Journal of Computer-Aided Molecular Design, **14:** 71–82, 2000.
KLUWER/ESCOM

Modeling the interactions of a peptide-major histocompatibility class I ligand with its receptors. II. Cross-reaction between a monoclonal antibody and two αβ T cell receptors

Didier Rognan[a,*], Jan Engberg[b], Anette Stryhn[c], Peter Sejer Andersen[c] & Søren Buus[c]
[a]*Department of Pharmacy, Swiss Federal Institute of Technology, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland;* [b]*Department of Pharmacology, The Royal Danish School of Pharmacy, Universitetsparken 2, DK-2200 Copenhagen N, Denmark;* [c]*Department of Experimental Immunology, Institute of Medical Microbiology and Immunology, Panum 18.3.22, University of Copenhagen, Blegdamsvej 3C, DK-2200 Copenhagen N, Denmark*

Received 26 January 1999; Accepted 23 June 1999

Key words: antibody, flexible docking, homology modeling, MHC, peptide

Summary

The recombinant antibody, pSAN13.4.1, has a unique T cell like specificity; it binds an Influenza Hemagglutinin octapeptide ($Ha_{255-262}$) in an MHC (H-2K^k)-restricted manner, and a detailed comparison of the fine specificity of pSAN13.4.1 with the fine specificity of two $Ha_{255-262}$-specific, H-2K^k-restricted T cell hybridomas has supported this contention. A three-dimensional model of pSAN13.4.1 has been derived by homology modeling techniques. Subsequently, the structure of the pSAN13.4.1 antibody in complex with the antigenic Ha-K^k ligand was derived after a flexible and automated docking of the MHC-peptide pair into the Fab combining site. Interestingly, the most energetically favored binding mode shows numerous analogies to the recently determined recognition of class I MHC-peptide complexes by αβ T cell receptors (TCRs). The pSAN13.4.1 also binds diagonally across the MHC binding groove but is more deeply anchored to the peptide-MHC (pep/MHC) ligand than TCRs, notably through numerous interactions of its heavy chain. The present model accounts well for the experimentally determined binding affinity of a set of 144 single amino acid substituted Ha analogues and the observed shared specificity between the pSAN antibody and two different T cell receptors for the Ha-K^k antigenic ligand. Analogies and differences between Fab and TCR recognition are explained by dissecting the binding role of each chain of the immune receptors as well as the contribution of all peptide amino acids.

Introduction

Immunoglobulins (Igs) and T cell receptors (TCRs) play a central role in our immune system by selectively recognizing foreign molecules. Although the three-dimensional fold of both protein classes has been structurally conserved along evolution [1, 2], they bind to totally different ligands. Igs contact entire target antigens [3] whereas TCRs only recognize small antigenic peptides bound to proteins of the Major Histocompatibility Complex (MHC) [4]. Thus, MHC-restricted peptide-specific crossrecognition of a single MHC-peptide pair by TCRs and Igs is rare. Even if recognition of MHC molecules by Igs occurs, the binding is generally localized to exposed parts of the MHC molecule [5] with a minor contribution of the bound peptide [6]. Using a phage-display library approach we have identified a unique T cell like monoclonal antibody, pSAN13.4.1 [7], which shares fine specificity with two different αβ TCRs (HK8.3-5H3, HK8.3-6F8) specific for the same antigenic ligand, a complex between the $Ha_{255-262}$ octapeptide (FESTGNLI, one-letter amino acid code) and the class I MHC H-2K^k molecule (K^k) [8]. Several lines of evidence suggest that the two TCRs and the pSAN Fab should interact similarly with the Ha-K^k complex:

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

(i) for all three receptors the main anchors of the Ha peptide are found at the same positions: P3, P4, P5 and P7 (Pn standing for position n), (ii) systematic single-amino acid substitutions of the Ha peptide lead to analogous variations of the binding to the three Ha-K^k receptors (the only discrepancy relates to the binding role of the peptide N-terminus (Phe1), which is crucial for TCR recognition, but not important for pSAN recognition).

We have previously modeled both the HK8.3-5H3 and the HK8.3-6F8 T cell receptor in association with the Ha-K^k ligand [9]. To delineate analogies and(or) differences between the pSAN-Ha-K^k ternary complex and two TCR-Ha-K^k ternary complexes, we have in this paper modeled the pSAN13.4.1 antibody in association with the Ha-K^k ligand. As the three-dimensional (3D) fold of Igs is highly conserved [10], we initially built an homology model of the pSAN Fab. The Ha-K^k ligand was then docked into the Fab combining site using an automated procedure, allowing for peptide side chain flexibility. The suggested optimal docking of the pSAN with the Ha-K^k ligand showed many structural similarities to the recently determined TCR-peptide-MHC structures supporting the contention that pSAN binds to the Ha-K^k ligand like an αβ T-cell receptor. Thus, the present model provides a rational explanation for the observed similarities, as well as differences, between the Fab and the two TCRs.

Materials and methods

Primary sequence alignment of the pSAN13.4.1 antibody

An in-house 3D database of 51 high resolution X-ray structures of Igs was first searched for selecting the light and heavy chains that present the highest sequence identity to the pSAN antibody, using the COMPOSER module [11] of the SYBYL software (TRIPOS Assoc., Inc., release 6.3). Primary sequences were aligned using the Needleman-Wunsch algorithm [12] and a modified Dayhoff PAM250 matrix [13]. A maximal insertion/deletion sequence of 8 amino acids was tolerated and a 70% identity cut-off was used to retrieve primary sequences. The search was restricted to the variable domains of the pSAN light and heavy chains (V_L: residues 1–109; V_H: residues 1–113).

Modeling of the peptide-MHC complex

The complex between the $Ha_{255-262}$ octapeptide and K^k has been previously built and refined by molecular dynamics (MD) simulations [9]. The energy-minimized time-averaged MD conformation was directly selected for further docking purpose.

