Journal of Computer-Aided Molecular Design, **14:** 53–69, 2000.
KLUWER/ESCOM

Modeling the interactions of a peptide-major histocompatibility class I ligand with its receptors. I. Recognition by two αβ T cell receptors

Didier Rognan[a,*], Anette Stryhn[b], Lars Fugger[c,d], Stig Lyngbæk[c], Jan Engberg[e], Peter Sejer Andersen[b] & Søren Buus[b]

[a]*Department of Pharmacy, Swiss Federal Institute of Technology, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland;* [b]*Department of Experimental Immunology, Institute of Medical Microbiology and Immunology, Panum 18.3.22, University of Copenhagen, Blegdamsvej 3C, DK-2200 Copenhagen, Denmark;* [c]*Department of Clinical Immunology, National University Hospital, Copenhagen, Denmark;* [d]*Department of Clinical Immunology, Aarhus University Hospital, Denmark;* [e]*Department of Pharmacology, The Royal Danish School of Pharmacy, Universitetsparken 2, DK-2200 Copenhagen, Denmark*

Received 26 January 1999; Accepted 30 May 1999

Key words: homology modeling, MHC, molecular dynamics, peptides, T cell receptor

Summary

A three-dimensional model of the complex between an Influenza Hemagglutinin peptide, $Ha_{255-262}$, and its restricting element, the mouse major histocompatibility complex (MHC) class I molecule, K^k, was built by homology modeling and subsequently refined by simulated annealing and restrained molecular dynamics. Next, three-dimensional models of two different T cell receptors (TCRs) both specific for the $Ha_{255-262}/K^k$ complex were generated based on previously published TCR X-ray structures. Finally, guided by the recently published X-ray structures of ternary TCR/peptide/MHC-I complexes, the TCR models were successfully docked into the $Ha_{255-262}/K^k$ model. We have previously used a systematic and exhaustive panel of 144 single amino acid substituted analogs to analyze both MHC binding and T cell recognition of the parental viral peptide. This large body of experimental data was used to evaluate the models. They were found to account well for the experimentally obtained data, lending considerable support to the proposed models and suggesting a universal docking mode for αβ TCRs to MHC-peptide complexes. Such models may also be useful in guiding future rational experimentation.

Introduction

MHC class I molecules play a key role in the defense against intracellular pathogens. Their function is to present antigenic peptides (T cell epitopes) to $CD8^+$ cytotoxic T cells [1]. Whereas biological data are available to describe the mechanisms of peptide processing [2], and binding to MHC class I molecules [3], much less is known about the T-cell recognition of peptide-MHC (pep/MHC) ligands [4]. Whereas numerous three-dimensional structures of antibodies and of class I MHC are currently available, the structure of the ternary TCR/peptide/MHC complex has long been obscure. Early comparison of TCR and antibody variable region sequences suggested a close structural similarity [5], confirmed several years later by the first crystal structures of TCR variable domains [6, 7]. Within each TCR chain, three complementary determining regions (CDRs) ensure diversity and were suggested to define the peptide-specific, MHC-restricted combining site [8]. The large sequence variation observed in TCR V regions, as well as the relatively low pairwise amino acid identity to Igs (about 25%), however, limited the use of Ig crystal structures as templates for building reliable three-dimensional models of TCRs. In the last two years, several X-ray structures of TCRs have been published [4]. The recently reported crystal structures of two Vαβ TCRs in com-

*To whom correspondence should be addressed. E-mail: didier@pharma.ethz.ch.

plex with pep/MHC ligands [9–12] are particularly relevant to the work presented here. In both cases, the Vαβ TCR binds diagonally across the MHC binding groove, thereby optimizing non-bonded contacts. As expected from previous site-directed mutagenesis data and binding studies [13], the CDR3α and CDR3β regions are positioned over the center of the MHC and constitute the contact area for the MHC-bound peptide, whereas CDR1-2α and CDR1-2β loops are positioned over the ends of the α1 and α2 helical regions of the class I MHC molecule. Whether the crystallographically determined binding mode may be extended to all peptide-MHC class I complexes is still a matter of debate [14, 15].

To generalize the experimental observations one would require the analysis of more X-ray structures and preferably include directed alterations (such as mutations) of one or more of the involved components. Until such experiments have been conducted, it is reasonable to build TCRs by homology modeling, and in any event, to be guided by such models in a rational approach to experimental work. In this paper, we have used existing crystal structures as templates to generate models of two new TCR/pep/MHC structures representing two different T cells, both specific for the same pep/MHC pair. These models have been examined in the light of a large body of experimental data obtained with an exhaustive set of 144 epitope analogs [16]. This extensive feedback to experimental data supports the validity of the models and suggests a conserved recognition of MHC-peptide ligands by T cells.

Materials and methods

Previously published experimental data

The experimental data has been published previously [16, 17]. In brief, a complete set of single substituted amino acid analogs of the parental Influenza virus Haemagglutinin peptide, positions 255–262 ($Ha_{255-262}$, in single letter amino acid code: FESTGNLI), was synthesized by standard Fmoc peptide chemistry. The mouse MHC class I K^k was purified from detergent lysates of a B cell lymphoma (AKTB-1b) by immuno-affinity chromatography. A K^k-restricted peptide from the Influenza virus nucleoprotein, positions 50–57 (NP_{50-57}, SDYERGLI) was ^{125}I radiolabeled and used to monitor binding to K^k. Binding to K^k of the panel of $Ha_{255-262}$ analogs was determined as the ability to inhibit binding of labeled NP_{50-57} to K^k. Two K^k-restricted T cell hybridomas, HK8.3-5H3 (5H3) and HK8.3-6F8 (6F8), both specific for $Ha_{255-262}$, were developed and their ability to recognize the panel of $Ha_{255-262}$ analogs was ascertained by *in vitro* stimulation assays.

Sequencing the T cell receptors

The two T cell hybridomas, 5H3 and 6F8, were screened by cDNA-PCR using a collection of Vα and Vβ oligonucleotides in conjunction with a Cα and a Cβ oligonucleotide, respectively [18]. Following cloning of the PCR products and sequencing of several individual clones, unambiguous identification of sequences representing functional transcripts could be made. These sequences have been submitted to the EMBL/Genebank databases (accession numbers: 5H3α: Y17471, 5H3β: Y17472, 6F8α: Y17473, 6F8β: Y17474).

Primary sequence alignments

Sequence alignments were performed using standard settings of the multiple sequence alignment MSA algorithm [19]. To accommodate for the different lengths of the TCR Vα and Vβ sequences a maximal insertion/deletion of 8 amino acids was tolerated during the alignment. Within the hypervariable TCR CDR loops, the alignment was manually corrected according to X-ray structures of the three templates [10, 12]. Pairwise amino acid identities were calculated by the ALIGN program [20].

Modeling of the peptide-MHC complex

The $Ha_{255-262}$-K^k complex was modeled by homology to the known K^b X-ray structure (pdb code 2vaa) [21]. The rotameric states for the side chains, which differed between K^k and K^b, were assigned according to a three-dimensional (3-D) database of 22 class I MHC X-ray structures [22]. Then the $Ha_{255-262}$ peptide was built from the K^b-bound X-ray structure [21] of the vesicular stomatitis nucleoprotein (VSV) octapeptide, positions 52–59 (VSV_{52-59}, RGYVYQGL, pdb code: 2vaa). The VSV_{52-59} to $Ha_{255-262}$ changes were realized using the BIOPOLYMER module of the SYBYL package (TRIPOS Assoc., Inc.), without altering the dihedral angles of the side chains. The MHC-peptide was then refined using the AMBER4.1 program [23]. The $Ha_{255-262}$ peptide was first

minimized and annealed for 100 ps using a recently reported procedure [24]. The complex was then solvated in a 10 Å-thick water shell and minimized again using a previously described protocol [25–27] and subjected to a restrained 200-ps Molecular Dynamics (MD) simulation at constant temperature (297 K) as already published [24]. In order to avoid large drifts from known class I MHC structures, a weak positional harmonic constraint of 0.5 kcal mol^{-1} $Å^{-1}$ was applied on MHC atoms located outside the peptide binding groove. Only peptide and protein amino acids participating in intermolecular interactions were allowed to freely move. As water molecules were here explicitly treated, a constant dielectric function ($\epsilon = 1$) was used. Non-bonded interactions were calculated as previously, using a twin cut-off (10–15 Å) and a time step of 1 fs was used throughout the MD simulations. Atomic velocities were randomly generated for a temperature of 50 K, and the temperature was linearly raised from 50 to 300 K over the first 5 ps. Productive simulation was then continued for 195 ps at 300 K. Coordinates, energies and velocities were stored every 500 steps (0.5 ps). For both complexes, a mean MD structure was time-averaged over the last 200 conformers, and futher energy-minimized by 1000 steps of steepest descent and by a conjugate-gradient procedure until the rms deviation of the potential energy was less than 0.1 kcal mol^{-1} $Å^{-1}$.

Modeling of the TCR/Ha$_{255-262}$/K^{k} complexes

The model of the 5H3 TCR is described and detailed below (the model of the 6F8 TCR has been generated along the same lines). α and β chains for the two TCRs were modeled by homology to the X-ray template for which the sequence identity was the higher. The 5H3 Vαβ dimer was constructed from the A6 α chain (pdb code: 1ao7) and the 2C β chain (pdb code: 2ckb). Amino acid mutations were done using the BIOPOLYMER module of the SYBYL package, and keeping the side chain directionality unchanged. A rapid conformational scan was performed when necessary to relax the side chain in its local environment. For the 5H3 TCR, five out of the six CDR regions (CDR1-3α, CDR1-2β) were built by homology to the corresponding loops of the same length in the A6 α chain and the 2C β chain (Figure 1b). For the 6F8 TCR, the CDR1-2α loops of the α chain were constructed by homology to that of the 2C structure whereas the CDR3α loop was obtained as for 5H3, from the CDR3α structure of the A6 TCR. The CDR1-2β regions, due to their unique length and sequence, were obtained by searching a loop of similar size and sequence using the LOOPSEARCH utility or SYBYL. In this procedure, a set of 837 high-resolution X-ray structures were searched for a loop of similar length and presenting a similar distance between Cα atoms of the residues delimiting the loop window. The loop showing the highest homology and the lowest rms deviations was futher selected for insertion.

