Design of enhanced agonists through the use of a new virtual screening method: Application to peptides that bind class I major histocompatibility complex (MHC) molecules

SERGIO MADURGA,¹ IGNASI BELDA,¹ XAVIER LLORÀ,²
AND ERNEST GIRALT^{1,3}

¹Institut de Recerca Biomèdica de Barcelona, Parc Científic de Barcelona, E-08028 Barcelona, Spain

²Illinois Genetic Algorithms Laboratory, Department of General Engineering, University of Illinois, Urbana, Illinois 61801, USA

³Departament de Química Orgànica, Universitat de Barcelona, E-08028 Barcelona, Spain

(RECEIVED January 11, 2005; FINAL REVISION May 12, 2005; ACCEPTED May 17, 2005)

Abstract

A new screening procedure is described that uses docking calculations to design enhanced agonist peptides that bind to major histocompatibility complex (MHC) class I receptors. The screening process proceeds via single mutations of one amino acid at the positions that directly interact with the MHC receptor. The energetic and structural effects of these mutations have been studied using fragments of the original ligand that vary in length. The results of these docking studies indicate that the mutant affinity ranking of long peptides can be practically reproduced with a screening approach performed using fragments of six residues. Fragments of four and five residues could mimic, in some cases, the structural arrangement of the side chains of the full-length peptide. We have compared the structural and energetic results of the docking calculations with experimental data using three unrelated ligand peptides that differ greatly in their affinity for the MHC complex. Analysis of the affinity of the fragments led to the identification of three important parameters in the construction of fragments that mimic the structural and energetic properties of the full-length ligand: the length of the fragment; its intermolecular energy; and the number and localization, internal or terminal, of the anchor residues. The results of this new peptide-design methodology have been applied to suggest new peptides derived from the MUC1–8 peptide that could be used as murine vaccines that trigger the immune response through the MHC class I protein H-2K^b.

Keywords: docking calculations; MHC–peptide interactions; virtual screening; peptide vaccine design; MUC1; tumor immunotherapy; H-2K^b; MHC class I

Major histocompatibility complex (MHC) molecules are highly polymorphic cell-surface proteins that present antigenic peptides for recognition by T-cell receptors

(TCRs), thereby triggering an immune response. In particular, the class I molecules (MHC-I) present peptides derived mainly from endogenously expressed proteins that are degraded by cytosolic proteases. MHC-I molecules are heterodimers of a polymorphic 45-kDa α -chain and the noncovalently associated invariant 12-kDa β_2 -microglobulin. Class I ligands are usually peptides of 8–10 amino acids that bind a groove formed by two α -helices on top of an antiparallel β -sheet (Bjorkman et al. 1987). The peptides are bound in an extended

Reprint requests to: Ernest Giralt, Institut de Recerca Biomèdica de Barcelona, Parc Científic de Barcelona, Josep Samitier 1-5, 08028 Barcelona, Spain; e-mail: egiralt@pcb.ub.es; fax: 34-93-4037126.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.051351605>.

conformation through conserved networks of hydrogen bonds and with the C- and N-terminal charges compensated by complementary MHC residues (Fremont et al. 1992). The hydrogen bonding and electrostatic interactions are relatively independent of the nature of the peptide side chains. In contrast, binding specificity is regulated by the interactions of the peptide side chains with the six pockets of the binding groove of the MHC molecule, designated A–F (Saper et al. 1991). Once the MHC–peptide complex is formed, it is transported to the cell surface and presented for recognition by the CD8⁺ TCRs of the cytotoxic T lymphocytes (CTLs). This specific recognition is mainly regulated by the peptide side chains that are oriented away from the β -sheet.

Peptide antigens specifically recognized by CD8⁺ TCRs in the context of MHC-I molecules have been used in a large number of cancer vaccine trials (Jager et al. 2002; Pouniotis et al. 2004). According to their amino acid sequence, these peptide antigens can be differentiated as self or analog peptides. The sequences of self peptides are frequently obtained directly from fragments of proteins that are linked to a certain pathological cellular behavior. In the case of analog peptides, their sequence is generally derived from these self peptides by introducing single amino acid substitutions. Positive immune responses have been obtained in some cancer patients with vaccinations of self peptides derived from sequences of MAGE-3, NY-ESO-1, Melan-A/MART-1, tyrosinase, and gp100 antigens (Jager et al. 1996a,b, 2000; Marchand et al. 1999). The latter three antigens are melanocyte lineage-specific proteins and are expressed in most melanoma cells. MAGE-3 and NY-ESO-1 are not expressed in nontransformed cells, with the exception of germ cells, and are frequently aberrantly expressed in tumor cells. Vaccinations with self peptides derived from overexpressed antigens, such as Her-2/neu, p53, or MUC1, have also produced positive immune responses (Disis and Cheever 1996; Disis et al. 1999; Vierboom et al. 1997). However, in several cases, vaccinations with analog peptides have yielded improved immunogenicity results. For example, analog peptides of the tumor-associated antigens of NY-ESO-1 (Chen et al. 2000) or gp100 (Rosenberg et al. 1998), modified at only one amino acid, induced enhanced stimulation of CTL. In addition, increased CTL activation caused by modifying HLA anchor residues has been demonstrated for other human melanoma-associated antigens (Kawakami et al. 1995, 2001; Valmori et al. 1998, 1999; Sliz et al. 2001), as well as for the prostate-specific antigen (Terasawa et al. 2002) and the tumor-associated antigen MUC-1 (Tsang et al. 2004). These studies indicate that analog peptides obtained with single residue mutations in the MHC anchor residues of a relatively

weakly immunogenic self-antigen can be used to activate specific T cells more efficiently than the native antigen (Apostolopoulos et al. 2002a). This may be explained by the fact that most of the antigens that are overexpressed in tumor cells are also expressed by normal tissues. Thus, the CTL repertoire of high-affinity peptides is likely to be deleted, while that for low-affinity epitopes may be maintained. In these cases, it might be more appropriate to use analogs of low-affinity binding peptides for immunization. The application of this strategy to vaccine design implies the mutation of buried, non-TCR-contacting amino acid residues of low-affinity binding peptides at the normal anchor positions to enhance their effectiveness in tumor immunotherapy, since the stability of the peptide–MHC complex usually correlates with overall immunogenicity. Methods such as random phage display and high-throughput screening of synthetic combinatorial libraries can be both costly and time consuming, due to the large number and diverse nature of MHC alleles and candidate peptides. In contrast, *in silico* virtual design of agonist molecules based on the three-dimensional structures of target proteins can reduce time and cost.

