Analysis of relationships between peptide/MHC structural features and naive T cell frequency in humans

Jean-Baptiste Reiser, François Legoux, Stéphanie Gras, Eric Trudel, Anne Chouquet, Alexandra Léger, Madalen Le Gorrec, Paul Machillot, Marc Bonneville, Xavier Saulquin, et al.

Analysis of Relationships between Peptide/MHC Structural Features and Naive T Cell Frequency in Humans

Jean-Baptiste Reiser,[*,†,‡] François Legoux,[§,1] Stéphanie Gras,[*,†,‡,2] Eric Trudel,[*,†,‡] Anne Chouquet,[*,†,‡] Alexandra Léger,[§] Madalen Le Gorrec,[*,†,‡] Paul Machillot,[*,†,‡] Marc Bonneville,[§] Xavier Saulquin,[§,¶] and Dominique Housset[*,†,‡]

The structural rules governing peptide/MHC (pMHC) recognition by T cells remain unclear. To address this question, we performed a structural characterization of several HLA-A2/peptide complexes and assessed in parallel their antigenicity, by analyzing the frequency of the corresponding Ag-specific naive T cells in $A2^+$ and $A2^-$ individuals, as well as within $CD4^+$ and $CD8^+$ subsets. We were able to find a correlation between specific naive T cell frequency and peptide solvent accessibility and/or mobility for a subset of moderately prominent peptides. However, one single structural parameter of the pMHC complexes could not be identified to explain each peptide antigenicity. Enhanced pMHC antigenicity was associated with both highly biased TRAV usage, possibly reflecting favored interaction between particular pMHC complexes and germline TRAV loops, and peptide structural features allowing interactions with a broad range of permissive CDR3 loops. In this context of constrained TCR docking mode, an optimal peptide solvent exposed surface leading to an optimal complementarity with TCR interface may constitute one of the key features leading to high frequency of specific T cells. Altogether our results suggest that frequency of specific T cells depends on the fine-tuning of several parameters, the structural determinants governing TCR–pMHC interaction being just one of them. *The Journal of Immunology*, **2014, 193: 5816–5826.**

T cells play a central role in vertebrate immune systems by protecting the organism against the proliferation of different infectious agents such as viruses, bacteria, or protozoa as well as tumor cells. To achieve these goals, T cells display at their surface specific receptors, the αβ TCRs able to discriminate between peptide fragments derived from self or foreign proteins, presented by proteins coded by MHC genes. Since the structure determination of the first TCR-peptide/MHC (pMHC) complex (1, 2), ~80 structures of such ternary complexes are now available in the Protein Data Bank, including murine and human TCRs, class I and class II MHC molecules, syngeneic, allogeneic, and xenogenic MHC molecules as well as self- and non–self-Ag peptides. They provide us with a structural insight into several key features of TCR-pMHC recognition such as TCR plasticity, MHC restriction, cross-reactivity, alloreactivity, immunodominance, and autoimmune TCR recognition. Unfortunately and because of the extreme diversity of the TCRs, the versatility of the TCR for docking onto the pMHC, and the exquisite sensitivity of this recognition event, these available structural data did not allow yet establishing general rules governing cognate TCR–pMHC interactions. As an illustration, only 3 MHC class I residues (65, 69, and 155) were identified as being always involved in contacts with the TCR in all the TCR-pMHC-I structures determined so far (3), except in one very recently determined ternary complex (4). However, the contribution of these residues to the interaction varies from one TCR to another, and these residues do not constitute a conserved energetic hotspot for TCR docking (5). They rather constitute MHC triads that the TCR can barely avoid to contact, given the structural constraints governing the TCR-pMHC docking.

Another possible approach used to shed light on the largely unknown rules governing TCR-pMHC recognition is to look at correlates between general structural characteristics of pMHC complexes and the nature of the T cell repertoire specific for a given pMHC. It may allow obtaining a much broader view of the impact of the pMHC structure features on the T cell response because looking at a few specific clones is clearly not sufficient to provide a detailed picture of the TCR-pMHC recognition signature. Several indicators of the efficacy of the T cell immune response, such as the ability of the antigenic peptide to be efficiently displayed by MHC molecules, the frequency of T cells reactive to a given pMHC complex or the avidity of T lymphocytes for the cells presenting a given Ag are potentially correlated with structural and thermodynamic properties of pMHC. By comparing the diversity of a T cell immune repertoire subset specific for different peptides presented by $H\text{-}2D^b$, Doherty

*Université de Grenoble Alpes, Institut de Biologie Structurale, F-38044 Grenoble, France; †Commissariat à l'énergie atomique et aux énergies alternatives, Direction des sciences du vivant, Institut de Biologie Structurale, F-38044 Grenoble, France; ‡Centre national de la recherche scientifique, Institut de Biologie Structurale, F-38044 Grenoble, France; §Institut national de la santé et de la recherche médicale, Unité mixte de recherche 892, Centre de Recherche en Cancérologie Nantes Angers, F-44000 Nantes, France; and ¶Université de Nantes, F-44000 Nantes, France

[1]Current address: Massachusetts General Hospital and Harvard Medical School, Charlestown, MA.

[2]Current address: Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria, Australia.

This work was supported by the platforms of the Grenoble Instruct centre (Integrated Structural Biology, Grenoble; Unité mixte de service 3518, Centre national de la recherche scientifique–Commissariat à l'énergie atomique et aux énergies alternatives–Université Joseph Fourier–European Molecular Biology Laboratory) with support from French Infrastructure for Integrated Structural Biology (FRISBI) (Grant ANR-10-INSB-05-02) and Grenoble Alliance for Integrated Biology (Grant ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology. This study was supported in part by the Agence Nationale de la Recherche (Grants ANR-05-MIIM-019 and ANR-11-BSV3-006-01) and by the Association pour la Recherche sur le Cancer.

Address correspondence and reprint requests to Dr. Dominique Housset, Institut de Biologie Structurale, 71, avenue des Martyrs, CS 10090, F-38044 Grenoble, France. E-mail address: dominique.housset@ibs.fr

Abbreviations used in this article: pMHC, peptide/MHC; r.m.s.d., root mean square difference.

and collaborators have observed that a featureless peptide from influenza, NP_{366}, was selecting a limited T cell repertoire, along with the emergence of "public" (i.e., present in numerous unrelated individuals) TCR β-chains. By contrast, a much more prominent influenza peptide, PA_{224}, was shown to select a diverse and "private" T cell repertoire. Furthermore, analysis of a PA_{224} mutant, in which the protruding arginine in position 7 was replaced by an alanine, showed that the reduced protuberance of the variant was also associated with a reduced diversity of the selected T cell repertoire (6). These results were in line with previous observations made on buried and featureless influenza epitope presented by HLA-A2, which selected an highly biased TCR repertoire (7). However, the analysis of two immunodominant epitopes from the human CMV protein pp65 presented by HLA-B*3508 showed that a prominent 12-mer peptide was selecting a public highly restricted TCR repertoire, whereas a nonprominent 8-mer peptide was selecting a much more diverse TCR repertoire (8). Therefore, the connection between pMHC structural characteristics and the nature of the specific T cell repertoire is complex and remains to be established.

Until recently, little was known about the characteristics of Ag-specific naive T cell repertoire, essentially because of the low frequency (in the 10^{-5}–10^{-6} range) of most specific T lymphocytes. Using a new approach based on pMHC oligomers staining and allowing ex vivo measurement of the frequency of naive T cells specific for a given pMHC in mice, Jenkins and collaborators recently showed that the frequency of murine naive $CD4^+$ T cells was positively correlated with both the repertoire diversity and response magnitude (9). This observation was then confirmed on murine $CD8^+$ T cells (10, 11). However, a consistent relationship between the frequency of naive T cells and the magnitude of the immune response is not always observed (12, 13) and indicate that other factors play a significant role such as the prevalence of high-avidity T cells in the naive repertoire (14). Moreover, although chromophore-labeled pMHC tetramers staining is recognized as a valuable approach to accurately estimate the frequency pMHC-specific T cells, in a few cases, functionally competent pMHC-specific T cells are not stained by tetramers bearing the same pMHC (15). Such phenomenon, thought to concern preferentially $CD4^+$ T cells and self-Ag (15, 16), may be accounted for by either different affinity threshold required for pMHC tetramer binding to T cells and functional T cell activation (17) or significant pMHC tetramer dissociation rates (16). In these cases, the pMHC tetramer staining technique may lead to underestimate the frequency of pMHC specific cases, whereas in other cases, it provides a fair estimate pMHC-specific T cell frequency (18). Nevertheless, the approach first described by Moon et al. (9) and then adapted to human (19, 20) opens up new investigations of possible relationship between the frequency of naive T cells specific for a given pMHC and the overall structural characteristics of this pMHC. This is further supported by the fact that a higher frequency of naive specific T cells goes along with a higher diversity (6, 9, 21) and the diversity and the possible bias in TCR gene segment usage of the naive T cell repertoire specific for a given Ag correlate well with the diversity observed in the selected T cell repertoire after primary immune response (9, 10, 12).

In this study, we present ex vivo frequency measurements of T cells specific for several viral and tumoral peptides, presented by the HLA-A*0201 MHC class I molecule (referred to as A2 later on) and structural characterization of the corresponding pMHC complexes. Because the naive repertoires specific for the set of peptides studied cover a wide range of frequency (10^{-3}–10^{-6}), we have gathered a unique data set of pMHC Ags to tackle the structural bases of pMHC immunogenicity. By analyzing the frequency of T cells specific for given pMHC class I complexes not only within $CD8^+$ but also $CD4^+$ T cell peripheral subsets from both $A2^+$ and $A2^-$ individuals, we should in principle be able to unveil the relationships between pMHC structural features and several parameters, such as selection by the restricting HLA allele versus intrinsic TCR cross-reactivity for a given pMHC, and possibly establish rules that govern pMHC antigenicity. Overall, our data showed that there is no straightforward correlation between a single structural feature of the peptide–MHC complex and the frequency of specific naive T cells. They suggest instead that beyond thymic and peripheral selection processes, pMHC antigenicity is determined by a mixed contribution of several structural parameters, such as favored interaction with CD8 coreceptor or germline TCR loops, peptide prominence, and/or mobility. Even if the preferred usage of a given TRAV segment is often associated with high frequency of naive specific T cells, we demonstrated that the ability of the non–TRAV-encoded TCR loops to form productive contact with the peptide is also essential.