The β chain (lacking the CDR3β loop, see 'Results' for justification) was paired to the α chain according to the crystal structure of the A6 TCR (pdb code: 1ao7) [11] and hydrogen atoms were added using SYBYL standard geometries. After minimizing the hydrogen atoms by 100 steps of steepest descent, the five hypervariable CDR regions considered were relaxed and submitted to a 100-ps simulated annealing (SA) under an already published protocol [24], while maintaining the amino acids of the framework fixed to their starting coordinates. The K^{k}-Ha complex was then docked into the refined TCR model (still lacking the CRD3β loop) according to the X-ray structure of A6 in complex with the Tax/HLA-A∗0201 ligand (pdb code: 1ao7). Once the MHC-peptide pair was docked into the TCR, the ternary complex was relaxed by 1000 steps of steepest descent and the CDR3β loop was finally built by the loop searching utility of the SYBYL program as previously described. Here, in addition to the previously described criteria (similarity in length and amino acid sequence), the steric and electrostatic complementarity to the TCR-peptide-MHC environment was used to restrain the loop selection procedure. Once the last loop was inserted, it was relaxed by 500 steps of steepest descent. A 20 Å-thick cap of TIP3P water molecules was then centered on the Ha peptide in the ternary complex and the solvent atoms were minimized. The solvated TCR-MHC-peptide complex was then refined and submitted to a restrained 500-ps MD simulation as previously described [27]. To save computational time, only the TCR-peptide-MHC interface was allowed to freely move. This interface was defined by the whole Ha peptide, the 6 CDRs of the 5H3 TCR and the α1, α2 helical regions (residues 56 to 83, 149 to 173) of the MHC protein that contact the TCR. All other atoms of the ternary complex were restrained to their starting location by using a weak harmonic constraint of 5 kcal mol^{-1} $Å^{-1}$.

a

α_1 domain

```
                1       10        20        30        40        50        60        70        80        90
                .       .         .         .         .         .         .         .         .         .
H-2K^k          GPHSLRYPHTAVSRPGLGKPRFISVGYVDDTQFVRFDSDAENPRYEPRVRWMEQVEPEYWERNTQIAKGNEQIFRVNLRTALRYYNQSAG
H-2K^b          GPHSLRYFVTAVSRPGLGEPRYMEVGYVDDTEFVRFDSDAENPRYEPRARWMEQEGPEYWERETQKAKGNEQSFRVDLRTLLGYYNQSKG
H-2D^b          GPHSMRYFETAVSRPGLEEPRYMEVGYVDNKEFVRFDSDAENPRYEPRAPWMEQEGPEYWERETQKAKGQEQWFRVSLRNLLGYYNQSAG
H-2L^d          GPHSMRYFETAVSRPGLGEPRYISVGYVDNKEFVRFDSDAENPRYEPQAPWMEQEGPEYWERITQIAKGQEQWFRVNLRTLLGYYNQSAG
H-2D^d          GSHSLRYFVTAVSRPGFGEPRYMEVGYVDNTEFVRFDSDAENPRYEPRARWIEQEGPEYWERETRRAKGNEQSFRVDLRTALRYYNQSAG
                        a         g  abb       c                  g     gg      d  b      g   b   e e     ff

$\alpha_2$ domain

                91      100       110       120       130       140       150       160       170       180
                .       .         .         .         .         .         .         .         .         .
H-2K^k          GSHTFQRMYGCEVGSDWRLLRGYEQYAYDGCDYIALNEDLKTWTAADMAALITKHKWEQAGDAERDRAYLEGTCVEWLRRYLQLGNATLPRT
H-2K^b          GSHTIQVISGCEVGSDGRLLRGYQQYAYDGCDYIALNEDLKTWTAADMAALITKHKWEQAGEAERLRAYLEGTCVEWLRRYLKNGNATLLRT
H-2D^b          GSHTLQQMSGCEVGSDWRLLRGYLQFAYEGRDYIALNEDLKTWTAADMAAQITRRKWEQSGAAEHYKAYLEGECVEWLHRYLKNGNATLLRT
H-2L^d          GTHTLQWMYGCDVGSDGRLLRGYEQFAYDGCDYIALNEDLKTWTAADMAAQITRRKWEQAGAAEYYRAYLEGECVEWLHRYLKNGNATLLRT
H-2D^d          GSHTLQWMAGCDVESDGRLLRGYWQFAYDGCDYIALNEDLKTWTAADMAAQITRRKWEQAGAAERDRAYLEGECVEWLRRYLKNGNATLLRT
                    g c c *     f       b                                   g   e       *         gg     g
```

b

α Chain

```
                1       10        20        30        40        50        60        70        80        90        100
                .       .         .         .         .         .         .         .         .         .         .
2C              QSVTQPDARVTVSEGASLQLRCKYSYSATPYLFWYVQYPRQGLQLLLKYYSGDPV-VQGVNGFEAEFSKSNSSFHLRKASVHWSDSAVYFCAVS-----GF
A6              KEVEQNSGPLSVPEGAIASLNCTYSDRGSQSFFWYRQYSGKSPELIMSIYSNGDK-EDG--RFTAQLNKASQYVSLLIRDSQPSDSATYLCAVT-----TDS
B7              QQVKQNSPSLSVQEGRISILNCDYTNSMFDYFLWYKKYPAEGPTFLISISSIKDKNEDG--RFTVFLNKSAKHLSLHIVPSQPGDSAVYFCAAMEG-----
5H3             QKVQQSPESLIVPEGGMASLNCTSSDRNVDYFWWYRQHSGKSPKMLMSIFSNGEK-EEG--RFTVHLNKASLHTSLHIRDSQPSDSALYLCAAD-----DSG
6F8             DSVTQTEGQVALSEEDFLTIHCNYSASGYPALFWYVQYPGEGPQFLFRASRDKEK-GSSR-GFEATYNKETTSFHLQKASVQESDSAVYYCALG-----VGD
                                      *  CDR1α                  CDR2α                                        *    CDR3α

                101     110       120
                .       .         .
2C              ASALTFGSGTKVIVLPYIQN
A6              WGKLQFGAGTQVVVTPDIQN
B7              AQKLVFGQGTRLTINPNIQN
5H3             YNKLTFGKGTVLLVSPDIQN
6F8             NSKLIWGLGTSLVVNPDIQN
```

β Chain

```
                1       10        20        30        40        50        60        70        80        90        100
                .       .         .         .         .         .         .         .         .         .         .
2C              EAAVTQSPRNKVAVTGGKVTLSCNQ-TNNHNNMYWYRQDTGHGLRLIHYSYGAGST--EKGDIPDG-YKASRPSQENFSLILELATPSQTSVYFCASGGGG--
A6              --GVTQTPKFQVLKTGQSMTLQCAQ-DMNHEYMSWYRQDPGMGLRLIHYSVGAGIT--DQGEVPNG-YNVSRSTTEDFPLRLLSAAPSQTSVYFCASRPGLAG
B7              --GVTQTPKFQVLKTGQSMTLQCAQ-DMNHEYMSWYRQDPGMGLRLIHYSVGAGIT--DQGEVPNG-YNVSRSTTEDFPLRLLSAAPSQTSVYFCASSYPGGG
5H3             EAAVTQSPRNKVAVTGGKVTLSCNQ-TNNHNNMYWYRQDTGHELRLIHYSYGAGST--EKGDIPDG-YKASRPSQENFSLILELATPSQTSVYFCASGDVQ-G
6F8             GAIVYQYPRRTICKSGTSMRMECQAVGFQATSVARYRQSPQKTPELIALSTVNSAIKYEQNFTQEK-FPISHPNLSFSSMTVLNAYLEDRGLYLCGARRDW-G
                                      *  CDR1β                  CDR2β                                        *  CDR3β

                101     110       120
                .       .         .
2C              ----TLYFGAGTRLSVLEDLR
A6              GRP-EQYFGPGTRLTVTEDLK
B7              F--YEQYFGPGTRLTVTEDLK
5H3             ARA-EQFFGPGTRLTVLEDLR
6F8             SSY-EQYFGPGTRLTVLEDLR
```

Figure 1. Alignment of the primary sequences of target proteins to that of X-ray templates. (a) Alignment of the antigen-binding domains (1–182) of K^k to that of four X-ray structure templates. Solid boxes represent identity to the K^k primary sequence. Letters (*a*–*f*) indicate polymorphic residues whose rotameric states have been assigned according to the X-ray structure of (a) HLA-B∗2705 (pdb code: 1hsa) [51], (b) H-2L^d (pdb codes: 1ld9, 1ldp) [35, 36], (c) HLA-A∗0201 (pdb code: 3hla) [52], (d) HLA-B∗5301 (pdb code: 1a1m) [53], (e) H-2D^d (pdb code: 1bii) [37], and (f) H-2D^b (pdb code: 1hoc) [33], respectively. *g* indicates a position for which the mutation does not occur in any X-ray template. Cysteines involved in disulfide bridging are noticed by a star. (b) Multiple sequence alignment of the primary sequences of the Vαβ regions (1–120) of 5H3 and 6F8 TCRs to that of three X-ray structure templates (2C, A6). The sequences are numbered according to Kabat [54]. Hypervariable regions (CDRs) are defined according to Chothia [5] and enclosed by solid boxes. Cysteines involved in disulfide bridging are noticed by a star.