Peptide candidates for MHC binding have been studied *in silico* using techniques that can be broadly classified into two categories: sequence-based or structure-based approaches. Sequence-based approaches, such as sequence motifs (Falk et al. 1991), matrix models (Gulukota et al. 1997; Godkin et al. 1998; Rammensee et al. 1999), Artificial Neural Network (Milik et al. 1998; Buus et al. 2003), Hidden Markov Model (Mamitsuka 1998; Brusica et al. 2002), and Support Vector Machine (Zhao et al. 2003), only consider the information held within the amino acid sequence to obtain an estimation of the binding affinity of new candidates; they do not take the structural data of the complex into account. Coverage of sequence-based techniques is limited to subsets of binding peptides that belong to the most extensive groups and cannot generate reliable data for peptides that are more poorly represented in the data set. On the other hand, structure-based approaches, such as Monte Carlo (Caflisch et al. 1992), molecular dynamics simulations (Zacharias and Springer 2004), dynamic programming (Sezerman et al. 1993; Gulukota 1996), free-energy mapping (Sezerman et al. 1996), and procedures based on the initial determination of the conformation of the terminal amino acids (Rosenfeld et al. 1995; Tong et al. 2004), can reproduce the conformation of the MHC-binding amino acids quite well, with root mean square deviation (RMSD) values between 1 and 2 Å, although generally with poor estimation of the binding free energy. However, good correlations between calculated and experimental free energies have been obtained using different free-energy scoring functions (Rognan et al. 1999; Liu et al. 2004; Schafroth and Floudas 2004) for

families of peptides that bind the same MHC receptor. With respect to the design methodology for new peptide ligands, it is worth noting the build-up approach applied to three members of the caspase family of cysteine proteases (Budin et al. 2001). This procedure requires the knowledge of a seed from which it iteratively grows a polymeric ligand.

In this article, we present a novel computational screening procedure using the AutoDock3 program to design enhanced agonist peptides that bind to MHC-I receptors (Morris et al. 1998). First, relative binding affinities of the unrelated peptides MUC1-8, VSV8, and dEV8 are compared with experimental values. Modifications in the intramolecular energy term of the free energy scoring function of AutoDock3 have been introduced to obtain improved structural and energetic results. Second, the screening process is applied to the low-affinity peptide MUC1-8 through single mutations of one amino acid at the positions that directly interact with the MHC receptor. Finally, we have determined the minimum number of amino acids of the ligand fragments required to yield results in the screening methodology that are similar to the full-length ligand. The application of this methodology will allow the screening process to be applied to longer peptides with only a small increase in the computational cost.

We have applied this new peptide design methodology to the murine MHC class I protein H-2K^b complexed with a derived MUC1 peptide. MUC1 is an important target for immunotherapeutic strategies because it is overexpressed on the cell surface of many human adenocarcinomas, such as breast and ovarian cancers (Brossart et al. 2001). Novel peptides with a similar amino acid sequence to MUC1-8 but with increased affinity to H-2K^b have been obtained using a screening procedure over the anchor positions.

Results and Discussion

Comparison of docking results with experimental data

The structural information on MUC1-8, VSV8, and dEV8 peptides complexed with murine H-2K^b was obtained through the Protein Data Bank (PDB) codes 1G7Q (Apostolopoulos et al. 2002b), 1KPU (Fremont et al. 1992), and 1LEG (Luz et al. 2002). They are high-resolution crystal structures with resolutions between 1.50 and 1.75 Å (Table 1). At 4°C, 23°C, and 37°C, the available values for the dissociation constants (Apostolopoulos et al. 2002b; Luz et al. 2002) indicate that the affinity order for these three ligands is dEV8 > VSV8 > MUC1-8, with no temperature-dependent variation. The MUC1-8 (SAPDTRPA) ligand has the lowest affinity value because

Table 1. Experimental dissociation constants of peptides binding to H-2K^b

Peptide	Sequence	PDB ^a	Res. ^a	K _D ^b		
				4°C	23°C	37°C
MUC1-8	SAPDTRPA	1G7Q	1.60	433	877	37000
MUC1-8-5F	SAPDFRPA			460	130	4900
MUC1-8-8L	SAPDTRPL			150	250	8900
MUC1-8-5F8L	SAPDFRPL			160	60	300
VSV8	RGYVYQGL	1KPU	1.50	34	27	163
dEV8	EQYKFYSV	1LEG	1.75		11	90

^a PDB codes and resolution (Å) of the available crystal structures for the complexes are indicated.

^b Experimental values of the dissociation constants (nM) were taken from the literature (Apostolopoulos et al. 2002b; Luz et al. 2002).

it does not contain the amino acids that are usually associated with high-affinity binding to H-2K^b (Falk et al. 1991; Rammensee et al. 1995). These residues are either Phe/Tyr at amino acid position P5 or Leu at P8. Comparison of the stability order of the MUC1-8 peptide with its three related sequences, on the other hand, indicates different affinity orders according to the temperature. At 23°C and 37°C, the affinity order is MUC1-8-5F8L > MUC1-8-5F > MUC1-8-8L > MUC1-8. At 4°C, a slightly different distribution is observed: MUC1-8-5F8L ≈ MUC1-8-8L > MUC1-8-5F ≈ MUC1-8. The common pattern is that the peptide with the double substitution, MUC1-8-5F8L, always has a higher affinity than the MUC1-8 peptide.

We have performed docking calculations based on these six H-2K^b-binding octapeptides using AutoDock3. The relative free energy differences obtained from the calculations ($\Delta\Delta G_{\text{DOCK}}$) are shown in Table 2, together with the relative free energies derived from the experimental dissociation constants. As can be seen, the relative stabilization of MUC1-8, VSV8, and dEV8 obtained

Table 2. Relative experimental free energies ($\Delta\Delta G_{\text{EXP}}$) and relative free energies obtained from the docking calculations ($\Delta\Delta G_{\text{DOCK}}$) of peptides binding to H-2K^b

Peptide	$\Delta\Delta G_{\text{EXP}}^a$			$\Delta\Delta G_{\text{DOCK}}^a$	RMSD ^{a,b}
	4°C	23°C	37°C		
MUC1-8	0	0	0	0	0.8
MUC1-8-5F	0.0	-1.1	-1.2	-0.6	0.7
MUC1-8-8L	-0.6	-0.7	-0.9	-1.1	0.7
MUC1-8-5F8L	-0.5	-1.6	-3.0	-2.1	1.0
VSV8	-1.4	-2.1	-3.4	-4.8	0.9
dEV8		-2.6	-3.7	-5.3	1.3

^a Units: free energy, kcal/mol; RMSD, Å.

^b RMSD values are calculated for the ligand heavy atoms of the lowest free energy solution, superimposed onto the experimental structure.

from the docking calculations is the same as that observed experimentally. Although the relative stabilization of VSV8 and dEV8 is slightly overestimated, the magnitude of the overestimation is similar to the experimental differences between temperatures. The experimental trend for the peptide with the double amino-acid substitution to be more stable than the unmutated MUC1-8 peptide is also reproduced by the docking calculations. The free energy results of docking calculations for peptides with single MUC1-8 amino acid modifications are sorted similarly to the experimental tendency observed at 4°C.

From a structural perspective, the docked conformations of MUC1-8, VSV8, and dEV8 peptides are similar to those obtained from the crystal structures. RMSD values for these three peptides ranged between 0.8 and 1.3 Å (Table 2). To compare the structures of the MUC1-8-related peptides, the RMSD values corresponding to peptides with single and double amino acid substitutions of MUC1-8 were calculated using the structure of only the conserved amino acids of the MUC1-8 peptide. The obtained RMSD values ranged between 0.7 and 1.0 Å. These values indicate that the introduction of the anchor amino acids to the MUC1-8 peptide does not cause a conformational change in the remaining amino acids. Analysis of the energetic and structural results indicates that the docking calculations are able to distinguish between peptides of different affinity.