Materials and Methods

Protein expression and purification of peptide-HLA-A2

Both HLA-A*0201 H chains and β_2-microglobulin were produced separately as described previously (22–24). In brief, the A245V mutant of the HLA-A*0201 H chain tagged at the C terminus with a biotinylation sequence was cloned into pHN1 expression vector as described in Ref. 22. Recombinant proteins were produced as inclusion bodies in the XA90F'LaqQ1 *Escherichia coli* strain. The inclusion bodies were solubilized in 8 M urea, 50 mM MES (pH 6.5), 0.1 mM DTT, and 0.1 mM EDTA, incubated overnight at 277 K, and centrifuged 30 min at 100,000 × *g*. The supernatant was collected and frozen at 193 K. 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 synthetic peptide (GL Biochem) into 350 ml 100 mM Tris (pH 8), 400 mM, L-arginine HCl, 2 mM EDTA, 5 mM reduced glutathione, 0.5 mM oxidized glutathione, 2 Complete EDTA-free mixture 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 a fast protein liquid chromatography system equilibrated in a 10 mM Tris (pH 8) buffer. It was eluted with 100–150 mM NaCl and concentrated with Amicon-10 or Amicon-30 devices to reach a final protein concentration of 2.5–5 mg/ml.

Determination of tetramer-specific T cell frequencies and repertoire overlaps

The frequencies of naive human T cell directed against the various pA2 complexes used in this study were determined by peptide-MHC tetramer staining as described previously (19). Briefly, PBMCs were obtained from human CMV, hepatitis C virus, and HIV-seronegative donors (Etablissement Français du Sang, Nantes, France) by Ficoll density gradient centrifugation (LMS, eurobio, Les Ulis, France). A total of 10^8 freshly isolated PBMCs were incubated with 200 μl dye-conjugated tetramers (PE- and/or APC-labeled, 10 μg/ml) for 30 min at room temperature, and then washed with 15 ml ice-cold sorter buffer (PBS plus 0.5% BSA plus 0.2% EDTA). The tetramer-stained cells were then enriched using anti-dye Ab-coated immunomagnetic beads. The resulting enriched fractions were stained with anti-CD14, -CD16, -CD19, -CD3, -CD4, -CD8, -CD45RO, -CD11a, and -CD27 labeled mAb and 7-aminoactinomycin D (BD Biosciences). Stained samples were then collected on a Canto II flow cytometer (BD Immunocytometry Systems) and analyzed using FlowJo software (Tree Star). Counting beads (Life Technologies) were used to make sure that the entire cell sample had been collected. The number of naive ($CD45RO^-$, $CD11a^{low}$,$CD27^{bright}$) $tetramer^+$ CD8 or CD4 T cells was divided by the total number of naive CD8 or CD4 T cells within 10^8 PBMCs to calculate the frequency of Ag-specific naive T cells.

Crystallization, diffraction data collection, processing, and refinement

Crystals of the different peptide/HLA-A2 binary complexes were grown by the hanging drop technique at 293K. Crystals were first obtained in drops of 2 μl of the 3.5–5 mg/ml protein solution and 2 μl of the reservoir solution (9–15% PEG 6000, 0.1 M trisodium citrate [pH 6.5] or 0.1 M MES [pH 6.5]

or 0.1 M sodium cacodylate [pH 6.5], 0–0.1 M NaCl). Diffracting crystals were obtained by microseeding technique: freshly prepared seeds were first made from peptide/HLA-A2 crystals with seed bead (Hampton Research) and a stabilization solution (mother liquor with PEG 6000 concentration increased to 20%) and then diluted in the same solution, and 1 μl of the seed solution was added on the crystallization drop after 24 h.

Prior to being flash frozen in liquid nitrogen for diffraction experiment, crystals were soaked in a cryoprotectant solution (mother liquor with PEG 6000 concentration increased to 30%). All data were collected at 100K on beamlines either ID14-eh2, ID14-eh3, ID14-eh4, ID23-eh1, ID23-eh2, or ID29 of the European Synchrotron Radiation Facility at different wavelengths and using the specific charge-coupled device detector depending of the beamline (MARCCD 165 mm or ADSC Quantum 4 or ADCS Quantum 135R).

Data processing was performed using XDS (25) and is summarized in Supplemental Table I. Peptide-HLA-A2 complex crystals belong to the monoclinic space group $P2_1$ with a = 51–53 Å, b = 79–81 Å, c = 55–57 Å, and β=112–114° as unit cell parameters.

Structure determination of the different HLA-A2 binary complexes was performed by molecular replacement with AMoRe (26) and using NVL-HLA-A2 crystal structures (Ref. 24, Protein Data Bank entry code 3MRC; http://www.rcsb.org/pdb/home/home.do) as initial model. The refinement protocol used for all complex structures included several cycles of refinement with REFMAC (27), followed by manual model rebuilding with Coot (28), until no interpretable electron density could be identified in the residual map. Final refinement statistics and Protein Data Bank entry are summarized in Supplemental Table I. Solvent exposed surfaces were calculated with AREAIMOL program (27), using a 1.4-Å diameter probe.

Results

Structure determination of viral and tumor peptides in complex with HLA-A2

To identify possible general relationships between structural properties of a pMHC complex and the frequency of specific naive T cells, we have solved the crystal structure of eight viral and tumoral peptides and/or some of their variants in complex with A2 (Supplemental Table I). The list of peptides includes an immudominant epitope of human CMV pp65 tegument protein, $pp65_{495–503}$ (**NLV**PMVATV, from now on referred to as NLV), one epitope from EBV, $BMLF1_{280–288}$ (**GLC**TLVAML referred to as GLC), one epitope from hepatitis C virus, $NS3_{1073–1081}$ (**CIN**GVCWTV referred to as CIN), a modified decapeptide from Melan-A MART1 protein $MART1_{26–35}$ (**ELA**GIGILTV referred to as ELA), and a decapeptide from the human PG transporter, $pgt_{178–187}$ (**LLA**GIGTVPI, referred to as LLA), identified as a Melan-A MART1 analog (29). Overall, the crystal structures of eight pMHC complexes have been refined at resolutions ranking from 1.1 Å (GLC-HLA-A2) to 2.35 Å (ELA_{I5W}-A2) allowing the determination of structural characteristics such as the peptide solvent exposed surface, accurate positioning of the side chains, and thermal motion assessment (see Supplemental Table I for details). For fruitful comparisons, we also included in the analysis the structures of wild-type peptides or variants already published and present in the Protein Data Bank (wild-type NLV: 3GSO (12); wild-type [EAA] and modified [ELA] Melan-A/MART1: 2GT9 (30), 1JF1 (31)) (Table I). The overall structure of the A2 molecule is very similar in all the peptide-A2 complexes determined in this study (root mean square difference [r.m.s.d.] ranging from 0.12 to 0.45 Å for the α1α2 HLA domain and with the 3GSO pdb structure used as a reference) and comparable to those already present in the Protein Data Bank. For all the structures, there was no ambiguity in interpreting the electron density for the peptide backbone. According to the conformation of the peptide backbone in the A2 peptide–binding groove, the structures can be split into three families of conformation for nonapeptides (1: CIN, CIN_{N3S}, CIN_{C6V}, that are similar to NLV; 2: GLC, 3: NLV_{V6C}, NLV_{V6G}; see Fig. 1A) and one family for decapeptides (ELA_{I5W}, LLA, that are similar to ELA, see Fig. 1B). Among all the peptide variants studied, most of them adopted a conformation very similar to that of the wild-type peptide (r.m.s.d. ranging from 0.18 to 0.61 Å), except for two of them, NLV_{V6C} and NLV_{V6G}, which adopted a different conformation at position 4-6 (r.m.s.d. of 1.37 and 0.94 Å, see Fig. 1C). The latter two variants define the third family of backbone conformations (see above). Indeed, when NLV Val6 is changed to either glycine or cysteine (amino acid at position 6 in CIN), the NLV conformation changed in such a way that the methionine at position 5 becomes buried in the MHC groove, whereas the side chain at position 6 becomes solvent exposed. Such a change was unexpected as in the CIN peptide the cysteine at position 6 is buried in the cleft and acts as an anchor residue.

To widen our studies, we have included in our analysis three well studied T cell epitopes for which the crystal structure was already known: $MP_{58–66}$ from influenza [GILGFVFTL, pdb entries 1HHI (32) and 2VLL (33)], HIV p17 GAG [SLYNTVATL, pdb entries 1T21 (34) and 2V2W (35)] and minor histocompatibility Ag HA-1 [VLHDDLLEA, pdb entry 3FT3 (36)]. The backbones of HIV p17 GAG and HA-1 peptides have a conformation similar to NLV (Fig. 1E), whereas $MP_{58–66}$ has a conformation closer to that adopted by GLC (Fig. 1F).