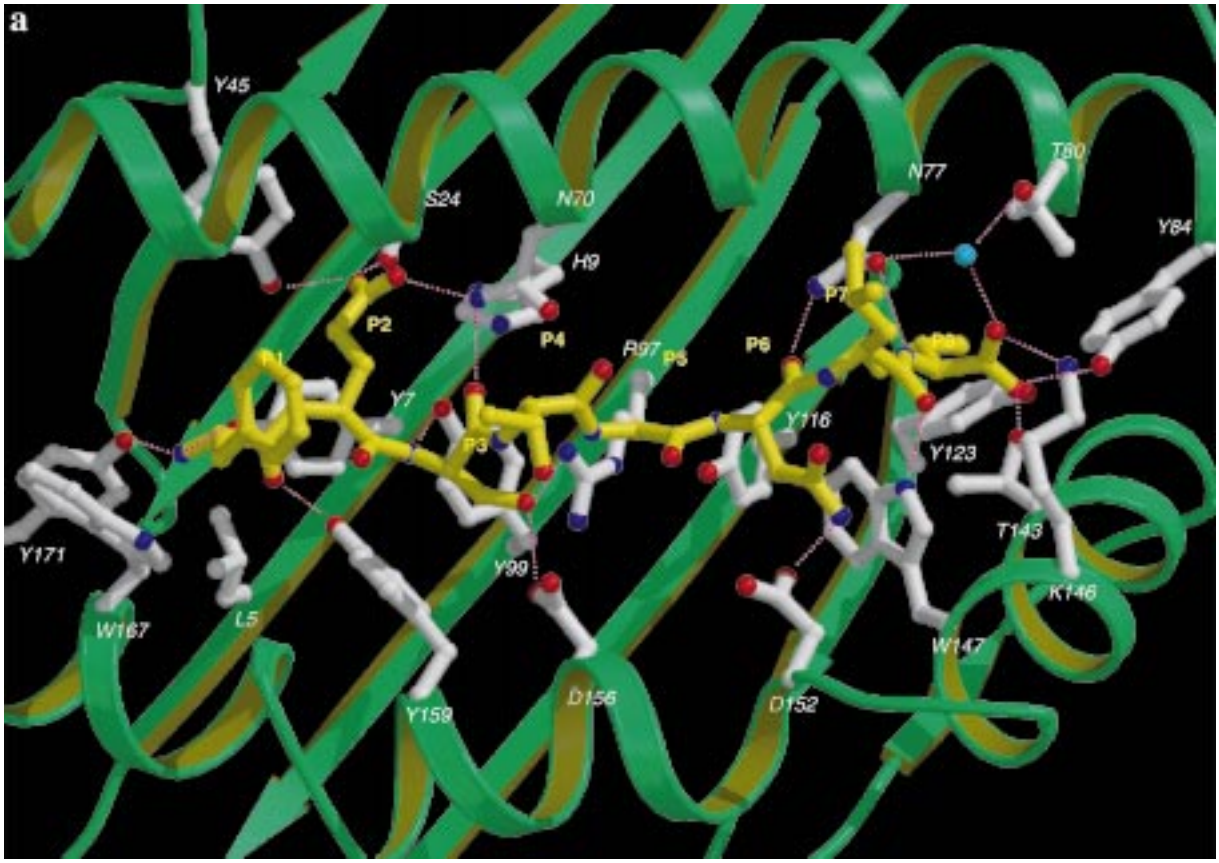

Figure 2. Three-dimensional model of the K^k-Ha complex. (a) Zoom into the binding cleft of K^k in complex with the Ha peptide (energy minimized time-averaged MD conformation). K^k backbone atoms are represented by a green solid ribbon. Only MHC side chains pointing towards the Ha peptide are displayed. Intermolecular hydrogen bonds are displayed as magenta broken lines. P1 to P8 indicate the Cα atoms of the bound peptides from position 1 to 8. The following color coding has been used: white, carbon atoms of K^k; yellow, carbon atom of the Ha peptide; blue, nitrogen; red, oxygen; cyan, water. The figure has been obtained using MOLSCRIPT [55] and rendered with Raster3D [56]. (b) Antigenic identity of the K^k-bound Ha peptide. The K^k protein and the bound Ha peptide are displayed by a solid surface and a stick model, respectively. Peptide side chains at P1, P4, P6 and P7 project out of the binding groove. The picture has been realized with the GRASP [57] and PREPI programs [58].

Comparison of the two TCR/pep/MHC ternary complexes

All dynamical properties of the two ternary complexes were analyzed by in-house routines and the CARNAL program of the AMBER4.1 software [23]. The stereochemical reliability of each energy-minimized time-averaged MD conformation was assessed by the PROCHECK program [28]. Correct folding of each MD model was analysed with the ProsaII software [29] by computing energy profiles from β-carbon interactions [30]. Surface areas and non-bonded interactions were examined using the ACCESS and CONTACT programs of the CCP4 package [31].

Results and discussion

Modeling the Ha-K^k binary complex

The primary sequence of K^k (α1-α2 antigen-binding domain, residues 1 to 182) was first aligned to those of four murine class I MHC proteins (H-2K^b, H-2D^b, H-2L^d, H-2D^d) for which crystal structures [21, 32–37] have been deposited in the Protein Data Bank [38] (Figure 1a). Of the four mouse MHC class I templates considered, the K^b had the highest identity to K^k, exhibiting a 86% pairwise amino acid identity to K^k in the regions corresponding to the antigen-binding domain (residues 1–182). Furthermore, both K^b and K^k proteins lack structural features (‘hydrophobic ridge’ in D^b [33] and L^d [35–36], ‘tryptophan wall’ in D^d

Table 1. Peptide-MHC hydrophobic non bonded contacts

$Ha_{255-262}$	K^k	Contacts[a]	Buried area[b]	Exposure[c]
Phe1	Arg62, Ile66, Trp167	10	197	26
Glu2	Tyr45, Ile66, Tyr159	7	150	15
Ser3	Ile66, Arg97, Tyr99, Asp156, Tyr159	10	99	18
Thr4	Ile66	1	61	58
Gly5	Arg15599	2	51	38
Asn6	Ile73, Trp147, Asp152, Arg155	6	111	28
Leu7	Val76	1	81	55
Ile8	Asn77, Tyr84, Phe95, Tyr123, Thr143 Lys146, Trp147	14	229	4

[a]Non-bonded contacts have been recorded, from the energy-minimized time-averaged MD structure, for carbon atoms distant by less than 4 Å, using the CONTACT program [31].
[b]Buried Surface Area ($Å^2$), calculated using the ACCESS program [31] using a probe radius of 1.4 Å.
[c]Percentage of residue exposure, calculated using the ACCESS program as previously described [25].

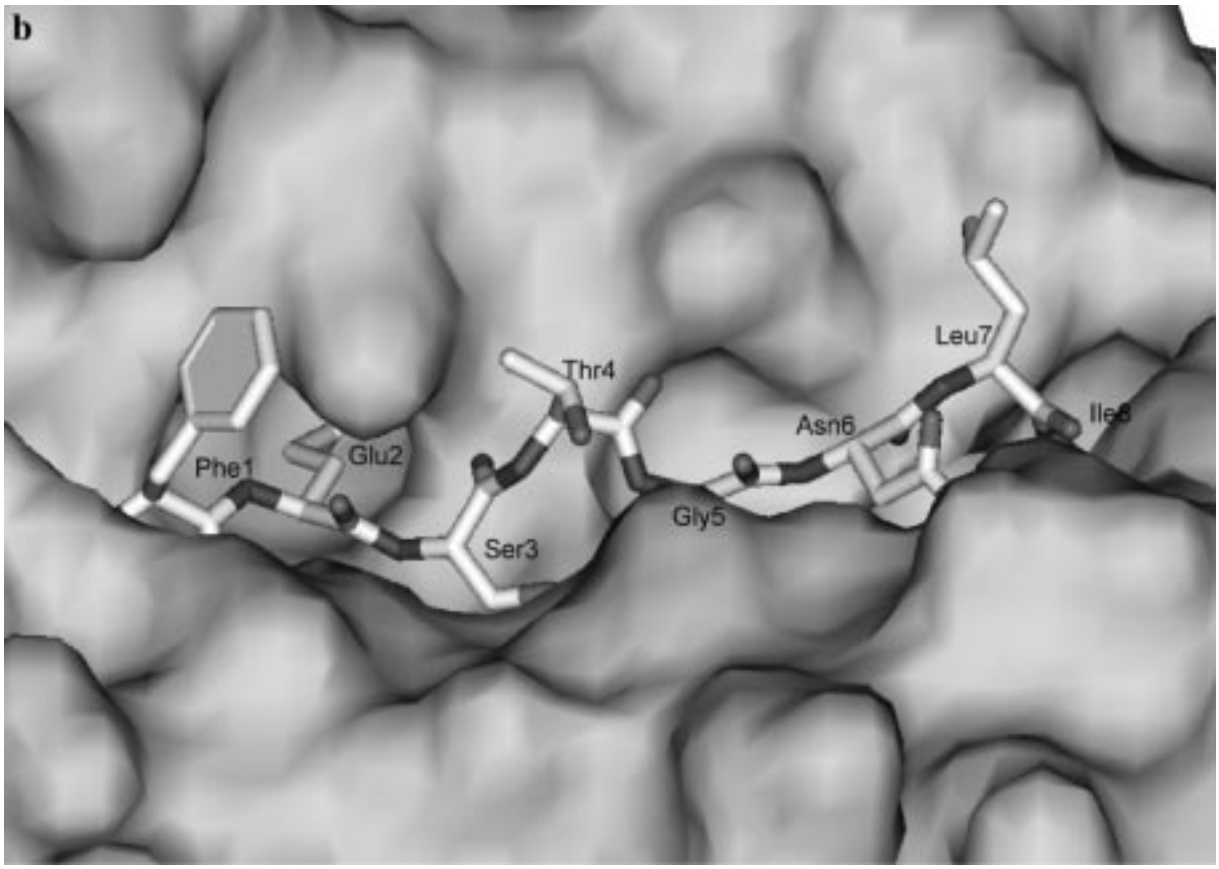

Figure 2. (continued).

[37]) that influences the conformation of the bound peptide. Thus, the antigen binding domain of K^k was modeled by homology to K^b and the Ha peptide was built from the K^b-bound X-ray structure of the vesicular stomatitis nucleoprotein [52–59] octapeptide [21]. The refined model shows the bound peptide having an overall extended conformation very similar to that found for all peptides co-crystallized with class I MHC proteins [3]. Both peptide termini are tightly nested into the binding cleft by several hydrogen bonds with conserved MHC residues (Figure 2a). A total of 17 direct intermolecular hydrogen bonds and 51 close van der Waals contacts could be identified in the energy-minimized time-averaged conformation (Table 1).