Regulation of the induced fit of the ligands

The coefficients of the scoring function used in AutoDock3 modulate the van der Waals, electrostatic, hydrogen bond, torsional, and desolvation energetic terms. These coefficients were derived using linear regression analysis with a large number of protein-inhibitor complexes for which both the structure and inhibition constants were known. In this adjustment, the same coefficients for intermolecular and intramolecular van der Waals interactions were assumed. It is worth noting that the intramolecular energy of the ligand is only considered in the calculation of the docking energy. The binding free energy is obtained from the intermolecular and the torsional energetic terms. Thus, variations in the intramolecular coefficient will directly affect the ability to explore different ligand conformations, but not the final energy of binding. Consequently, the determination of a different set of values for coefficients of the intermolecular and intramolecular van der Waals energy terms will be troublesome to obtain by regression analysis.

Previous docking calculations performed with the MUC1-8 peptide using the default values of the coefficients of AutoDock3 yielded docked structures with different orientation of the side chain of the Arg residue

compared with the crystallographic structure. In the docked structures, the side chain of this residue adopts a strained turn. In this turn, the C $_{\alpha}$ -C $_{\delta}$ and C $_{\gamma}$ -C $_{\zeta}$ distances are 3.3 Å and 3.4 Å, whereas in the crystallographic structure they are 3.8 Å and 3.7 Å, respectively. The contribution of the internal energy of this Arg residue is 0.8 kcal/mol in the strained turn conformation and -0.1 kcal/mol in the crystallographic structure. In spite of this destabilizing energy contribution (0.9 kcal/mol), the structure of the peptide with the strained turn conformation of the Arg residue is the best solution obtained with the docking calculations because of the greater stabilization of the intermolecular energy contribution (1.3 kcal/mol). To reproduce the crystallographic structure, we used a coefficient of 1.0 for the van der Waals C-C intramolecular interactions. Thus, these interactions were modeled with the unreduced AMBER parameters. With this value, the intramolecular energy difference between these two docked structures is increased to 4.7 kcal/mol, and the strained turn conformation is not obtained for any of the 10 docking solutions. The crystallographic structure is reproduced with an overall RMSD of 0.8 Å.

The side chain of the Arg residue of the MUC1-8 peptide obtained from the crystallographic structure extends away from the binding groove. However, the results of the docking calculations using the default parameters of AutoDock3 localize this residue inside a small cavity between the Gln114 and Glu152 residues, which does not correspond to any of the A to F pockets of the H-2K^b MHC molecule. The corresponding residues (P6) of the VSV8 and dEV8 are Glu and Tyr, respectively. For these two amino acids, and for the full peptides, no significant differences are observed using the default values or the modified C-C intramolecular van der Waals coefficients. This correspondence could be explained by the lower flexibility of the Glu and Tyr side chains. In particular, the Glu side chain is smaller and the Tyr side chain is more rigid than that of Arg. Thus, in the docking calculations for peptides with long, flexible side chain residues that could be introduced in small protein cavities, a different set of intramolecular parameters that reduce the importance of the induced fitting of the ligand could be necessary.

Docking of peptide fragments

For each of the MUC1-8, VSV-8, and dEV8 octapeptides, all possible peptide fragments from one amino acid up to eight were constructed to perform calculations on docking to the H-2K^b MHC molecule. In order to classify structures as “well-docked,” a 2.2 Å RMSD threshold was used. This value has proved to be a good compromise between geometric accuracy and the pro-

vision of a sufficient number of allowed structures to span the whole receptor.

The fragments of the MUC1–8 peptide that dock with a heavy atom RMSD value <2.2 Å from the crystallographic coordinates are indicated in Table 3. Only one of a total of eight single amino acid fragments is correctly docked. The other parts of the binding groove or the surface of the MHC molecule have higher affinity for the other single residues. In contrast, the obtained structure for Ala8 is very similar to its crystal structure. This result indicates that the binding groove of this amino acid is very specific for the hydrophobicity of the side chain and the electrostatic interactions of the terminal carboxyl group. On the other hand, all of the longer fragments are well docked when the first two residues of the MUC1–8 peptide are included. Thus, for short fragments of MUC1–8, the specificity of the ligand is governed mainly by these N-terminal residues.

Well-docked fragments (RMSD <2.2 Å) of VSV8 and dEV8 peptides are shown in Tables 4 and 5, respectively. For each peptide, only one single amino acid fragment is correctly docked. As in the case of the MUC1–8 peptide, these fragments correspond to the C-terminal hydrophobic residues. However, an increase in the number of fragments of different lengths that could be properly docked is now obtained for the high-affinity VSV8 and dEV8 peptides. For MUC1–8, VSV8, and dEV8, the total number of well-docked fragments is 11, 18, and 17, respectively. To obtain a probability of $>50\%$ that fragments of the same length will be well docked, fragments of six amino acids are needed for the low-affinity peptide MUC1–8. In contrast, for high-affinity peptides, the length of these fragments is reduced to four amino acids. Thus, the construction of the docked structure of a full-length peptide can be

Table 3. Peptide fragments of MUC1–8 that dock with a heavy atom RMSD value <2.2 Å from the experimental structure of the ligand

Length ^a	Ratio ^b	Fragments ^c
1	1/8	A ₈
2	1/7	SA _{1–2}
3	1/6	SAP _{1–3}
4	2/5	SAPD _{1–4} , TRPA _{5–8}
5	1/4	SAPDT _{1–5}
6	2/3	SAPDTR _{1–6} , PDTRPA _{3–8}
7	2/2	SAPDTRP _{1–7} , APDTRPA _{2–8}
8	1/1	SAPDTRPA _{1–8}

^a Length is the number of amino acids of the MUC1–8 fragments.

^b Ratio indicates the proportion of fragments of each length that dock with a heavy atom RMSD <2.2 Å.

^c Subindices indicate the amino acid position of the peptide fragments in MUC1–8.

Table 4. Peptide fragments of VSV8 that dock with a heavy atom RMSD value <2.2 Å from the experimental structure of the ligand

Length ^a	Ratio ^b	Fragments ^c
1	1/8	L ₈
2	2/7	RG _{1–2} , GL _{7–8}
3	2/6	YVY _{3–5} , QGL _{6–8}
4	3/5	RGYV _{1–4} , GYVY _{2–5} , YQGL _{5–8}
5	4/4	RGYVY _{1–5} , GYVYQ _{2–6} , YVYQG _{3–7} , VYQGL _{4–8}
6	3/3	RGYVYQ _{1–6} , GYVYQG _{2–7} , YVYQGL _{3–8}
7	2/2	RGYVYQG _{1–7} , GYVYQGL _{2–8}
8	1/1	RGYVYQGL _{1–8}

^a Length is the number of amino acids of the VSV8 fragments.

^b Ratio indicates the proportion of fragments of each length that dock with a heavy atom RMSD <2.2 Å.

^c Subindices indicate the amino acid position of the peptide fragments in VSV8.

obtained with fragments of different sizes according to the affinity of the ligand peptide. Figure 1 shows the reconstruction of the MUC1–8 ligand with the superposition of two six-amino-acid fragments. In the case of the VSV8 ligand, a good reconstruction of the structure is obtained with the superposition of three fragments of only four amino acids (Fig. 2).