Frequency measurements of naive T cells specific for the selected peptide-A2 complexes

The frequency of naive T cells specific for NLV-A2, CIN-A2, CIN_{N3S}-A2, CIN_{C6V}-A2, ELA-A2, ELA_{I5W}-A2, LLA-A2, GIL-A2, SLY-A2 and VLH-A2 were assessed for $A2^+$ and $A2^-$ individuals and for $CD8^+$ and $CD4^+$ T cells as described by Legoux et al. (Table I, Fig. 2) (19). Therefore, these frequency data enable us to investigate the impact of thymic selection, the presence of CD8 coreceptor and possible T cell repertoire cross-reactivity. Because naive $CD4^+$ T cells were not selected by MHC class I molecule in thymus, the measured frequency of naive $CD4^+$ T cells, ranging from 5×10^{-7} to 7×10^{-6} (Table I), is a fair estimate of the basal T cell response to any pMHC complex, regardless of the coreceptor and thymic selection. The impact of CD8 coreceptor on the frequency could be assessed by comparing $CD4^+$ and $CD8^+$ T cell frequencies in $A2^-$ individuals, avoiding any bias possibly induced by the presence of the peptide during the selection process. An increase by a factor ranging from 5 to 25 could be accounted for by the presence of the CD8, depending on the Ag (Fig. 2C, 2D). The comparison of the $CD8^+$ T cells frequency in $A2^+$ and $A2^-$ individuals allowed to assess the impact of the MHC restricting allele presence during the thymic selection process. Heterogeneous results were obtained among the peptides studied. For ELA and LLA, the presence of the A2 haplotype was associated with a 10 times increase in the frequency, although it had no impact on the percentage of high-avidity specific T cells (19). This suggested that ELA and LLA could be present in the thymus, albeit at low level as no deletion of high avidity T cell clones is observed. For NLV and CIN_{N3S}, a slight or a significant frequency decrease was observed in $A2^+$, as compared with $A2^-$ individuals, suggesting that neither the NLV peptide nor a close homolog is present in the thymus. Moreover, the frequency decrease could be explained by the fact that T cell clones specific for some pMHC (e.g., NLV-A2 or CIN_{N3S}-A2) could be more cross-reactive with other self- or environmental peptides Ags. This cross-reactivity could lead to the negative selection of many specific T cell clones in the thymus (lower frequency) and to the activation of numerous specific T cells in the periphery (lower number of naive T cells in individuals that never encountered the pathogen). Overall, the frequency ratio between $A2^+$ and $A2^-$ individuals was correlated with the percentage of naive cells among T cells specific for a given pMHC and could be an indicator of the cross-reactivity

Table I. Peptide solvent accessible surface, overall mobility, and frequency of HLA-A2/peptide–specific T cell in the naive repertoire

Peptide (Sequence)	PDB Entry Code	Peptide/A2 or TCR/Peptide/A2 Complex	Peptide Solvent Accessible Surface Å^2	A2 α1α2 Helices Mobility (Average Main Chain B/ Average Side Chain B) Å^2	Peptide Mobility (Average Main Chain B/ Average Side Chain B) Å^2	Peptide Relative Mobility (Main Chain/Side Chain)	A2$^+$ Individuals: Peptide-Specific Naive CD8$^+$ T Cell Frequency ($\times 10^{-5}$)	A2$^+$ Individuals: Peptide-Specific Naive CD4$^+$ T Cell Frequency ($\times 10^{-5}$)	A2$^-$ Individuals: Peptide-Specific Naive CD8$^+$ T Cell Frequency ($\times 10^{-5}$)	A2$^-$ Individuals: Peptide-Specific Naive CD4$^+$ T Cell Frequency ($\times 10^{-5}$)
pp65$_{495–503}$ (**NLV**PMVATV)	3GSO	NLVWT-A2	360	15.2/18.5	16.2/17.9	−0.12/0.16	0.35 ± 0.30	0.20 ± 0.25	0.64 ± 0.64	0.14 ± 0.17
	3MRC	*NLV$_{V6C}$-A2**	277	17.6/21.4	24.7/27.5	0.68/1.06				
	3MRD	*NLV$_{V6G}$-A2**	251	15.5/20.4	26.3/26.3	1.10/1.10				
	3GSN	RA14-NLV$_{wt}$-A2	40 (368)	44.4/44.4	37.1/37.3	−0.95/−0.93				
NS3$_{1073–1081}$ (**CIN**GVCWTV)	*3MRG*	*CIN-A2*	252	11.5/14.5	16.6/17.4	0.69/0.87	4.0 ± 2.3	0.13 ± 0.084	2.1 ± 1.9	0.16 ± 0.13
	3MRH	*CIN$_{N3S}$-A2*	323	41.4/42.8	62.9/63.8	2.62/2.73	0.45 ± 0.32	0.18	1.1 ± 0.79	0.046 ± 0.039
	3MRL	*CIN$_{C6V}$-A2*	295	29.4/31.6	50.6/51.0	2.36/2.42	8.5 ± 3.9	0.24 ± 0.12	7.3 ± 3.8	0.31 ± 0.17
BMLF1$_{280–288}$ (**GLC**TLVAML)	*3MRE*	*GLC-A2*	327	8.3/9.6	12.0/13.5	45/40				0.44 ± 0.3
MART1$_{26–35}$ (**ELA**GIGILTV)	1JF1	ELA-A2	322	19.2/23.5	20.7/23.7	−0.26/0.09	60 ± 30	0.68 ± 0.27	12 ± 8.3	0.52 ± 0.20
	3MRO	*ELA$_{I5W}$-A2*	386	14.4/15.4	16.0/16.3	0.17/0.22	1.06 ± 0.85	0.038 ± 0.029	0.39 ± 0.28	0.031 ± 0.019
	3HG1	MEL5-ELA-A2	53 (309)	33.1/36.3	34.0/33.9	−0.22/-0.14				
PGT$_{178–187}$ (**LLA**GIGTVPI)	*3MRR*	*LLA-A2*	306	29.9/32.7	35.1/34.3	0.53/0.42	43 ± 27	0.37 ± 0.13	11.8 ± 7.0	0.63 ± 0.26
MP$_{58–66}$ (**GIL**GFVFTL)	1HHI	GIL-A2	250	21.7/22.6	27.2/25.7	0.62/0.43	0.63 ± 0.49	0.064 ± 0.044	2.3 ± 3.4	0.12 ± 0.10
	2VLL	GIL-A2	251	20.7/22.9	20.4/22.2	−0.21/0.06				
	1OGA	JM22-GIL-A2	35 (255)	23.4/27.2	17.8/19.2	−0.96/−0.79				
HIV p17 gag (**SLY**NTVATL)	1T21	SLY-A2	319	20.3/21.8	37.5/41.9	1.95/2.51	0.35 ± 0.20	0.055 ± 0.057	0.31 ± 0.20	0.057 ± 0.042
	2V2W	SLY-A2	307	20.7/23.0	23.5/23.7	0.26/0.29				
HA-1 (**VLH**DDLLEA)	3FT3	VLH-A2	386	18.7/20.2	22.0/23.5	0.48/0.76	0.23 ± 0.17	0.02 ± 0.01	0.46 ± 0.53	0.045 ± 0.050
Meloe-1 (**TLN**DECWPA)	unpublished structural data[a]	Meloa1-A2	338	18.9/20.1	18.2/21.1	−0.34/0.08	6[b]			

The total peptide solvent accessible surfaces were calculated with areaimol from ccp4 with a probe radius of 1.4Å. Overall peptide mobilities were evaluated through the normalized relative B factors calculated as $(< B >_{peptide} - < B >_{HLA\text{-}A2\ \alpha1\alpha2}) / \sigma(B_{HLA\text{-}A2\ \alpha1\alpha2})$. Protein Data Bank codes in italics indicate HLA-A2/peptide structures from the current study. *Variants for which a conformational change was observed, in comparison with the wild-type peptide.

[a] J.-B. Reiser and D. Housset, unpublished structural data.

[b] Meloe-1–specific CD8$^+$ T cell frequency has been estimated to be 1/10 of that of ELA-specific CD8$^+$ T cell, according to Godet et al. (37).

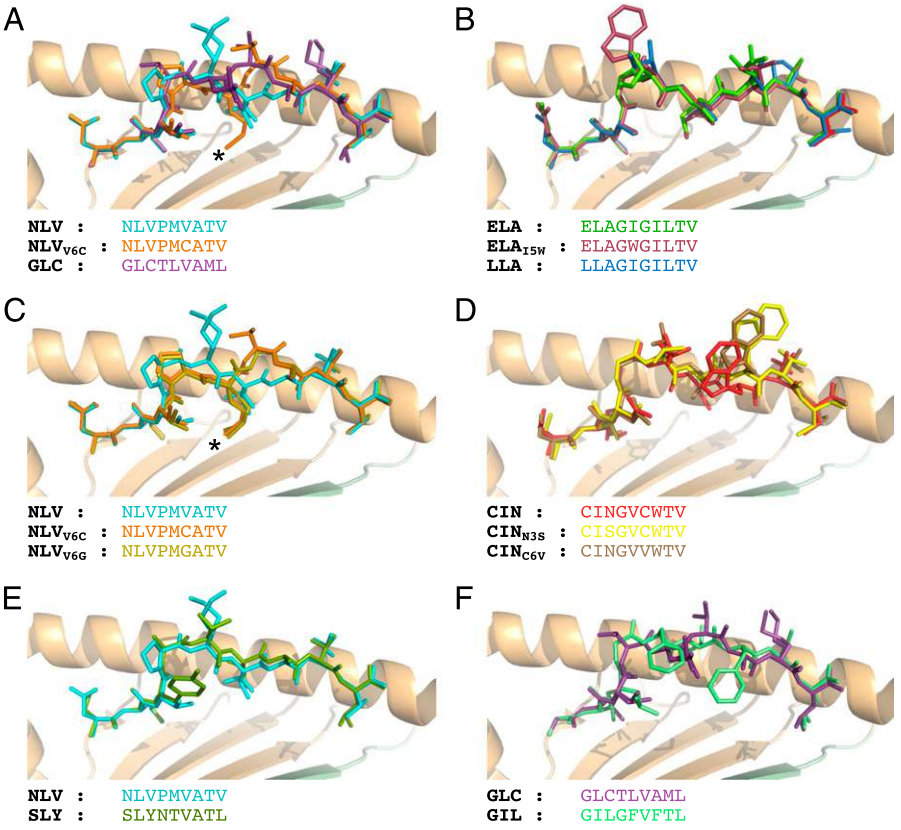

FIGURE 1. (**A**) Superposition of the NLV (in cyan), NLV$_{V6C}$ (in orange) and GLC (in magenta) peptides. They represent the 3 conformations of nonapeptides encountered in this study. (**B**) Superposition of the ELA (in green), ELA$_{I5W}$ (in red), and LLA (in sky blue) decapeptides. (**C**) Superposition of the NLV peptide (in cyan) with its two variants, NLV$_{V6C}$ (in orange) and NLV$_{V6G}$ (in olive green). (**D**) Superposition of the CIN peptide (in red) with its two variants, CIN$_{N3S}$ (in pale yellow) and CIN$_{C6V}$ (in deep-olive green). (**E**) Superposition of the NLV (in cyan) and the SLY (in lime green) peptides. (**F**) Superposition of the GLC (in magenta) and the GIL (in light lime green) peptides. The flip of the NLV$_{V6C}$ and NLV$_{V6G}$ variant is indicated with a black star (A and C).