The present model of the $Ha_{255-262}$-K^k is in very good agreement with experimental binding data on Ha peptide analogs and peptide libraries [16, 17]. Glu2 and Ile8 have been experimentally determined as the primary anchor residues. The model shows Glu2 deeply buried in the polar pocket B (Figures 2a, b) where it establishes a set of 3 H-bonds with polar side chains. The side chain of the other primary anchor, Ile8, is entirely buried in the apolar pocket F (Figures 2a, b). It is the amino acid which has the most apolar contacts with the binding cleft (Table 1) and the minimal exposure to solvent (or any TCR). Ile and Val are the preferred amino acids at this position [17]. The slightly decreased binding of Ha analogues bearing a bulkier aromatic residue (F, Y, W) at position 8 (P8) is likely to be explained by the presence of Tyr116 which restricts the depth of pocket F. Phe1, Ser3 and Asn6 are partially buried in the binding groove (Figures 2a, b, Table 1). Aromatic residues are preferred at P1 [17], as they can interact with Trp167 forming a wall delineating pocket A. The side chain of Phe1 projects out of the binding groove with about 26% of its surface accessible to solvent (Table 1). The side chain of Ser3 is hydrogen-bonded to polar side chains at the rim of pocket D (Figure 2a). Pocket D is opened to the central part of the binding cleft so that long hydrophobic residues bind to this pocket. Experimentally, nearly all amino acids can be accommodated at P3 [16, 17]. Asn6 interacts with pocket E (Figures 2a, b) with about 28% of its surface accessible to solvent (Table 1), thus explaining the variability observed at this position from binding data of Ha analogues and peptide libraries [16, 17]. Gly5 is facing two large amino acids (Arg97, Arg155) which entirely bury its Cα atom (Figure 2b). The significant proportion (38%) of its accessible surface is contributed mainly by its carbonyl group that faces solvent (Figure 2b). This observation explains why small residues (G, A) have been found to be the most favored amino acids by positional scanning peptide combinatorial libraries [17]. However, a large variability is still possible at P5. The likely explanation of these experimental data is the conformational flexibility of the middle part (P4–P7) of the bound peptide, already noticed in several MHC-peptide crystal structures [39]. Among the less buried peptide residues are Thr4 and Leu7 (Table 1), in agreement with the finding that virtually all 20 naturally occurring amino acids are accepted by K^k in these two positions.

Modeling the ternary 5H3 TCR-Ha-K^k complex

Three X-ray templates representing ternary complexes, 2C (pdb codes: 1tcr, 2ckb) [9, 10], A6 (pdb code: 1ao7) [11] and B7 (pdb code: 1bd2) [12] were chosen for a multiple sequence alignment to 5H3 and to determine the best template for each chain. X-ray structures of unliganded Vα [6] or Vβ chains [7], as well as Vαβ TCRs bound to antibodies [40, 41] were not taken in consideration for two main reasons: (i) the conformation of the highly variable CDR loops is dependent on the presence and nature of the TCR ligand [40], (ii) the side chains participating in the VαVβ interface have different orientations in monomers and αβ dimers [9].

Computation of sequence identity clearly shows that the A6 Vαβ TCR [11] is the best candidate for building the Vα chain (residues 1–120) of 5H3 (58.3% identity), whereas 2C [9] was by far the best model for the Vβ chain (84.2% identity, Figure 1b). As the conformation of the most hypervariable CDR3β loop (a suspected major contact area for the MHC-bound peptide [10–12]) is highly dependent on the TCR ligand [4, 34], we decided to dock the K^k-Ha complex to a 5H3 model lacking the β3 loop first and then finalize the ligand-restrained modeling of the CDR3β loop. Thus, the β chain lacking the CDR3 loop was paired to the α chain according to the A6 crystal structure. At least two slightly different VαVβ pairing modes have been described in the literature [36]. The first one, conserved in four (A6, 2C, N15, KB) out of the five TCRs for which an X-ray structure has been described, has been utilized here. The alternative pairing, currently present in a single structure (B7) [12], has been discarded in the present homology models as it would have minored contacts between the αβ dimer and a peptide position (P7) known to bind to

the 5H3 and 6F8 clones [16, 17]. Thus, the Vαβ 5H3 TCR was energy minimized and docked into the K^k-Ha complex according to the A6-TCR-MHC-peptide X-ray template. The rationale for that choice comes from the suggestion that the length and conformation of the CDR3β loop dictate the orientation of the pep-MHC ligand with respect to the TCR [11]; note the rather similar length of the CDR3β loop in both target TCRs and A6 (10 amino acids for 5H3 and 6F8, versus 11 for A6 and 6 for 2C, see Figure 1b). The docking mode observed in the third ternary complex with B7 [12] was discarded as it is highly dependent on the special Vαβ interface of this TCR.

Until now the residues surrounding the missing CDR3β loop had been kept fixed. Once the MHC-peptide pair had been docked into the TCR, the missing CDR3β loop was built by a loop searching procedure (see Materials and methods). Modeling the 5H3 CDR3β loop by homology the retrieved loop (from the Thaematin A protein) enabled the two polymorphic 5H3 TCR positions (Val97β, Gln98β) to be directed towards peptide residues (P7, P4) of complementary physico-chemical properties and of high relevance for 5H3 recognition [6]. It should be pointed out that the various procedures selected all along the construction of the ternary complex lead to a 3-D model of reasonable topology, prior to any energy refinement. The complex was finally solvated, energy minimized and submitted to a restrained 500-ps MD simulation. The refined MD model differs mainly from the starting homology derived model in the orientation of the 6 hypervariable loops with respect to the pep/MHC ligand. Root-mean square deviations from the starting structure were significantly higher for the CDR loops (from 1.14 to 1.53 Å) than for the MHC molecule (0.924 Å) or for the bound Ha peptide (0.457 Å). The benefit of the MD refinement over the pure homology derived model concerns the interaction interface between the two proteins and the peptide ligand. In the present case, MD allows an optimization of intramolecular interactions (notably H-bonds). This may be easily explained by the choice of the CDR loop building procedure, based on TCR templates (A6 α chain, 2C β chain) whose MHC-pep environment is different from that of the target model. It may be also noted that the above-described rms deviations from the starting structure as well as positional atomic fluctuations (usually between 0.4 and 0.5 Å) were constant after 250 ps, so that time-averaging of internal coordinates was indeed possible.

Figure 3. Three-dimensional model of the 5H3-K^k-Ha ternary complex. (a) Overall topology of the three-dimensional structure. The backbone structure of the Vα,Vβ chains of 5H3 and of the α1-α2 antigen binding domain of K^k are displayed as ribbon tubes. The bound Ha peptide is represented as a stick model. P1 to P8 indicate the Cα positions of the peptide sequence. The figure has been obtained using MOLSCRIPT [55] and rendered with Raster3D [56]. (b) Zoom into the 5H3-Ha-K^k interaction interface. The color coding is the following: cyan, carbon atom of the TCR Vα chain; green, carbon atom of the TCR Vβ chain; yellow, carbon atom of the MHC heavy chain; white, carbon atom of the bound peptide; blue, nitrogen; red, oxygen. Water molecules and hydrogen bonds are displayed as magenta balls and white broken lines, respectively. The figure has been obtained using MOLSCRIPT [55] and rendered with Raster3D [56].

In the relaxed time averaged conformation of the ternary complex (Figures 3a, b), the 5H3 Vαβ has a total of 15 direct hydrogen bonds (Table 1) and 40 close van der Waals contacts (Table 2) with its MHC-peptide ligand. The CDR1α loop is directed towards the two α-helices of the antigen-binding domain. The CDR2α hypervariable region contacts an exposed segment of the K^k α2 helix and the CDR3α region interacts with the peptide central part (Thr4, Gly5) and an exposed part of the K^kα1 helix. The CDRs 1β and 2β have nearly no contacts with the MHC-peptide complex (Figure 3a). Of the 5H3 Vβ chain, only the CDR3β

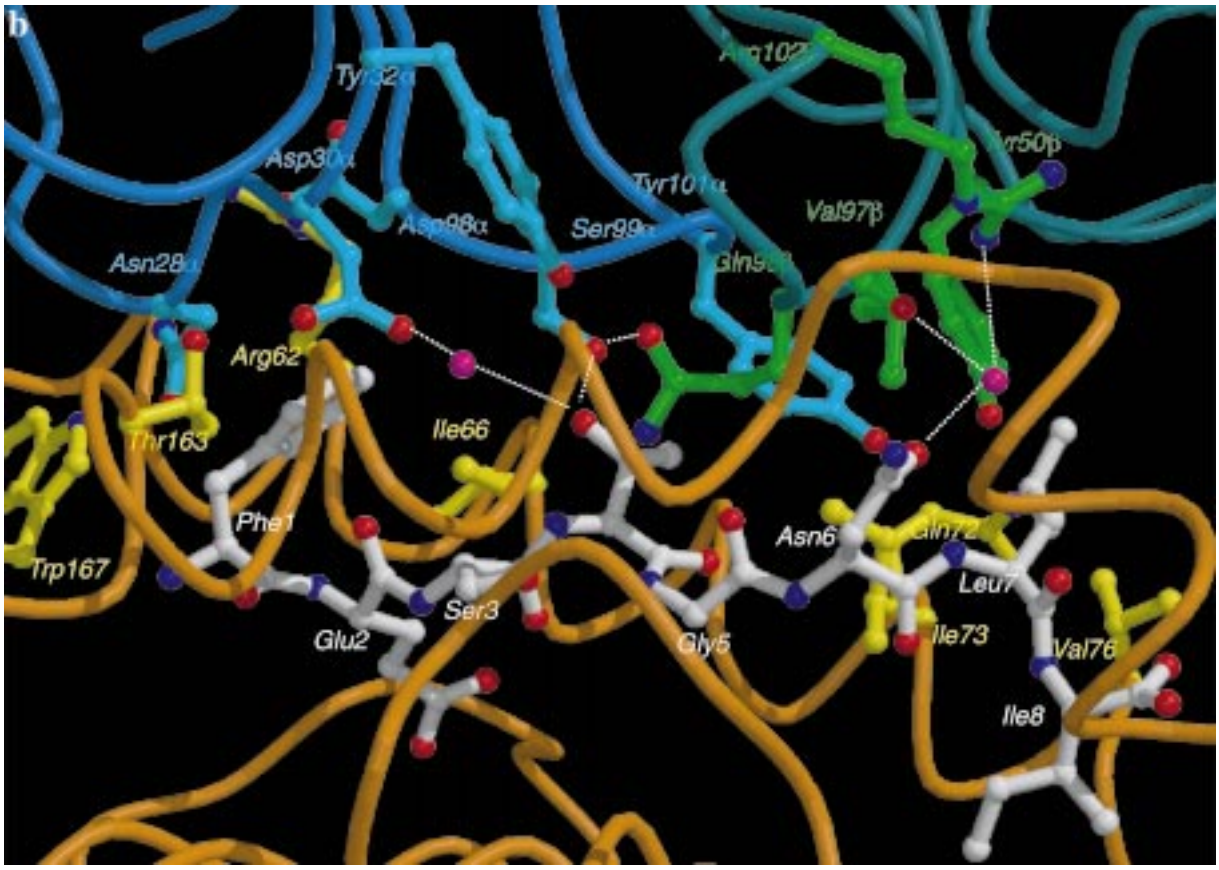

Figure 3. (continued).

loop interacts with the peptide (Asn6, Leu7). From the point of view of the peptide, only 9 close van der Waals contacts and 2 direct H-bonds engage the 5H3 TCR directly. In addition, 4 H-bonds with the combining site are mediated by two strongly-bound water molecules (Figure 3b). Atomic positional fluctuations of the 178 water molecules within the peptide-centered cap varied rather broadly from 0.39 to 7.24 Å, but none of them was more than 5 Å away from the ternary complex. The need for solvation is explained by the rather low complementarity between the surface of the TCR and that of the pep/MHC ligand which leads to 'holes' in the structure of the ternary complex that are filled by the solvent. Thus, bound water molecules contribute to enhance the complementarity between the TCR and the pep/MHC ligand.