An analysis of the intermolecular energy of each amino acid was performed to identify the minimum length of the peptide fragment that could dock in the same way as the full ligand. The contribution of the intermolecular energy of each amino acid was obtained from the docked structures of the MUC1–8, VSV8, and dEV8 peptides (Table 6). As can be seen, the terminal amino acids, which include the carboxyl and amino terminal groups, have the greatest contribution. Part of this contribution, from -1 to -2 kcal/mol, comes from the electrostatic energy. The energetic importance of the terminal amino acids in the stability of the complex has been recognized in several experimental studies (Fahnestock et al. 1992; Bouvier and Wiley 1994; Molano et al. 1998). Molano et al. generated site-directed mutants affecting the peptide binding groove of H-2K^b. They suggested that the relative energetic importance of the anchor residues may be smaller than the interaction of the terminal amino acids. As can be seen from the values shown in Table 6, this behavior is common for high- and low-affinity ligands, although the interaction energy of each amino acid of the low-affinity MUC1–8 peptide is, in general, less favorable than that corresponding to the higher-affinity peptides. Combining the results of Table 6 with those of Tables 3, 4, and 5, it can be inferred that when the average intermolecular energy per residue of a peptide fragment is less than -2.5 kcal/mol, the probability of obtaining a well-docked structure is 46% (3-mer), 80% (4-mer and 5-mer), and

Table 5. Peptide fragments of dEV8 that dock with a heavy atom RMSD value < 2.2 Å from the experimental structure of the ligand

Length ^a	Ratio ^b	Fragments ^c
1	1/8	V ₈
2	1/7	SV ₇₋₈
3	3/6	EQY ₁₋₃ , QYK ₂₋₄ , YKF ₃₋₅
4	3/5	EQYK ₁₋₄ , QYKF ₂₋₅ , FYSV ₅₋₈
5	3/4	EQYKF ₁₋₅ , YKFYS ₃₋₇ , KFYSV ₄₋₈
6	3/3	EQYKFY ₁₋₆ , QYKFYS ₂₋₇ , YKFYSV ₃₋₈
7	2/2	EQYKFYS ₁₋₇ , QYKFYSV ₂₋₈
8	1/1	EQYKFYSV ₁₋₈

^a Length is the number of amino acids of the dEV8 fragments.

^b Ratio indicates the proportion of fragments of each length that dock with a heavy atom RMSD < 2.2 Å.

^c Subindices indicate the amino acid position of the peptide fragments in dEV8.

100% (6-mer, 7-mer, and 8-mer peptides). This tendency is also consistent with a greater promiscuity of short peptides in their binding to different parts of the MHC surface.

Design of an enhanced agonist peptide

Novel peptides with similar amino acid sequence to MUC1–8 but with increased affinity to H-2K^b were obtained with a screening procedure over the amino acids whose side-chain atoms only interact with the MHC receptor. The determination of these amino acids was performed by selecting the residues with C_α–C_β bonds pointing toward the β-sheet surface. These residues belong to the P2, P3, P5, and P8 positions of the MUC1–8 peptide. However, the P3 position was not included in the screening process because it corresponds to a Pro, which has a more rigid backbone conformation. The other positions were screened to obtain MUC1–8 analogs with increased affinity to the MHC molecule. The tested MUC1–8 analogs possess only one changed amino acid in order to enhance the probability that its interaction with the TCR will be maintained.

For the P2, P5, and P8 ligand positions, the screening amino acid X covered all the natural amino acids except Pro. As can be seen in Table 7, all 57 peptides can dock with an RMSD value < 2.2 Å. These RMSD values were calculated for the heavy atoms of the conserved amino acids of MUC1–8. Based on comparison of the relative free energy values of all tested peptides with that of MUC1–8, the best amino acids substitutions of MUC1–8 were obtained with Arg, Trp, and Arg for P2, P5, and P8 ligand positions, respectively. The SRPDTRPA peptide, with a free energy difference of -3.4 kcal/mol compared with MUC1–8, was obtained as the best analog peptide. The SRPDTRPA peptide is 1.0 kcal/mol more stable

than the best peptides obtained in the screening of P5 (SAPDWRPA) and P8 (SAPDTRPR). A comparison of the docked structure of the SRPDTRPA peptide with the crystal MUC1–8 peptide is shown in Figure 3. The positive charge of the Arg residue of P2 is compensated by the Glu24 residue of the MHC receptor. The importance of this amino acid at this position in the binding to H-2K^b has also been reported in the literature on the basis of Ala substitution experiments in the SRDHSRTPM (YEA9) nonapeptide (Apostolopoulos et al. 2002a). On the other hand, the screening procedure performed at the P5 position indicates that the Trp substitution in MUC1–8 will be more stabilizing than the stabilization generated by the normal anchor residues, Tyr and Phe. The C-terminal position of MUC1–8 will be optimized with the Arg residue, which establishes a coulombic interaction with the Asp77 residue of the binding groove.

Ligand fragment approach to the screening of peptides

To apply the screening methodology to longer peptide ligands, a similar screening methodology but using peptides of smaller size would be computationally advantageous. Also, in cases in which structural data for the reference peptide are lacking, the number of amino acids of the screening peptide will need to be reduced in order to use a fully flexible model in the docking calculations.

We have determined the optimum length of peptide fragments that can mimic the structural and energetic properties of full-length peptides by comparing the screening results of the octapeptide with those obtained with the 4-mer, 5-mer, and 6-mer peptides. From a structural point of view, the comparison of the docking results is carried out by analyzing the number of screened peptides that can dock with a conformation near to the

Figure 1. Comparison of the predicted structures of SAPDTRPA₁₋₆ (yellow) and PDTRPA₃₋₈ (blue) MUC1–8 hexapeptide fragments with the MUC1–8 crystal conformation (green) binding to H-2K^b. The 6-mer peptide fragments were taken from Table 3.

Figure 2. Comparison of the predicted structures of RGYV₁₋₄ (yellow), GYVY₂₋₅ (blue), and YQGL₅₋₈ (magenta) VSV8 tetrapeptide fragments with the VSV8 crystal conformation (green) binding to H-2K^b. The 4-mer peptide fragments were taken from Table 4.

crystal structure of the unmutated fragments. For this, we used a criterion of RMSD < 2.2 Å for the heavy atoms of the conserved amino acids of MUC1–8. In addition, the reproducibility of the free energy ranking of the full-length octapeptides was analyzed by calculating the linear correlation coefficient (r) for each set of the screened fragments. Only the fragments that dock with RMSD values < 2.2 Å are considered in the calculation of r . Good structural results with the 4-mer peptides were obtained in the screening performed at position P2, obtaining 16 out of 19 mutations with an RMSD < 2.2 (Table 8). However, the correlation of the free energy differences of these 16 fragments with the corresponding mutations of the reference octapeptide is very weak (0.52). It also can be seen that neither of the two best substitutions of the SXPDT fragments (His and Lys) coincides with the best substitution predicted using the screening of the full octapeptide (Arg). The screening of 4-mer peptides at the P8 position yielded fewer structures docked with an RMSD < 2.2 Å. The correlation coefficient for the free energy values of these structures is also low, indicating that the two best substitutions for this fragment (TRPX) do not coincide with the best substitution of the SAPDTRPX peptide. As can be seen in Table 3, the small fragments of the central part of the MUC1–8 peptide do not dock with an RMSD < 2.2 Å. Consequently, the screening of 4-mer peptides at the P5 position was not undertaken. Comparison of the correlation coefficients calculated for the 4-mer, 5-mer, and 6-mer peptides indicates that six amino acids is the best fragment size to rank the peptides according to their free energy. In addition, the reproducibility of the best substitutions of the octapeptides is only achieved in all three screened positions by using hexapeptide fragments. Screening position P2 is reproduced with the second-best solution, whereas P5 and P8 were reproduced with the

best solution. It should be noted that 4-mer and 5-mer fragments could be a good option with which to obtain the structure when these fragments are part of the ends of the reference peptide, because they correspond to the fragments with the greatest intermolecular energy contribution (Table 6).

