunit in case of homodimerization. Alternatively, the
10 catalytic residue(s) can be stabilized by a protein partner (or a second subunit in the case of a
11 dimer) when they are already present in an otherwise flexible part of the active site [3,4].

12 With the genome-sequencing era, the presence of previously uncharted GTPases
13 widely conserved in bacteria expanded our vision of this family [5]. Among them, nine
14 proteins related to the Era/Obg subfamily were identified and six turned out to be essential for
15 *Bacillus subtilis* growth [6], namely Bex* (Era), Obg (ObgE), YsxC (YihA), YphC (YfgK),
16 YlqF and YqeH (these two latter proteins have no homolog in *E. coli*). Although their
17 function was initially elusive, a wealth of biochemical data has been accumulated over the
18 past few years showing that Era [7,8], Obg [9,10,11], YsxC [12,13,14], YlqF [15,16] and
19 YqeH [17,18] are very likely involved in ribosome assembly, although some of them might
20 also participate in other cellular processes [19,20]. EngA (Essential neisserial GTPase) also
21 termed YphC or Der for 'double Era' [21,22], has been also shown to participate in the
22 ribosome biogenesis [13,23,24]. Furthermore, its encoding gene has been proposed to be part
23 of the minimal bacterial genome required to sustain life [25]. Intriguingly, *E. coli* EngA has
24 also been recently reported to interact with membrane and/or membrane proteins [26].

25 EngA is unique among GTPases because it bears two consecutive G domains, GD1
26 and GD2, followed by a C-terminal KH domain. Each G domain lacks the catalytic glutamine
27 found in many GTPases (e.g. Ras) and therefore EngA belongs to a class of GTPases termed
28 HAS-GTPases (Hydrophobic Amino acid Substituted for catalytic glutamine GTPases) [27].
29 Both G domains are required for interaction of EngA with the ribosomes *in vitro* [28], and for
30 the *in vivo* function of the protein [29]. Crystal structures of EngA from *Thermotoga maritima*
31 and *B. subtilis* with two phosphate ions bound to GD1 in the former and a GDP molecule
32 bound to GD1 in the latter revealed two different arrangements of the G domains. Based on
33 this, it has been suggested that a large conformational change might take place upon GTP
34 hydrolysis by GD1, leading to a different positioning of this domain relative to GD2 and KH
35 domains [30,31]. This conformational change has been proposed to play a key role for the
36 interaction of EngA with the ribosome.

37 Unlike most eukaryotic GTPases, which usually require additional proteins to
38 hydrolyze and to exchange nucleotides rapidly, several bacterial GTPases involved in
39 ribosome assembly show a relatively high intrinsic GTPase activity even in the absence of
40 additional protein partners [32,33]. Nevertheless, additional co-factors might be required for
41 efficient catalysis. Indeed, Inouye and colleagues have previously reported that potassium
42 stimulates the activity of the two bacterial GTPases MnmE and EngA from *T. Maritima*,
43 although the mechanism underlying this activation remained unknown at that time [21,34].
44 More recently, the chemical basis for this activation was unraveled for MnmE [35]. Potassium
45 promotes a GTP-dependent dimerization of MnmE leading to a stabilization of the transition
46 state for GTP hydrolysis. In fact, the 3-D structure of MnmE revealed that K⁺ mimics the
47 arginine finger provided in *trans* by a GAP (e.g. for GAP/Ras complexes), and K⁺ was
48 therefore proposed to act as a 'chemical' GAP [35]. A similar positioning and co-ordination
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 mode of the K⁺ ion was recently observed for the ferrous iron transporter, FeoB, indicating
2 the general character of this activation mechanism for these two GTPase families [36].

3 Here, we report the characterization of the *B. subtilis* EngA (YphC) and of its two
4 shortened EngA variants called GD1 and GD2-KH. Our results indicate that potassium
5 strongly stimulates the activity of both domains and congruently that of the whole enzyme.
6 Yet, the activity of GD1, either basal or potassium-activated, is always far higher than that of
7 GD2, and is similar to that of EngA. Accordingly, mutations targeting conserved motifs in
8 GD1 affect more drastically the GTPase activity of EngA than similar mutations introduced
9 into GD2. In contrast to MnmE, potassium-dependent activation is not associated to a
10 dimerization of EngA, and thus potassium presumably helps to ideally position the catalytic
11 residue(s) and/or the nucleotide thereby leading to a more efficient GTPase activity. We also
12 obtained two new crystal structures of *B. subtilis* EngA with different nucleotide occupancy
13 of each G-domain. One of them provides the first structure of a EngA GD1 in its apo form
14 while the second structure contains a GTP analogue bound to GD2. Comparison of our
15 structures with other structures of bacterial GTPases stimulated by potassium suggests that
16 K⁺-activated GTP hydrolysis by GD2 is associated with a significant structural reorganization
17 of its active site. Taken together, our data shed some light on the chemical basis of the high
18 GTPase activity of GD1 as opposed to GD2.
19
20
21
22

23 Results

24 Members of the EngA family are widely spread in the bacterial kingdom and a
25 sequence alignment of EngA proteins from different bacteria is shown in Figure S1. This
26 enzyme family is unique among GTPases as it bears two consecutive GTP-binding domains,
27 GD1 and GD2, each containing the four conserved signature sequence motifs typically found
28 in all GTPase proteins, G1 to G4, followed by a C-terminus KH domain (Fig. S1). A fifth
29 signature, G5, is less obvious to detect from the sequence inspection but can be easily
30 identified by structure-based alignments.
31

32 *EngA and GD1 possess a high GTPase activity, in contrast to GD2.*

33 In order to characterize the *B. subtilis* EngA and to understand the respective role of
34 its two GTP-binding domains, full-length EngA (named hereafter as EngA) and its two
35 shortened EngA variants, GD1 and GD2-KH were overexpressed in *E. coli* and purified to
36 near homogeneity (Fig. 1). A previous report on EngA from *Thermotoga maritima* had shown
37 that its GTPase activity was strongly stimulated, about 4-fold, by K⁺ [21]. Therefore, we
38 checked the effect of K⁺ on the GTPase activity of *B. subtilis* EngA, and test whether K⁺ was
39 able to affect the GTPase activity of either variant. For EngA, a slow GTPase activity was
40 measured in the absence of K⁺ ($k_{\text{cat}} \sim 0.3\text{-}0.5 \text{ min}^{-1}$) which was strongly stimulated by
41 addition of 300 mM potassium (Fig. S2). The GTPase activity of EngA increased to reach
42 approximately a 15-fold stimulation in the presence of 1 M KCl (Fig. 1A). A similar effect
43 was observed when K⁺ was replaced by rubidium, a slightly larger ion with similar chemical
44 properties that often substitutes efficiently for K⁺ in biological processes [37]. In contrast, the
45 use of the smaller sodium ion had no effect on the GTPase activity of EngA. GTPase
46 activities of shortened EngA variants were also studied (Fig 1B). In the absence of K⁺, GD1
47 has a GTPase activity similar to that obtained for EngA and about 150 times higher than
48 GD2-KH ($\sim 0.6 \text{ min}^{-1}$ vs 0.004 min^{-1}). Interestingly, both domains showed a strong
49 stimulation by K⁺, with a maximal GTPase activity for GD1 in the presence of 500 mM K⁺,
50 while that of GD2-KH increased steadily until the highest K⁺ concentration tested (1M). At
51 1M KCl, the GTPase activity of GD1 is $\sim 3.4 \text{ min}^{-1}$ and that of GD2-KH is $\sim 0.45 \text{ min}^{-1}$.
52
53
54
55
56

57 The kinetic parameters of GTP hydrolysis by *B. subtilis* EngA were further
58 investigated in the presence or absence of potassium. Rates of GTP hydrolysis were plotted
59 against increasing concentrations of GTP and, in both cases, the best fit of the data were
60
61
62
63
64
65

1 obtained using Henri-Michaelis-Menten equations (Fig. 2). The fitted parameters were $K_M =$
2 $75 \mu\text{M} \pm 9$, $V_{\text{max}} = 6.1 \pm 0.2 \text{ nmol/min/mg protein}$ ($k_{\text{cat}} \sim 0.3 \text{ min}^{-1}$) and $K_M = 153 \mu\text{M} \pm 11.6$,
3 $V_{\text{max}} = 144.5 \pm 3.4 \text{ nmol/min/mg protein}$ ($k_{\text{cat}} \sim 7.4 \text{ min}^{-1}$) in the absence or presence of 300
4 mM K^+ , respectively (Table S1). Thus, K^+ appears to impact much more on the catalytic
5 efficiency of EngA (20-fold increase) than on its K_M for GTP (2 fold increase). The K_M value
6 found here for the *B. subtilis* EngA agrees with those reported previously for EngA of other
7 species with values varying between 39, 110 and 143 μM for *M. smegmatis*, *T. maritima* and
8 *E. coli* enzymes, respectively [21,29,38]. Yet, the k_{cat} values measured at 400 mM K^+ were
9 quite lower for *T. maritima* (0.87 min^{-1}) and *E. coli* enzymes (1.16 min^{-1}) [21,29]. For *M.*
10 *smegmatis* EngA and in the absence of K^+ , a k_{cat} of 0.005 min^{-1} was reported [38]. A
11 comparison with other bacterial GTPases reveals that MnmE has a somewhat higher GTPase
12 activity than *B. subtilis* EngA in the presence of K^+ (k_{cat} values vary between 7.8 and 26 min^{-1})
13 while its activity without K^+ was similar to that reported here ($k_{\text{cat}} = 0.33 \text{ min}^{-1}$) [20], and
14 FeoB in the presence of potassium has a $k_{\text{cat}} = 0.258 \text{ min}^{-1}$ [39]. For the other bacterial
15 enzymes namely Era, Obg, YihA, YchF and HflX they all had a lower or even a much lower
16 GTPase activity than *B. subtilis* EngA, but some of them might require K^+ to be fully active
17 [20].

18 *EngA, GD1 and GD2 have similar affinities for nucleotides and K^+ does not affect them.*

19 To confirm the marginal effect of K^+ on the affinity of *B. subtilis* EngA for
20 nucleotides, fluorescence experiments were performed. *B. subtilis* EngA contains two Trp
21 residues in its sequence (Fig. S1 and S3) and, upon excitation at 295 nm, the protein displayed
22 a classical Trp fluorescence spectra (Fig. S4). As shown in Figure 3, addition of increasing
23 GDP concentrations produced an increase in the intrinsic Trp fluorescence of EngA, which
24 was more pronounced in the absence of K^+ . As observed for the kinetic experiments, a single
25 class of binding sites was sufficient to obtain a reliable fit of the data, whether or not K^+ was
26 present. The two K_D values were in the same range: $8.0 \pm 1.0 \mu\text{M}$ in the absence of K^+ and
27 $16.3 \pm 2.3 \mu\text{M}$ in the presence of K^+ (Table S2). Affinities for GMPPNP, a non-hydrolysable
28 GTP analogue, or GDP plus AlFx, a powerful inhibitor of GTPases [40], were also measured
29 by the same technique, and no major differences were found when K^+ was added (Table S2).
30 Fluorescence resonance energy transfer was also analyzed between the tryptophan residues
31 and the MANT moiety of the MANT-GDP, and again, the presence of K^+ did not affect the
32 K_D of EngA for this nucleotide analogue (Table S2).

33 Given the presence of two potential GTP-binding sites in EngA, it was rather
34 unexpected to detect only one class of binding sites for each nucleotide by fluorescence
35 measurements. Possible explanations for this are (i) GD1 and GD2 have quite different
36 affinities for one type of nucleotide and one binding site escapes the detection limits because
37 a limited range of nucleotide concentration could only be used here ($< \text{mM}$), (ii) both domains
38 have similar affinities for a given class of nucleotide, or (iii) the modification of fluorescence
39 is due to nucleotide binding to only one of the two sites (i.e. binding of nucleotide to the
40 second site does not produce a significant modification of Trp fluorescence). To answer this
41 question, the same approach was applied to isolated shortened EngA variants, GD1 and GD2-
42 KH, as they each contain a single Trp residue (Fig. S3) and they both displayed a classical
43 Trp fluorescence spectra (Fig. S4). Upon nucleotide binding, each domain showed a marked
44 increase in Trp fluorescence indicating that the conformation and/or environment of each Trp
45 residue is affected by this process. Therefore, these two Trp residues are ideal reporter probes
46 for this binding step. As shown in Figure 4, in the presence of 300 mM NaCl GD1 binds GDP
47 and GMPPNP with K_D values of $2.7 \pm 0.4 \mu\text{M}$ and $17.8 \pm 2.4 \mu\text{M}$, respectively (Table S2). In
48 the presence of 300 mM KCl, these values were slightly increased to $4.5 \pm 0.9 \mu\text{M}$ and $35.7 \pm$
49 $4.9 \mu\text{M}$ for GDP and GMPPNP, respectively. For GD2-KH, similar values were obtained for
50 the two nucleotides with K_D values of $1.6 \pm 0.2 \mu\text{M}$ ($2.9 \pm 0.8 \mu\text{M}$ in the presence of K^+) and
51

1 15.2 ± 1.3 μM (6.8 ± 2.0 μM in the presence of K⁺) for GDP and GMPPNP, respectively
2 (Fig. 4 and Table S2). Therefore, the presence of K⁺ does not impact significantly on the
3 affinity of EngA, GD1 or GD2-KH for the different nucleotides and, overall, the shortened
4 EngA variants bind nucleotides with affinities in the same range than the whole enzyme.