The Phe1 side chain is the first TCR contact residue. In the proposed model, it is centered in a subsite composed of the K^k pocket A (Arg62, Ile66, Thr163, Trp167) and 5H3 residues of the CDR1α,3α loops (Asn28α, Asp30α, Asp98α). Only carbon atoms of the Vα chain delimit this subsite which is closed by three MHC-TCR H-bonds. The proposed binding mode of Phe1 is compatible with its important binding role to 5H3 [16]. Assuming a conserved binding mode for Ha analogues, a Tyr amino acid at P1 will not fit as well between 1α and 3α loops of 5H3 since the hydroxyl group would be in too close contact with the carbon atoms of Asp30α and Asp98α side chains. A Trp residue would be too bulky, and an Ile, Val or Leu side chain would only partially fill the TCR part of the subsite. The second TCR-contact residue is Thr4. Its side chain, in the hypothetical model, is hydrogen-bonded to the CDR3α loop (Ser99α). Two additional H-bonds mediated by a water molecule are also engaged to both 1α (Asp30α) and 3β (Gln98β) CDRs. The methyl group of the Thr side chain points towards the Tyr101α aromatic ring and remains close to two K^k hydrophobic residues (Ile66, Ile73), and thus enhances the exquisite specificity of the 5H3-K^k complex for Thr4. This interaction model accounts well for the stringent requirements observed at position 4 of the

Table 2. The 5H3-K^k-Ha complex: intermolecular hydrogen bonds

5H3	K^k	$Ha_{255-262}$	Frequency[a] (%)
Asp26α (OD1)	Arg62 (NH2)		100
Arg27α (NH2)	Glu166 (OE2)		98
Arg27α (O)	Trp167 (NE1)		48
Asn28α (OD1)	Asn63 (ND2)		90
Asp30α (OD2)	Thr163 (OG1)		100
Asn52α (ND2)	Glu166 (OE1)		100
Lys68α (NZ)	Thr163 (OG1)		98
Lys68α (NZ)	Glu166 (OE2)		100
Asp93α (OD2)	Arg62 (NE)		50
Tyr101α (OH)	Gln72 (NE2)		36
Tyr50β (OH)	Gln72 (NE2)		44
Gly100β (N)	Asp152 (OD1)		32
Ala101β (N)	Gly151(O)		96
Ser99α (O3G)		Thr4 (OG1)	98
Asp31α (OD1)		Thr4 (OG1)[b]	
Gln98β (NE2)2		Thr4 (OG1)[b]	
Gln98β (NE2)		Gly5 (O)	94
Val97β (O)		Asn6 (OD1)[b]	
Arg102β (NH2)		Asn6(OD1)[b]	

[a]Interactions were statistically monitored throughout the MD trajectory for a total of 250 conformations.
[b]Water-mediated H-bond.

Table 3. 5H3-K^k-Ha complex: non-bonded hydrophobic contacts

5H3	K^k	$Ha_{255-262}$	# of contacts[a]
Gln1α	Arg62		1
Arg27α	Trp167		4
Asn28α	Arg62		3
		Phe1	3
Phe50α2	Ala158		1
Ser51α	Ala158		1
Lys68α	Glu166		2
Asp93α	Arg62		4
		Phe1	1
Tyr101α2	Gly69		4
		Thr4	2
Tyr50β		Leu7	2
Val97β	Ile73		1
Val97β		Leu7	2
Gln98β	Arg155		2
Gln98β		Thr4, Asn6	2
Gly100β	Gly151, Asp152		4
Ala101β	Glu154		1

[a]Non-bonded contacts have been recorded, from the energy-minimized time-averaged MD structure, for carbon atoms distant by less than 4 Å, using the CONTACT program [31].

Ha peptide [16]. Only a conservative substitution (Ser for Thr) is compatible with 5H3 binding. This peptide mutation suppresses only the apolar component of the pep/MHC interaction with the TCR and results in a 10-fold loss in 5H3 recognition. In comparison to the first two TCR-anchors (Phe1, Thr4), Asn6 and Leu7 provide much weaker interactions to the 5H3 TCR model. Asn6 is hydrogen-bonded through a water molecule to two residues of the CDR3β loop (Val97β backbone, Arg102β side chain). However, the motion of the Asn6 side chain is limited by an already existing H-bond to K^k (Asp152 side chain, Figure 2a). The Ile7 side chain is directed towards hydrophobic residues of the β chain (Tyr50β, Val97β). Its side chain is loosely interacting with a rather flexible subsite, delimited by some K^k hydrophobic amino acids (Ile73, Val76). Even if the proposed interaction of Ile7 with the β chain is rather weak, it supports the Vαβ pairing chosen in the construction of the TCR dimer. Hence, the alternative pairing observed in the B7 structure [12] would not allow any contacts between Ile7 and the β chain. The specificity of the 5H3 model for positions 6 and 7 is much less stringent than that observed for P1 and P4. It could explain why large peptide variations are tolerated at P6 and P7 by 5H3 [15]. As expected, no contacts between major MHC anchors (Glu2, Ile8) and the 5H3 TCR could be identified. The only apparent disagreement between the present model and binding data relates to the role of position 3 (Ser) of the Ha peptide. In our model, Ser3 is interacting with the upper polar part of pocket D of the K^k (Figure 2b), and no contacts to 5H3 were identified. Binding experiments show that Ser3 may be replaced by several amino acids without losing affinity for H2-K^k while fewer variations (notably small amino acids) are compatible with 5H3 recognition [16]. A plausible explanation to these peculiar structure-binding relationships would be an indirect effect of Ser3 on 5H3 binding in which a mutation (notably by bulky amino acids) may locally affect the conformation of K^k in the vicinity of pocket D. Mutation of MHC-anchoring residues leading to TCR antagonism [42] would support this hypothesis.

Modeling of the 6F8 TCR-Ha-K^k complex

The framework of the 6F8 Vαβ TCR was constructed by homology to the three existing X-ray structures (Figure1b), starting from the 2C TCR for the α chain (45.7% identity for the Vα part) and from the A6 TCR for the β chain (27.8% identity). The Vβ of the 6F8 chain presents a much lower sequence identity to

Table 4. 6F8-K^k-Ha complex: intermolecular hydrogen bonds

6F8	K^k	$Ha_{255-262}$	Frequency[a] (%)
Asp1α (OD1)	Arg62 (NH2)		100
Asp1α (OD32)	Arg62 (NE)		88
Asp1α (O2D1)	Arg62 (NH1)		100
Gly28α (O)	Trp167 (NE1)		72
Asp52α (OD1)	Arg157 (NE)		86
Asp52α (O2D1)	Arg157 (NH2)		100
Asp52α (OD22)	Arg157 (NH2)		26
Lys68α (NZ)	Glu166 (OE1)		100
Ser102α(OG)	Gln65 (OE1)		100
Asn52β (O)	Gln72 (NE2)		58
Tyr103β (OH)	Gly51 (O)		100
Asp100α (OD1)		Thr4 (OG1)	58
Asp100α (OD2)		Thr4 (OG1)	98
Asn101α (ND2)		Thr4 (OG1)	64
Gly100β (N)		Asn6 (OD1)[b]	

[a]Interactions were statistically monitored throughout the MD trajectory for a total of 250 conformations.
[b]Water-mediated H-bond.

Table 5. The 6F8-K^k-Ha complex: non-bonded hydrophobic contacts

6F8	K^k	$Ha_{255-262}$	# of contacts[a]
Asp1α	Arg62		1
Ser27α	Arg62		2
Gly28α		Phe1	3
Tyr29α	Thr163		1
Tyr29α		Phe1	2
Arg51α	Glu154		2
Asn101α	Gly69		1
Ser102α	Arg62		1
Asn52β	Gln72, Ile73		2
Trp98β		Leu7	4
Gly100β	Arg155		1

[a]Non-bonded contacts have been recorded, from the energy-minimized time-averaged MD structure, for carbon atoms distant by less than 4 Å, using the CONTACT program [31].