In Table 8, the letter code of the residues whose side chains point to the β -sheet of the MHC binding groove is underlined (anchor positions). It can be seen that better results are obtained when these residues constitute the terminal residues of the fragments. For each screened position, most structures docked with RMSD < 2.2 Å are obtained using the SXPDT, PDXRPA, and PDTRPX families of peptides. These fragments have three anchor positions. In almost all cases, the ends of the three fragments correspond to one anchor position, except for the Ser residue of SXPDT, which corresponds to the N terminus of the MUC1–8 ligand.

In summary, the ligand fragment approach to the screening of peptides that bind MHC molecules can be studied using hexapeptides, taking into account the first solutions. An important feature in the construction of the fragments in order to obtain the greatest number of fragments docked similarly to the full-length peptide is that the side chains of the terminal amino acids of the fragments point to the β -sheet of the MHC molecule.

Conclusions

We have described a new screening procedure applied to the design of new peptide ligands with high affinity to MHC class I receptors. The application of this procedure to low-affinity peptides derived from proteins that are overexpressed in certain cancers and presented by a known MHC receptor is of particular interest due to the potential for use of the enhanced agonists in vaccine therapy. Initially, the selection of the residue positions of the ligand sequence that could improve the interaction of the peptide with the MHC receptor with-

Table 6. Intermolecular energy of each amino acid of MUC1–8, VSV8, and dEV8 peptides binding to H-2K^b

Position	MUC1–8 ^a	VSV8 ^a	dEV8 ^a
P1	–4.2	–6.8	–6.2
P2	–2.4	–2.2	–4.6
P3	–2.3	–4.2	–4.3
P4	–1.6	–1.8	–3.2
P5	–1.8	–3.6	–3.9
P6	–2.7	–2.7	–0.6
P7	–2.2	–1.2	–2.0
P8	–4.9	–6.6	–6.3

^a Energies in kcal/mol.

Table 7. Structural and energetic results of docking calculations screening the P2, P5, and P8 positions of the MUC1–8 peptide

Fragment	Screening position	Ratio of structures with RMSD < 2.2 Å ^a	The best substitution	$\Delta G_{\text{Best}} - \Delta G_{\text{MUC1-8}}$ ^b
SXPDTRPA	P2	19/19	R	–3.4
SAPDXRPA	P5	19/19	W	–2.4
SAPDTRPX	P8	19/19	R	–2.4

The screening procedure is performed with all the natural amino acids except proline.

^aNumber of docked structures from the total with an RMSD value < 2.2 Å from the heavy atoms of the conserved amino acids of MUC1–8.

^bThe substitution X that yields the best free energy and its free energy difference with respect to the MUC1–8 peptide are indicated. Free energy is shown in kcal/mol.

out affecting the interaction with the CTL CD8⁺ molecules have to be determined. In this selection, prolines are left unchanged due to their particular effect on the backbone structure. Then, each of the selected positions is screened independently by substitution of all natural amino acids except Pro followed by docking calculations that include flexibility for all side chains. Finally, the substitution that yields the greatest increase in affinity is selected.

As an example, binding of the MUC1–8 ligand to the H-2K^b receptor was selected because of the high-resolution crystal data available for several H-2K^b complexes and the availability of affinity values of several related MUC1–8 ligand sequences. Comparison between calculations and experimental data allowed the identification of an exaggerated induced fit of the ligand that can be solved by increasing the intramolecular van der Waals interactions. With this modification, the relative stabilization of MUC1–8, VSV8, and dEV8 obtained by the docking calculations is the same as that observed experimentally, with RMSD values ranging from 0.8 to 1.3 Å. The energetic and structural results indicate a good capacity of the docking calculations to distinguish between peptides of different affinity.

The crystal structures of the full-length peptides have been reproduced with the superposition of fragments of variable length of MUC1–8, VSV8, and dEV8. The length of these fragments can be smaller for peptides with higher affinity values. We were able to estimate the minimum length of the fragments by calculating the intermolecular energy per residue.

Novel peptides with similar amino acid sequence to MUC1–8 but with increased affinity to H-2K^b have been obtained with a screening procedure over the P2,

P5, and P8 ligand positions. Among the entire ensemble of studied peptides, SRPDTRPA, with a free energy difference of –3.4 kcal/mol compared with MUC1–8, was found to provide the most stable complex structure.

The minimum number of amino acids required to construct fragments of the full-length ligand peptides for which the screening methodology can practically reproduce the mutant affinity ranking is six. Thus, to obtain a good estimation of the free energy differences, the application of the ligand fragment approach to the screening of peptides that bind MHC molecules should be performed with hexapeptide fragments. In addition, the best results are obtained when the side chains of the terminal amino acids point to the β -sheet of the MHC molecule.

The use of short ligands for systematic screening of certain positions will be computationally advantageous in cases in which the full ligand is very long, or when experimental structural data are not available for the reference peptide and/or the MHC receptor. In these cases, structures generated by homology modeling from similar MHC–ligand complexes could, perhaps, be used. However, in these situations, introducing a certain flexibility of the ligand backbone dihedral angles would be advisable.

Figure 3. Comparison of crystal (SAPDTRPA) and mutated structures of the MUC1–8 peptide (SRPDTRPA) binding to H-2K^b. The crystal structure is shown in green and the mutated structure is color coded according to atom type. (A) Top view. (B) Side view.

Table 8. Structural and energetic results of docking calculations of the screening procedure with fragments of the MUC1-8 peptide

Fragment ^a	Screening position	Ratio of structures with RMSD < 2.2 Å ^b	<i>r</i> ^c	The two best substitutions ^d
SXPD	P2	16/19	0.52	H, K
SXPDT	P2	19/19	0.53	H, K
SXPDT ^u R	P2	8/19	0.77	H, R
PDXRP	P5	2/19	— ^e	W, R
PDXRPA	P5	19/19	0.99	W, Y
TRPX	P8	6/19	0.65	M, T
DTRPX	P8	6/19	0.91	R, K
PDTRPX	P8	16/19	0.92	R, K

^a Residues whose side chains point to the β-sheet of the MHC binding groove are underlined.

^b Number of docked structures from the total with an RMSD value < 2.2 Å from the heavy atoms of the conserved amino acids of MUC1-8.

^c The correlation *r* between free energy differences of the screening fragments and the corresponding octapeptides.

^d The substitutions X that yield the two best free energies are indicated. X represents all natural amino acids except proline.

^e Not shown because of the availability of only two points.