Homology modeling of the pSAN antibody

The pSAN light chain shows a 96.7% sequence identity to the neuramidinase NC41 light chain (pdb code: 1nca) whereas the heavy chain exhibits a 81.7% amino acid identity to the murine antiphenylarsonate antibody heavy chain (pdb code: 6fab). Both X-ray structures were then chosen as templates for building the framework of the pSAN antibody. The first five hypervariable loops (L1-3, H1-2) were also built directly from the reference X-ray structures due to very high sequence conservation. The hypervariable H3 loop, for which no canonical conformation could have been defined [10], was built after searching a set of 118 X-ray structures of H chains available in the Protein Data Bank [14] for H3 regions of similar length (11 amino acids). The H3 region (from the anti-ouabain Fab, pdb code 1ibg) showing the highest sequence homology to that of pSAN was finally selected (Table 1).

Automated flexible docking of the Ha-K^k complex to the pSAN Fab region

The Ha-K^k complex was systematically docked into the Fab binding site using the AutoDock 2.4 program [15] which uses a combined Monte Carlo (MC) /Simulated Annealing (SA) docking approach. To simplify and fasten the docking MCSA procedure, only the Ha peptide was docked, starting from its K^k-bound model [9] and allowing conformational flexibility only for the side chains located between P4 and P8 positions. First, a three-dimensional grid (size: 22.5 Å × 22.5 Å × 22.5 Å, resolution: 0.375 Å) was centered on the pSAN combining site. Steric and electrostatic interaction energies were pre-computed before the docking phase, at each point of the three-dimensional lattice between the Fab binding site and several probe atoms (C, N, O, H) representing the different atom types of the Ha peptide ligand. AMBER4.1 non-bonded parameters and atomic charges [16] were used for that purpose. A sigmoidal distance-dependent dielectric function [17] was used for computing electrostatic interactions. Second, the Ha peptide was docked into the Fab starting from its K^k-bound conformation [9]

Table 1. Construction of the six complementary determining regions (CDRS) of the pSAN antibody

CDR[a]	Origin	PDB code	Homology[b]
L1	N9 Neuramidinase-NC41 Fab L Chain (26-32)	1nca	100
L2	N9 Neuramidinase-NC41 Fab L Chain (50-52)	1nca	100
L3	N9 Neuramidinase-NC41 Fab L Chain (91-96)	1nca	83
H1	Anti-phenylarsonate H Chain (26-32)	6fab	85
H2	Anti-phenylarsonate H Chain (52_A-55)	6fab	75
H3	Anti-ouabain H Chain (96-101)	1ibg	42

[a]CDR regions were assigned as defined by Chothia [10] using the Martin numbering [33].
[b]Homology calculated after sequence alignment using the Needleman-Wunsch algorithm [13] and a PAM250 matrix.

but using random dihedral angles for pSAN-binding side chains (P4-P8) [8]. 50 independent docking runs of each 50 temperature-reduction cycles (from 500 to 40 K) were realized starting from a random orientation of the Ha peptide within the three-dimensional grid. Within each SA cycle, random MC modification of the peptide conformation (side chain dihedral angles of P4 to P8 positions) were performed. Once 3000 MC steps were either accepted or rejected according to the classical Metropolis criterion [18], the Boltzmann temperature factor was then decreased, and another cycle of MC docking steps performed starting from the most energetically favored state of the previous cycle. Thus, a minimal number of $50 \times 50 \times 3000$ (7 500 000) bound states have been scanned by the flexible docking procedure. The most favored bound state for each independent docking run was saved, and the resulting 50 docking proposals were examined by a cluster analysis based on rms deviations from the starting orientation of the bound ligand. Two different orientations were clustered in a same family when rms deviations of all ligand atoms were within 1.0 Å. For the 43 families found after clustering, each peptide location was further used as a guide to merge the entire peptide-MHC complex into the Fab binding site. Out of these 43 possibilities, 40 were topologically impossible because numerous steric clashes were observed between Fab and MHC backbone atoms. From the remaining three solutions, the docking state presenting the best Fab-pep/MHC interaction energy was selected as the final hit. It was the only one for which all peptide side chains were orientated according to experimental binding studies [8], either towards the MHC protein or in the direction of the Fab combining site.

Energy refinement and molecular dynamics of the pSAN-Ha-K^k complex

Energy refinement and restrained MD simulation of the Fab-peptide-MHC ternary complex were performed using the AMBER95 force field [19], as previously described for the ternary complex between the same MHC-peptide pair and two different T cell receptors [9].

Analysis of the refined MD model

All dynamical properties of the ternary complex were analyzed by in-house routines and the CARNAL program of the AMBER4.1 software [16]. The stereochemical reliability of each energy-minimized time-averaged MD conformation was assessed by the PROCHECK program [20]. Correct folding of each MD model was analysed with the ProsaII software [21] by computing pairwise residue interaction energy using a C_β-C_β pair potential [22] and a 50-residue window for energy averaging at each residue position. Surface areas and non-bonded interactions were examined using the ACCESS and CONTACT programs of the CCP4 package [23].