both crystal templates than that of 5H3. Therefore, all CDRβ loops were built by searching a 3D database of 837 X-ray structures for loops of similar length and composition. As for the 5H3, the CDR3β loop was built afterwards, once the Ha-K^k complex was docked to 6F8. Docking, energy refinement and molecular dynamics simulation of the solvated complex were realized using conditions identical to those used for the 5H3 TCR-Ha-K^k ternary complex. It may be noted that rms deviations of the refined MD structure from the starting homology derived model were very similar to that reported previously for the first ternary complex. Again, MD refinement ensures a tighter interaction of the CDR loops with the pep/MHC ligand. Docking the 6F8 αβ TCR to the $Ha_{255-262}/K^k$ ligand according to the A6/tax/HLA-A2 structure is also topologically feasible. In the energy-minimized time averaged conformation of the ternary complex (Figures 4a, b), the 5H3 Vαβ has a total of 14 direct hydrogen bonds (Table 4) and 22 close van der Waals contacts (Table 5) with its MHC-peptide ligand. As previously observed for 5H3, the selected 6F8 CDR β3 loop orients two of its side chains (Ser101β, Trp98β) towards peptide amino acids (Thr4, Ile7) known to interact with the 6F8 TCR [16]. As for 5H3 binding, the main 6F8 TCR anchor positions of the K^k-bound peptide are Phe1, Thr4 and Ile7. In the proposed model, the Phe 1 side chain is directed almost perpendicularly to the aromatic ring of Tyr29α, thus supporting the experimentally determined fine specificity of this TCR for Phe1-containing peptides [16]. The Vα chain delimits the upper part of the subsite that Phe1 is contacting (Figure 4a, b). Gly28α backbone atoms, in close contacts with the Phe1 side chain (Table 5), restrict the depth of the subsite and thus preclude the docking of bulkier aromatic amino acids (Tyr, Trp). In our complex with 6F8, the Thr4 side chain is hydrogen-bonded to the only CDR3 loop of the Vα chain (Asp100α, Asn101α; Figure 4a, b). No direct and stable interactions between Thr4 and the β chain could be observed all along the MD trajectory of the solvated ternary complex. The preferential recognition of Thr with respect to Ser at P4 [16] could not be explained by looking at peptide-TCR interactions but probably by the close proximity of a K^k residue (Ile73) whose methyl group is directed towards the methyl group of Thr4 (Figure 4b). The less stringent specificity of 6F8 compared to 5H3 for Asn at P6 is compatible with the present model in which no direct interactions to the TCR could be depicted. Only one water-mediated H-bond to Gly100β (CDR3β loop) could be found (Figure 4b). Interestingly, binding data show that the Asn to Trp mutation is still compatible with 6F8 recognition [16]. Hence, a Trp side chain at P6 would be located within the van der Waals interaction area of Trp98β, a hypervariable residue of the β chain. The last TCR anchoring amino acid is Leu7. As in the 5H3-bound state, its side chain is directed towards a flexible subsite composed of K^k residues (Ile73, Ile76) and TCR atoms of the CDR3β loop (Trp98β). As this subsite is entirely made of apolar residues, it accounts

well for the experimentally described preference of the 6F8 clone for hydrophobic amino acids at P7. Bulky hydrophobic side chains (Phe, Tyr, Trp) are however not allowed at P7, probably because some of the side chain atoms would come too close to both K^k (Ile73) and 5H3 (Trp98β) residues. Surprisingly, a non conservative peptide mutation (Leu to Arg or Lys) is still compatible with 6F8 binding [16]. A long and positively charged side chain at P7 could readily interact with a residue (Asn52β) located in the CDR2β loop (Figure 4b). Again, no real TCR binding role could be assigned to Ser3. No H-bond acceptor or donor atom of the TCR is located within 4 Å of the Ser3 hydroxyl moiety. Thus, we postulate an indirect effect of this position on 6F8 binding.

Comparative analysis of the two MD models with existing TCR/pep/MHC crystal structures

Obviously, MD models cannot be compared to high-resolution X-ray structures for a pure qualitative and quantitative analysis of receptor-ligand interactions. Therefore, caution should be given to numerical values discussed in the present section, especially since motion of the highly flexible TCR combining site may affect accessible and buried surface areas of each partner in the TCR-peptide MHC assembly. However, the models may be used as rough guides to depict general features of ligand recognition by αβ TCRs.

The energy-minimized MD time-averaged structures of the two modeled TCR/pep/MHC complexes were analyzed first by the PROCHECK program [28] for their stereochemical quality. The overall quality of the two models, although lower than that of the three available X-ray structures [10–12], is however very acceptable and lies in the range of expected values for molecular models [43]. As expected for a protein model, standard deviations of the peptide bonds (ω angle) is significantly higher for the two MD models (about 9 deg.) than for the X-ray structure template (about 1.5 deg.), which have not been further refined here. This discrepancy is easily explained by the rather low torsional energy barrier assigned to the ω dihedral in the AMBER95 force field [44]. Enhancing this force constant to higher values (50 kcal mol^{-1}) has been recently shown to decrease the standard deviation of ω angles [45]. The G-factor, a score based on the observation of stereochemical parameters (covalent geometry, torsion angles), reveals also an overall good quality of the two MD models (G-factor values above −0.5). For example, only 2 to 5 residues

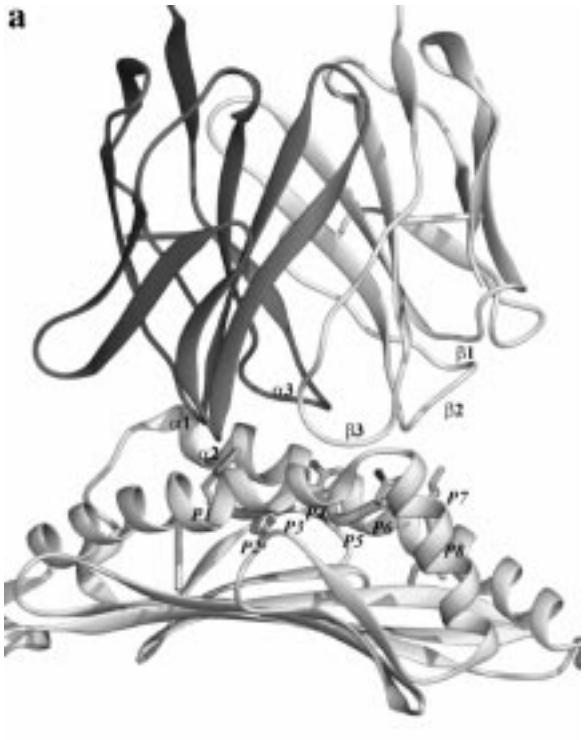

Figure 4. Three-dimensional model of the 5H3-K^k-Ha ternary complex (see Figure 3 for the legend). (b) Zoom into the 6F8-Ha-K^k interaction interface.

(out of about 420) were assigned to Ramachandran disallowed regions. They were basically located at the junction between the restrained and free regions of the ternary complexes (see Materials and methods). The folding of each model was then checked using a knowledge-based potential of mean force [30] by calculating pairwise residue interactions using β-carbons (Table 6). Energy profiles for the two MD models were very similar to that of the reference X-ray structures (data not shown). The quality of the protein fold, as determined by the z-score of pair energy (Cβ interactions) [30], was satisfactory for both models (Table 6).

As the orientations of both TCR models with respect to K^k were based on the A6/Tax/HLA-A∗0201 crystal structure, the global TCR/pep/MHC interactions of the two models presented here are, of course, more similar to those observed in the A6 template [11] than to those in the other two X-ray structures (2C and B7). The majority (about two third) of the intermolecular interactions of the TCR involves the K^k heavy chain. Due to the diagonal binding of the TCR across the MHC-peptide binding site, the Vα chain develops

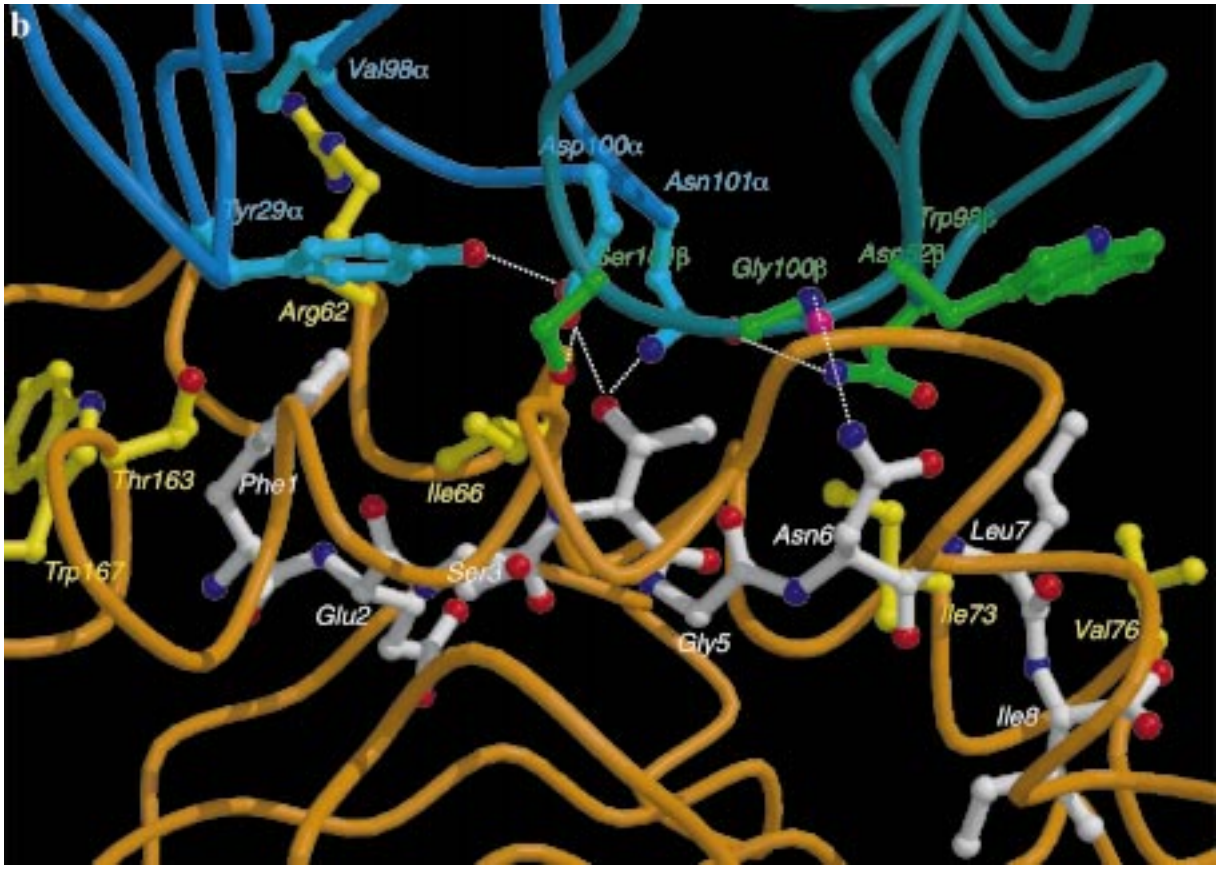

Figure 4. (continued).