of the T cell repertoire (19). However, the analysis of the frequency of CD8$^+$ and CD4$^+$ naive T cells specific for the A2/peptide complexes studied in this paper indicated that although the absolute value of frequencies differed, the hierarchy of frequencies between peptide Ags was overall well conserved for CD8$^+$ and CD4$^+$ T cells, as well as for A2$^+$ and A2$^-$ individuals (Fig. 2) (19), suggesting a significant role of intrinsic pMHC structural properties. Thus, the structural and frequency data presented in this study should allow us to identify a few parameters underlying the frequency of naive T cells specific for a given Ag.

No systematic correlation between peptide prominence and frequency of specific naive T cells

We have first investigated whether any general coincidence between peptide prominence from the MHC structural groove and the diversity of the naive T cell repertoire specific for a given pMHC could be observed, as suggested by Turner et al. (6). As the frequency of naive-specific T cell has been shown to be a good indicator of the diversity of such a repertoire (9), we have evaluated the frequency of naive specific T cells for several peptides presented by HLA-A2, in both the CD4 and CD8 compartment and both A2$^+$ and A2$^-$ individuals. These frequency data have been gathered in Table I with different structural characteristic of the peptide/MHC complexes derived from crystallographic structures. As shown in Fig. 2A, there is no obvious global correlation between the solvent exposed surface of the peptide and the frequency of CD8$^+$ naive T cells specific for a given peptide in A2$^+$ individuals. The same observation is made for A2$^-$ individuals (Fig. 2C). High-frequency values are observed for ELA and LLA, whereas their exposed surfaces are within the average for nona- and decapeptides (300–320 Å^2). Low-frequency values are observed for both the most buried (GIL, 250 Å^2) and the most bulgy peptide (VLH, 386 Å^2). At first glance, this observation does not support the assumption that a larger solvent-exposed surface does recurrently constitute a significant advantage (6) or drawback (8) for a peptide to be recognized by T cells. A more in depth analysis of Fig. 2A indicates that we can define two groups of pMHC Ags, with different behaviors. One group includes 8 pMHC for which the frequency of naive specific CD8$^+$ T cells in A2$^+$ individuals is rather low ($\leq 10^{-5}$, GIL, SLY, CIN$_{N3S}$, NLV, SSL, KLV, VLH, ELA$_{I5W}$). For this group, there seems to be no impact on the prominence of the peptide on the frequency. The second group includes the four pMHC for which the frequency is significantly above 10^{-5} (ELA, LLA, CIN, and CIN$_{C6V}$). For this one, the solvent exposed surface appears to be reasonably correlated with the frequency, with an optimal value around 320 Å^2, as observed for ELA. Interestingly, the T cell repertoire specific for ELA and LLA exhibit a strong bias on Vα segment usage. Indeed, a prevalence of TRAV12-2 is observed in TCRs specific for these peptides, reaching 80% of the repertoire for ELA and 40% for LLA. The frequency of naive specific T cells for a tumoral Ag, Meloe-1, presented by A2 has also been found to be quite high, and estimated to 1/10 of the one observed for ELA (37). Meloe-1–specific T cells also exhibited a strong bias in the usage of one Vα segment (TRAV19). These data suggest two main features: first, the ability of a given TRAV segment to form favorable interaction with a given pMHC complex significantly contribute to the high frequency of T cell specific for this pMHC complex. Second, within a conserved TCR docking mode constrained by the TRAV-pMHC interface, an optimal peptide solvent exposed surface may statistically favors an adequate molecular fit with TCRs.

Impact of substituting bulky residues at positions exposed to the TCR

To further investigate the impact of the peptide solvent exposed surface, we aimed at increasing the prominence of a peptide for which the frequency of specific T cells is already high by substituting a small side chain residue for a bulkier one at positions that were

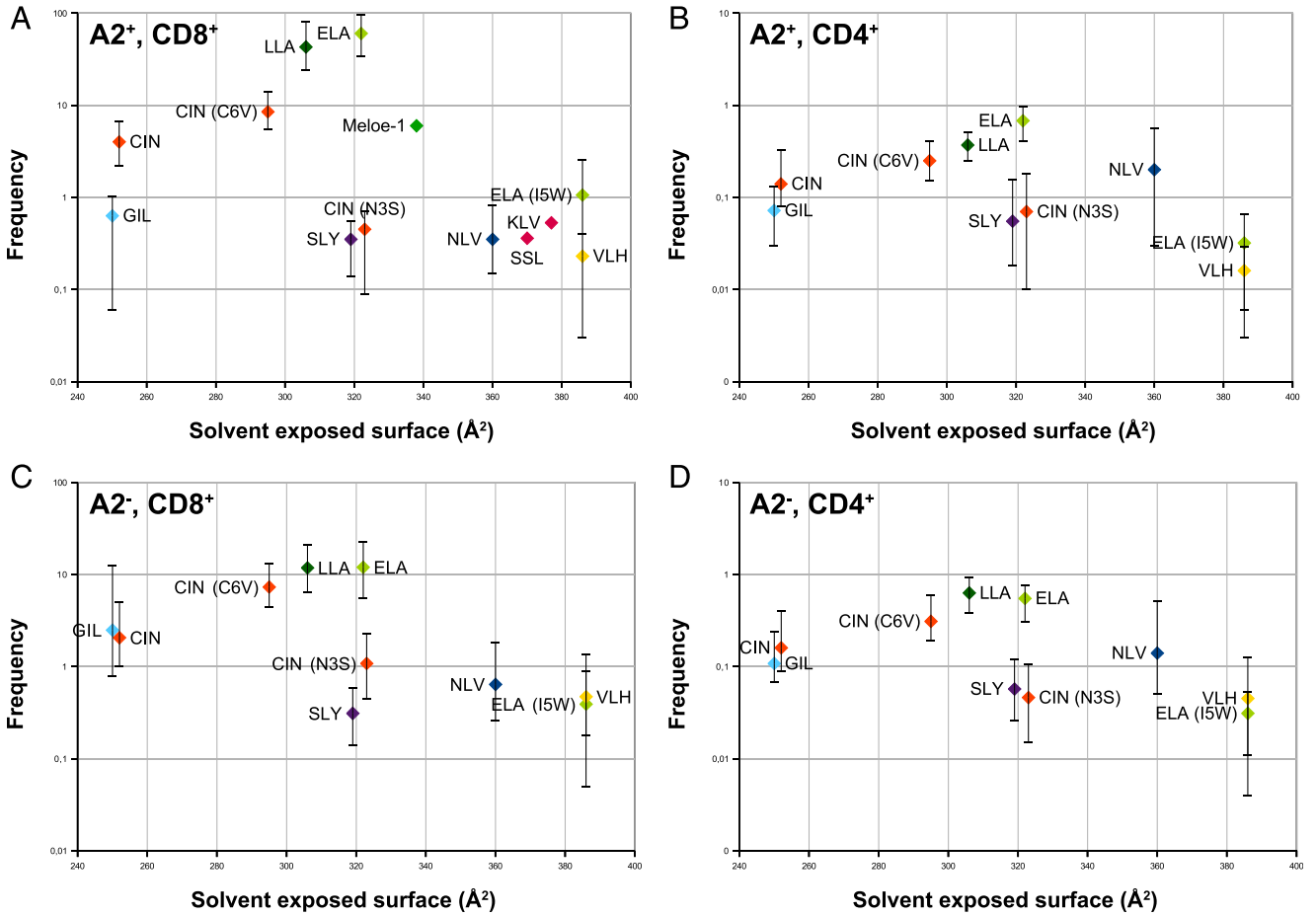

FIGURE 2. Diagram of the frequency of naive specific T cells, as a function of peptide-solvent exposed accessible surface. (**A**) Frequency of naive $CD8^+$ T cell in $A2^+$ individuals. (**B**) Frequency of naive $CD4^+$ T cell in $A2^+$ individuals. (**C**) Frequency of naive $CD8^+$ T cell in $A2^-$ individuals. (**D**) Frequency of naive $CD4^+$ T cell in $A2^-$ individuals.