Results and discussion

Homology modeling of the pSAN antibody

The V_L and V_H sequences of the pSAN antibody share a very high level of identity with known Igs for which a crystal structure has been reported. Hence, the N9-neuramidinase NC41 Fab was selected for building the pSAN light chain (93.6% identity for the L chain, Figure 1). No loop searching procedure was necessary for the three L variable regions as: (i) L1 and L2 loops

Light Chain

```
            1        10        20        30        40        50        60        70        80        90        100
            .         .         .         .         .         .         .         .         .         .         .
pSAN        DIQMTQSPKFMSTSVGDRVSITCKASQDVSTAVAWYQQKPGQSPKLLIYWASTRHTGVPDRFTGSGSGTDYTLTISSVQAEDLALYYCQQHYSTPYTFGG
1nca        DIVMTQSPKFMSTSVGDRVTITCKASQDVSTAVVWYQQKPGQSPKLLIYWASTRHIGVPDRFAGSGSGTDYTLTISSVQAEDLALYYCQQHYSPPWTFGG
                                  *    L1                   L2                                                  *    L3
            101     109
            .         .
pSAN        GTKLEIKRA
1nca        GTKLEIKRA
```

Heavy Chain

```
            1        10        20        30        40        50  53     60        70        80  83    90
            .         .         .         .         .         .   .     .         .         .   .     .
pSAN        QVQLQQSGAELAKPGSSVKMSCKASGSTFTTHWMHWVKQRPGQGLEWIGFINPSTGYADYNQNFKDKATLTADKSSNTAYMQLSSLTSEDSAVYFCTRNP
6fab        EVQLQQSGVELVRAGSSVKMSCKASGYTFTSNGINWVKQRPGQGLEWIGYNNPGNGYIAYNEKFKGKTTLTVDKSSSTAYMQLRSLTSEDSAVYFCARSE
                                  *       H1                  H2                                              *

               100   101       110
                .     .         .
pSAN        TVVARYWYFDVWGAGTTVTVST
6fab        YYGGSYKF-DYWGQGTTLTVSS
              H3
```

Figure 1. Alignment of the primary sequences of the $V_{L,H}$ regions (1-109L, 1-113H) of pSAN to those of two X-ray structure templates (pdb codes: 1nca, 6fab). Primary sequences are numbered according to Martin [33]. Hypervariable regions (L1-3, H1-3) are defined according to Chothia [10] and enclosed by solid boxes. Underlined residues, labeled as the preceding non-underlined amino acid with additional Arabic letters (e.g., Pro52$_A$H), correspond to insertion sequences. Cysteines involved in disulfide bridging are noticed by a star.

are identical to that of the X-ray template, (ii) the conformation of the L3 hypervariable region could be unambiguously classified into the canonical structure type 1 as described in the literature [10]. The V_H chain of pSAN also presents a high sequence identity (82% for the entire H chain, 75% for the framework region) to the antiphenylarsonate V_H domain. It is thus largely sufficient for building a reliable model according to knowledge-based homology modeling rules [11]. Because of high sequence identity and conserved length, H1 and H2 CDR loops were also modeled by homology to the antiphenylarsonate H chain. The most hypervariable H3 loop was built by homology to another crystal structure template (anti-ouabain Fab, Table 1).

Flexible docking of the Ha-K^k complex to the pSAN antibody

To allow all possible binding modes, the Ha-K^k complex was systematically docked into the Fab binding site using a combined MCSA docking approach [15]. The MCSA technique allows a flexible docking of small ligands to a macromolecule, always converging for each independent docking attempt towards energetically favored docking states. It has been successfully used for reproducing Ig-bound conformations of peptide ligands [24] as well as for predicting binding modes of small molecules prior to X-ray structure determination of the corresponding protein-ligand

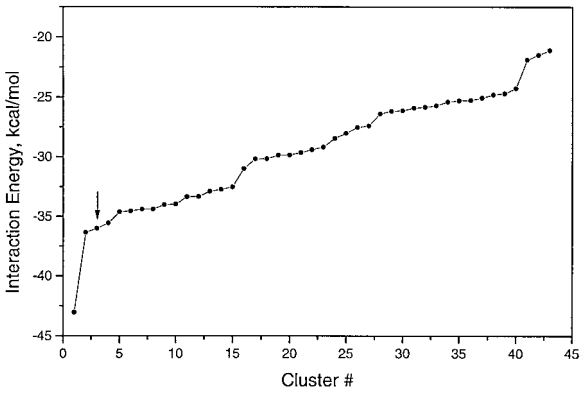

Figure 2. Fab-peptide interaction energy for the 50 best docking modes classified into 43 clusters, found by AutoDock [15]. The interaction energy (E_{int}, kcal/mol) is here the sum of the intramolecular interaction energy (as defined in the experimental section) plus the internal energy of the ligand. The selected cluster is indicated by an arrow.

complexes [25]. The 50 best orientations proposed by AutoDock [15] on the basis of Fab-peptide interaction energy could be clustered into 43 families (Figure 2). Although the interaction energy differences are relatively tiny between the best 15 solutions, the proposed orientations differ a lot once the bound peptide structure is further used to orient the whole Ha-K^k complex (Table 2). Hence, most of the proposed orientations are absolutely incompatible with a proper anchoring of the MHC protein in the final ternary complex because of numerous steric clashes between the pep/MHC lig-

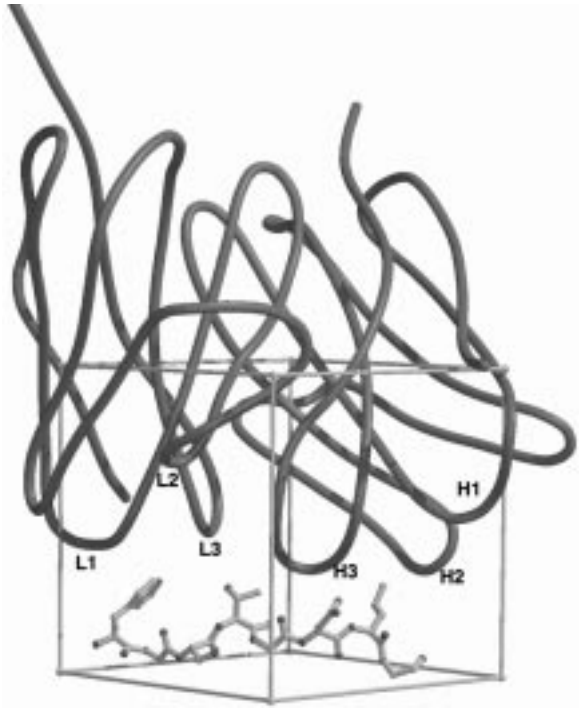

Figure 3. Three-dimensional box encompassing the pSAN combining site, used for the flexible docking of the Ha peptide. The pSAN backbone is represented by tube ribbons (L, light chain; H, heavy chain). The six hypervariable regions are labeled (L1-3, H1-3). The most favored pSAN-bound conformation of the HA peptide (ball and stick model) found by AutoDock [15] is displayed. The picture has been obtained using the MOLSCRIPT program [34] and rendered with Raster3D [35].