5 To further assess the binding of nucleotides to EngA and to its isolated domains, the
6 Thermal Shift Assay (TSA) was also used. In this technique, above a given temperature
7 threshold, the protein starts to unfold thus exposing its hydrophobic patches to the solvent.
8 These newly exposed areas bind the fluorescent probe thereby boosting its fluorescence
9 intensity. Consequently, the unfolding process can be monitored by the fluorescence increase
10 as a function of temperature, with the mid-point between the folded and unfolded protein
11 states defining the melting temperature, *T_m* [41,42]. This technique can also be used to
12 monitor the protein stability induced by various effectors [41,43]. In the absence of
13 nucleotide, EngA had a *T_m* of 41.5 °C, and this value increased to 48.8 °C and 59.4 °C in the
14 presence of 1 mM of GMPPNP or GDP, respectively (Fig. 5A and Table S3). Likewise, apo
15 GD1 and GD2-KH which had a *T_m* of 51 °C and 42.8 °C, respectively, were significantly
16 protected against thermal denaturation by either GMPPNP (~ 4.5 °C *T_m* increase for both
17 domains) or GDP (~ 13 °C *T_m* increase for both domains; Fig. 5B, 5C and Table S3). GDP
18 appears then to be much more efficient than GMPPNP to increase the thermal stability of
19 each protein construct, consistent with a seemingly tighter GDP binding, as observed for other
20 bacterial GTPases.
21

22 *K⁺ does not promote the dimerization of EngA or GD1.*

23
24 GDP-Al-F_x and GTP were previously shown to trigger the dimerization of the two
25 GTPase domains of MnmE in a potassium dependant manner [35]. Therefore, we checked
26 whether K⁺ might act in the same manner for EngA. First, gel filtrations were performed in
27 the presence, or not, of potassium and different nucleotides: GMPPNP, GDP and GDP-Mg-
28 AlF_x; this latter complex is a transition state analogue [40] that was mandatory to stabilize the
29 interaction between the two GTPase domains of MnmE in the presence of potassium [35].
30 Regardless of the conditions used, EngA always eluted from the gel filtration with an
31 estimated molecular weight corresponding to a monomer (Fig. S5). Yet, transient dimers
32 might not have withstood gel filtration so analytical ultracentrifugation experiments were also
33 performed. Clearly, the major species obtained (~ 90%) in the presence of GDP-Mg-AlF_x plus
34 K⁺ for either EngA (at 3.67 S) or GD1 (at 2.09 S) corresponded to a monomer (Fig. 6).
35 Predicted values are 3.67 S and 2.05 S, respectively, for globular monomers *versus* 5.82 S and
36 3.25 S for dimers. Importantly, this result was unaffected by varying the protein concentration
37 between 1.4 and 4.7 mg/ml for EngA, and between 1.2 and 4.5 mg/ml for GD1. Similar
38 results were also obtained in the presence of K⁺ alone, or K⁺ and GDP or GMPPNP (not
39 shown). Therefore, in contrast to MnmE, potassium does not seem to promote the
40 dimerization of EngA or GD1 alone.
41

42 *Equivalent mutations are more drastic when they target GD1 rather than GD2.*

43 To further address the role of GD1 and GD2 in the context of the whole protein,
44 identical mutations were introduced into the G1 or G4 motifs of either G domain (Fig. S1). In
45 the G1 motifs of EngA (GX₄GKS), the two invariant lysine residues are known to coordinate
46 the phosphate of nucleotide and mutation of this residue to an alanine into GD1, as opposed to
47 GD2, strongly reduced the GTPase activity of *E. coli* EngA [29]. Concerning the G4 motif
48 (NKXD), mutation of the invariant aspartate to an asparagine was previously shown to switch
49 the specificity in other GTPases, e.g. rasP21, EF-Tu or FtsY [44,45,46], thereby favoring the
50 binding of xanthine nucleotide over guanine nucleotide. Here, mutation of the conserved Lys
51 residue to Ala in the G1 motif of GD1 (K16A) drastically altered the GTPase activity of
52 EngA while the equivalent mutation in GD2 (K188A) reduced the activity to ~ 50% (Fig.
53 7A). The mutation Asp to Asn in the G4 motif decreased significantly the GTPase (to ~ 40%)
54
55
56
57
58
59
60
61
62
63
64
65

1 when introduced into GD1 (D122N) while it modestly affected the activity (to ~ 80%) when
2 introduced into GD2 (D298N). Because these mutations were expected to modify the
3 nucleotide specificity, their possible gain-of-function for the XTPase activity was also
4 studied. At 200 μ M nucleotide concentration, the GTPase activity of the wild-type EngA was
5 about three times higher than its XTPase activity (Fig. 7B). The same trend was observed for
6 the D298N mutant although less pronounced than for the wild-type. Importantly, a reverse
7 result was found with the D122N mutant with a XTPase activity more than two times higher
8 than its GTPase activity. Overall, these results suggest that in EngA, GD1 has a much higher
9 GTPase or XTPase activity than GD2 because (i) equivalent mutations have a more drastic
10 effect, especially for the lysine residue of the G1 motif, when introduced into GD1 than into
11 GD2 and (ii) a higher XTPase than GTPase activity was obtained only when the mutation
12 switching the nucleotide specificity was introduced into GD1. However, the 50% reduced
13 GTPase activity for the EngA K188A mutant is at odds with the activity of the GD1 domain
14 that is estimated to be ~ 40 times higher than that of the GD2 in the conditions used for the
15 assay. Therefore, in the context of the whole enzyme, mutations in GD2 somehow also impact
16 the activity of the GD1 domain.

17
18
19 *Structure of B. subtilis EngA with GD1 in its apo form or with GMPPCP in GD2.*

20
21 Previously, the structure of *B. subtilis* EngA was solved with one GDP molecule in
22 each G domain [30] (PDB code: 2HJG). Here, attempts to crystallize EngA or its isolated
23 domains in different states were made. While we have not been able to obtain diffracting
24 crystals for either GD1 or GD2-KH so far, suitable crystals of EngA were obtained in two
25 different conditions. The first structure, EngA1, was obtained in the presence of ammonium
26 sulfate and GDP and was refined to a resolution of 2.25 Å. The second structure, EngA2, was
27 obtained without ammonium sulfate and in the presence of GMPPCP, and refined to a
28 resolution of 2.6 Å. The two structures were obtained from crystals belonging to the same
29 space group and with similar unit cell dimensions and crystal packing, when compared to the
30 2HJG entry (Table S4). Both structures had unambiguous electron density for the whole
31 polypeptide chain, except for residues 30-40 (switch I of GD1), 60-67 (part of GD1 switch
32 II), 122-131 (part of the GD1 G4 motif) and 206-211 (switch I of GD2). These regions were
33 also disordered in the *B. subtilis* EngA structure determined previously [30]. The content of
34 nucleotide-binding sites was also well defined in the electron density maps. The EngA1
35 structure contains one sulfate ion in GD1 and one GDP in GD2. In the EngA2 structure, the
36 GD2 nucleotide-binding site is fully occupied by a GMPPCP molecule whereas the GD1
37 nucleotide-binding site remains empty. Despite the different nature of the molecule bound to
38 the two G domains, the overall structures obtained are very similar to that previously
39 published [30], with r.m.s differences of 0.42 Å (379 Ca atoms) and 0.31 Å (378 Ca atoms)
40 for EngA1 and EngA2, respectively. When the GD1 in EngA1 or EngA2 are superposed to
41 the GD1 in 2HJG, the position of the GD2 and the KH domains do not differ by more than 0.9
42 Å in translation and 1.3° in rotation. This clearly indicates that no major structural change
43 occurred in its global arrangement (Fig S3). The conformation of the GD1 in either its apo-
44 (EngA2) or GDP-bound forms (2HJG) are nearly identical (r.m.s. difference of 0.3 Å) and the
45 presence of a sulfate ion that takes the place of the β -phosphate (EngA1) results in a GD1
46 conformation also very similar to the diphosphate nucleotide bound state. (Fig. S6A).
47 However and quite unexpectedly, the presence of a GTP analogue bound to the GD2 in
48 EngA2 structure did not lead to any significant conformational changes (r.m.s. difference of
49 0.25 Å). A 1 Å shift in the switch II region (residues 236-242) is the only significant
50 structural difference, and is likely induced by the presence of the GMPPCP γ phosphate close
51 to it (Fig. S6B). Apparently, the binding of GMPPCP is not sufficient to trigger the expected
52 conformational change in the GD2 switch regions. The crystallization conditions and/or the
53 crystal packing may be partly responsible for maintaining the *B. subtilis* EngA in a GDP-
54
55
56
57
58
59
60
61
62
63
64
65

1 bound like state. The EngA1 and EngA2 structures also suggest that the GD2 binds either
2 GDP or GMPPCP more efficiently than the GD1, corroborating observation made for the *T.*
3 *maritima* EngA structure where a GDP molecule was found in G2 while no nucleotide was
4 added during the purification and the crystallization of the enzyme [31]. However, as
5 indicated above, the GD1 and GD2 affinity for GDP and GTP are similar in solution. Thus,
6 this apparent inconsistency may come from different kinetics of binding or from the fact that
7 the crystallization process may preferentially select EngA molecules that contain a nucleotide
8 bound to GD2. Attempts to obtain an EngA structure with either a bound Mg^{2+} or a bound K^+
9 ion in the presence of different nucleotides led to nearly identical structures than those
10 reported here with no identifiable electron density for Mg^{2+} or K^+ ions. These ions were also
11 not reported in the two other EngA structure previously published, despite the presence of
12 Mg^{2+} [30] or K^+ [31] in the crystallization conditions.
13
14

15 Discussion

16 Many enzymes often require co-factors for an optimal functioning and monovalent cations,
17 especially potassium, have long been known to be required for a wide variety of biochemical
18 pathways [47]. Early on, these monovalent cations were foreseen to form a ternary complex
19 with the enzyme and substrate [48]. This has been recently demonstrated for one bacterial
20 GTPase, MnmE, where K^+ elegantly substituted for an arginine finger that is usually provided
21 in *trans* by a GAP in many eukaryotic GTPases [35]. MnmE works as a homodimer and K^+ in
22 the presence of GTP (or GDP/AlFx) promotes the dimerization of the two GTP-binding
23 domains provided by each monomer. For EngA, we found no evidence of K^+ -promoted
24 dimerization but the idiosyncratic presence of two GTP-binding domains within a single
25 polypeptide chain makes this enzyme more complicated to study. Thus, to elucidate the role
26 of potassium, we have studied GD1 and GD2-KH separately. Clearly, K^+ strongly stimulates
27 the GTPase activity of either protein construct showing that it binds independently to both
28 GD1 and GD2. While our study was in progress, three other families of bacterial GTPases
29 were shown to be also stimulated by potassium: the circularly permuted GTPase
30 (cpGTPase) and RbgA, both being involved in ribosome biogenesis [49,50], and the G
31 protein-coupled ferrous iron transporter B, FeoB [39]. Interestingly, K^+ did not seem to
32 promote the dimerization of cpGTPase [49], while the oligomerization state remains unclear
33 for FeoB [51] and this property was not investigated for RbgA [50]. Therefore, our data show
34 that K^+ is a genuine EngA GTPase-activating element sufficient for activating the GTP
35 hydrolysis of both GD1 and GD2 domains independently of a dimerization process. All K^+ -
36 sensitive bacterial GTPases studied so far including EngA require few hundreds millimolar
37 potassium to be fully activated and this is consistent with the intracellular concentrations
38 found in *E. coli* (~ 200 mM) or in *Bacillus subtilis* (~ 400 mM) [52,53]. Apart from
39 cpGTPase, combined structural and phylogenetic analyses have suggested that several
40 members of the TrmE-Era-EngA-Septin like superfamily [54] are possibly K^+ -sensitive
41 GTPases [20,35,39]. Importantly, the comparison of sequences and structures of members of
42 this superfamily identifies two conserved asparagine residues involved in the binding of
43 potassium (see Fig. 1 in [39]). The first asparagine located within the G1 motif
44 (GXXNXGKT/S) makes a direct contact with K^+ in MnmE and FeoB [35,36]. The second
45 asparagine located 5 residues downstream of the conserved lysine residue in the G1 motif
46 makes two hydrogen bonds with the switch I motif thereby favoring the stabilization of a K-
47 loop binding site [36], and this K-loop was only observed in the GTP-bound conformation of
48 FeoB [39]. Although cpGTPases have a unique topology among GTPases, YqeH does possess
49 these two conserved asparagine residues at the expected position in its sequence (see Fig. 1 in
50 [49]). In EngA, the switch I and therefore the putative K-loop is disordered in the crystal
51 structures available so far, even in our structure of *B. subtilis* EngA with a GMPPCP bound to
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 GD2. Yet, the strict conservation of these two asparagine residues in either GD1 or GD2 (see
2 Fig. S1) combined with our observation that the GTPase activity of each G-domain of EngA
3 is truly activated by K^+ , suggest that this activation occurs through a mechanism similar to
4 that described for MnmE and FeoB, with each switch I of EngA adopting a K-loop
5 conformation when GTP is bound to the two G domains.