Materials and methods

Selection of template structures

High-resolution MHC-I structures complexed with peptide ligands were retrieved from the Protein Data Bank (PDB). The H-2K^b allele was chosen as a model system because of the availability of affinity values for related and unrelated peptides alongside the availability of high-resolution complex structures of the unrelated peptides (resolution values < 1.8 Å). The PDB entries 1G7Q, 1KPU, and 1LEG, which contain the MHC H-2K^b complexed with the peptide ligands MUC1-8, VSV8, and dEV8, respectively, were used as templates for the docking calculations.

Preparation of MHC and ligand structures

The structural models of the three MHC-peptide complexes were initially prepared with Xplor (Brunger 1992), using the heavy and light chain of the MHC and the peptide ligand molecules.

The phosphate ions and the N-acetyl-D-glucosamine, fucose, and 2-methyl-2,4-pentanediol molecules were removed, as they do not directly interact with the binding groove. The crystallographic waters were also removed. For the 1KPU PDB entry, the model A was selected, and in the structure of the MHC molecule of the 1LEG PDB entry, the S-hydroxycysteine residue was replaced with cysteine, as it is far away from the binding pocket. The hbuild function of Xplor was used to add hydrogen atoms to the ligand and MHC molecules. Then, the MHC-ligand structures were minimized with the positions of all heavy atoms restrained to the X-ray coordinates using a harmonic potential with a force constant of 2 kcal/(mol Å²). Finally, nonpolar hydrogens of the final structures were removed to be consistent with the united-atom approximation of AutoDock3. The protonation state of the histidines was considered neutral. The polar hydrogen of the His imidazol ring was located on the δ nitrogen, except when a hydrogen bond involving the ε nitrogen as a donor could be formed. In those cases, the hydrogen was placed on the ε nitrogen.

Docking calculations

Kollman united-atom partial charges were assigned to protein and ligand molecules, and atomic fragmental volumes of the protein atoms were assigned using the Addsol utility of AutoDock3. For the three MHC-ligand systems, a grid map of 70 × 70 × 70 points with a grid-point spacing of 0.375 Å was prepared using AutoDock Tools. Because the location of the ligands in the complex was known, the maps were centered on the ligand binding site. The grid maps were calculated using AutoGrid, version 3.0. Docking calculations were performed by defining flexibility in the side chains of the ligands with the backbone conformation of the peptides taken from the crystallographic structures. The Autotors utility was used to define the rotatable bonds in the ligand. Ten dockings were performed with the Lamarckian Genetic Algorithm using a population size of 50 individuals with a total of 10⁷ energy evaluations. The crystallographic coordinates of the ligands were used as the reference structure to evaluate the predicted docked conformations.

Scoring function

AutoDock3 has a regression-based scoring function that consists of van der Waals, electrostatic, hydrogen bonding, and desolvation energy terms, and an entropic term that measures the loss of torsional degrees of freedom upon binding. This scoring function was parameterized on a set of 30 protein-ligand complexes. In the docking calculations, the solutions are initially ranked according to the docking energy, Δ*E*_{DOCK}, which includes the energy contribution of the intramolecular van der Waals interactions of the ligand. The Δ*E*_{DOCK} is calculated as

$$\begin{aligned} \Delta E_{\text{DOCK}} = & f_{\text{vdW}} \left(\sum_{i>j}^{\text{intra}} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + \sum_{i,k}^{\text{inter}} \left(\frac{A_{ik}}{r_{ik}^{12}} - \frac{B_{ik}}{r_{ik}^6} \right) \right) \\ & + f_{\text{hbond}} \sum_{i,k}^{\text{inter}} E(t) \left(\frac{C_{ik}}{r_{ik}^{12}} - \frac{D_{ik}}{r_{ik}^{10}} + E_{\text{hbond}} \right) \\ & + f_{\text{elec}} \sum_{i,k}^{\text{inter}} \frac{q_i q_k}{\epsilon(r_{ik}) r_{ik}} \\ & + f_{\text{sol}} \sum_{i_c,k}^{\text{inter}} S_{i_c} V_k \exp(-r_{i_c,k}^2 / 2\sigma^2) \end{aligned} \quad (1)$$

After the final docked structures of the ligands are obtained, the relative free energy of binding, Δ*G*_{Binding}, is calculated as

$$\begin{aligned} \Delta G_{\text{Binding}} = & f_{\text{vdW}} \sum_{i,k}^{\text{inter}} \left(\frac{A_{ik}}{r_{ik}^{12}} - \frac{B_{ik}}{r_{ik}^6} \right) \\ & + f_{\text{hbond}} \sum_{i,k}^{\text{inter}} E(t) \left(\frac{C_{ik}}{r_{ik}^{12}} - \frac{D_{ik}}{r_{ik}^{10}} + E_{\text{hbond}} \right) \\ & + f_{\text{elec}} \sum_{i,k}^{\text{inter}} \frac{q_i q_k}{\epsilon(r_{ik}) r_{ik}} \\ & + f_{\text{sol}} \sum_{i_c,k}^{\text{inter}} S_{i_c} V_k \exp(-r_{i_c,k}^2 / 2\sigma^2) \\ & + f_{\text{tor}} N_{\text{tor}} \end{aligned} \quad (2)$$

where f_{vdW} , f_{hbond} , f_{elec} , f_{tor} , and f_{sol} are coefficients empirically determined using linear regression analysis of Equation 2 with values of 0.1485, 0.0656, 0.1146, 0.3113, and 0.1711, respectively. The subindex i, j runs over the interaction sites of the ligand, and k over the interaction sites of the protein. The expression multiplied by the f_{vdW} coefficient contains the Lennard-Jones dispersion/repulsion energy terms. In the docking energy, Equation 1, the intramolecular interactions of the ligand are also considered. This summation is extended to all pairs of atoms in the ligand that are separated by three or more bonds. The summation multiplied by the f_{hbond} coefficient takes into account the directional 12–10 hydrogen bonding term. $E(t)$ is a directional weight based on the angle, t , between the probe and the target atom, and E_{hbond} is the estimated average energy of hydrogen bonding of water with a polar atom. This constant term, E_{hbond} , models desolvation of polar atoms. The expression that multiplies the f_{elec} coefficient considers the screened Coulombic electrostatic potential using a sigmoidal distance-dependent dielectric function, $\epsilon(r_{ik})$, to model solvent screening (Mehler and Solmajer 1991). For all carbon atoms in the ligand, i_c , their corresponding solvation parameters, S_{i_c} , are multiplied by the fragmental volumes, V_k , of the surrounding protein atoms and weighted by an exponential function. In Equation 2, N_{tor} is the number of rotatable bonds in the ligand that rotate heavy atoms. The parameters of the dispersion/repulsion energies have been taken from the AMBER force field.

The intramolecular interaction energy of the ligand is not included in the calculation of the binding free energy, since it is assumed that the free energy variation due to the conformational change of the ligand in passing from solution to the complexed form is small. However, intramolecular interaction energy must be taken into consideration in the docking energy to avoid unrealistic superpositions of ligand atoms. The magnitude of the intramolecular interactions regulates the degree of the induced fit made by the protein. The default values of the coefficients of the AutoDock3 program are identical for intermolecular and intramolecular van der Waals interactions (f_{vdW}). In this study, we modified the default value of f_{vdW} that is applied to the interactions between carbon atoms of the ligand. The f_{vdW} will be the default value, except when i and j indexes of the summation of the intramolecular interactions represent carbon atoms, assigning, in this case, $f_{\text{vdW}} = 1$. Therefore, the interaction between carbon atoms will be now computed using the standard AMBER values.