and and its receptor. Only three docking modes out of the 10 top docking solutions (#2, #3 and #7; Table 2) are topologically possible from the MHC point of view. Out of these three possibilities, the docking mode #3 was finally selected as it corresponds to the only one for which the experimentally determined Fab-anchoring side chains of the peptide (P3, P4, P6, P7) [8] were all orientated towards the Fab combining site. All solutions after the 10th proposal were found to be unrealistic from the MHC point of view (Table 2).

Interestingly, the most favored docking state to the pSAN Fab (Figures 3, 4) is very analogous to that observed in TCR-MHC-peptide X-ray structures [26–30] giving thus a very first explanation for the experimentally determined shared specificity between the two TCRs and the pSAN antibody for the same Ha-K^k ligand [8]. By comparison to the TCR binding mode depicted in several crystal structures, the pSAN Fab also binds diagonally across the binding groove but the entire protein is translated by about 3 Å in the direction of the MHC groove in order to optimize interactions of the H3 loop with the pep/MHC ligand (Figure 4). The pSAN Fab footprint on the Ha-K^k complex is thus different from that reported for 6 Fabs for which binding to the K^b molecule occurs to discrete α helical segments of the MHC protein [5].

The pSAN-Ha-K^k ternary complex

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 Fab-Ha-K^k complex was finally solvated, energy minimized and submitted to a restrained 500-ps molecular dynamics simulation as previously described [9]. The refined MD model of the pSAN Fab did not differ significantly 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 of the six hypervariable loops from the starting structure were significantly lower (from 0.34 to 0.69 Å) than those observed after a similar refinement of TCR loops [9]. The conserved structure of the Fab loops may be explained by a much tighter docking of the Fab to the pep/MHC ligand with respect to the docking model hypothesized for two αβ TCRs [9].

The benefit of the MD refinement over the pure homology derived model concerns here the interaction interface between the Fab and the MHC molecule. The K^k protein was significantly modified (rms deviations of 1.05 Å for the α1 α-helix) upon MD refinement. 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.

Analysing the stereochemical parameters of the refined MD model of the pSAN antibody indicates no major distortion of the structure (Table 3). Five residues (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). Potential misfolds were then searched using a knowledge-based potential of mean force [22] by calculating pairwise residue interactions from β-carbons (Table 3). The energy profile of the MD model suggests a reliable fold for the pSAN antibody because pairwise residue interaction energies are negative throughout the sequence (Figure 5). The overall

Table 2. Comparison of the 50 most favored docking modes

Cluster[a]	E_{int}^{b}	Orientations[c]	Rmsd[d]	Docking of the whole Ha-K^k complex?
1	−38.66	4	3.81	impossible: steric clashes between MHC (α2 helix) and Fab ($V_{L\eta}$ chains)
2	−35.35	1	5.58	possible: no clashes, but peptide position P7 orientated towards solvent
3	−36.01	2	1.12	possible: no clashes, all peptide positions well orientated
4	−35.57	2	5.07	impossible: steric clashes between MHC (α1 helix) and Fab (V_L chain)
5	−34.65	1	3.82	impossible: steric clashes between MHC (α2 helix) and Fab ($V_{L\eta}$ chains)
6	−34.41	3	5.58	impossible: steric clashes between MHC (β pleated sheet) and Fab ($V_{L\eta}$ chains)
7	−34.39	1	5.93	possible: no clashes, but peptide position 4 orientated towards solvent
8	−34.09	1	2.89	impossible: steric clashes between MHC (α2 helix) and Fab ($V_{L\eta}$ chains)
9	−33.99	2	4.71	impossible: steric clashes between MHC (α2 helix) and Fab ($V_{L\eta}$ chains)
10	−33.37	1	4.28	impossible: steric clashes between MHC (α2 helix) and Fab ($V_{L\eta}$ chains)
11 to 43	−33.34 to 21.08	1	2.12 to 7.23	impossible: steric clashes between MHC and Fab

[a]Cluster number, classified by increasing Fab-peptide interaction energy.
[b]Mean interaction energy (kcal/mol) for all complexes of the same cluster: Intermolecular interaction energy + peptide internal energy.
[c]Number of similar orientations in the cluster.
[d]rms deviations (in Å) of the Ha peptide (all atoms) from its TCR-bound position (expected location when bound to an αβ TCR).

quality of the antibody fold, as determined by the negative z-scores of pair energy [22] was also satisfactory for the MD model (Table 3).

Numerous intermolecular contacts between the MHC-peptide pair and its Fab receptor are observed in our refined model (19 H-bonds, 53 van der Waals contacts; see Tables 4, 5). The L1 region mainly interacts with the C-terminal part of the α2 helix of K^k and the protruding Arg62 side chain of the α1 helix (Figure 6). The L2 loop lies over the central part of the K^k α2 helix and develops almost exclusively apolar contacts to the MHC molecule. The central L3 and H3 regions bind to both the bound peptide and the K^k molecule (Tables 4, 5). The L3 hypervariable region contacts the peptide N-terminal amino acids (P1 to P4) whereas the H3 loop interacts with the C-terminal half (P4 to P7) of the Ha antigen. H2 and H1 regions have lesser contacts to the pep/MHC ligand, mainly to the α1 helix of K^k and the peptide C-terminal amino acids, respectively.