6 In a previous report, the GTPase activities of GD1 and GD2 moieties of EngA from *T.*
7 *maritima* have been investigated [31]. In agreement with our study, the GTPase activity of *T.*
8 *maritima* GD1 approximated to that of the full-length protein. However, *T. maritima* GD2-
9 KH retained about one third of the GTPase activity of the full-length enzyme. This contrasts
10 with the *B. subtilis* GD2-KH that has only a marginal GTPase activity but this difference
11 might be attributable to the intrinsically low GTPase activity of the thermostable enzyme
12 (0.11 min^{-1} in the presence of 0.4 M KCl vs $\sim 5 \text{ min}^{-1}$ for *B. subtilis* EngA at the same salt
13 concentration). Remarkably, the activity of *B. subtilis* EngA without K^+ is similar to that of *T.*
14 *maritima* at 400 mM KCl. For the *B. subtilis* GD2-KH domain, the activity without K^+ is very
15 low and similar to that of human Ras [55]. In the presence of 300 mM K^+ , it reaches 0.11 min^{-1}
16, a value comparable to that of full-length *T. maritima* EngA or some bacterial GTPases, e. g.
17 Era or Obg [20]. Shortened EngA variants have been also used to decipher the multiple
18 interactions of EngA with ribosomes [28,38], and it has been shown for instance that the KH
19 domain alone can bind to the 30S subunit while GD2-KH in the presence of GMPPNP binds
20 preferentially to the 50S subunit [28]. For the whole EngA, the interactions with the
21 ribosomal subunits seem to depend on the nature of the nucleotide bound to each G domain
22 [28] (see below).

23 Since the affinity for nucleotides may impact the GTPase activity, we have
24 characterized the binding abilities of GD1 and GD2 for different nucleotides. In a previous
25 study, the binding affinity of *S. typhimurium* EngA for nucleotides was studied using a
26 microcalorimetry approach [56]. While the two G-domains were found to have a similar
27 affinity for GDP ($K_D \sim 1\text{-}3 \mu\text{M}$), they exhibited a quite dissimilar affinities for GTP: one with
28 a K_D of $\sim 12 \mu\text{M}$ and the other with a $K_D > 100 \mu\text{M}$. Lamb and colleagues proposed that the
29 binding site with the highest affinity reflected binding to GD1 whereas GD2 would bind GTP
30 with a much lower affinity, and that this low affinity of GD2 for GTP might explain the small
31 participation of GD2 in the overall GTPase activity of the enzyme [56]. Here, in the presence
32 of Mg^{2+} , we found rather similar affinities for GD1 and GD2 for either GDP or GMPPNP.
33 Therefore, the low GTPase activity of GD2 is clearly not due to a poor binding ability of GTP
34 in the presence of Mg^{2+} . Rather, this must reflect a different catalytic proficiency of the two
35 G-domains.

36 Because ultracentrifugation analysis showed that K^+ did not promote the dimerization
37 of GD1, residues within this domain must be self-sufficient to carry out GTP hydrolysis, and
38 in a much more efficient way than in GD2. In MnmE, a glutamate residue in the switch II
39 motif (Glu282, see Fig 8A) is poised to stabilize a nucleophilic water molecule via a bridging
40 water molecule, and this residue has been proposed to act as a general base to hydrolyze GTP
41 [35]. In *B. subtilis* GD1 structures, this switch II is partially disordered but it is well resolved
42 in GD1 of *T. maritima* where the Asp61 points towards a phosphate bound in the active site.
43 Interestingly, although Asp61 from *T. maritima* EngA and Glu282 from MnmE are not
44 superimposed, they are located on each side of the nucleotide and at a similar distance from
45 the position of the γ -phosphate of GTP (Fig. 8A). Asp61 is not strictly conserved in GD1 from
46 different species but acidic residues (e.g. Asp62, Asp65 and Glu66 in *B. subtilis*, all being
47 disordered in *B. subtilis* structures) are always found in the switch II of GD1 in the EngA
48 family (Fig. 8A). This is in sharp contrast with the switch II of GD2 that contains no acidic
49 residues (Fig. 8B). Interestingly, in the case of FeoB no catalytic base could be identified, and
50 instead, it has been proposed that the attacking water molecule is aligned only by the protein
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 backbone [36]. Thus, due to the different nature of the switch II in GD1 and GD2, a different
2 mechanism might take place in each G-domain to activate a water molecule leading to GTP
3 hydrolysis.

4 In the switch II of GD2, Lys236 has been shown to contact both the α - and β -
5 phosphate of GDP [30]. In fact, the NZ group of this lysine also occupies the same position
6 than the arginine finger guanidinium group in the Ras-RasGAP complex structure (1wq1 pdb
7 entry, [57]) or the K^+ ion in the MnmE structure [35] (Fig. S7). Assuming that K^+ will bind to
8 GD2 in the same way as in MnmE and FeoB, the side chain of this lysine must be reoriented
9 to leave a free access to K^+ . Moreover, the whole switch II region of GD2 should move away
10 in order to allow the G2 Thr210 to interact with the Mg^{2+} ion and permit the switch I to adopt
11 the K-loop conformation. Last, as also observed in *T. maritima* EngA, the aspartic acid
12 residue of the G3 motif (Asp229 for *B. subtilis*), is not pointing towards the nucleotide, but
13 towards helix $\alpha 8$ instead, more than 10 Å away from the Mg^{2+} ion location (Fig. 8B). Overall,
14 our data suggest that the GD2 domain may rely on two different GTP hydrolytic mechanisms
15 *in vitro*: a low intrinsic activity without potassium due to an incomplete active site lacking the
16 contribution of genuine G2 and G3 motifs, highlighted by the different available *B. subtilis*
17 EngA GD2 structures, and a ~ 100 -fold increased GTPase activity upon K^+ binding thereby
18 leading to the proper positioning of the switch I and II regions.

19 Unexpectedly, we found that the binding of a GTP-like analogue (GMPPCP) to *B.*
20 *subtilis* EngA GD2 led to a conformation very similar to that of the GDP-bound state. This is
21 in sharp contrast with the major domain rearrangement found between the GTP-like
22 conformation and GDP-bound states of GD1 [30]. At first glance, one could conclude that
23 there is no substantial conformational change associated with the GTP bound state of the GD2
24 domain. However, the binding of a GTP molecule or an analogue is not always sufficient for
25 switching the GTPase in its “on” conformation, e.g. for FeoB [58,59]. As observed for Ras
26 and MnmE, the presence of a GAP (Ras-GAP for ras or K^+ for MnmE) is required for
27 obtaining the activated conformation of the G-domain. Since our structure was obtained
28 without K^+ , it may still retain the biological “off” conformation and presumably represents the
29 GTP bound state in a low activity state.

30 Although the two GTP-binding domains of EngA are not equivalent in term of
31 GTPase activity, they were reported to be both essential *in vivo*. Indeed, identical mutations
32 targeting the conserved serine residue in the G1 motif and introduced in either GD1 or GD2
33 of *E. coli* EngA both resulted in a severe growth defect [29]. Unexpectedly, the N to D
34 mutation introduced separately in the NKXD G4 motif of either domain, led to a different
35 result. At low temperature (30°C), both functional GD1 and GD2 domains were required for
36 bacterial growth while at higher temperature (42°C), either one of the two functional domains
37 was dispensable [23]. Hence, the authors of this study concluded that with only a single
38 functional domain, either GD1 or GD2, *E. coli* growth was maintained at high temperature.
39 This suggests that G1 and G2 can work in an independent manner, and this is consistent with
40 the fact that we, and others, have been unable to observe any cooperativity in GTP hydrolysis
41 [21,29]. Recently, a model was put forward to explain how the nature of the nucleotide bound
42 to each G-domain might modulate the interaction with ribosomes [28] (see also [38] for a
43 slightly modified model). In the original model, although GD1 and GD2 can work
44 independently, they act *in vivo* in a concerted manner to bind to either the 50S subunit, the
45 30S subunit or the whole ribosome [28]. Congruent with this, some communication between
46 the two G-domains is likely to occur because mutation of the conserved lysine in the G1 motif
47 of GD2, and to a lesser extent the mutation of the G4 motif in GD2, were found here to
48 significantly impact the overall rate of GTPase activity of the *B. subtilis* EngA. Likewise, the
49 mutation in *E. coli* EngA of the invariant lysine or serine in the G1 motif of GD2 led to a
50 lower GTPase activity (<50%) with a dramatically reduced K_M for GTP [29]. Thus, additional
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

work will clearly be needed to unravel the mechanism by which these two sites talk to each other.

In conclusion, we have shown that each G domain of *B. subtilis* EngA is able to bind K⁺ and their respective GTPase activity is strongly activated by this cation. In the whole enzyme, the GTPase activity seems to primarily originate from the G1 domain although both domains need to be in a functional state (i.e. not mutated) to get a fully active EngA. Crystal structure of EngA with an apo G1 and a GDP bound to G2 has a global conformation similar to that previously published for *B. subtilis* EngA with GDP in its two G domains. More surprisingly, even when a GTP analogue is bound to G2, the same overall structure is found, in contrast to the *T. maritima* EngA which presumably reflects the GTP bound state of G1. Additional structures and further biochemical characterization of this very unusual family of bacterial GTPases will be required to get a full understanding of its functioning mechanism.

Footnote: *The *B. subtilis* nomenclature is used here with the *E. coli* name given in parentheses.

Materials and Methods

Cloning and site-directed mutagenesis. *YphC (engA)* was cloned into a pET15b vector and overexpressed as previously described [13]. GD1 and GD2-KH were amplified by PCR using the following oligonucleotides:

5'-GGGCATATGGGTAAACCTGTCGTAGCCATTGTC-3' and

5'-GCTGAGCAATGTTTTTAAAATGCTCTGCAAC-3' primers for GD1 and

5'-CATATGGAAGTTATTCAATTCTGTCTGATC-3' and

5'-GGGGCTCAGCTTATTTTCTAGCTCTTGCAAATAT-3' for GD2-KH as 5' and 3' primers, respectively, and cloned into a pCR Blunt vector (Invitrogen). The gene was then sub-cloned into a pET15b (Novagen) expression vector using NdeI and BlnI restriction sites for insertion. Mutants were generated by the Quickchange protocol (Stratagene) with the pET15b-*engA* used as a template and the following oligonucleotides bearing the desired mutations and restriction sites, when indicated, for restriction screening.

K16A : 5'-CGGGAGACCAAATGTAGGAGCTAGCACAATCTTTAACCGG-3', introducing a NheI restriction site,

K16A-rev : 5'-CCGGTTAAAGATTGTGCTAGCTCCTACATTTGGTCTCCCG-3',

N119D : 5'-GCCTGTTGTTTTAGCGGTTGATAAGCTTGATAACACAGAAATG-3', introducing a HindIII restriction site,

N119D-rev : 5'-CATTTCTGTGTTATCAAGCTTATCAACCGCTAAAACAACAGGC-3',

D122N : 5'-GCCTGTTGTTTTAGCGGTTAATAAGCTTAATAACACAGAAATG-3', introducing a HindIII restriction site,

D122N-rev : 5'-CATTTCTGTGTTATTAAGCTTATTAACCGCTAAAACAACAGGC-3',

K188A : 5'-GACGTCCAAATGTCGGAGCTAGCTCACTTGTGAATGCG-3', introducing a NheI restriction site

K188A-rev : 5'-CGCATTCAACAAGTGAGCTAGCTCCGACATTTGGACGTC-3',

D297N : 5'-GTCGTAAACAAATGGAATGCTGTTGACAAAGATGAGAGC-3', no restriction sites was either introduced or deleted,

D297N-rev : 5'-GCTCTCATCTTTGTCAACAGCATTCCATTTGTTTACGACG-3'.

The DNA template was then digested using DpnI enzyme and amplified vectors were transformed into the *E. coli* XL1 blue strain. Positive clones were selected on LB-agar plates supplemented with ampicillin, and verified after plasmid preparation by restriction mapping and DNA sequencing (Genome express).