accessible to the TCR. With this aim, the isoleucine at position 5 of the ELA peptide was replaced by a tryptophan. The structure of the ELA_{I5W} variant confirmed that the tryptophan side chain was actually exposed to the solvent, leading to a 64 $Å^2$ (i.e., 20%) increase of the peptide solvent exposed surface (Fig. 1B, Table I). Unexpectedly, this modification leads to a 60 times decrease of the frequency of naive specific T cells. At first glance, it suggests that the Ile to Trp substitution impairs the ability of most TRAV12-2 bearing T cells to recognize the pMHC, the T cell frequency for this variant being close to the basal level of T cell frequency for any given pMHC ($0.5\ 10^{-6}$–$0.5\ 10^{-5}$, in the presence of the proper coreceptor, i.e., CD8). In previously published data, the same ELA variant was shown to be able to activate five of seven ELA-A2–specific T cell clones (38), indicating that some TRAV12-2 containing TCR could productively bind both ELA- and ELA_{I5W}-A2 complexes, although ELA- and ELA_{I5W}-specific T cell repertoires show a limited overlap (Fig. 3). In the crystal structure of the Mel5-TCR-ELA-A2 complex (39), the isoleucine side chain in position 5 is caged between CDR1α, CDR2α, and CDR3β. When the ELA_{I5W} variant structure is superimposed to the ELA-A2 moiety of the Mel5-TCR-ELA-A2 complex, the tryptophan side chain engenders steric clashes either with the CDR1α, CDR3α, CDR3β, or with the MHC α2 helix (Supplemental Fig. 1). The same observation is made with the two other TCR-ELA-A2 structures published recently (40) and also by modeling the other plausible rotamers of the tryptophan side chain. According to the three available TCR-ELA-A2 ternary complexes, the footprint of the Vα domain is quite conserved, with the CDR1α forming a pivotal point. Therefore, activation of some ELA-A2–specific T cells by the ELA_{I5W} variant may be explained by structural adaptation of both CDR3s to create a larger pocket at CDR3 loops interface to fit the larger Trp side chain. However, the nature and the length of the CDR3β may dictate the degree of tolerance for amino acid substitution at position 5. The drastic drop in the frequency suggests that this type of promiscuous clones only represents a small fraction of the ELA-specific T cell repertoire. Therefore, the high frequency of ELA-A2–specific T cell does not seem to be only accounted for by a favorable interaction between Vα germline encoded loops and the pMHC. An optimally permissive peptide central region, easily fitted by both CDR3s also seems to play a very significant role, further supporting the proposed hypothesis of an optimal peptide solvent exposed surface favoring high frequency of specific T cells.

Does the mobility of the peptide impact on Ag-specific T cell frequency?

Another parameter that can influence the antigenicity of a given pMHC is the mobility of the peptide within the MHC peptide–binding groove. The interaction between the TCR and a given pMHC is often associated with significant conformational changes of the CDR loops and a negative entropic contribution to the free energy of binding (41–44). In some cases, changes in the peptide conformation are also observed upon TCR binding (45). As the affinity of the TCR for a cognate pMHC is rather moderate [i.e., in

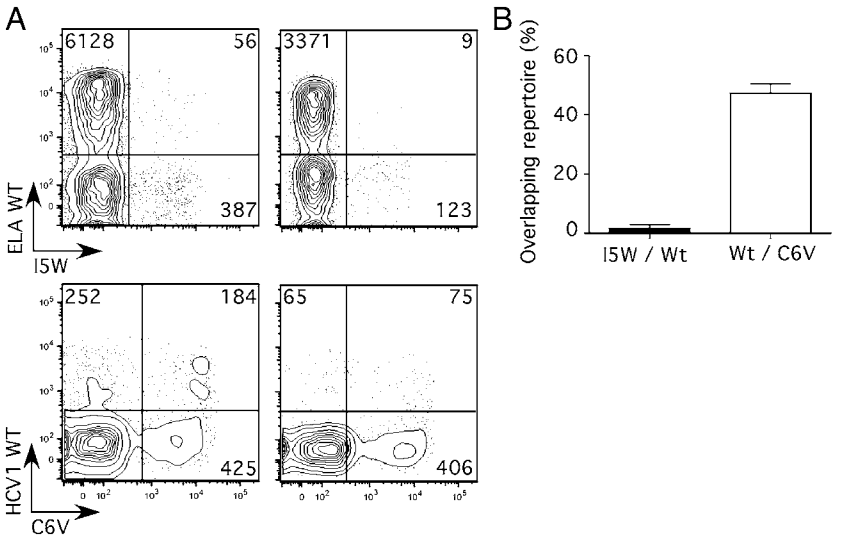

FIGURE 3. Ex vivo analysis of the overlap between T cell repertoires directed against wild type and variants ELA and CIN-A2 complexes. (**A**) Representative dot blots showing the overlaps between T cell repertoires directed against ELA/A2 versus ELA_{I5W}/A2 and CIN/A2 versus CIN_{C6V}/A2 in 10^8 PBMC of two donors. Analysis was performed on dump$^-$CD3$^+$CD8$^+$ gated events. The dump channel corresponds to pooled events positive for 7-actinomycine D (dead cells), CD19 (B cells), CD16 (NK cells) or CD14 (monocytes). (**B**) Percentage of ELA_{I5W} and CIN specific T cell repertoires overlapping with ELA and CIN_{C6V} specific T cell repertoires, respectively. Results based on the analysis of ELA/A2-, ELA_{I5W}/A2 CIN/A2-, and CIN_{C6V}/A2-specific T cell repertoires in three donors.

the 1–100 μM range (46, 47)], the energetic cost of stabilizing part of the TCR–pMHC interface can represent a significant contribution to the global free energy of binding. Therefore one can expect that a less mobile peptide could favor TCR binding, as it would reduce the entropic cost of further stabilizing the peptide. Whether this contribution might play a significant role remained to be assessed. The comparison of the average thermal motion of the A2 α-helices and the peptide may indicate how strongly the peptide is anchored in the MHC peptide-binding groove. However, B factors derived from crystallographic structure Protein Data Bank files cannot be directly compared, because they strongly depend on the resolution of the structure, on the overall B factor of the crystal and on the stereochemical restraint scheme used for the refinement process. To overcome these difficulties, we used a normalized relative B factor as defined in Table I, which is positive if the peptide is more mobile than the α1-α2 A2 helices, and negative if it is less mobile. A relative B factor value of 1 means the B factor is 1 σ above the mean B factor the HLA-A2 structure, σ being the SD of the HLA-A2 B factors distribution. For the three peptide-A2 complexes for which the structure of a complex with a TCR is available in the PDB and for which we have collected frequency data [JM22-MP58-66-A2 (7), RA14-NLV-A2 (24), and Mel5-ELA-A2 (39)], the relative B factor of the peptide is negative, showing that in the ternary complex the peptide mobility is lower than that of the α1-α2 A2 helices. For the corresponding binary complexes, the relative B factor is always significantly higher, ranging from −0.21 to 0.62 for the peptide main chain (Table I). The relative B-factor is therefore a good indicator of the stabilization of the peptide upon TCR binding, even in the absence of significant conformational changes. By plotting the overall relative B factor of the peptide as a function of the frequency of naive T cell specific for the given peptide, we observed no clear correlation between the peptide mobility and the frequency of pMHC-specific T cells. Most peptides have a quite low average relative B factor, in the −0.3 to +0.6 range, whereas two peptides are significantly more mobile than the others: CIN_{N3S} and CIN_{C6V} (Fig. 4). The CIN_{N3S} variant is not surprisingly the most mobile (relative B factor of 2.62; Table I, Fig.3), as it was designed to disrupt two hydrogen bonds that tight the CIN peptide to the Gln155 residue by a side chain–side chain interaction (Supplemental Fig. 2). However, the peptide basically keeps the wild-type conformation but a positional shift toward the outside of the MHC groove is observed that ranges from 0.3 to 0.7 Å for extremities (residues 1, 2, 3, 9) up to 1.46 Å at position 6, associated with a significant increase of the peptide solvent exposed surface (323 Å^2 instead of 252 Å^2; Fig. 1D). The mutation is associated with a drastic decrease of the frequency of specific T cells (10 times in comparison with the wild-type peptide) and becomes similar to that of NLV-specific T cells. This drop in frequency is thus likely due, at least in part, to this significant higher mobility that can impact both the TCR-pMHC complex stability and the pMHC half-life. CIN_{C6V} is also quite more mobile, albeit to a lesser extent (relative B of 2.36). However and quite surprisingly, the frequency of naive-specific T cell for this peptide is twice the one of T cells specific for the wild-type peptide, clearly indicating that other parameters can overwhelm the energetic cost of an increased mobility (see below).

Bases for the different frequencies of T cells specific for either NLV or CIN

In our approach to understand the bases that govern the immunogenicity of different peptide-A2 complexes, we have compared into more details the NLV-A2 and CIN-A2 complexes. NLV and CIN are nonapeptides, which adopt a similar conformation in the HLA-A2 peptide-binding groove (Supplemental Fig. 3), the CIN peptide being just slightly more buried (Table I), especially in the central region of the peptide where positional difference reaches 0.9 Å (Cα distance at position 6). Therefore, the difference in frequency of naive specific T cells observed for these two peptides (NLV: 0.35×10^{-5}, CIN 4.0×10^{-5}) is likely to originate from the nature of the side chains. We thus designed two variants of these two peptides by substituting CIN residues for NLV's and vice versa, at the secondary anchoring position 6. The NLV_{V6C}-A2 structure showed that the peptide adopted a completely different conformation in comparison with the wild-type Ag (see above and Fig. 1C). This suggests that the conformational stability of NLV relies on the buried valine at position 6, presumably because exposing it to the solvent would be energetically too costly. This observation highlights the significance of structural characterization in evaluating the impact of peptide amino acid changes on immunogenicity. In this study, the amino-acid modifications has changed the register of side chains that is presented to the TCR, strongly impacting a large portion of the epitope and preventing any simple analysis of the role of the side chain type at NLV position 6 for TCR recognition.