much more interactions than the Vβ chain (Table 7). In fact, the two TCRs examined in the current study do not make extensive use of their β chains. As proposed by Garboczi, the length of the CDR3β loop (10 amino acids) probably precludes a tight docking of the β chain to the MHC binding groove, so that the respective contributions of β1 and β2 loops are really minor (Table 7). This feature is supported by the binding mode of the 2C TCR to the K^b-dEV83 complex [9, 10], for which a smaller β3 loop (6 amino acids, Figure 1b) allows a better engagement of the entire β chain. About 80% of the peptide surface is already deeply buried into the binding cleft of K^k prior to TCR binding (Table 1) leaving only about 20% accessible to TCR interaction. After TCR binding another about 17% of the peptide surface is buried in the TCR combining site, leaving only 5% of the Ha peptide surface (about 50 $Å^2$) accessible in the final ternary complex (Figure 5a). It follows that the interaction of the peptide with the TCR is much weaker than the interaction of the peptide with the MHC. These figures are very similar to those found in the two X-ray structure templates, suggesting that these values should be independent of the TCR and the pep/MHC ligand. As expected, the proportion of the peptide ligand buried in the TCR combining site is higher for a nonamer (Tax peptide) than for the octameric Ha peptide (Figure 5a). Hence, the surface of the peptide that is bulging out of the MHC binding groove is generally higher for a nonamer than for an octameric ligand. In any case, a very small fraction of the peptide surface remains free in all ternary complexes. Thus, peptide atoms may be basically classified into either MHC-anchors or TCR-anchors, and every peptide atom not involved in MHC binding should be engaged in TCR recognition and vice versa. Some secondary MHC anchors (P1 and to a lesser extent P3), partially buried in the MHC binding groove, play a dual binding role and are as important for MHC binding as for TCR recognition.

We have previously observed that the 6F8 is less sensitive than the 5H3 T cell hybridoma; a relationship which is further compounded by the 6F8

Table 6. Structural analysis of the two MD models

Structure	X-ray			MD models	
	1ao7[a]	2ckb[b]	1bd2[c]	5H3[d]	6F8[e]
Ramachandran plot results[f]					
Most favoured,%	86.7	71.8	86.9	79.2	74.9
Additional allowed,%	12.8	26.5	12.5	19.5	22.1
Generously allowed,%	0.5	1.4	0.6	0.6	1.6
Disallowed,%	0.0	0.3	0.0	0.6	1.4
Bad contacts <2.5 Å	2	44	3	31	1
ω angle standard deviation	1.5	1.2	0.7	9.0	10.1
G-factor[g]	0.19	-0.08	0.18	-0.39	-0.38
z-score[h]					
MHC	-6.01	-5.39	-6.21	-6.01	-5.50
TCRα chain	-4.18	-3.37	-3.77	-3.56	-3.32
TCRβ chain	-2.53	-3.53	-3.00	-2.77	-2.73

[a]A6/HLA-A2/Tax complex [10].
[b]2C/K^b/d-EV8 complex [11].
[c]B6/HLA-A2/Tax complex [12].
[d]5H3-Ha-K^k energy-minimized time-averaged MD model.
[e]6F8-Ha-K^k energy-minimized time-averaged MD model.
[f]Stereochemical analysis by PROCHECK [28]. The percentages of residues in 4 different regions of a Ramachandran plot are indicated.
[g]Score based on the observation of stereochemical parameters [28]. High resolution protein structures usually have a G-factor above -0.5. A score lower than -1.0 needs further refinement of the structure.
[h]Sequence-structure compatibility calculated from β-carbon pair interactions using ProsaII [29]. Correctly folded structures always have negative z-scores [30].

being CD8 dependent, whereas the 5H3 is CD8 independent. Although accessory factors independent on TCR-peptide-MHC interactions might explain the higher sensitivity of the 5H3 CD8-independent clone [46], we postulate that the isolated 5H3 TCR should have a much higher affinity for the $Ha_{255-262}/K^k$ ligand than the isolated 6F8 TCR has. The models suggest that the overall complementarity between 6F8 and the $Ha_{255-262}/K^k$complex is looser than that observed between 5H3 and the $Ha_{255-262}/K^k$complex (see Tables 4 and 5 vs. 2 and 3). The weaker recognition of the 6F8 may be explained by weaker interactions with the K^k protein rather than with the bound peptide (Figure 5b, Table 7). Note that the only conserved interaction between the 5H3, 6F8 and the A6-Tax-A2 ternary complexes, is the salt bridge between Lys68α (HV4 loop of the α chain) and Glu166 (K^k). As both residues are present in the vast majority of TCR [5] and class I MHC [47] sequences, it is likely that this interaction is a driving force to the docking of a TCR to a class I MHC-peptide pair (and should be an obvious target for mutational studies). Interestingly, water molecules as well as backbone atoms of the CDR loops also participate in binding to the pep/MHC ligand (Tables 2, 4) for both TCRs, a feature recently depicted in the refined X-ray structure of the 2C/dEV8/K^b complex [10], from which our models have not been derived. With respect to specificity, there is a noticeable difference between the 5H3 and the 6F8 models relating to peptide position 6 (Asn6), which was only interacting with the 6F8 TCR through one water-mediated H-bond. Experimentally, the 5H3 TCR has a much more stringent specificity for Asn6 than 6F8. Thus, the models fit nicely with the experimental observations, explaining both affinity and specificity differences between the 5H3 and the 6F8. Interestingly, the peptide contributes only to about 30% of the total surface recognized by the TCRs (Figure 5b). Whether this minor contribution is sufficient for a single TCR to discriminate among different peptides is an open question. It is tempting to speculate that the peptide dictates the TCR specificity whereas the class I MHC molecules would contribute more to the stability of the tertiary complex and dissociation of the pep/MHC ligand from its receptor. Ignoring the stabilization caused by CD8 accessory molecules [48],

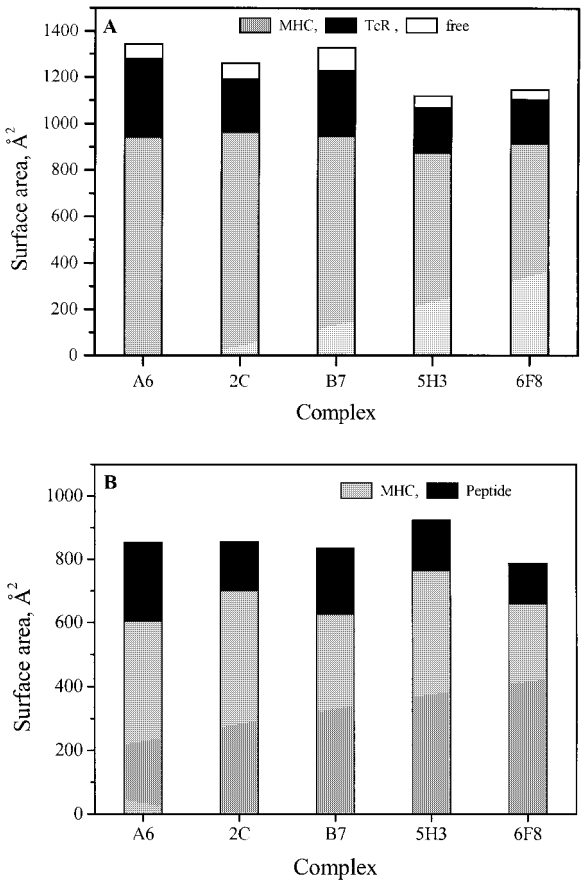

Figure 5. Buried surface areas of the TCR-MHC-peptide complexes, in three X-ray structures (A6, 2C, B7) and two MD models (5H3, 6F8). (a) Surface area of the peptide buried by the class I MHC molecule (gray), the TCR (dark). The total accessible surface area of the bound peptide is indicated in white. (b) Surface area of the TCR buried by the MHC protein (gray) and the peptide antigen (dark).

the dissociation constant of TCR/pep/MHC complexes usually lies in the μM range. For all TCR-MHC-peptide complexes, 90% of the binding (about 900 $Å^2$) is ensured by the six CDR loops (Table 7). Whether this property may be directly related to the binding affinity or the dissociation constant of the pep/MHC ligand is presently not known. Finally, the models proposed here are in agreement with the majority of the mutagenesis data [13, 49, 50] and do not support that the two TCRs studied here can be oriented oppositely relative to the pep/MHC as recently suggested for other ternary complexes [14, 15].

Table 7. Surface area of CDR variable loops buried by the MHC/PEP ligand

TCR region	X-ray[a]	MD models	
		5H3	6F8
α1,%[b]	17–26	27	23
α2,%	10–16	12	15
α3,%	16–23	20	17
β1,%	1–18	1	0
β2,%	0–19	5	13
β3 3,%	8–33	26	26
Framework,%	7–8	9	7
Vα, $Å^2$	453–557	621	492
Vβ, $Å^2$	248–422	338	344
Vα contribution,%	51–66	65	59
Vαβ, $Å^2$	845–915	959	836
Total, $Å^2$	923–979	1049	897

[a] $2C/K^b$/d-EV8 complex (pdb code: 2ckb) [11], A6/HLA-A2/Tax complex (pdb code: 1ao7) [10], B6/HLA-A2/Tax complex (pdb code: 1bd2) [12].
[b] Surface areas have been calculated, from the energy-minimized time-averaged MD structure, for carbon atoms distant by less than 4 Å, using the ACCESS program [31] with a probe radius of 1.4 Å. The contribution of each CDR loop is given as a percentage of the total buried surface area.

Conclusions

Homology models are proposed for a peptide/MHC class I complex as well as for two different T cell receptors both specific for this particular peptide/MHC pair, and for the two resulting ternary TCR/peptide/MHC complexes. The proposed models account well for a large body of experimental data. Thus, the present study suggests that the fine mechanisms of peptide selection by MHC molecules and TCR can be successfully modeled at a detailed molecular level. The approach taken here required that the target TCR possessed enough sequence identity to one or more template TCRs for which an X-ray structure has been determined. Such models should be validated by comparison to experimental data, i.e., involving variations in either the peptide, the MHC or the TCR molecule. The two ternary models are largely similar to available X-ray structures, supporting the notion of generalized T cell receptor binding mode. Combining experimental and computational approaches should enhance our understanding of MHC-restricted, peptide-specific T cell responses, and might allow for rational manipulation of the immune system in the future.