1 *Protein overexpression and purification.* Shortened EngA variants and single point mutants
2 were overexpressed and purified as for the wild type. C41(DE3) *E. coli* strain transformed
3 with the vector of interest were grown until the optical density at 600 nm reached a value of
4 0.6-0.8, then 1 mM isopropyl- β -thiogalactopyranoside (Euromedex) was added. After
5 overexpression for 4 hours at 37°C, cells were harvested by a low speed centrifugation (4,000
6 g for 10 min), resuspended in a lysis buffer (50 mM Tris-HCl pH 8, 250 mM NaCl, 10 mM
7 imidazole, 1 mM PMSF, 5 μ M leupeptine and 5 μ M pepstatine A) and lysed using a French
8 Press at 16,000 psi. The lysate was then centrifuged for 30 min at 20,000 g at 4°C to remove
9 cell debris. The supernatant was applied onto a Ni-IDA agarose column equilibrated with the
10 lysis buffer. The resin was washed with 10 column volumes using 50 mM Tris-HCl pH 8, 250
11 mM NaCl, 2.5 mM β -mercaptoethanol, 20 mM imidazole and 10% glycerol, and the protein
12 was eluted using the same buffer except that 250 mM imidazole was used. EngA was further
13 purified by size exclusion chromatography using a Superdex G75 16/60 (GE Healthcare)
14 equilibrated in 20 mM Tris-HCl pH 8, 250 mM NaCl, 2.5 mM β -mercaptoethanol. Fractions
15 containing the protein of interest were pooled and 10% glycerol was added before a flash-
16 frozen step in liquid nitrogen. Purified proteins were stored at -80°C. Before use, the protein
17 samples were thawed, applied onto a PD10 desalting column (GE healthcare) and
18 concentrated using a vivaspin filter with a molecular weight cut-off of 30 kDa. The protein
19 concentration was calculated using an extinction coefficient at 280 nm of 33350 M⁻¹cm⁻¹ for
20 EngA, 17420 for GD2-KH and 14440 for GD1, as obtained from the ProtParam tool software
21 (<http://web.expasy.org/protparam/>).
22
23
24
25
26

27 *GTP hydrolysis assay.* GTPase activities were determined using a coupled-enzymatic assay
28 containing pyruvate kinase and lactate dehydrogenase as previously described [60]. The
29 proteins (2 μ M or 8 μ M of EngA in the presence of KCl or NaCl, respectively, and 4 μ M for
30 GD1 and 12.5 μ M for GD2-KH) and the coupled assay system containing 50 mM Tris/HCl
31 pH 8, 4 mM phosphoenol pyruvate, 0.4 mM NADH, 1 mM MgCl₂, 30 μ g pyruvate kinase and
32 15 μ g lactate dehydrogenase, were mixed into a quartz cuvette (final volume of 750 μ l) in the
33 presence of GTP (up to 1 mM) and salts, either KCl or NaCl, as indicated in the figure
34 legends. When the GTPase and XTPase activities of the wild-type EngA and the single point
35 mutants were studied (Fig. 7), the assay medium (final volume of 500 μ l) contained 3 mM
36 MgCl₂ and the pyruvate kinase was increased to 120 μ g to ensure an immediate conversion of
37 XDP back into XTP. The reaction was started by adding 3.2 μ M or 0.8 μ M of wild-type
38 EngA or single point mutants in the absence or presence of 300 mM K⁺, respectively. GTP or
39 XTP hydrolysis was monitored by following the change in absorbance at 340 nm upon
40 conversion of NADH to NAD⁺, for 10 min at 37°C, using either a Hitachi U3350 or a Safas
41 UVmc² spectrophotometer. The slope measured (Δ DO/min) was used to calculate the activity
42 of the protein taking into account the $\epsilon_{\text{NADH}_{340\text{nm}}} = 6220 \text{ l}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$. Control experiments
43 were systematically performed in the absence of nucleotides (protein alone) or in the absence
44 of proteins (nucleotides alone) and the activities measured in all cases were negligible.
45 Between many different EngA preparations, the specific activities measured at 1 mM GTP
46 varied between 100-130 nmoles/min/mg protein.
47
48
49
50
51
52

53 *Fluorescence spectroscopy.* Fluorescence measurements were performed essentially as
54 described previously for the unmodified nucleotides or the *N*-methylantraniloyl-GDP
55 (MANT-GDP) analogue using a PTI QuantaMaster 4 spectrofluorimeter [61,62]. Dissociation
56 constants (K_D) for nucleotides were calculated by monitoring intrinsic tryptophan
57 fluorescence variation using 295 nm excitation wavelength, with a slit width of 2 nm, and
58 scanning emission fluorescence between 310 and 400 nm (for the unmodified nucleotide) or
59
60
61
62
63
64
65

1 400 and 500 nm (for MANT-GDP), with a slit width of 4 nm. For the intrinsic tryptophan
2 fluorescence, inner filter effect was corrected using *N*-Acetyltryptophanamide (NATA) as a
3 reference. 1 μ M NATA or 0.5 μ M of YphC were mixed with 50 mM Tris/HCl pH 8, 2 mM
4 MgCl₂, NaCl or KCl and increasing amount of nucleotide (GDP, GMPPNP or MANT-GDP),
5 as indicated in the figure legends. The fluorescence was normalized respective to the
6 integrated value of the fluorescence peak obtained in the absence of unmodified nucleotide
7 (between 310 and 400 nm) or MANT-GDP (between 400 and 500 nm). The data were fitted
8 using the GraFit 5.0.11 software with the following equation:
9

$$10 \quad F = F_{\min} + \{(F_{\max} - F_{\min})[(E + L + K_D) - ((E + L + K_D)^2 - 4EL)^{1/2}]\}/2E$$

11 where F is the relative fluorescence intensity in the presence of nucleotide, F_{\min} is the relative
12 fluorescence in the absence of nucleotide, F_{\max} is the fluorescence intensity at saturating
13 concentration of nucleotide, L is the nucleotide concentration, E is the protein concentration
14 and K_D is the dissociation constant of the protein-nucleotide complex [63].
15
16
17
18

19 *Fluorescence-based Thermal Shift Assay (TSA)*. Assays were performed into a 96-well thin-
20 wall PCR plate (BioRad) using an IQ5 96-well format real-time PCR instrument (Bio-Rad)
21 over a temperature range from 20°C to 90°C with temperature increments of 0.2°C [64].
22 Assay samples (23 μ l) contained 50 mM Tris/HCl pH 8, 0.5 mg/ml of EngA or 1.5 mg/ml of
23 GD1 or GD2-KH, and concentration of salt and/or nucleotides as indicated, and 2 μ l of
24 SYPRO Orange solution (Molecular Probes, Eugene, OR, 500x in DMSO diluted 5 times in
25 water) was added. The plates were heated and changes in fluorescence for the
26 environmentally sensitive dye SYPRO Orange were monitored through a charge-coupled
27 device (CCD) camera [42]. The melting temperature (T_m) values for each protein were
28 determined graphically from the fluorescence first derivative of the melting curves. The
29 wavelengths for excitation and emission were 470 nm and 570 nm, respectively.
30
31
32
33

34 *Ultracentrifugation analysis*. Sedimentation velocity experiments were performed as
35 previously described using a Beckman XL-I analytical ultracentrifuge equipped with an AnTi
36 rotor [65]. Samples and buffers (400 μ l and 420 μ l, respectively) were loaded into their
37 respective channels in double-sector ultracentrifuge cells and run at 45 000 r.p.m. at 4°C.
38 Scans were recorded at 280 nm. The SEDFIT program continuous distribution $c(s)$ analysis
39 (Sedfit: <http://www.analyticalultracentrifugation.com>) [66,67] was used to fit the data by
40 generating the sedimentation distribution profiles. Mass identification of each peak was
41 carried out considering EngA and GD1 as globular proteins, with Stokes radius R_h being 1.25
42 times the anhydrous radius calculated from chemical composition [68].
43
44
45

46 *Crystallization and structure determination*. Prior to crystallization, samples of EngA were
47 applied onto a size exclusion chromatography column (HiLoad 200 16/60, GE healthcare)
48 equilibrated in 20 mM Tris pH 8, 250 mM NaCl, 2.5 mM β -mercaptoethanol and fractions
49 containing pure protein were pooled and concentrated up to 8-9 mg/ml using a vivaspin
50 concentrator with a 30 kDa cut-off. EngA samples were then mixed with either GDP or
51 GMPPCP to reach final protein concentrations of 5 mg/ml (GDP) or 2.5 mg/ml (GMPPCP)
52 and ligand concentrations of 10 mM (GDP) or 0.5 mM (GMPPCP). Crystals were obtained
53 using hanging drop vapor-diffusion technique and by mixing 2 μ L of EngA/nucleotide mix
54 with 2 μ L of the crystallization solution. Two different types of crystals were obtained:
55 EngA1 (EngA/GDP mix in 4 to 10% PEG 3350, 100 mM Mes pH 6.5, 300 mM ammonium
56 sulfate), EngA2 (EngA/GMPPCP mix in 8% PEG 3350, 100 mM MES pH 6.5, 300 mM
57 NaCl).
58
59
60
61
62
63
64
65

1 For X-ray diffraction data collection, crystals were first cryo-protected by soaking
2 them in the crystallization solution to which 30% glycerol was added and then by directly
3 frozen in liquid nitrogen. All crystallographic data were collected at -160°C on beamlines
4 ID23-eh1 or ID23-eh2 of the European Synchrotron Radiation Facility (E.S.R.F.) in Grenoble
5 at wavelengths of 1.282 Å or 0.873 Å using an ADSC Quantum 315R CCD detector. Data
6 processing was performed using XDS package [69] and statistics are summarized in Table S1.
7 The complex structures were determined by molecular replacement method with AMoRe [70]
8 from the CCP4 suite [71] and using the EngA structure already available as search model
9 (pdb code: 2HJG, [30]). The final models were obtained after several rounds of alternate
10 manual building with the coot software [72] and maximum likelihood refinement with
11 Refmac5 of the CCP4 suite [73]. The nature and the occupancy of the ligand in GD1 and
12 GD2 domain were determined with the help of (mFobs – Fcalc) residual omit maps. EngA1
13 contains a sulfate ion in the GD1 nucleotide-binding site and a GDP at full occupancy in the
14 GD2 nucleotide-binding site. For EngA2, the GD1 nucleotide-binding site is empty whereas
15 the GD2 nucleotide-binding site is fully occupied by a GMPPCP molecule. The final
16 refinement statistics are summarized in Table S1. The PyMol software was used to generate
17 3D structure figures. The structure factors and coordinates have been deposited in the Protein
18 Data Bank (PDB) under the accession numbers 4DCS and 4DCV.
19
20
21
22

23 References

- 24
- 25 1. Bourne HR, Sanders DA, McCormick F (1990) The GTPase superfamily: a conserved
26 switch for diverse cell functions. *Nature* 348: 125-132.
- 27 2. Vetter IR, Wittinghofer A (2001) The guanine nucleotide-binding switch in three
28 dimensions. *Science* 294: 1299-1304.
- 29 3. Sprang SR (1997) G protein mechanisms: insights from structural analysis. *Annu Rev*
30 *Biochem* 66: 639-678.
- 31 4. Gasper R, Meyer S, Gotthardt K, Sirajuddin M, Wittinghofer A (2009) It takes two to
32 tango: regulation of G proteins by dimerization. *Nat Rev Mol Cell Biol* 10: 423-429.
- 33 5. Caldon CE, Yoong P, March PE (2001) Evolution of a molecular switch: universal
34 bacterial GTPases regulate ribosome function. *Mol Microbiol* 41: 289-297.
- 35 6. Morimoto T, Loh PC, Hirai T, Asai K, Kobayashi K, et al. (2002) Six GTP-binding
36 proteins of the Era/Obg family are essential for cell growth in *Bacillus subtilis*.
37 *Microbiology* 148: 3539-3552.
- 38 7. Inoue K, Alsina J, Chen J, Inouye M (2003) Suppression of defective ribosome assembly in
39 a *rbfA* deletion mutant by overexpression of Era, an essential GTPase in *Escherichia*
40 *coli*. *Mol Microbiol* 48: 1005-1016.
- 41 8. Sharma MR, Barat C, Wilson DN, Booth TM, Kawazoe M, et al. (2005) Interaction of Era
42 with the 30S ribosomal subunit implications for 30S subunit assembly. *Mol Cell* 18:
43 319-329.
- 44 9. Jiang M, Datta K, Walker A, Strahler J, Bagamasbad P, et al. (2006) The *Escherichia coli*
45 GTPase CgtAE is involved in late steps of large ribosome assembly. *J Bacteriol* 188:
46 6757-6770.
- 47 10. Sato A, Kobayashi G, Hayashi H, Yoshida H, Wada A, et al. (2005) The GTP binding
48 protein Obg homolog ObgE is involved in ribosome maturation. *Genes Cells* 10: 393-
49 408.
- 50 11. Scott JM, Ju J, Mitchell T, Haldenwang WG (2000) The *Bacillus subtilis* GTP binding
51 protein obg and regulators of the sigma(B) stress response transcription factor
52 cofractionate with ribosomes. *J Bacteriol* 182: 2771-2777.
- 53
54
55
56
57
58
59
60
61
62
63
64
65