The CIN_{C6V} variant was designed to substitute the valine amino-acid found at position 6 in NLV to the cysteine of CIN. If the extremities remain basically unchanged (shifts of 0.2–0.3 Å for

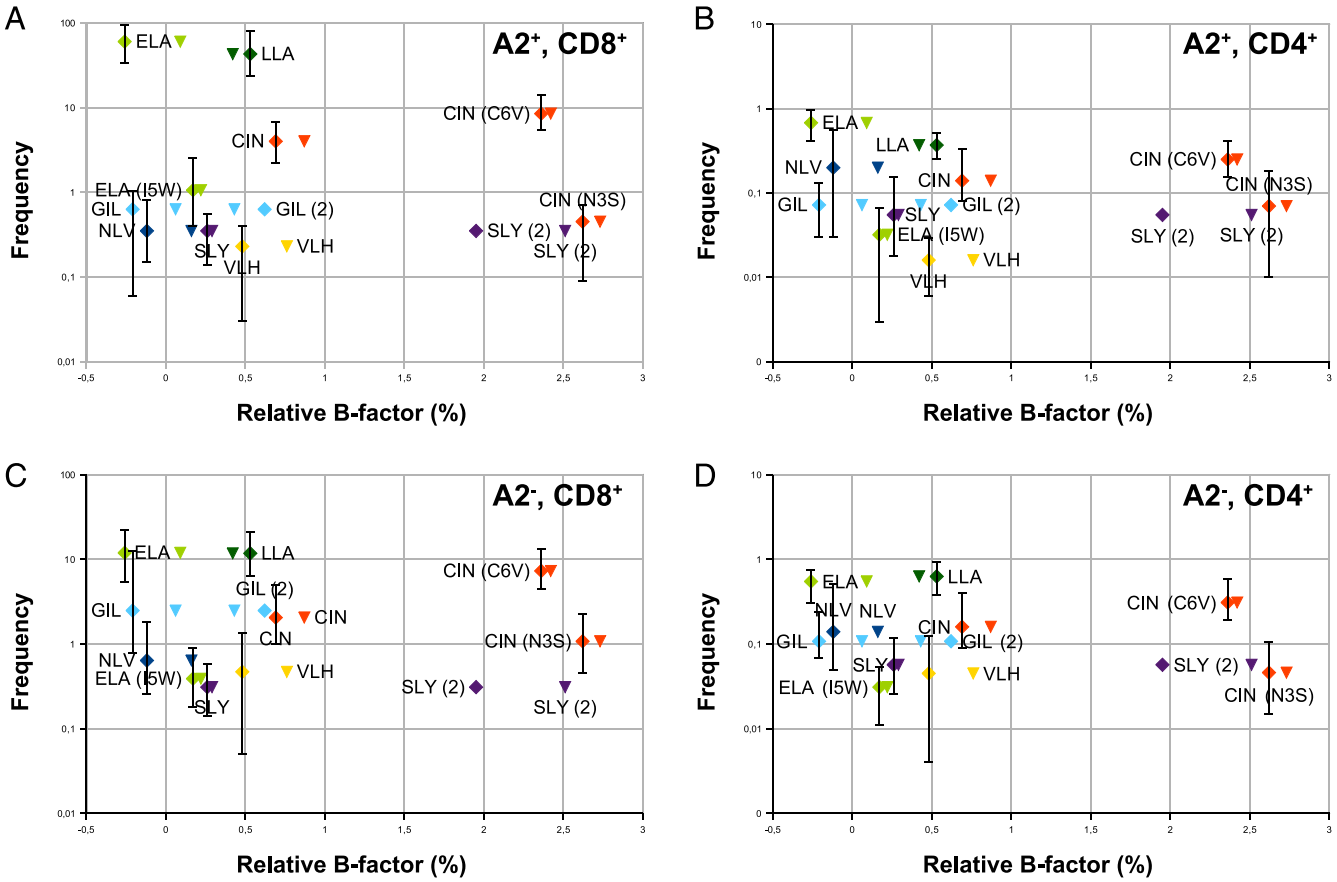

FIGURE 4. Diagram of the frequency of naive specific T cells, as a function of peptide average relative B factor (see Table I for definition). Squares represent the peptide average relative B factor for main chain atoms, triangles the average for side chain atoms. (**A**) Frequency of naive $CD8^+$ T cell in $A2^+$ individuals. (**B**) Frequency of naive $CD4^+$ T cell in $A2^+$ individuals. (**C**) Frequency of naive $CD8^+$ T cell in $A2^-$ individuals. (**D**) Frequency of naive $CD4^+$ T cell in $A2^-$ individuals.

residues 1, 2, 3, and 9), we observe a 1.4 Å displacement of residues 6 and 7 toward the outside of the groove to avoid steric hindrance with A2 residues 69, 70, and 73. The path of NLV peptide backbone falls in between that of CIN peptide (more buried) and CIN_{C6V} peptide (less buried), but remains more similar to CIN. Both the CIN_{C6V} solvent exposed surface (295 $Å^2$ instead of 252 $Å^2$ for the wild-type CIN peptide) and mobility (mean relative B factor of 2.36) are increased compared with the CIN, albeit to a lesser extent than the one observed for the CIN_{N3S} variant. Unexpectedly, the frequency of T cells specific for this peptide variant is not closer to that of NLV specific T cells, but instead twice higher than the one observed for the wild-type CIN peptide. Moreover, ~50% of the CIN-specific T cell repertoire is also cross-reactive for CIN_{C6V} indicating that many TCRs can adapt to both the wild type and the C6V variant CIN epitope (Fig. 3). How can moderate changes in peptide position lead to drastic modification of specific T cell frequency? First, the position of the Trp7 side chain is the most significantly impacted by the mutation at positions 3 and 6. The overall displacement of the side chain outside of the groove reaches 2.6 and 4.6 Å for the CIN_{C6V} and CIN_{N3S} variants, respectively, and is further illustrated by the increased solvent exposed surface of the side chain (Fig. 1D, Table I). Thus, a more solvent exposed Trp side chain may facilitate the interaction with TCRs, as observed previously for pMHC class II murine complexes (21). The increased prominence harbored by the CIN_{C6V} variant possibly represents an optimum for productive interactions with the TCR CDR3 loops, whereas the one of CIN_{N3S} may be beyond this optimum. The fact that half of T cells that recognize CIN also recognize CIN_{C6V} suggests a similar TCR docking mode on these two closely related epitopes. Second, the CIN_{C6V} variant has its extremities significantly more stable than the central region, indicating a more conserved anchoring in the MHC peptide–binding groove in comparison with CIN_{N3S}. Thus, the increased CIN_{C6V} mobility may be sufficient to favor TCR interaction possibly throughout an increased plasticity at the TCR–pMHC interface, but not enough to impaired pMHC stability. Third, the A2 Gln155 also adopts different position. In the wild-type CIN-A2 structure, the Gln155 side chain forms two hydrogen bonds with the side chain of CIN Asn3 and is also at hydrogen bond distance from the Trp7 Nε1 atom. In the CIN_{C6V}, there is a 2.3 Å shift of the amine group. One hydrogen bond with CIN Asn3 is maintained whereas the one with Trp7 is lost (Supplemental Fig. 2). In CIN_{N3S}, the Gln155 conformer is completely different, with its amine group pointing outside of the peptide binding groove and forming no hydrogen bonds with the peptide. As Gln155 is one of the three MHC class I residue that have been almost always seen to be contacted by the TCR (5), one cannot exclude that the repertoire of specific T cell may also depend on the conformation of this central A2 residue.

Discussion

One of the much-debated questions regarding TCR recognition of peptide–MHC complexes is whether some structural signatures of an efficient T cell response can be identified. In the current study, we have looked for possible connections between structural characteristics of a few different class I pMHC complexes and the

frequency of specific naive T cells in the whole peripheral T cell repertoire, reflecting their antigenicity. On the one hand, the most buried peptide (CIN) presenting a quite featureless pMHC surface to the TCR, is associated with a rather high frequency of naive specific T cells. On the other hand, the highest frequency is observed for peptides (ELA, LLA) for which the protuberance is significantly lower than the one of the least antigenic peptides (NLV, HA-1). Moreover, introducing a bulkier residue on ELA central region, increasing the solvent exposed surface by 20% led to a decrease of the frequency of specific T cells by a factor of 60. Our data seem at odds with the concept that the number of TCR able to productively interact with a featureless pMHC surface might be significantly lower, because of the need of specific amino acids and a specific spatial distribution to interact with a buried peptide, as previously illustrated by the recognition of the influenza Ag MP_{58-66}-A2 by the immunodominant public TCR JM22 (7). The data obtained by Wynn et al. (8) on two hCMV peptides presented by HLA-B*3508 also showed that a quite prominent 12-mer peptide (427 $Å^2$ of peptide solvent exposed surface) selected much less diverse highly biased T cell–specific repertoire, whereas a quite buried 8-mer peptide (220 $Å^2$ of peptide solvent exposed surface) selected a diverse T cell repertoire. In this context, the observations made by Turner et al. (6) on two influenza peptides presented by the murine MHC molecule H-$2D^b$ cannot be generalized. Instead, the whole set of data suggests that an optimal prominence, corresponding to a peptide solvent exposed surface in the range of 280–330 $Å^2$, may statistically favors productive interaction with TCRs and could more easily induce high frequency of specific T cells. The different conformations and solvent exposures of the Trp7 in CIN and their variants as well as the data obtained on the ELA_{I5W} variant may also support this concept of an optimal solvent exposed surface, that enable a maximum number of interatomic contacts with the TCR, without detrimental steric hindrance. Either more prominent or more featureless peptide may require a more specific set of TCRs, illustrated by a combination of biased usage in TRAV or TRBV gene segment and/or CDR3 loops, to form a productive interaction, thus leading to the selection of a restricted T cell repertoire. However, the peptide solvent exposed surface is clearly not the only structural parameters correlated with the frequency of specific T cells.