The Ha peptide preferentially contacts the L3 and H3 loops (Figures 4, 6). Very few interactions between the peptide N-terminal side chain (Phe1) and the L1-L2 loops could be depicted. When bound to pSAN, Phe1 side chain only contacts the Tyr92L aromatic ring. The main restriction for the P1 side chain is thus given by pocket A of the K^k molecule. This peculiar binding mode accounts well for the broad specificity of the pSAN antibody which can accommodate a large variety of amino acids at position P1 of the peptide [8]. The lack of van der Waals interactions at P1 is however compensated by two H-bonds between P1-

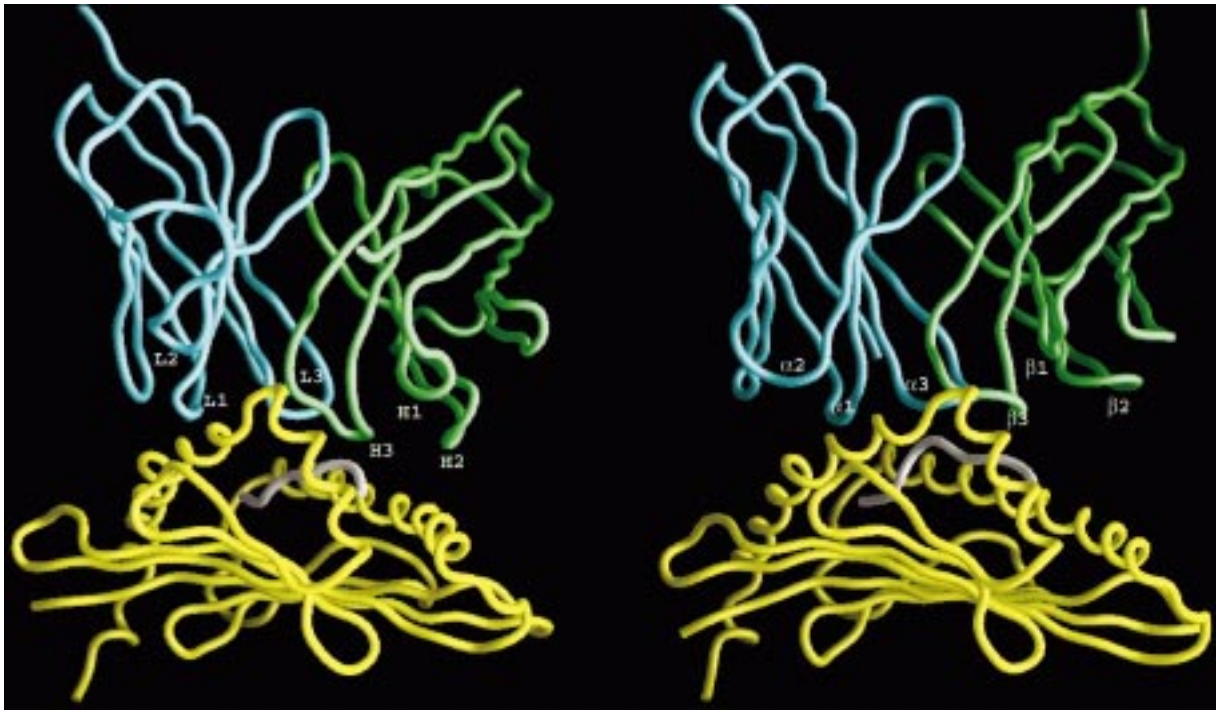

Figure 4. Comparison of the binding mode of the pSAN Fab (left) and of the 5H3 TCR (right) to the Ha-K^k ligand. The backbone traces of the ternary complexes are displayed as tube ribbons (cyan, V_L chain of pSAN, Vα chain of 5H3; green, V_H chain of pSAN, Vβ chain of 5H3; yellow, K^kantigen-binding domain; white, $Ha_{255-262}$ peptide). Hypervariable regions are labeled for the Fab (L1-3, H1-3) and the TCR (α1-3, β1-3). The picture has been obtained using the MOLSCRIPT program [34] and rendered with Raster3D [35].

Table 3. Structural analysis of the PSAN-HA-KK model

Ramachandran plot results[a]	
Residues in most favoured regions, %	79.7
Residues in additional allowed, %	16.6
Residues in generously allowed, %	2.5
Residues in disallowed regions, %	1.2
Bad contacts <2.5 Å	0
ω Angle standard deviation	8.3
G-factor[b]	−0.24
z-score[c]	
MHC	−5.83
Fab L chain	−3.06
Fab H chain	−5.66

[a]Stereochemical analysis by PROCHECK [20]. The percentages of residues in 4 different regions of a Ramachandran plot are indicated.
[b]Score based on the observation of stereochemical parameters [20]. High resolution protein structures usually have a G-factor above −0.5. A score lower than −1.0 needs further refinement of the structure.
[c]Sequence-structure compatibility calculated from β-carbon pair interactions using Prosa II [21]. A positive z-score of pair energy interaction indicates a probable misfold [22].

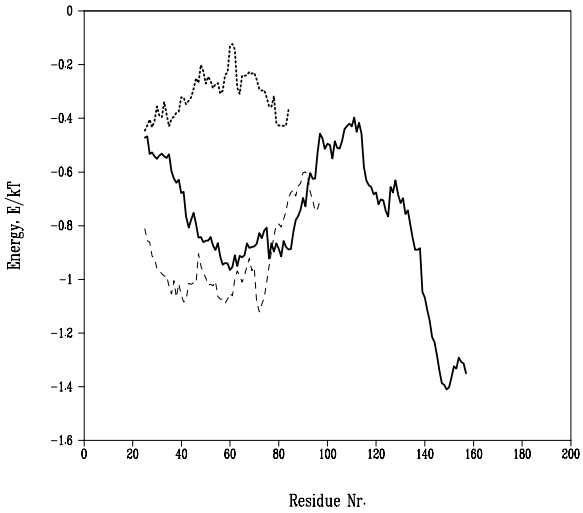

Figure 5. Energy profile of the pSAN-Ha-K^kMD model. The pairwise residue interaction energy is computed from a knowledge-based potential of mean force [21, 22] using C_β carbon interactions, and is given as E/kT. A 50-residue window has been used for energy averaging at each residue position. Solid, dotted and dashed lines represent energy profiles for the K^k MHC protein, the pSAN light chain and the pSAN heavy chain, respectively.