12. Cooper EL, Garcia-Lara J, Foster SJ (2009) YsxC, an essential protein in *Staphylococcus aureus* crucial for ribosome assembly/stability. *BMC Microbiol* 9: 266.
13. Schaefer L, Uicker WC, Wicker-Planquart C, Foucher AE, Jault JM, et al. (2006) Multiple GTPases participate in the assembly of the large ribosomal subunit in *Bacillus subtilis*. *J Bacteriol* 188: 8252-8258.
14. Wicker-Planquart C, Foucher AE, Louwagie M, Britton RA, Jault JM (2008) Interactions of an essential *Bacillus subtilis* GTPase, YsxC, with ribosomes. *J Bacteriol* 190: 681-690.
15. Matsuo Y, Morimoto T, Kuwano M, Loh PC, Oshima T, et al. (2006) The GTP-binding protein YlqF participates in the late step of 50 S ribosomal subunit assembly in *Bacillus subtilis*. *J Biol Chem* 281: 8110-8117.
16. Uicker WC, Schaefer L, Britton RA (2006) The essential GTPase RbgA (YlqF) is required for 50S ribosome assembly in *Bacillus subtilis*. *Mol Microbiol* 59: 528-540.
17. Loh PC, Morimoto T, Matsuo Y, Oshima T, Ogasawara N (2007) The GTP-binding protein YqeH participates in biogenesis of the 30S ribosome subunit in *Bacillus subtilis*. *Genes Genet Syst* 82: 281-289.
18. Uicker WC, Schaefer L, Koenigsnecht M, Britton RA (2007) The essential GTPase YqeH is required for proper ribosome assembly in *Bacillus subtilis*. *J Bacteriol* 189: 2926-2929.
19. Britton RA (2009) Role of GTPases in bacterial ribosome assembly. *Annu Rev Microbiol* 63: 155-176.
20. Verstraeten N, Fauvart M, Versees W, Michiels J (2011) The universally conserved prokaryotic GTPases. *Microbiol Mol Biol Rev* 75: 507-542.
21. Hwang J, Inouye M (2001) An essential GTPase, Der, containing double GTP-binding domains from *Escherichia coli* and *Thermotoga maritima*. *J Biol Chem* 276: 31415-31421.
22. Mehr IJ, Long CD, Serkin CD, Seifert HS (2000) A homologue of the recombination-dependent growth gene, *rdgC*, is involved in gonococcal pilin antigenic variation. *Genetics* 154: 523-532.
23. Hwang J, Inouye M (2006) The tandem GTPase, Der, is essential for the biogenesis of 50S ribosomal subunits in *Escherichia coli*. *Mol Microbiol* 61: 1660-1672.
24. Shajani Z, Sykes MT, Williamson JR (2011) Assembly of bacterial ribosomes. *Annu Rev Biochem* 80: 501-526.
25. Gil R, Silva FJ, Pereto J, Moya A (2004) Determination of the core of a minimal bacterial gene set. *Microbiol Mol Biol Rev* 68: 518-537.
26. Lee R, Aung-Htut MT, Kwik C, March PE (2011) Expression phenotypes suggest that Der participates in a specific, high affinity interaction with membranes. *Protein Expr Purif* 78: 102-112.
27. Mishra R, Gara SK, Mishra S, Prakash B (2005) Analysis of GTPases carrying hydrophobic amino acid substitutions in lieu of the catalytic glutamine: implications for GTP hydrolysis. *Proteins* 59: 332-338.
28. Tomar SK, Dhimole N, Chatterjee M, Prakash B (2009) Distinct GDP/GTP bound states of the tandem G-domains of EngA regulate ribosome binding. *Nucleic Acids Res* 37: 2359-2370.
29. Bharat A, Jiang M, Sullivan SM, Maddock JR, Brown ED (2006) Cooperative and critical roles for both G domains in the GTPase activity and cellular function of ribosome-associated *Escherichia coli* EngA. *J Bacteriol* 188: 7992-7996.
30. Muench SP, Xu L, Sedelnikova SE, Rice DW (2006) The essential GTPase YphC displays a major domain rearrangement associated with nucleotide binding. *Proc Natl Acad Sci U S A* 103: 12359-12364.

- 1 31. Robinson VL, Hwang J, Fox E, Inouye M, Stock AM (2002) Domain arrangement of Der,
2 a switch protein containing two GTPase domains. *Structure* 10: 1649-1658.
- 3 32. Karbstein K (2007) Role of GTPases in ribosome assembly. *Biopolymers* 87: 1-11.
- 4 33. Brown ED (2005) Conserved P-loop GTPases of unknown function in bacteria: an
5 emerging and vital ensemble in bacterial physiology. *Biochem Cell Biol* 83: 738-746.
- 6 34. Yamanaka K, Hwang J, Inouye M (2000) Characterization of GTPase activity of TrmE, a
7 member of a novel GTPase superfamily, from *Thermotoga maritima*. *J Bacteriol* 182:
8 7078-7082.
- 9 35. Scrima A, Wittinghofer A (2006) Dimerisation-dependent GTPase reaction of MnmE:
10 how potassium acts as GTPase-activating element. *Embo J* 25: 2940-2951.
- 11 36. Ash MR, Maher MJ, Guss JM, Jormakka M (2011) The initiation of GTP hydrolysis by
12 the G-domain of FeoB: insights from a transition-state complex structure. *PLoS One*
13 6: e23355.
- 14 37. Rodríguez-Navarro A (2000) Potassium transport in fungi and plants. *Biochim Biophys*
15 *Acta* 1469: 1-30.
- 16 38. Agarwal N, Pareek M, Thakur P, Pathak V (2012) Functional characterization of
17 EngA(MS), a P-loop GTPase of *Mycobacterium smegmatis*. *PLoS One* 7: e34571.
- 18 39. Ash MR, Guilfoyle A, Clarke RJ, Guss JM, Maher MJ, et al. (2010) Potassium-activated
19 GTPase reaction in the G Protein-coupled ferrous iron transporter B. *J Biol Chem* 285:
20 14594-14602.
- 21 40. Chabre M (1990) Aluminofluoride and beryllifluoride complexes: a new phosphate
22 analogs in enzymology. *Trends Biochem Sci* 15: 6-10.
- 23 41. Lo MC, Aulabaugh A, Jin G, Cowling R, Bard J, et al. (2004) Evaluation of fluorescence-
24 based thermal shift assays for hit identification in drug discovery. *Anal Biochem* 332:
25 153-159.
- 26 42. Pantoliano MW, Petrella EC, Kwasnoski JD, Lobanov VS, Myslik J, et al. (2001) High-
27 density miniaturized thermal shift assays as a general strategy for drug discovery. *J*
28 *Biomol Screen* 6: 429-440.
- 29 43. Vedadi M, Niesen FH, Allali-Hassani A, Fedorov OY, Finerty PJ, Jr., et al. (2006)
30 Chemical screening methods to identify ligands that promote protein stability, protein
31 crystallization, and structure determination. *Proc Natl Acad Sci U S A* 103: 15835-
32 15840.
- 33 44. Hwang YW, Miller DL (1987) A mutation that alters the nucleotide specificity of
34 elongation factor Tu, a GTP regulatory protein. *J Biol Chem* 262: 13081-13085.
- 35 45. Zhong JM, Chen-Hwang MC, Hwang YW (1995) Switching nucleotide specificity of Ha-
36 Ras p21 by a single amino acid substitution at aspartate 119. *J Biol Chem* 270: 10002-
37 10007.
- 38 46. Shan SO, Walter P (2003) Induced nucleotide specificity in a GTPase. *Proc Natl Acad Sci*
39 *U S A* 100: 4480-4485.
- 40 47. Di Cera E (2006) A structural perspective on enzymes activated by monovalent cations. *J*
41 *Biol Chem* 281: 1305-1308.
- 42 48. Suelter CH (1970) Enzymes activated by monovalent cations. *Science* 168: 789-795.
- 43 49. Anand B, Surana P, Prakash B (2010) Deciphering the catalytic machinery in 30S
44 ribosome assembly GTPase YqeH. *PLoS One* 5.
- 45 50. Achila D, Gulati M, Jain N, Britton RA (2012) Biochemical characterization of the
46 ribosome assembly GTPase RbgA in *Bacillus subtilis*. *J Biol Chem*
47 <http://www.jbc.org/cgi/doi/10.1074/jbc.M111.331322>.
- 48 51. Wittinghofer A, Vetter IR (2011) Structure-function relationships of the G domain, a
49 canonical switch motif. *Annu Rev Biochem* 80: 943-971.

- 1 52. Meury J, Kepes A (1981) The regulation of potassium fluxes for the adjustment and
2 maintenance of potassium levels in *Escherichia coli*. *Eur J Biochem* 119: 165-170.
- 3 53. Teixeira de Mattos MJ, Neijssel OM (1997) Bioenergetic consequences of microbial
4 adaptation to low-nutrient environments. *J Biotechnol* 59: 117-126.
- 5 54. Leipe DD, Wolf YI, Koonin EV, Aravind L (2002) Classification and evolution of P-loop
6 GTPases and related ATPases. *J Mol Biol* 317: 41-72.
- 7 55. Gideon P, John J, Frech M, Lautwein A, Clark R, et al. (1992) Mutational and kinetic
8 analyses of the GTPase-activating protein (GAP)-p21 interaction: the C-terminal
9 domain of GAP is not sufficient for full activity. *Mol Cell Biol* 12: 2050-2056.
- 10 56. Lamb HK, Thompson P, Elliott C, Charles IG, Richards J, et al. (2007) Functional
11 analysis of the GTPases EngA and YhbZ encoded by *Salmonella typhimurium*.
12 *Protein Sci* 16: 2391-2402.
- 13 57. Scheffzek K, Ahmadian MR, Kabsch W, Wiesmuller L, Lautwein A, et al. (1997) The
14 Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic
15 Ras mutants. *Science* 277: 333-338.
- 16 58. Guilfoyle A, Maher MJ, Rapp M, Clarke R, Harrop S, et al. (2009) Structural basis of
17 GDP release and gating in G protein coupled Fe²⁺ transport. *Embo J* 28: 2677-2685.
- 18 59. Koster S, Wehner M, Herrmann C, Kuhlbrandt W, Yildiz O (2009) Structure and function
19 of the FeoB G-domain from *Methanococcus jannaschii*. *J Mol Biol* 392: 405-419.
- 20 60. Jault JM, Di Pietro A, Falson P, Gautheron DC (1991) Alteration of apparent negative
21 cooperativity of ATPase activity by alpha-subunit glutamine 173 mutation in yeast
22 mitochondrial F1. Correlation with impaired nucleotide interaction at a regulatory site.
23 *J Biol Chem* 266: 8073-8078.
- 24 61. Galinier A, Lavergne JP, Geourjon C, Fieulaine S, Nessler S, et al. (2002) A new family
25 of phosphotransferases with a P-loop motif. *J Biol Chem* 277: 11362-11367.
- 26 62. Jault JM, Fieulaine S, Nessler S, Gonzalo P, Di Pietro A, et al. (2000) The HPr kinase
27 from *Bacillus subtilis* is a homo-oligomeric enzyme which exhibits strong positive
28 cooperativity for nucleotide and fructose 1,6-bisphosphate binding. *J Biol Chem* 275:
29 1773-1780.
- 30 63. Divita G, Goody RS, Gautheron DC, Di Pietro A (1993) Structural mapping of catalytic
31 site with respect to alpha-subunit and noncatalytic site in yeast mitochondrial F1-
32 ATPase using fluorescence resonance energy transfer. *J Biol Chem* 268: 13178-
33 13186.
- 34 64. Attali C, Frolet C, Durmort C, Offant J, Vernet T, et al. (2008) *Streptococcus pneumoniae*
35 choline-binding protein E interaction with plasminogen/plasmin stimulates migration
36 across the extracellular matrix. *Infect Immun* 76: 466-476.
- 37 65. Karst JC, Foucher AE, Campbell TL, Di Guilmi AM, Stroebel D, et al. (2009) The
38 ATPase activity of an 'essential' *Bacillus subtilis* enzyme, YdiB, is required for its
39 cellular function and is modulated by oligomerization. *Microbiology* 155: 944-956.
- 40 66. Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation velocity
41 ultracentrifugation and lamm equation modeling. *Biophys J* 78: 1606-1619.
- 42 67. Dam J, Schuck P (2004) Calculating sedimentation coefficient distributions by direct
43 modeling of sedimentation velocity concentration profiles. *Methods Enzymol* 384:
44 185-212.
- 45 68. Lebowitz J, Lewis MS, Schuck P (2002) Modern analytical ultracentrifugation in protein
46 science: a tutorial review. *Protein Sci* 11: 2067-2079.
- 47 69. Kabsch W (1993) Automatic processing of rotation diffraction data from crystals of
48 initially unknown symmetry and cell constants. *J Appl Crystallogr* 26: 795-800.
- 49 70. Navaza J (1994) AMoRe: an automated package for molecular replacement. *Acta*
50 *Crystallogr A* 50: 157-163.