Fragment approach

The structure of MUC1–8, VSV8, and dEV8 ligands is divided into all possible fragments, from one amino acid up to eight, yielding 36 fragments per ligand. The amino and carboxyl groups are only included in the fragments that contain one of the terminal amino acids. The distribution of charges for each fragment is taken from the corresponding full-length MUC1–8, VSV8, and dEV8 ligands.

Construction of new peptides for the screening procedure

Initially, the full MUC1–8 ligand was taken as a template. The side-chain atoms of one of the amino acids that directly interact with the β -sheet of the MHC (here P2, P5, or P8) were removed. The rest of the atoms were held fixed and new atoms corresponding to new amino acids were introduced

and optimized with Xplor. Pro residues are not considered in either the selection or in the screening process. For structural analyses of docked conformations, the reference structure was taken from the ligand crystallographic coordinates without consideration of the screening amino acid.

Acknowledgments

We acknowledge financial support from MCYT-FEDER (Bio2002–2301) and Fundación BBVA.

References

- Apostolopoulos, V., Yu, M., Corper, A.L., Li, W., McKenzie, I.F., Teyton, L., Wilson, I.A., and Plebanski, M. 2002a. Crystal structure of a non-canonical high affinity peptide complexed with MHC class I: A novel use of alternative anchors. *J. Mol. Biol.* **318**: 1307–1316. [Erratum: 2002. *J. Mol. Biol.* **322**: 478.]
- Apostolopoulos, V., Yu, M., Corper, A.L., Teyton, L., Pietersz, G.A., McKenzie, I.F., Wilson, I.A., and Plebanski, M. 2002b. Crystal structure of a non-canonical low-affinity peptide complexed with MHC class I: A new approach for vaccine design. *J. Mol. Biol.* **318**: 1293–1305. [Erratum: 2002. *J. Mol. Biol.* **322**: 477.]
- Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L., and Wiley, D.C. 1987. Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* **329**: 506–512.
- Bouvier, M. and Wiley, D.C. 1994. Importance of peptide amino and carboxyl termini to the stability of MHC class I molecules. *Science* **265**: 398–402.
- Brossart, P., Schneider, A., Dill, P., Schammann, T., Grunebach, F., Wirths, S., Kanz, L., Buhring, H.J., and Brugger, W. 2001. The epithelial tumor antigen MUC1 is expressed in hematological malignancies and is recognized by MUC1-specific cytotoxic T-lymphocytes. *Cancer Res.* **61**: 6846–6850.
- Brünger, A.T. 1992. *XPLOR Version 3.1: A system for X-ray crystallography and NMR*. Yale University Press, New Haven, CT.
- Brusic, V., Petrovsky, N., Zhang, G., and Bajic, V.B. 2002. Prediction of promiscuous peptides that bind HLA class I molecules. *Immunol. Cell Biol.* **80**: 280–285.
- Budin, N., Majeux, N., Tenette-Souaille, C., and Caflisch, A. 2001. Structure-based ligand design by a build-up approach and genetic algorithm search in conformational space. *J. Comput. Chem.* **22**: 1956–1970.
- Buus, S., Lauemoller, S.L., Worning, P., Kesmir, C., Frimurer, T., Corbet, S., Fomsgaard, A., Hilden, J., Holm, A., and Brunak, S. 2003. Sensitive quantitative predictions of peptide-MHC binding by a 'Query by Committee' artificial neural network approach. *Tissue Antigens* **62**: 378–384.
- Caflisch, A., Niederer, P., and Anliker, M. 1992. Monte Carlo docking of oligopeptides to proteins. *Proteins* **13**: 223–230.
- Chen, J.L., Dunbar, P.R., Gileadi, U., Jager, E., Gnjatich, S., Nagata, Y., Stockert, E., Panicali, D.L., Chen, Y.T., Knuth, A., et al. 2000. Identification of NY-ESO-1 peptide analogues capable of improved stimulation of tumor-reactive CTL. *J. Immunol.* **165**: 948–955.
- Disis, M.L. and Cheever, M.A. 1996. Oncogenic proteins as tumor antigens. *Curr. Opin. Immunol.* **8**: 637–642.
- Disis, M.L., Grabstein, K.H., Sleath, P.R., and Cheever, M.A. 1999. Generation of immunity to the HER-2/neu oncogenic protein in patients with breast and ovarian cancer using a peptide-based vaccine. *Clin. Cancer Res.* **5**: 1289–1297.
- Fahnestock, M.L., Tamir, I., Narhi, L., and Bjorkman, P.J. 1992. Thermal stability comparison of purified empty and peptide-filled forms of a class I MHC molecule. *Science* **258**: 1658–1662.
- Falk, K., Rotzschke, O., Stevanovic, S., Jung, G., and Rammensee, H.G. 1991. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature* **351**: 290–296.
- Fremont, D.H., Matsumura, M., Stura, E.A., Peterson, P.A., and Wilson, I.A. 1992. Crystal structures of two viral peptides in complex with murine MHC class I H-2Kb. *Science* **257**: 919–927.
- Godkin, A.J., Davenport, M.P., Willis, A., Jewell, D.P., and Hill, A.V. 1998. Use of complete eluted peptide sequence data from HLA-DR and -DQ molecules to predict T cell epitopes, and the influence of the nonbinding terminal regions of ligands in epitope selection. *J. Immunol.* **161**: 850–858.