The mobility of the peptide can undoubtedly influence peptide-specific T cell frequency. The affinity of the peptide for the MHC molecule and the half-life of the pMHC complex have been previously shown to impact the selection of a T cell repertoire (10, 48, 49). In this study, the destabilized CIN_{N3S} peptide anchoring may have induced the observed significant frequency decrease. Moreover, TCR docking is always associated with a significant decrease of the peptide mobility, especially in the peptide central region (Table I), highlighting the entropic cost for stabilizing TCR–pMHC interface. Nevertheless, there is no systematic correlation between frequency and peptide mobility, because the peptides for which the naive-specific T cells are the more frequent are either very rigid (ELA, LLA) or quite mobile (CIN_{C6V}) (Table I). Thus, one cannot exclude that a significant albeit limited mobility could provide the peptide with some plasticity that may widen the repertoire of specific T cells, as suggested by the data on CIN_{C6V} variant. Our data combined with previous observation confirm that the mobility of the peptide remains a genuine parameter governing the antigenicity, albeit with a contribution that can be counterbalanced by other parameters.

In the examples studied in this paper, with the exception of CIN_{C6V} for which the information is not available, a high frequency of naive specific T cells is always associated with a preferred usage of a given TRAV segment. ELA-A2 and LLA-A2 clearly possess a structural motif involving both the MHC and the peptide that forms stabilizing interactions with the TRAV12-2 encoded segment, as suggested by the biased usage of this gene segment and by the several TCR-ELA-A2 structures available to date (39, 40). Therefore, the high frequency of T cells specific for these pMHC may originate from a structural predisposition of all TRAV12-2 encoded TCRs to interact with ELA-A2 and LLA-A2. However, the fact that the $CD8^+$ naive–specific T cell frequency for ELA and LLA is similar, whereas the contribution of T cell clones using TRAV12-2 segment differ significantly (80% for ELA and 40% for LLA) suggests that such a bias is not the exclusive cause of frequent specific T cells. Moreover, we demonstrated that the very high frequency of ELA-specific naive T cell is not solely accounted for by the favorable interactions between the TRAV12-2 and the ELA-A2 complex but also by an optimal fit between the central region of the peptide and the CDR3 loops. Indeed, the frequency data obtained on the ELA_{I5W} variant (see above) show that a more prominent side chain can drastically reduce the number of TCR able to engage a productive interaction, even if the protruding residue is a tryptophan, an amino-acid known to favor interaction with TCR (21). Thus, if the predisposition of a pMHC to interact with a given TRAV-encoded segment should logically favor interaction between this pMHC and all the TCRs using this TRAV segment, our data suggest that the rest of the pMHC epitope also plays a significant role.

Our data demonstrated that thymic selection also affects the frequency of pMHC-specific T cells. The presence of the peptide in thymus seems to positively impact the frequency, as observed for ELA and LLA. Such an effect should logically be restricted to self-peptide (ELA or LLA) or to foreign peptides that possess a significant homology with a self-peptide involved in positive selection (as it is possibly the case for CIN, but the homologous self peptide remains to be determined). As previously proposed (19), the presence in low amounts of the cognate peptide in the thymic APCs may favor the positive selection of sufficiently high-avidity specific T cells, biasing the selected repertoire for the recognition of this self peptide. The propensity of naive T cells specific for a given pMHC to cross-react with other self- or environmental peptide Ags could clearly downmodulate the overall frequency of naive-specific T cells. Indeed, the cross-reactivity of the specific T cell repertoire may lead to either negative selection in the thymus or to activation of a significant part of this repertoire in the periphery. However, such a cross-reactivity is yet impossible to predict. Anyway, the ratio between $CD8^+$ naive specific T cell frequencies in $A2^+$ and $A2^-$ individuals, when above one, may indicate the degree of homology with a putative selecting peptide and may be used to guide peptide modifications that aims at increasing its antigenicity.

In the current study, the presence of the proper coreceptor, CD8 for class I MHC molecules, is always associated with an increase of the frequency. As CD8 interacts with the MHC $\alpha 3$ domain, the presence of CD8 constitutes an additional interface that stabilizes the TCR-pMHC complex. However, the amplitude of this increase is quite heterogeneous. The largest increase (factor between 20 and 25) is observed for ELA, LLA, GIL, and two CIN_{N3S} and CIN_{C6V} variants. Intermediate values (10–15 times) are observed for CIN, ELA_{I5W} and VLH. Lowest values (5-fold increase) are observed for NLV and SLY. According to what is known about CD8–MHC interactions, an influence of the pMHC structure on the ability of the CD8 to interact with the $\alpha 3$ domain of the MHC is unlikely. A more significant impact of the CD8 presence may be explained by a higher ratio of moderate or low-affinity TCR within the specific T cell repertoire as suggested by the work of Holler and Kranz on mice TCR (50) and Laugel et al. on human

TCR (17): in both species, a TCR-pMHC $K_D < 3$ μM lead to CD8 independent T cell activation, whereas for TCR-pMHC $K_D > 30$ μM, the presence of CD8 was essential for full T cell activation. The frequency of T cells specific for the most mobile CIN_{N3S} and CIN_{C6V} peptides in this study is among the most impacted by the presence of CD8 (increase by a factor ∼25). Such an observation is consistent with the idea that more flexible peptides may overall interact with lower affinity with TCRs because of the entropic cost of stabilizing the peptide. The percentage of naive T cells in the specific repertoire may also impact this increase factor because the presence of CD8 is likely more significant for naive T cells than for already high-avidity memory T cells. However, this remains to be experimentally confirmed. On a structural point of view, the impact of the coreceptor on the frequency is reminiscent of that of a favorable interaction between a germline encoded TCR region and the pMHC as described for ELA and LLA: both mechanisms offer a partial but recurrent contribution to the TCR–pMHC complex stability.

Our data show that even though the thymic selection process and the presence of the antigenic peptide or related ones in the thymus influence the frequency of pMHC-specfic T cells, one is able to identify a few clues on how pMHC structural properties impact the frequency of naive-specific T cells. A modest but recurrent contribution to matching interfaces, such as the CD8/α3 or the TRAV12-2-ELA/A2 interface, clearly enhances the frequency of specific naive T cells. However, statistically optimal fit of the peptide central region with the CDR3α/CDR3β pocket also contributes to a higher frequency, as suggested by our data on ELA and CIN. Indeed, decreasing the peptide solvent exposed surface may sometimes lead to improve the pMHC antigenicity (ELA versus ELA_{I5W}). Furthermore, we showed that increasing the peptide mobility in its central region may also increase the frequency of specific naive T cells (CIN versus CIN_{C6V}) because it likely enhances the spectra of TCR able to interact with it without reducing the affinity of this interaction below a certain threshold. Thus the solvent exposed surface of the peptide and the plasticity of its central region appear to be genuine parameters governing the frequency of specific T cells. Finely tuning these parameters, albeit challenging, may truly open ways to efficiently modulate pMHC antigenicity.

Acknowledgments

We thank the staff from ID14-2, ID14-3, ID14-4, ID23-1, ID23-2, and ID29 beamlines for help with synchrotron data collections at the European Synchrotron Radiation Facility (E.S.R.F.) (Grenoble, France).

Disclosures

The authors have no financial conflicts of interest.