Acknowledgements

This work is supported by the Schweizerischer Nationalfonds zur Förderung der wissenschaftlichen Forschung (Project #31-45504.95) and by the Danish Medical Research Council (Project #9601615, #9305480 and #9400901). D.R. wishes to thank the calculation center of the ETHZ for allocation of computing time on the CRAY-J90 cluster and the Intel Paragon machine. The authors wish to thank P. Gosh and I. Wilson for making X-ray coordinates of MHC-peptide-TCR structures available prior to PDB release. Energy-minimized time-averaged coordinates of both ternary complexes (5H3-Ha-K^k, 6F8-Ha-K^k) are available in PDB format at the following URL: http://www.pharma.ethz.ch/ didier/jcamd

References

1. Townsend, A. and Bodmer, H., Annu. Rev. Immunol., 7 (1989) 601.
2. Heemels, M.T. and Ploegh, H.L., Annu. Rev. Biochem., 64 (1995) 643.
3. Batalia, M.A. and Collins E.J., Biopolymers, 43 (1997) 281.
4. Wilson, I.A. and Garcia, K.C., Curr. Opin. Struct. Biol., 7 (1997) 839.
5. Chothia, C., Boswell, D.R. and Lesk, A.M., EMBO J., 7 (1988) 3745.
6. Bentley, G.A., Boulot, G., Karjalainen, K. and Mariuzza, R.A., Science, 267 (1995) 1984.
7. Fields, B.A., Ober, B., Malchiodi, E.L., Lebedeva, M., Braden, B.C., Ysern, X., Kim, J.-K., Shao, X., Ward, S. and Mariuzza, R.A., Science, 270 (1995) 1821.
8. Davis, M.M. and Bjorkman, P.J., Nature, 344 (1988) 395.
9. Garcia, K.C., Degano, M., Stanfield, R.L., Brunmark, A., Jackson, M.R., Peterson, P.E., Teyton, L. and Wilson, I.A., Science, 274 (1996) 209.
10. Garcia, K.C., Degano, M., Pease, L.R., Huang, M., Peterson, P.A., Teyton, L. and Wilson, I.A., Science, 279 (1998) 1666.
11. Garboczi, D.N., Ghosh, P., Utz, U., Fan, Q.R., Biddison, W.E. and Wiley, D.C., Nature, 384 (1996) 134.
12. Ding, Y.H., Smith, K.J., Garboczi, D.N., Utz, U., Biddison, W.E. and Wiley, D.C., Immunity, 8 (1998) 403.
13. Jorgensen, J.L., Esser, U., Fazekas de St. Groth, B., Reay, P.A. and Davis, M.M., Nature, 355 (1992) 224.
14. Vasmatzis, G., Cornette, J., Sezerman, U. and DeLisi, C., J. Mol. Biol., 261 (1996) 72.
15. Sim, B.-C., Travers, B.J. and Gascoigne, N.R., J. Immunol., 159 (1997) 3322.
16. Stryhn, A., Andersen, P.S., Pedersen, L.O., Svejgaard, A., Holm, A., Thorpe, C.J., Fugger, L., Buus, S. and Engberg, J., Proc. Natl. Acad. Sci. USA, 93 (1996) 10338.
17. Stryhn, A., Pedersen, L.O., Romme, T., Holm, C.B., Holm, A. and Buus, S., Eur. J. Immunol., 26 (1996) 1911.
18. Casanova, J.R., Romero, P., Widmann, C., Kourilsky, P. and Maryanski, J.L., J. Exp. Med., 174 (1991) 1371.
19. Lipman, D.J., Altschul, S.F. and Kececioglu, J.D., Proc. Natl. Acad. Sci. USA, 86 (1989) 4412 (web site: http://www.ibc.wustl.edu/ibc/msa.html).
20. http://www2.igh-cnrs.fr/bin/align-guess.cgi
21. Fremont, D.H., Matsumura, M., Stura, E.A., Peterson, P.A. and Wilson, I.A., Science, 257 (1992) 919.
22. Rognan, D., Molecular modelling of protein-peptide interactions. Application to major histocompatibility complex proteins and design of nonpeptide ligands. Habilitationsschrift, ETH Zürich, 1997.
23. Pearlman, D.A., Case, D.A., Caldwell, J.C., Ross, W.S., Cheatham III, D.E., Fergusson, D.E., Seibel, G.L., Singh, U.C., Weiner P.K. and Kollman, P.A., AMBER 4.1, University of California, San Francisco, CA, USA, 1995.
24. Garcia, F., Rognan, D., Lamas, J.R., Marina, M. and López de Castro, J.A., Tissue Antigens, 51 (1998) 1.
25. Rognan, D., Scapozza, L., Folkers, G. and Daser, A., Biochemistry, 33 (1994) 11476.
26. Rognan, D., Scapozza, L., Folkers, G. and Daser, A., Proc. Natl. Acad. Sci. USA, 92 (1995) 753.
27. Rognan, D., Krebs, S., Kuonen, O., Lamas, J.R., López de Castro, J.A. and Folkers, G., J. Comput.-Aided Mol. Design, 11 (1997) 463.
28. Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M., J. Appl. Crystallogr., 26 (1993) 9283.
29. ProsaII v-3.0: http://www.came.sbg.ac.at/
30. Sippl, M., Proteins Struct. Funct. Genet., 17 (1993) 355.
31. Collaborative computational project, Number 4: Acta Crystallogr., D50 (1994) 760.
32. Fremont, D.H., Stura, E.A., Matsumura, M., Peterson, P.A. and Wilson, I.A., Proc. Natl. Acad. Sci. USA, 92 (1995) 2479.
33. Young, A.C.M., Zhang, W., Sacchettini, J.C. and Nathenson, S.G., Cell, 76 (1994) 39.
34. Zhang, W., Young, A.C.M., Imarai, M., Nathenson, S.G. and Sacchettini, J.C., Proc. Natl. Acad. Sci. USA, 89 (1992) 8043.
35. Balendiran, G.K., Solheim, J.C., Yong, A.C.M., Hansen, T.H., Nathenson, S.G. and Sacchettini, J.C., Proc. Natl. Acad. Sci. USA, 94 (1997) 6880.
36. Speir, J., Garcia, K.C., Brunmark, A., Degano, M., Peterson, P.A., Teyton, L. and Wilson, I.A., Immunity, 8 (1998) 553.
37. Achour, A., Persson, K., Harris, R.A., Sundbäck, J., Sentman, C.L., Lindqvist, Y., Schneider, G. and Kärre, K., Immunity, 9 (1998) 199.
38. Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer Jr., E.F., Brice, M.D., Rodgers, J.M., Kennard, O., Shimanouchi, T. and Tasumi, M., J. Mol. Biol., 112 (1977) 535.
39. Madden, D.R., Garboczi, D.N. and Wiley, D.C., Cell, 75 (1993) 693.
40. Housset, D., Mazza, G., Gregoire, C., Piras, C., Malissen, B. and Fontecilla-Camps, J.C., EMBO J., 16 (1997) 4205.
41. Wang, J.H., Lim, K., Smolyar, A., Teng, M.K., Liu, J.H., Tse, G.D., Liu, J., Hussey, R.E., Chishti, Y., Thomson, C.T., Sweet, R.M., Nathenson, S.G., Chang, H.-C., Sacchettini, J.C. and Reinherz, E.L., EMBO J., 1998 (17) 10.
42. Dong, T., Boyd, D., Rosenberg, W., Alp, N., Takiguchi, M., McMichael, A. and Rowland-Jones, S., Eur. J. Immunol., 26 (1996) 335.
43. Mosimann, S., Meleshko, R. and James, M.N.G., Proteins Struct. Funct. Genet., 23 (1995) 301.
44. Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz Jr., K.M., Ferguson, D.M., Spellmeyer, D.M., Fox, T., Caldwell, J.W. and Kollman, P.A., J. Am. Chem. Soc., 117 (1995) 5179.
45. Lozano, J.J., López-de-Briñas, E., Centeno, N.B., Guigó and Sanz, F., J. Comput.-Aided Mol. Design, 11 (1997) 395.
46. Zamoyska, R., Curr. Opin. Immunol., 10 (1998) 82.
47. Bjorkman, P.J. and Parham, P., Annu. Rev. Biochem., 59 (1990) 253.

48. Garcia, K.C., Scott, C.A., Brunmark, A., Carbone, F.R., Peterson, I.A., Wilson, I.A. and Teyton, L., Nature, 384 (1996) 577.
49. Sant-Angelo, D.B., Waterbury, G., Preston-Hurlburt, P., Yoon, S.T., Medzhitov, R., Hing, S.C. and Janeway Jr., C.A., Immunity, 4 (1996) 367.
50. Chang, H.-C., Smolyar, A., Spoerl, R., Witte, T., Yao, Y., Goyarts, E.C., Nathenson, S.G. and Reinherz, E.L., J. Mol. Biol., 271 (1997) 278.
51. Madden, D.R., Gorga, J.C., Strominger, J.L. and Wiley, D.C., Cell, 70 (1992) 1035.
52. Smith, K.J., Reid, S.W., Harlos, K., McMichael, A.J., Stuart, D.I., Bell, J.I. and Jones, E.Y., Immunity, 4 (1996) 215.
53. Saper, M.A., Bjorkman, P.J. and Wiley, D.C., J. Mol. Biol., 219 (1991) 277.
54. Kabat, E.A., Wu, T.T., Reid-Miller, M., Pevey, H.M. and Gotteman, K.S., Sequence of Proteins of Immunological Interest (NIH, Bethesda, MD) ed 5., 1991.
55. Kraulis, P.J., J. Appl. Crystallogr., 24 (1991) 946.
56. Merritt, E.A. and Murphy, M.E.P., Acta Crystallogr., D50 (1994) 869.
57. Nicholls, A., Sharp, K.A. and Honig, B., Proteins Struct. Funct. Genet., 11 (1991) 281.
58. Islam, S.A. and Sternberg, M.J.E. (1998) PREPI Version 0.95, Biomolecular modelling Laboratory, Imperial Cancer Research Fund, London, U.K. (http://www.icnet.uk/bmm)
59. http://www.povray.org.