Figure 6. Close up into the pSAN-Ha-K^k interaction interface. The color coding is the following: cyan, carbon atom of the pSAN L chain; green, carbon atom of the pSAN H chain; yellow, carbon atom of the K^k heavy chain; white, carbon atom of the $Ha_{255-262}$ peptide; blue, nitrogen; red, oxygen. Intermolecular hydrogen bonds are displayed as white broken lines. The picture has been obtained using the MOLSCRIPT program [34] and rendered with Raster3D [35].

P2 carbonyl groups and the Tyr92L side chain. The second Fab anchor, in agreement with binding data on Ha analogues [8] is Ser3, whose hydroxyl group in our model is hydrogen-bonded to $Tyr100_B$ of the H chain. The close proximity of this aromatic amino acid may explain the experimentally observed specificity of the pSAN antibody for Ser or small amino acids (G, A). Larger side chain chains would be accommodated by the K^k binding groove but would be expelled from the Fab binding site because of steric hindrance with the L3 hypervariable loop. Surprisingly, the most important Fab-anchoring position (Thr4) determined by experimental binding studies [8] was found to be hydrogen-bonded to backbone atoms of the L3 and H3 loops. By analyzing the atomic trajectories of the ternary complex, two H-bonds to Tyr92L (L3 region) and Ala100H (H3 region) could be alternatively disclosed (Table 4). The experimentally determined fine specificity of the pSAN antibody for Thr at P4 may be explained not only by the previously described intermolecular hydrogen bonds (which would be also possible for Ser, the other residue allowed at P4) but also by the close proximity of the H3 loop which develops a total of 10 non-bonded contacts to the Thr4 side chain (Table 4). The third Fab-anchor residue, Asn6, is hydrogen-bonded to a polymorphic position of the heavy chain (Thr97H side chain, H3 loop). As previously proposed for TCR-Ha-K^k complexes [9], the Asn6 side chain is both bound to K^k (Asp152) and its corresponding immune receptor, in casu pSAN (Figure 6). However, Asn can be replaced by numerous amino acids at P6, without impairing Fab recognition. Binding of small apolar side chains (I,V), observed in in vitro binding experiments [8] may be explained by extensive hydrophobic contacts to the apolar part of the H3 loop (Val98-Val99) sitting near the P6 side chain. It is also not surprising that a small-sized hydrogen-bond acceptor (Ser, Thr) is recognized

Table 4. Intermolecular hydrogen bonds in the pSAN-HA-K^k complex

pSAN	K^k	$Ha_{255-262}$	Frequency[a] (%)
Asp1L (OD1)	Arg62 (NH2)		90
Gln27L (OE1)	Arg170 (NH2)		98
Asp28L (OD2)	Thr163 (OG1)		100
	Trp167 (NE1)		94
Ser30L (N)	Thr163 (OG1)		54
Ser30L (OG)	Thr163 (OG1)		100
Thr31L (OG1)	Ala158 (O)		30
Thr53L (OG1)	Glu154 (OE2)		100
Thr94L (O)	Arg62 (NH1)		68
Ser54H (O)	Arg79 (NH2)		72
$Tyr100_B$H (OH)	Asp156 (OD2)		40
$Trp100_C$H (NE1)	Asp152 (OD1)		100
Tyr92L (OH)		Phe1 (O)	30
Tyr92L (OH)		Glu2 (O)	100
$Tyr100_B$H (OH)		Ser3 (OG)	36
Tyr92L (O)		Thr4 (OG1)	92
Ala100H (O)		Thr4 (OG1)	50
Thr97H (OG1)		Asn6 (OD1)	100
		Leu7 (N)	38

[a]H-bonds were geometrically defined by an acceptor (A) to donor (D) distance less than 3.25 Å and a D-H..A angle higher than 120 deg. Interactions were statistically monitored during the whole MD trajectory for a total of 250 conformations.

Table 5. Non-bonded hydrophobic contacts in the pSAN-HA-K^k complex

pSAN	K^k	$Ha_{255-262}$	# of contacts[a]
Asp1L	Arg62		1
Ile2L	Arg62		2
Gln27L	Trp167		1
Asp28L	Thr163, Trp167		2
Trp50L	Glu154, Arg155, Ala158		12
Tyr92L	Tyr159, Thr163		2
		Phe1	2
Thr31H		Leu7	2
Thr54H		Leu7	2
Val98H	Ile73		3
		Thr4, Asn6, Leu7	5
Val99H	Gly69, Gln72, Ile73		4
Ala100H	Ile66		1
		Thr4	3
$Arg100_A$H		Thr4	1
$Tyr100_B$H	Arg155		5
		Thr4	3
$Trp100_C$H	Gly151, Asp152		2

[a]Non-bonded contacts have been recorded from the energy-minimized time-averaged conformation, between carbon atoms distant by less than 4 Å, using the CONTACT program [23]. L and H represent the light and the heavy chain of the pSAN Fab, respectively.

by the Fab binding site as the H-bond to Thr97H would be conserved. However, the present model cannot account for the recognition of Ha analogues bearing a negatively charged amino acid (Asp, Glu) at P6, that should normally be electrostatically disfavored by the proximity of Asp152 in the K^k binding groove. Thus, the most plausible explanation for the weak binding of these ligands is a conformational rearrangement of the K^k-bound peptides, allowing a longer distance between the two negatively charged amino acids. The last anchor residue, Leu7, is centered in an apolar subsite made of K^k residues (Ile73, Val76) and three amino acids of the pSAN heavy chain (Thr31H, Thr54H, Val98H). The binding pocket is not fully occupied and may thus accommodate apolar side chains of similar bulk (Ile) or even higher (Met, Phe) as suggested by single substitutions of the parental Ha epitope [8].