References

1. Garboczi, D. N., P. Ghosh, U. Utz, Q. R. Fan, W. E. Biddison, and D. C. Wiley. 1996. Structure of the complex between human T-cell receptor, viral peptide and HLA-A2. *Nature* 384: 134–141.
2. Garcia, K. C., M. Degano, R. L. Stanfield, A. Brunmark, M. R. Jackson, P. A. Peterson, L. Teyton, and I. A. Wilson. 1996. An αβ T cell receptor structure at 2.5 A and its orientation in the TCR-MHC complex. *Science* 274: 209–219.
3. 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.
4. Liu, Y. C., J. J. Miles, M. A. Neller, E. Gostick, D. A. Price, A. W. Purcell, J. McCluskey, S. R. Burrows, J. Rossjohn, and S. Gras. 2013. Highly divergent T-cell receptor binding modes underlie specific recognition of a bulged viral peptide bound to a human leukocyte antigen class I molecule. *J. Biol. Chem.* 288: 15442–15454.
5. Burrows, S. R., Z. Chen, J. K. Archbold, F. E. Tynan, T. Beddoe, L. Kjer-Nielsen, J. J. Miles, R. Khanna, D. J. Moss, Y. C. Liu, et al. 2010. Hard wiring of T cell receptor specificity for the major histocompatibility complex is underpinned by TCR adaptability. *Proc. Natl. Acad. Sci. USA* 107: 10608–10613.
6. Turner, S. J., K. Kedzierska, H. Komodromou, N. L. La Gruta, M. A. Dunstone, A. I. Webb, R. Webby, H. Walden, W. Xie, J. McCluskey, et al. 2005. Lack of prominent peptide-major histocompatibility complex features limits repertoire diversity in virus-specific $CD8^+$ T cell populations. *Nat. Immunol.* 6: 382–389.
7. 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.
8. Wynn, K. K., Z. Fulton, L. Cooper, S. L. Silins, S. Gras, J. K. Archbold, F. E. Tynan, J. J. Miles, J. McCluskey, S. R. Burrows, et al. 2008. Impact of clonal competition for peptide-MHC complexes on the $CD8^+$ T-cell repertoire selection in a persistent viral infection. *Blood* 111: 4283–4292.
9. Moon, J. J., H. H. Chu, M. Pepper, S. J. McSorley, S. C. Jameson, R. M. Kedl, and M. K. Jenkins. 2007. Naive $CD4^+$ T cell frequency varies for different epitopes and predicts repertoire diversity and response magnitude. *Immunity* 27: 203–213.
10. Kotturi, M. F., I. Scott, T. Wolfe, B. Peters, J. Sidney, H. Cheroutre, M. G. von Herrath, M. J. Buchmeier, H. Grey, and A. Sette. 2008. Naive precursor frequencies and MHC binding rather than the degree of epitope diversity shape $CD8^+$ T cell immunodominance. *J. Immunol.* 181: 2124–2133.
11. Obar, J. J., K. M. Khanna, and L. Lefrançois. 2008. Endogenous naive $CD8^+$ T cell precursor frequency regulates primary and memory responses to infection. *Immunity* 28: 859–869.
12. La Gruta, N. L., W. T. Rothwell, T. Cukalac, N. G. Swan, S. A. Valkenburg, K. Kedzierska, P. G. Thomas, P. C. Doherty, and S. J. Turner. 2010. Primary CTL response magnitude in mice is determined by the extent of naive T cell recruitment and subsequent clonal expansion. *J. Clin. Invest.* 120: 1885–1894.
13. Iglesias, M. C., O. Briceno, E. Gostick, A. Moris, C. Meaudre, D. A. Price, M.-N. Ungeheuer, A. Saez-Cirion, R. Mallone, and V. Appay. 2013. Immunodominance of HLA-B27-restricted HIV KK10-specific $CD8^+$ T-cells is not related to naive precursor frequency. *Immunol. Lett.* 149: 119–122.
14. Cukalac, T., J. Chadderton, W. Zeng, J. G. Cullen, W. T. Kan, P. C. Doherty, D. C. Jackson, S. J. Turner, and N. L. La Gruta. 2014. The influenza virus-specific CTL immunodominance hierarchy in mice is determined by the relative frequency of high-avidity T cells. *J. Immunol.* 192: 4061–4068.
15. Sabatino, J. J., Jr., J. Huang, C. Zhu, and B. D. Evavold. 2011. High prevalence of low affinity peptide-MHC II tetramer-negative effectors during polyclonal $CD4^+$ T cell responses. *J. Exp. Med.* 208: 81–90.
16. Stone, J. D., M. N. Artyomov, A. S. Chervin, A. K. Chakraborty, H. N. Eisen, and D. M. Kranz. 2011. Interaction of streptavidin-based peptide-MHC oligomers (tetramers) with cell-surface TCRs. *J. Immunol.* 187: 6281–6290.
17. Laugel, B., H. A. van den Berg, E. Gostick, D. K. Cole, L. Wooldridge, J. Boulter, A. Milicic, D. A. Price, and A. K. Sewell. 2007. Different T cell receptor affinity thresholds and CD8 coreceptor dependence govern cytotoxic T lymphocyte activation and tetramer binding properties. *J. Biol. Chem.* 282: 23799–23810.
18. Jenkins, M. K., and J. J. Moon. 2012. The role of naive T cell precursor frequency and recruitment in dictating immune response magnitude. *J. Immunol.* 188: 4135–4140.
19. Legoux, F., E. Debeaupuis, K. Echasserieau, H. De La Salle, X. Saulquin, and M. Bonneville. 2010. Impact of TCR reactivity and HLA phenotype on naive CD8 T cell frequency in humans. *J. Immunol.* 184: 6731–6738.
20. Alanio, C., F. Lemaitre, H. K. Law, M. Hasan, and M. L. Albert. 2010. Enumeration of human antigen-specific naive $CD8^+$ T cells reveals conserved precursor frequencies. *Blood* 115: 3718–3725.
21. Chu, H. H., J. J. Moon, A. C. Kruse, M. Pepper, and M. K. Jenkins. 2010. Negative selection and peptide chemistry determine the size of naive foreign peptide-MHC class II-specific $CD4^+$ T cell populations. *J. Immunol.* 185: 4705–4713.
22. 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.
23. 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.
24. Gras, S., X. Saulquin, J. B. Reiser, E. Debeaupuis, K. Echasserieau, A. Kissenpfennig, F. Legoux, A. Chouquet, M. Le Gorrec, P. Machillot, et al. 2009. Structural bases for the affinity-driven selection of a public TCR against a dominant human cytomegalovirus epitope. *J. Immunol.* 183: 430–437.
25. Kabsch, W. 1993. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* 26: 795–800.
26. Navaza, J. 1994. AMoRe: an automated package for molecular replacement. *Acta Crystallogr. A* 50: 157–163.
27. Collaborative Computational Project, Number 4. 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 50: 760–763.
28. Emsley, P., and K. Cowtan. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* 60: 2126–2132.
29. Dutoit, V., V. Rubio-Godoy, M. J. Pittet, A. Zippelius, P. Y. Dietrich, F. A. Legal, P. Guillaume, P. Romero, J. C. Cerottini, R. A. Houghten, et al. 2002. Degeneracy of antigen recognition as the molecular basis for the high frequency of naive A2/Melan-a peptide multimer$^+$$CD8^+$ T cells in humans. *J. Exp. Med.* 196: 207–216.
30. Borbulevych, O. Y., F. K. Insaidoo, T. K. Baxter, D. J. Powell, Jr., L. A. Johnson, N. P. Restifo, and B. M. Baker. 2007. Structures of MART-126/27-35 Peptide/HLA-A2 complexes reveal a remarkable disconnect between antigen structural homology and T cell recognition. *J. Mol. Biol.* 372: 1123–1136.

31. Sliz, P., O. Michielin, J. C. Cerottini, I. Luescher, P. Romero, M. Karplus, and D. C. Wiley. 2001. Crystal structures of two closely related but antigenically distinct HLA-A2/melanocyte-melanoma tumor-antigen peptide complexes. *J. Immunol.* 167: 3276–3284.
32. 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.
33. Ishizuka, J., G. B. Stewart-Jones, A. van der Merwe, J. I. Bell, A. J. McMichael, and E. Y. Jones. 2008. The structural dynamics and energetics of an immunodominant T cell receptor are programmed by its Vbeta domain. *Immunity* 28: 171–182.
34. 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.
35. 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.
36. Spierings, E., S. Gras, J. B. Reiser, B. Mommaas, M. Almekinders, M. G. Kester, A. Chouquet, M. Le Gorrec, J. W. Drijfhout, F. Ossendorp, et al. 2009. Steric hindrance and fast dissociation explain the lack of immunogenicity of the minor histocompatibility HA-1Arg Null allele. *J. Immunol.* 182: 4809–4816.
37. Godet, Y., J. Desfrançois, V. Vignard, D. Schadendorf, A. Khammari, B. Dreno, F. Jotereau, and N. Labarrière. 2010. Frequent occurrence of high affinity T cells against MELOE-1 makes this antigen an attractive target for melanoma immunotherapy. *Eur. J. Immunol.* 40: 1786–1794.
38. Douat-Casassus, C., N. Marchand-Geneste, E. Diez, N. Gervois, F. Jotereau, and S. Quideau. 2007. Synthetic anticancer vaccine candidates: rational design of antigenic peptide mimetics that activate tumor-specific T-cells. *J. Med. Chem.* 50: 1598–1609.
39. Cole, D. K., F. Yuan, P. J. Rizkallah, J. J. Miles, E. Gostick, D. A. Price, G. F. Gao, B. K. Jakobsen, and A. K. Sewell. 2009. Germ line-governed recognition of a cancer epitope by an immunodominant human T-cell receptor. *J. Biol. Chem.* 284: 27281–27289.
40. Borbulevych, O. Y., S. M. Santhanagopolan, M. Hossain, and B. M. Baker. 2011. TCRs used in cancer gene therapy cross-react with MART-1/Melan-A tumor antigens via distinct mechanisms. *J. Immunol.* 187: 2453–2463.
41. Garcia, K. C., M. Degano, L. R. Pease, M. Huang, P. A. Peterson, L. Teyton, and I. A. Wilson. 1998. Structural basis of plasticity in T cell receptor recognition of a self peptide-MHC antigen. *Science* 279: 1166–1172.
42. 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 αβ T cell receptors in antiviral immunity. *Immunity* 18: 53–64.
43. Reiser, J. B., C. Darnault, C. Grégoire, T. Mosser, G. Mazza, A. Kearney, P. A. van der Merwe, J. C. Fontecilla-Camps, D. Housset, and B. Malissen. 2003. CDR3 loop flexibility contributes to the degeneracy of TCR recognition. *Nat. Immunol.* 4: 241–247.
44. Reiser, J. B., C. Grégoire, C. Darnault, T. Mosser, A. Guimezanes, A. M. Schmitt-Verhulst, J. C. Fontecilla-Camps, G. Mazza, B. Malissen, and D. Housset. 2002. A T cell receptor CDR3β loop undergoes conformational changes of unprecedented magnitude upon binding to a peptide/MHC class I complex. *Immunity* 16: 345–354.
45. 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.
46. Rudolph, M. G., R. L. Stanfield, and I. A. Wilson. 2006. How TCRs bind MHCs, peptides, and coreceptors. *Annu. Rev. Immunol.* 24: 419–466.
47. Gras, S., S. R. Burrows, S. J. Turner, A. K. Sewell, J. McCluskey, and J. Rossjohn. 2012. A structural voyage toward an understanding of the MHC-I-restricted immune response: lessons learned and much to be learned. *Immunol. Rev.* 250: 61–81.
48. Sette, A., A. Vitiello, B. Reherman, P. Fowler, R. Nayersina, 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.
49. Chen, W., L. C. Antón, J. R. Bennink, and J. W. Yewdell. 2000. Dissecting the multifactorial causes of immunodominance in class I-restricted T cell responses to viruses. *Immunity* 12: 83–93.
50. Holler, P. D., and D. M. Kranz. 2003. Quantitative analysis of the contribution of TCR/pepMHC affinity and CD8 to T cell activation. *Immunity* 18: 255–264.