- 1 71. Collaborative Computational Project Number 4 (1994) The CCP4 Suite: Programs for
2 Protein Crystallography. Acta Crystallogr D Biol Crystallogr 50: 760-763.
3 72. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta
4 Crystallogr D Biol Crystallogr 60: 2126-2132.
5 73. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of Macromolecular Structures
6 by the Maximum-Likelihood Method. Acta Crystallogr D Biol Crystallogr 53: 240-
7 255.
8
9

10 **Figure legends**

11
12
13 **Figure 1. Salt effects on the GTPase activity of EngA, GD1 and GD2-KH.** *A*, the GTPase
14 activity of EngA was monitored in the presence of 1 mM of GTP and KCl (closed circles),
15 RbCl (closed squares), or NaCl (open squares); insert shows a SDS-PAGE of the purified
16 EngA (indicated by a black arrowhead) which migrated close to its predicted molecular
17 weight of 50.9 kDa). *B*, the GTPase activity of GD1 (closed circles, left scale) or GD2-KH
18 (closed squares, right scale) was monitored in the presence of KCl; insert shows a SDS-PAGE
19 of the purified proteins, indicated by black arrowheads: left lane, GD1 (predicted molecular
20 weight of 20.7 kDa) ; right lane, GD2-KH (predicted molecular weight of 32.1 kDa).
21
22
23

24 **Figure 2. GTPase activity of EngA as a function of increasing concentrations of GTP.** *A*,
25 in the absence of KCl; *B*, in the presence of 300 mM of KCl. K_M and V_{max} values were
26 determined from iterative nonlinear fits of the Michaelis-Menten equation to the experimental
27 data, using the GraFit 5.0.11 software (Erithacus software) and the fitted curve are shown.
28 The inserts show the Lineweaver-Burk plots of the data. One set of data is shown and similar
29 results were obtained from at least three independent experiments.
30
31
32

33 **Figure 3. Binding of GDP to EngA.** The binding of GDP was monitored by the increase in
34 tryptophan fluorescence of EngA at 25°C, in the absence (open triangles) or in the presence
35 (closed circles) of 300 mM KCl. The data were fitted using the GraFit 5.0.11 software and the
36 fitted curves are shown.
37
38

39 **Figure 4. Binding of GDP and GMPPNP to GD1 and GD2-KH.** The binding of GDP (*A*
40 and *C*) or GMPPNP (*B* and *D*) was monitored by the increase in tryptophan fluorescence of
41 GD1 (*A* and *B*) and GD2-KH (*C* and *D*) in the presence of 300 mM KCl (white triangles,
42 dashed lines for the fitted curves) or 300 mM NaCl (white circles, plain lines for the fitted
43 curves).
44
45

46 **Figure 5. Thermal denaturation of EngA, GD1 and GD2-KH and protective effects of**
47 **nucleotides.** The denaturation of EngA (*A*), GD1 (*B*) and GD2-KH (*C*) was monitored by the
48 Thermal Shift Assay in the absence (plain black lines) or in the presence of 1 mM GMPPNP
49 (plain grey lines) or 1 mM GDP (dotted grey lines).
50
51

52 **Figure 6. Ultracentrifugation experiments of EngA or GD1.** Superposition of selected
53 sedimentation profiles obtained at 278 nm and the corresponding fitted curves using
54 continuous size distribution analysis for EngA (*A*) and GD1 (*D*). Superposition of the
55 residuals (*B*, EngA; *E*, GD1). The fitted curves were used to calculate the sedimentation
56 coefficient distribution of the enzyme, $c(s)$ (*C*, EngA; *F*, GD1).
57
58
59
60
61
62
63
64
65

1 **Figure 7. GTPase or XTPase activities of wild-type or mutated EngA.** *A*, the GTPase
2 activity of the wild-type EngA (WT) or the mutants was measured in the presence of 1 mM
3 GTP using a coupled-enzymatic assay in the absence (black bars) or in the presence of 300
4 mM KCl (white bars); the results are expressed relative to that of the wild-type EngA in the
5 presence of 300 mM KCl (100% corresponding to 100 μ mol GTP hydrolyzed/min/mg
6 proteins). The activity of the K16A mutant is undetectable in the absence of potassium. *B*,
7 same as *A* but 200 μ M of nucleotides were used in the presence of 300 mM KCl (100%
8 corresponding to 77 μ mol GTP hydrolyzed/min/mg proteins); white bars, GTPase activity;
9 grey bars, XTPase activity. The values represent the mean of three experiments and the
10 standard deviation is indicated as error bars, in grey color.
11
12

13 **Figure 8. Superimposition of GD1 or GD2 domains and sequence conservation in the**
14 **switch 2 motifs.** *A*, the EngA GD1 domain is depicted in orange and the GDP molecule in
15 cyan. Gln61 from hRas is shown in turquoise, MnmE Glu282 and GDP-AIFx transition state
16 analogue are shown in green. *T. maritima* EngA Asp55 and Asp61 are shown in purple.
17 Structure superpositions were performed by superimposition of the G1, G4 and G5 motifs
18 using PyMol. *B*, the EngA GD2 domain is depicted in yellow with its bound GDP molecule
19 shown in cyan. *T. maritima* EngA Asp55 is shown in purple. Putative Mg²⁺ position, as
20 obtained from the superposed MnmE structure, is shown in pale green. Hydrogen bonds are
21 depicted with cyan dashed lines. Structure superpositions were performed by superimposition
22 of the G1, G4 and G5 motifs. One hundred EngA sequences were used to show the sequence
23 variation in the switch 2 following the G3 motif and using the WebLogo server
24 (<http://weblogo.berkeley.edu/>).
25
26
27
28

29 **Figure S1. Domain organization and sequence alignment of EngA.** *A*, The boundaries of
30 the three domains, GD1, GD2 and KH of EngA are indicated and the four motifs found in all
31 GTPases, G1 to G4, are shown as pink and green boxes in GD1 and GD2, respectively. The
32 mutants constructed in this study are also indicated. *B*, alignment of some representative
33 sequences of EngA using Clustal W (<http://npsa-pbil.ibcp.fr>) and ESPript
34 (<http://esript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi>). The abbreviations used are BACSU,
35 *Bacillus subtilis*; THEMA, *Thermotoga maritima*; ECOLI, *Escherichia coli*; PSEAE,
36 *Pseudomonas aeruginosa*; NEIGO, *Neisseria gonorrhoeae*; LISMO, *Listeria monocytogenes*;
37 STARA, *Staphylococcus aureus*, MYCTU, *Mycobacterium tuberculosis*. Blue boxes are
38 drawn when at least 50% of the residues were conserved (residues in red color with a yellow
39 frame), and fully conserved residues are shown in white color with a red frame. The residues
40 mutated are indicated by black arrowheads and the two Trp residues by stars. The G1 to G4
41 motifs are underlined in pink and green in GD1 and GD2, respectively and the secondary
42 structure of *B. subtilis* EngA deduced from its 3D structure is shown above the sequences.
43
44
45
46

47 **Figure S2. Illustration of the GTPase activity of EngA.** The GTPase activity of EngA was
48 monitored at 340 nm in the presence of 1 mM GTP and using a coupled-enzymatic assay to
49 follow the change in absorbance at 340 nm upon conversion of NADH to NAD⁺, in the
50 absence (black trace, 3.2 μ M of EngA) or in the presence of 300 mM K⁺ (red trace, 0.8 μ M of
51 EngA).
52
53
54

55 **Figure S3. Overall view of the *B. subtilis* EngA1 and EngA2 structures superposed to the**
56 **previously solved *B. subtilis* EngA structure (pdb entry 2HJG).** The EngA1 C α backbone
57 is depicted in pale orange (GD1), pale yellow (GD2) and pale green (KH domain). The
58 EngA2 C α backbone is depicted in orange (GD1), lime (GD2) and forest green (KH domain).
59 The 2HJG C α backbone and the GDP molecules (thin lines) bound to GD1 and GD2 are
60
61
62
63
64
65

1 depicted in grey. The GDP molecule bound to the EngA1 GD2 is shown with cyan sticks. The
2 GMPPCP molecule bound to the EngA2 GD2 is shown with turquoise sticks. The EngA1
3 sulfate ion present in the GD1 nucleotide-binding site is shown in yellow. The two tryptophan
4 residues present in the EngA sequence have been highlighted in hot pink. The minimal GDP-
5 tryptophan side chain distances are 13.0 Å and 8.4 Å for GD1 Trp48 and GD2 Trp296,
6 respectively.

7
8 **Figure S4. Tryptophan emission spectrum of EngA, GD1 and GD2-KH.** After excitation
9 at 295 nm, tryptophan emission spectrum of 0.5 μM EngA (*A*), 1 μM GD1 (*B*) and 1 μM
10 GD2-KH (*C*) are shown in 50 mM Tris/HCl pH 7.5 and 300 mM KCl (red curves) or 300 mM
11 NaCl (black curves).
12

13
14 **Figure S5. Analytical gel filtration profiles of EngA in the presence of different effectors.**
15 Analytical gel filtrations were performed using a Superdex 75 10/30 equilibrated with a buffer
16 containing 50 mM Tris/HCl pH 7.5 and 300 mM NaCl (*A*), 300 mM KCl (*B*) or 300 mM
17 KCl, 1 mM GDP, 1 mM AlCl₃ and 10 mM NaF (*C*) after a prior incubation of 0.5 mg of
18 EngA in the same buffer during 15 min on ice. Fractions of 0.5 ml were collected and the
19 elution volumes of albumin (65 kDa; red arrow), ovalbumin (43 kDa; blue arrow),
20 chymotrypsin (25 kDa; green arrow) and RNase (13.7 kDa; grey arrow) are indicated. V_o
21 corresponds to the void volume and V_t to the total volume of the column. A similar profile
22 was also obtained when EngA was incubated in the presence of 50 mM Tris/HCl pH 7.5, 300
23 mM KCl and 1 mM GMPPNP and submitted to a gel filtration equilibrated in the same buffer
24 (not shown).
25
26
27
28

29 **Figure S6. Superimposition of GD1 or GD2 nucleotide-binding site from different EngA**
30 **structures.** *A*, GD1 domains. The EngA1 is depicted in pale yellow, with the sulphate ion
31 shown in yellow (sulfur atom). The EngA3 structure is depicted in orange with the GDP
32 bound molecule shown in cyan. The EngA4 structure is depicted in green and *T. maritima*
33 EngA is depicted in purple, including its two bound phosphate ions in pink (phosphorus
34 atoms). *B*, G2 domains. The EngA4 GD2 is depicted in pale yellow with its bound GMPPCP
35 molecule shown in turquoise. The EngA3 GD2 is depicted in grey, with its bound GDP
36 molecule in cyan.
37
38
39

40 **Figure S7. View of the EngA GD2 switch II region and its superposition with MnmE,**
41 **FeoB and Ras-RasGAP.** The EngA GD2 is shown in yellow with the switch II region
42 highlighted in orange with its own GDP and GMPPCP, superimposed here, depicted in cyan.
43 The FeoB K⁺ and Mg²⁺ ions are shown as semitransparent pale green and purple spheres,
44 respectively. K-loops are shown in light blue (FeoB) and pink (MnmE) coils. Ras Gln61 and
45 RasGAP Arg789 (arginine finger) are shown in turquoise. Hydrogen bonds stabilizing either
46 the GD2 Lys236 NZ atom or the Mg²⁺ ion are depicted with cyan dashed lines. Structure
47 superimpositions were performed by superimposition of the G1, G4 and G5 motifs. The EngA
48 Met 233 carbonyl group almost superposed to the G2 Thr35 from FeoB, shown in sticks, and
49 could contribute to stabilize the Mg²⁺ ion in the absence of a structurally stable K-loop.
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Figure 1
[Click here to download high resolution image](#)