Comparative receptor-ligand analysis of pSAN vs. T cell receptors

Due to the deep docking of the pSAN Fab into the MHC binding groove, more intermolecular interactions (Fab-MHC, Fab-peptide) are found when the Ha-K^k ligand is bound to the Fab, than when it is bound to the two TCRs previously studied [9]. Interestingly, the number of H-bonds between K^k and the three receptors is conserved. The peptide surface buried by the Fab is twice as important as that buried by the two TCRs. It remains also higher than the corresponding surface buried in existing TCR-peptide-MHC X-ray structures (Table 6). It is difficult to ascertain if this increase is related to the observed decrease in the surface area buried by K^k or to a conformational rearrangement of the peptide. More three-dimensional structures of Fab-peptide-MHC complexes would be necessary to depict general rules if they exist.

From the viewpoint of the peptide, the proposed binding modes to two αβ TCRs and a Fab are different. In the TCR-bound state, by analogy to several reported crystal structures [26–30], the peptide N-terminus contacts the α1 loop whereas the central α3 and β3 regions interact with P4-P7 residues [9]. When bound to the pSAN Fab, the whole peptide only interacts with the L3 and H3 loops (corresponding to the α3 and β3 CDR regions, respectively; Figure 4). Notably, the peptide N-terminus (Phe1) has nearly no contacts to the L1 and L2 regions (Tables 4, 5). Thus the present model accounts well for the lack of preferred amino acids at the P1 position of the peptide recognized by pSAN [8]. This is a striking difference with the binding environment of Phe1 in its TCR-bound state [9], which is composed of amino acids of the CDR1α region and of pocket A of K^k. Altogether, they contribute to the fine specificity of two T cell clones for a Phe residue at position P1 [8].

One of the possible structural explanations for the proposed different binding mode relates to the re-

Table 6. Accessible and buried surface areas of the Ha peptide to three different receptors

Receptor	TCR		Fab
	X-ray[a]	MD[b]	pSAN
Free[c]	69–105	46–54	25
MHC-buried	944–966	878–914	819
TCR(Fab)-buried	224–232	188–190	376

[a]Range of values calculated from three X-ray structures of TCR-MHC-peptide complexes: A6/Tax/HLA-A2 [26], 2C/d-EV8/H-2K[b] [28], B7/Tax/HLA-A2 [29].
[b]Range of values calculated from energy-minimized time averaged MD models of TCR-Ha-K^k complexes [9].
[c]Surface areas have been calculated recorded from the energy-minimized time-averaged conformation using the ACCESS program [23] with a probe radius of 1.4 Å.

Table 7. Surface area of CDR variable loops buried by the Ha-K^k ligand

Receptor	TCR		Fab
	X-ray[a]	MD[b]	pSAN
L1 or α1,%[c]	17–26	23–27	8
L2 or α2,%	10–16	10–12	4
L3 or α3,%	16–23	17–20	12
H1 or β1,%	1–18	0–1	16
H2 or β2,%	0–19	5–13	11
H3 or β3,%	8–33	26	38
Framework,%	7–8	7–9	11
V_L or V_α, Å^2	453–557	492–621	283
V_H or V_β, Å^2	288–422	338–344	736
V_α/V_L contribution,%	51–66	59–65	28
V_{LH}/ $V_{\alpha\beta}$, Å^2	845–915	836–859	1029
Total, Å^2	923–979	897–1049	1148

[a]Range of values calculated from three X-ray structures of TCR-MHC-peptide complexes: A6/Tax/HLA-A2 [26], 2C/d-EV8/H-2K[b] [28], B7/Tax/HLA-A2 [29].
[b]Range of values calculated from energy-minimized time averaged MD models of TCR-Ha-K^k complexes [9].
[c]Surface areas have been calculated recorded from the energy-minimized time-averaged conformation using the ACCESS program [23] 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.

duced length of the L2 CDR region (3 amino acids) compared to the α2 loop of a TCR (9 amino acids, see Figure 4). The c"" β strand has different orientations in antibodies and TCRs [31], and induces different shapes of the L2 and α2 variable loops. As the α chain typically dictates binding of the TCR to the pep/MHC ligand (Table 7), it is not surprising to observe a shifted binding mode for the Fab. By computing the relative binding contribution of each CDR region, striking differences are noticed between TCRs and the Fab. The most spectacular one concerns the respective contribution of the two chains to the binding of the MHC-peptide pair (Table 6). Only 28% of the L chain (equivalent to the TCR α chain) is buried by the pep/MHC ligand, whereas 60 to 65% of the TCR α chain is engaged in ligand binding. More than a third of the total Ha-K^k surface is buried by the single H3 hypervariable region. The major contribution of the Fab H chain seems to be a hallmark of Fab-protein recognition [32] that has not been changed in the peculiar case of binding a pep/MHC ligand.

The total buried surface area of the Ha-K^k ligand is about 10% higher when its receptor is a Fab (Table 6) showing that the Fab is more deeply docked into the MHC binding groove than the two TCRs. Interestingly, the equilibrium dissociation constant of the Ha-K^k complex from the pSAN antibody (K_D between 10 and 100 nM, see [7]) is probably much lower than that from TCRs (in the μM range, unpublished data).