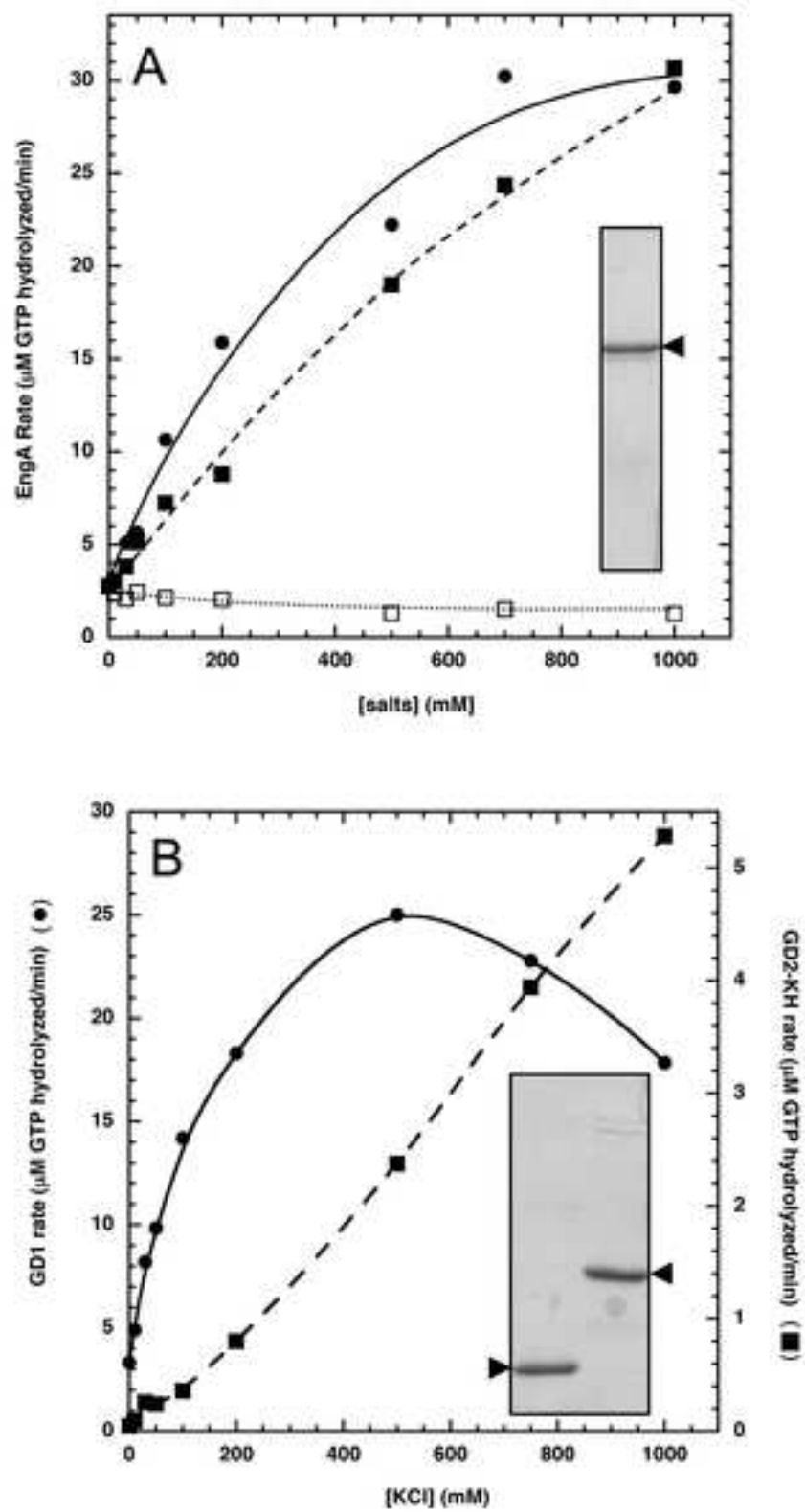


Fig. 1

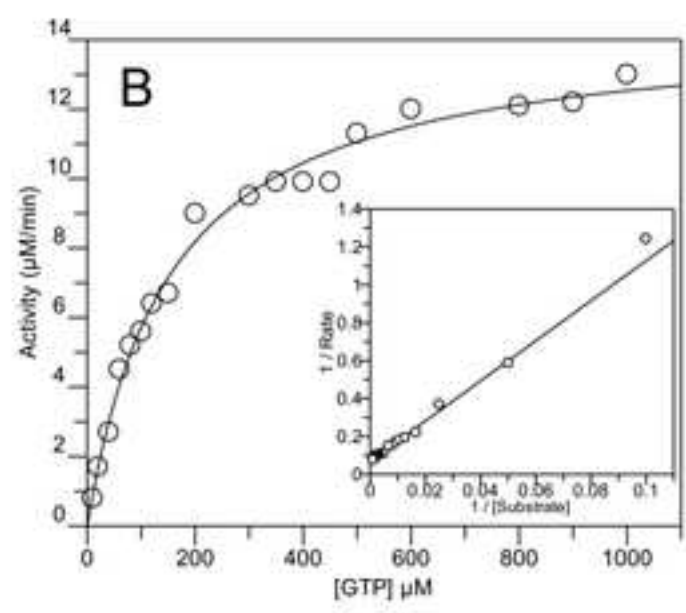
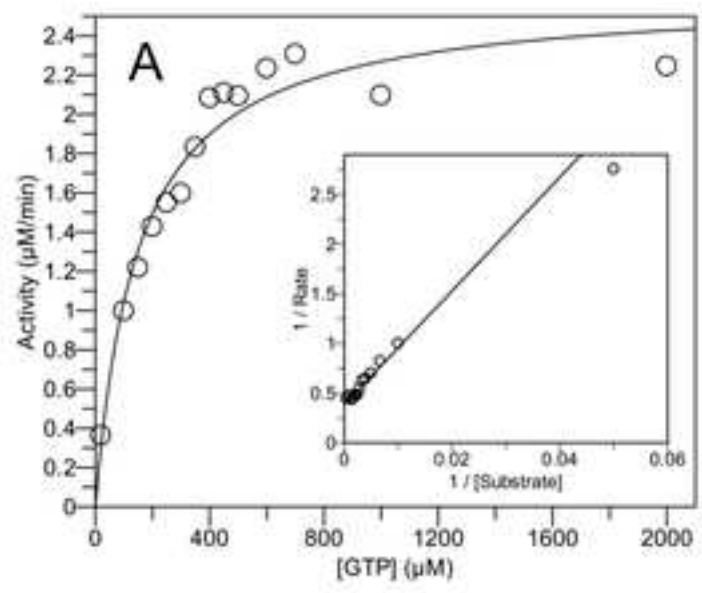


Fig. 2

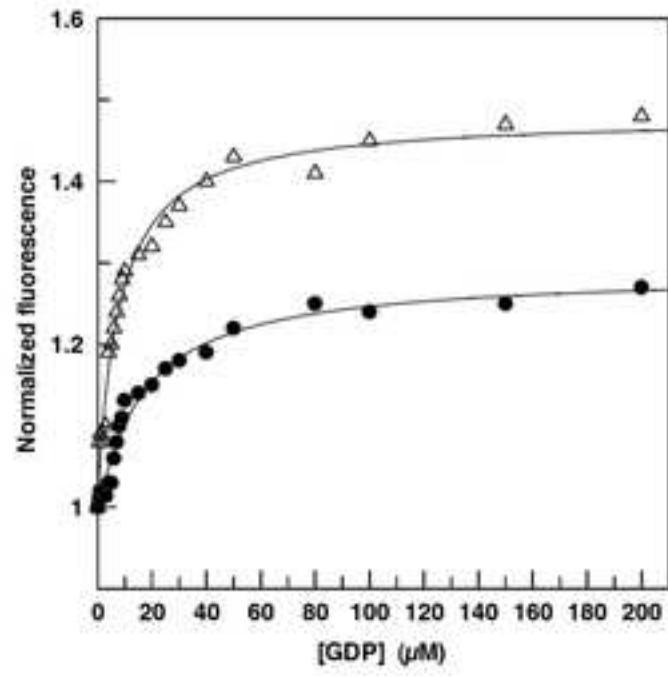


Fig. 3

Figure 4

[Click here to download high resolution image](#)

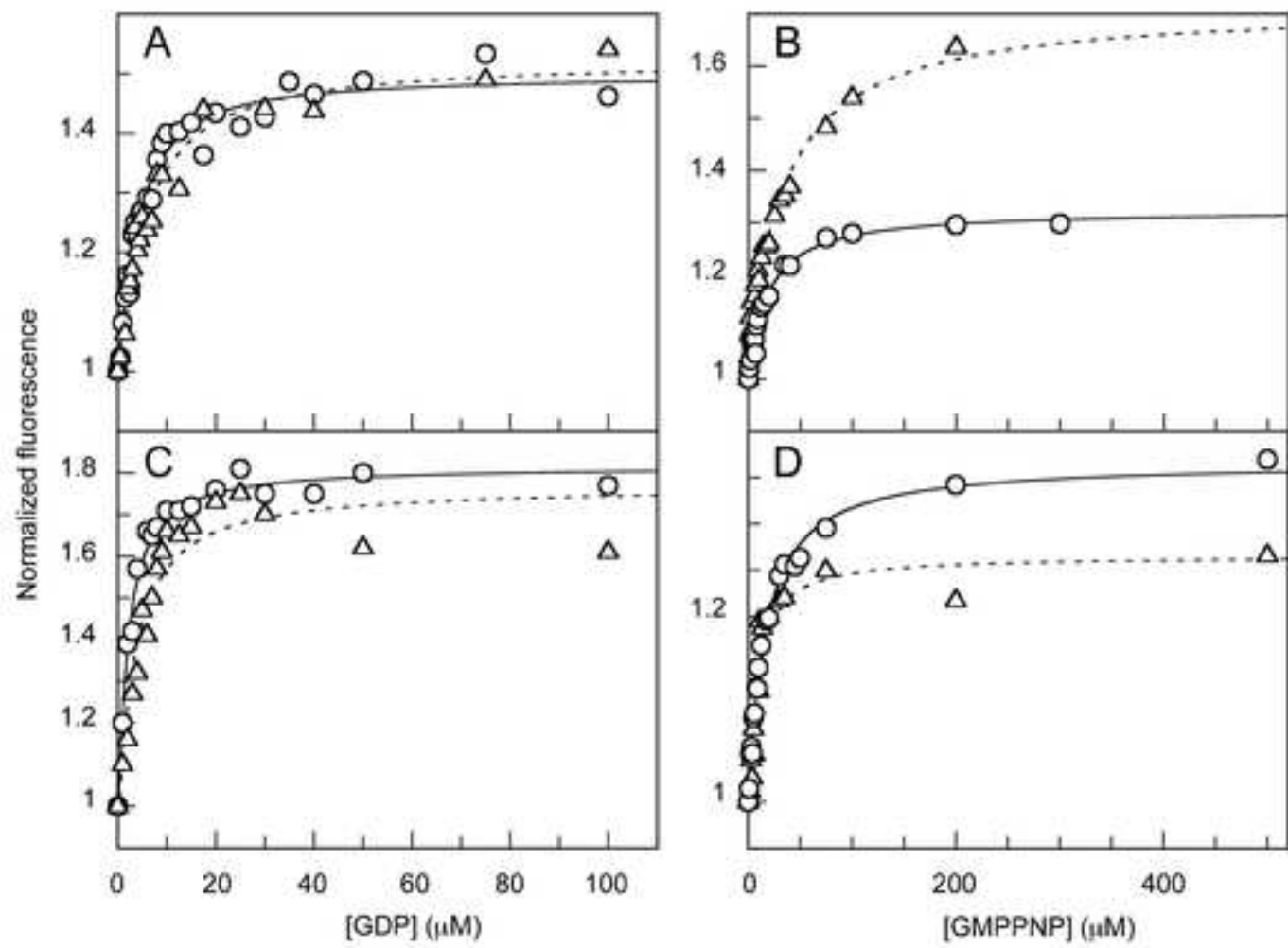


Fig. 4

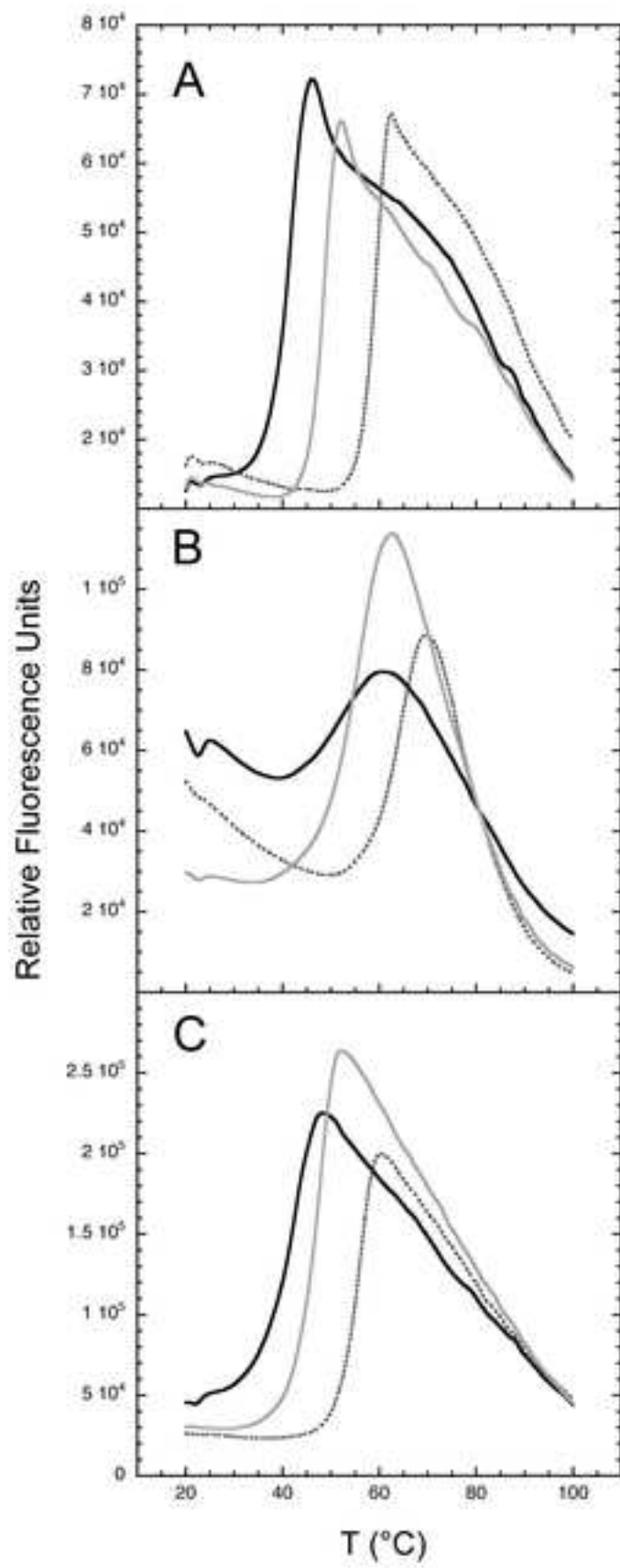


Fig. 5

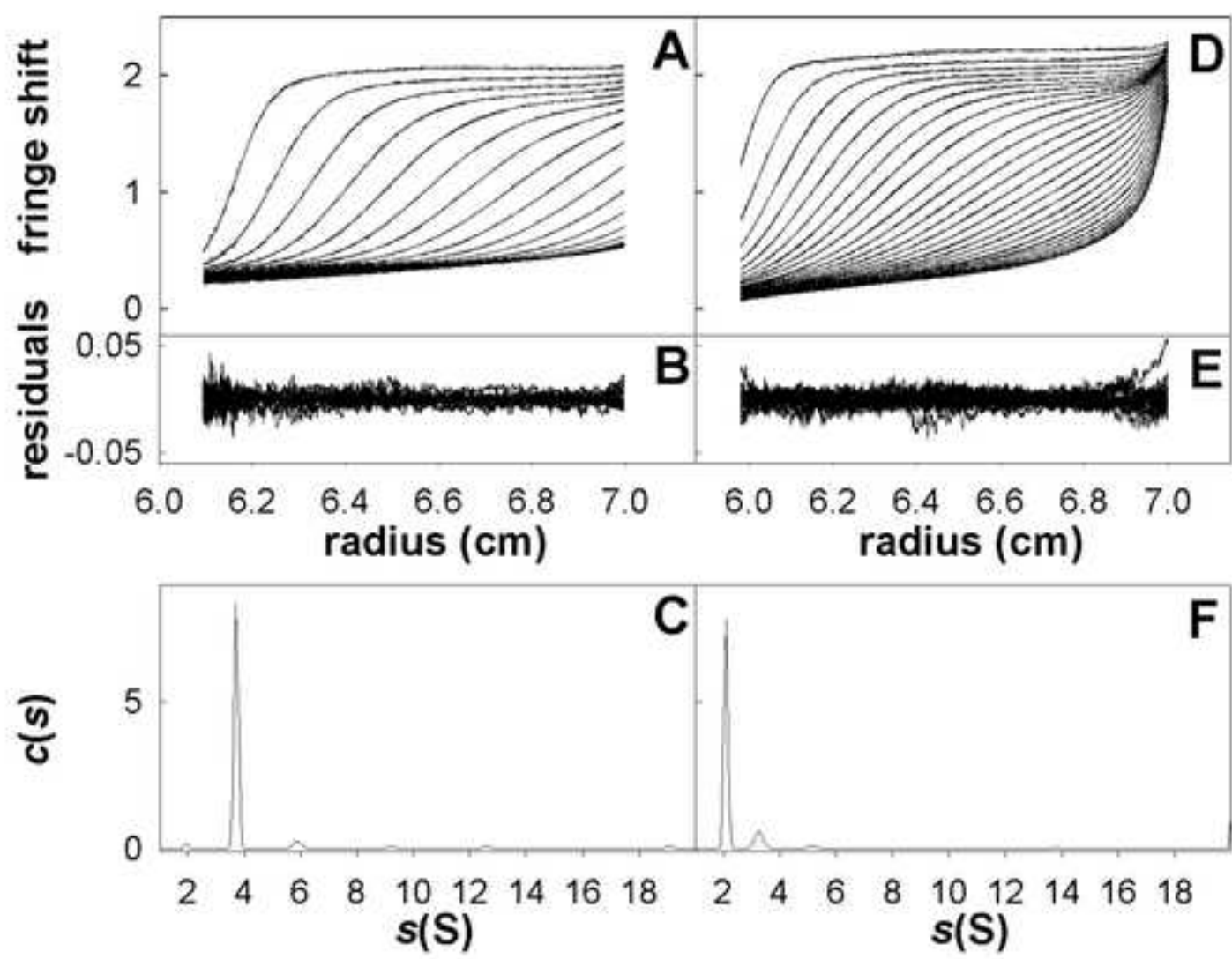


Fig. 6

Figure 7

[Click here to download high resolution image](#)

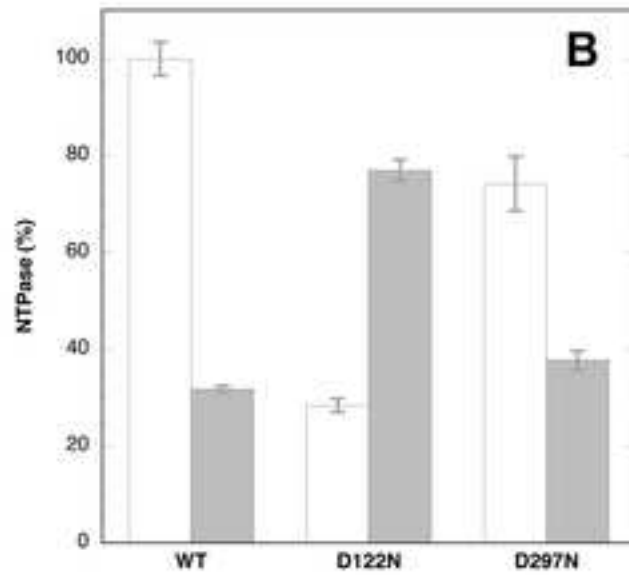
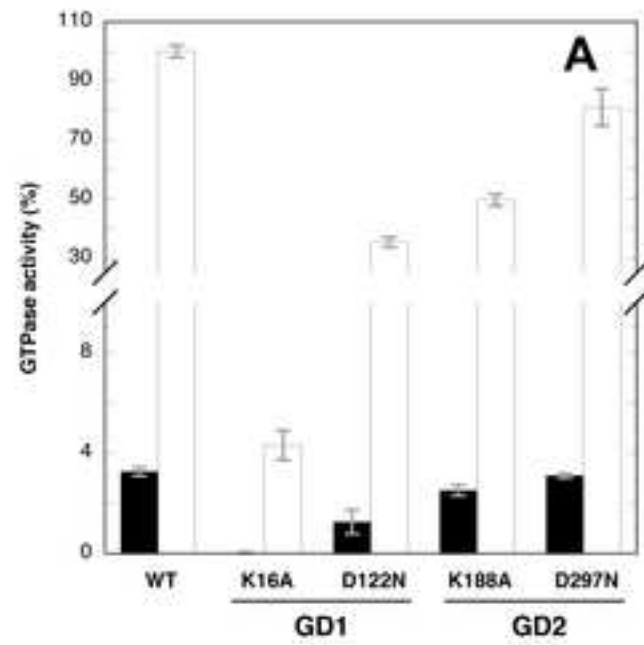


Fig. 7

Figure 8
[Click here to download high resolution image](#)

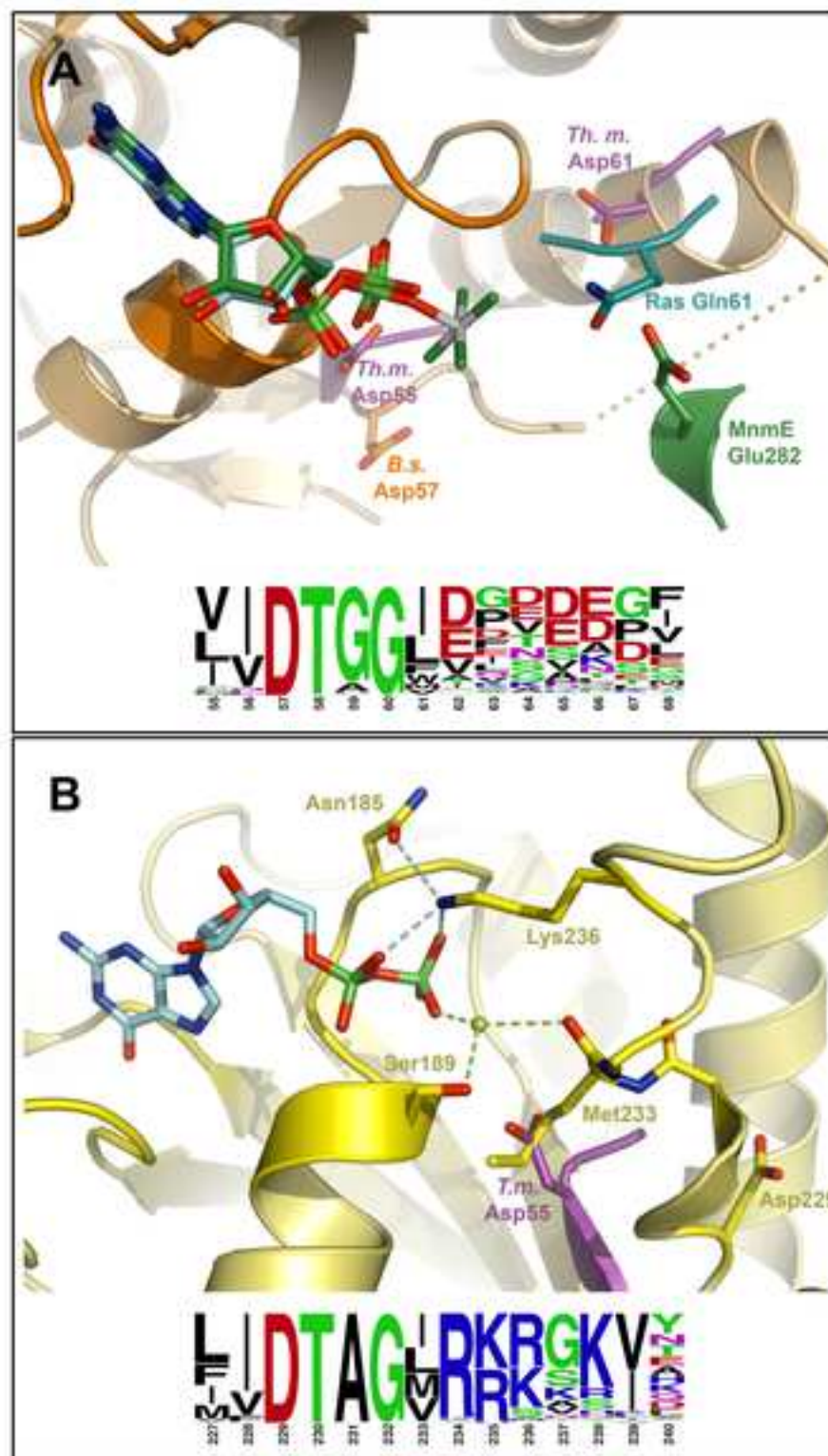


Fig. 8

Titre : Quelques exemples de biologie structurale

Résumé : La biologie structurale et la cristallographie des protéines en particulier s'offrent comme des méthodes de choix pour étudier à l'échelle atomique les interactions protéine-protéine et protéine-ligand qui sont au cœur des processus biologiques. Ce mémoire a pour objectif de donner différents exemples d'études qui montrent comment les données structurales peuvent répondre aussi bien à des questions fondamentales en rapport avec la santé humaine qu'à des questions sur la recherche de nouvelles molécules à visée thérapeutique. Les projets, que j'ai abordés, portent à la fois sur les mécanismes immunologiques de reconnaissance, de signalisation et de réponse des organismes hôtes face aux pathogènes et au soi-altéré ainsi que sur la recherche de nouvelles molécules anticancéreuses ou antibiotiques. En plus d'exposer rapidement mon parcours de biologiste structural, je m'attache ici à décrire les trois grands projets auxquels j'ai travaillé ces dernières années au sein de l'Institut de Biologie Structurale Jean-Pierre Ebel de Grenoble. Ils concernent :

- La reconnaissance et la réponse antivirales ou antitumorales des cellules T
- De nouvelles voies de la signalisation de la phagocytose
- Des protéines liant les nucléotides et la recherche de leurs inhibiteurs

Mots clés : biologie structurale, cristallographie, SPR, cellules T, TCR, HLA, antigène, phagocytose, cellules apoptotiques, DOCK, ELMO, protéines kinases, CK2, GTPases bactériennes, EngA, inhibiteurs, molécules anticancéreuses, molécules antibiotiques

Title: A few examples of structural biology

Abstract: Structural biology and more particularly protein crystallography are methods of choice to study at the atomic level protein-protein or protein-ligand interactions, which lie at the heart of biological processes. This thesis has the objective to give several examples of studied cases that show how structural data can give answer to fundamental questions related to human health as well as to questions on the research of new therapeutic molecules. The projects that I have treated concern immunological mechanisms of recognition, signalization and response by the host organism to pathogen and altered self, but also the research of new anticancer or antibiotic molecules. Besides exposing my structural biologist background, I present the three main projects on which I have worked during the last past years at the "Institut de Biologie Structurale Jean-Pierre Ebel". These projects concern:

- The antiviral or antitumor recognition and response by T cells
- The new phagocytosis signaling pathways
- The proteins able to bind nucleotides and the research of their inhibitors

Keywords: structural biology, crystallography, SPR, T cells, TCR, HLA, antigen, phagocytosis, apoptotic cells, DOCK, ELMO, protein-kinases, CK2, bacterial GTPases, EngA, inhibitors, anticancer and antibiotic molecules