Conclusions

A three-dimensional model is proposed for explaining the unique binding mode of a T cell like recombinant antibody specific for a class I MHC-peptide complex. The Fab binds diagonally across the MHC binding groove, as proposed for several αβ TCRs in recently determined X-ray structures of TCR-peptide-MHC complexes. The model accounts well for a large set of experimental binding data on 144 peptide analogues, and explains analogies and differences between the peptide specificity of the antibody with respect to that of two TCRs. The conserved three-dimensional fold between Igs and TCRs allows a similar global binding mode to a pep/MHC ligand. However, slight differences in their molecular architecture, notably at the CDR2α versus the L2 hypervariable regions, allow a tighter docking of the Fab with respect to the TCR.

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. The energy-minimized time-averaged coordinates of the pSAN-Ha-K^k ternary complex are available in PDB format at the following URL: http://www.pharma.ethz.ch/ didier/jcamd

References

1. Chothia, C., Boswell, D.R. and Lesk, A.M., EMBO J., 7 (1988) 3745.
2. Davis, M.M. and Bjorkman, P.J., Nature, 334 (1988) 395.
3. Finger, H. and Seeliger, H.P., In Kwapinsky, B.G. (Ed.), The Origin, Development and Significance of Imunoglobulins, Research in Immunochemistry and Immunology, Vol. 1. University Park Press, Baltimore, MD, 1972, pp. 3–70.
4. Zinkernagel, R.M and Doherty, P.C., Nature, 248 (1974) 701.
5. Witte, T., Smolyar, A., Spoerl, R., Goyarts, E.C., Nathenson, S.G., Reinherz, E.L. and Chang, H.-C., Eur. J. Immunol., 27 (1997) 227.
6. Hogquist, K.A., Grandea III, A.G. and Bevan, M.J., Eur. J. Immunol., 23 (1993) 3028.
7. Andersen, P.S., Stryhn, A., Hansen, B.E., Fugger, L., Engberg, J. and Buus, S., Proc. Natl. Acad. Sci. USA, 93 (1996) 1820.
8. 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.
9. Rognan, D., Stryhn, A., Fugger, L., Lyngbæk, S., Engberg, J., Sejer Andersen, P., and Buus, S., J. Comput.-Aided Mol. Design, 14 (2000) 53 (this issue).
10. Chothia, C., Lesk, A.M., Tramontano, A., Levitt, M., Smith-Gill, S.J., Air, G., Sheriff, S., Padlan, E.A., Davies, D., Tulip, W.R., Colman, P.M., Spinelli, S., Alzari, P.M. and Poljak, R.J., Nature, 342 (1989) 877.
11. Blundell, T.L., Carney, D.P., Gardner, S., Hayes, F.R.F., Howlin, B., Hubbard, T.J.P., Overington, J.P., Singh, D.A., Sibanda, B.L. and Sutcliffe, M., Eur. J. Biochem., 172 (1988) 513.
12. Needleman, S.B. and Wunsch, C.D., J. Mol. Biol., 48 (1970) 443.
13. Dayhoff, M.O., Schwartz, M.M. and Orcutt, B.C., Atlas of Protein Sequence and Structure, 5 (1979) 345.
14. 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.
15. Goodsell, D.S. and Olson, A.J., Proteins Struct. Funct. Genet., 8 (1990) 185.
16. 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.
17. Mehler, E.L. and Solmajer, T., Protein Eng., 4 (1991) 903.

18. Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E., J. Chem. Phys., 21 (1953) 1087.
19. 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.E., J. Am. Chem. Soc., 117 (1995) 5179.
20. Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M., J. Appl. Crystallogr., 26 (1993) 9283.
21. ProsaII v-3.0: http://www.came.sbg.ac.at/
22. Sippl. M., Proteins Struct. Funct. Genet., 17 (1993) 355.
23. Collaborative computational project, Number 4: Acta Crystallogr., D50 (1994) 760.
24. Friedman, A.R., Roberts, V.A. and Tainer, J.A., Proteins Struct. Funct. Genet., 20 (1994) 185.
25. Lunney, E.A., Hagen, S.E., Domagala, J.M., Humblet, C., Kosinski, J., Tait, B.D., Warmus, J.S., Wilson, M., Ferguson, D., Hupe, D., Tummino, P.J., Baldwin, E.T., Bhat, T.N., Liu, B. and Erickson, J.W., J. Med. Chem., 37 (1994) 2664.
26. Garboczi, D.N., Ghosh, P., Utz, U., Fan, Q.R., Biddison, W.E. and Wiley, D.C., Nature, 384 (1996) 134.
27. 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.
28. Garcia, K.C., Degano, M., Pease, L.R., Huang, M., Peterson, P.A., Teyton, L. and Wilson, I.A., Science, 279 (1998) 1666.
29. Ding, Y.H., Smith, K.J., Garboczi, D.N., Utz, U., Biddison, W.E. and Wiley, D.C., Immunity, 8 (1998) 403.
30. Teng, M.K., Smolyar, A., Tse, A.G.D., Liu, J.H., Hussey, R.E., Nathenson, S.G., Chang, H.C., Reinherz, E.L. and Wang, J.H., Curr. Biol., 8 (1998) 409.
31. Fields, B., Ober, B., Malchiodi, E., Lebedeva, M., Braden, B.C., Ysern, X., Kim, J.K., Shao, S., Ward, E.S. and Mariuzza, R.A., Science, 270 (1995) 1821.
32. Wilson, I. and Garcia, K.C., Curr. Opin. Struct. Biol., 7 (1997) 839.
33. Martin, A.C.R., Proteins Struct. Funct. Genet., 25 (1996) 130.
34. Kraulis, P.J., J. Appl. Crystallogr., 24 (1991) 946.
35. Merritt, E.A. and Murphy, M.E.P., Acta Crystallogr., D50 (1994) 869.