- Gulukota, K., Vajda, S., and Delisi, C. 1996. Peptide docking using dynamic programming. *J. Comput. Chem.* **17**: 418–428.
- Gulukota, K., Sidney, J., Sette, A., and DeLisi, C. 1997. Two complementary methods for predicting peptides binding major histocompatibility complex molecules. *J. Mol. Biol.* **267**: 1258–1267.
- Jager, E., Bernhard, H., Romero, P., Ringhoffer, M., Arand, M., Karbach, J., Ilsemann, C., Hagedorn, M., and Knuth, A. 1996a. Generation of cytotoxic T-cell responses with synthetic melanoma-associated peptides in vivo: Implications for tumor vaccines with melanoma-associated antigens. *Int. J. Cancer* **66**: 162–169.
- Jager, E., Ringhoffer, M., Dienes, H.P., Arand, M., Karbach, J., Jager, D., Ilsemann, C., Hagedorn, M., Oesch, F., and Knuth, A. 1996b. Granulocyte-macrophage-colony-stimulating factor enhances immune responses to melanoma-associated peptides in vivo. *Int. J. Cancer* **67**: 54–62.
- Jager, E., Gnjatic, S., Nagata, Y., Stockert, E., Jager, D., Karbach, J., Neumann, A., Rieckenberg, J., Chen, Y.T., Ritter, G., et al. 2000. Induction of primary NY-ESO-1 immunity: CD8+ T lymphocyte and antibody responses in peptide-vaccinated patients with NY-ESO-1+ cancers. *Proc. Natl. Acad. Sci.* **97**: 12198–12203.
- Jager, E., Jager, D., and Knuth, A. 2002. Clinical cancer vaccine trials. *Curr. Opin. Immunol.* **14**: 178–182.
- Kawakami, Y., Eliyahu, S., Jennings, C., Sakaguchi, K., Kang, X., Southwood, S., Robbins, P.F., Sette, A., Appella, E., and Rosenberg, S.A. 1995. Recognition of multiple epitopes in the human melanoma antigen gp100 by tumor-infiltrating T lymphocytes associated with in vivo tumor regression. *J. Immunol.* **154**: 3961–3968.
- Kawakami, Y., Wang, X., Shofuda, T., Sumimoto, H., Tupesis, J., Fitzgerald, E., and Rosenberg, S. 2001. Isolation of a new melanoma antigen, MART-2, containing a mutated epitope recognized by autologous tumor-infiltrating T lymphocytes. *J. Immunol.* **166**: 2871–2877.
- Liu, Z., Dominy, B.N., and Shakhnovich, E.I. 2004. Structural mining: Self-consistent design on flexible protein-peptide docking and transferable binding affinity potential. *J. Am. Chem. Soc.* **126**: 8515–8528.
- Luz, J.G., Huang, M., Garcia, K.C., Rudolph, M.G., Apostolopoulos, V., Teyton, L., and Wilson, I.A. 2002. Structural comparison of allogeneic and syngeneic T cell receptor-peptide-major histocompatibility complex complexes: A buried alloreactive mutation subtly alters peptide presentation substantially increasing V(β) interactions. *J. Exp. Med.* **195**: 1175–1186.
- Mamitsuka, H. 1998. Predicting peptides that bind to MHC molecules using supervised learning of hidden Markov models. *Proteins* **33**: 460–474.
- Marchand, M., van Baren, N., Weynants, P., Brichard, V., Dreno, B., Tessier, M.H., Rankin, E., Parmiani, G., Arienti, F., Humblet, Y., et al. 1999. Tumor regressions observed in patients with metastatic melanoma treated with an antigenic peptide encoded by gene MAGE-3 and presented by HLA-A1. *Int. J. Cancer* **80**: 219–230.
- Mehler, E.L. and Solmajer, T. 1991. Electrostatic effects in proteins: Comparison of dielectric and charge models. *Protein Eng.* **4**: 903–910.
- Milik, M., Sauer, D., Brunmark, A.P., Yuan, L., Vitiello, A., Jackson, M.R., Peterson, P.A., Skolnick, J., and Glass, C.A. 1998. Application of an artificial neural network to predict specific class I MHC binding peptide sequences. *Nat. Biotechnol.* **16**: 753–756.
- Molano, A., Erdjument-Bromage, H., Fremont, D.H., Messaoudi, I., Tempst, P., and Nikolic-Zugic, J. 1998. Peptide selection by an MHC H-2Kb class I molecule devoid of the central anchor (“C”) pocket. *J. Immunol.* **160**: 2815–2823.
- Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., and Olson, A.J. 1998. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **19**: 1639–1662.
- Pouniotis, D.S., Halton, J., and Apostolopoulos, V. 2004. Peptide based vaccine design for cancer immunotherapy. *Lett. Drug Des. Discovery* **1**: 203–210.
- Rammensee, H.G., Friede, T., and Stevanović, S. 1995. MHC ligands and peptide motifs: First listing. *Immunogenetics* **41**: 178–228.
- Rammensee, H., Bachmann, J., Emmerich, N.P., Bachor, O.A., and Stevanović, S. 1999. SYFPEITHI: Database for MHC ligands and peptide motifs. *Immunogenetics* **50**: 213–219.
- Rognan, D., Lauemoller, S.L., Holm, A., Buus, S., and Tschinke, V. 1999. Predicting binding affinities of protein ligands from three-dimensional models: Application to peptide binding to class I major histocompatibility proteins. *J. Med. Chem.* **42**: 4650–4658.
- Rosenberg, S.A., Yang, J.C., Schwartzentruber, D.J., Hwu, P., Marincola, F.M., Topalian, S.L., Restifo, N.P., Dudley, M.E., Schwarz, S.L., Spiess, P.J., et al. 1998. Immunologic and therapeutic evaluation of a synthetic peptide vaccine for the treatment of patients with metastatic melanoma. *Nat. Med.* **4**: 321–327.
- Rosenfeld, R., Zheng, Q., Vajda, S., and DeLisi, C. 1995. Flexible docking of peptides to class I major-histocompatibility-complex receptors. *Genet. Anal.* **12**: 1–21.
- Saper, M.A., Bjorkman, P.J., and Wiley, D.C. 1991. Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 Å resolution. *J. Mol. Biol.* **219**: 277–319.
- Schafroth, H.D. and Floudas, C.A. 2004. Predicting peptide binding to MHC pockets via molecular modeling, implicit solvation, and global optimization. *Proteins* **54**: 534–556.
- Sezerman, U., Vajda, S., Cornette, J., and DeLisi, C. 1993. Toward computational determination of peptide-receptor structure. *Protein Sci.* **2**: 1827–1843.
- Sezerman, U., Vajda, S., and DeLisi, C. 1996. Free energy mapping of class I MHC molecules and structural determination of bound peptides. *Protein Sci.* **5**: 1272–1281.
- Sliz, P., Michielin, O., Cerottini, J.C., Luescher, I., Romero, P., Karplus, M., and Wiley, D.C. 2001. Crystal structures of two closely related but antigenically distinct HLA-A2/melanocyte-melanoma tumor-antigen peptide complexes. *J. Immunol.* **167**: 3276–3284.
- Terasawa, H., Tsang, K.Y., Gulley, J., Arlen, P., and Schlom, J. 2002. Identification and characterization of a human agonist cytotoxic T-lymphocyte epitope of human prostate-specific antigen. *Clin. Cancer Res.* **8**: 41–53.
- Tong, J.C., Tan, T.W., and Ranganathan, S. 2004. Modeling the structure of bound peptide ligands to major histocompatibility complex. *Protein Sci.* **13**: 2523–2532.
- Tsang, K.Y., Palena, C., Gulley, J., Arlen, P., and Schlom, J. 2004. A human cytotoxic T-lymphocyte epitope and its agonist epitope from the nonvariable number of tandem repeat sequence of MUC-1. *Clin. Cancer Res.* **10**: 2139–2149.
- Valmori, D., Fonteneau, J.F., Lizana, C.M., Gervois, N., Lienard, D., Rimoldi, D., Jongeneel, V., Jotereau, F., Cerottini, J.C., and Romero, P. 1998. Enhanced generation of specific tumor-reactive CTL in vitro by selected Melan-A/MART-1 immunodominant peptide analogues. *J. Immunol.* **160**: 1750–1758.
- Valmori, D., Fonteneau, J.F., Valitutti, S., Gervois, N., Dunbar, R., Lienard, D., Rimoldi, D., Cerundolo, V., Jotereau, F., Cerottini, J.C., et al. 1999. Optimal activation of tumor-reactive T cells by selected antigenic peptide analogues. *Int. Immunol.* **11**: 1971–1980.
- Vierboom, M.P., Nijman, H.W., Offringa, R., van der Voort, E.I., van Hall, T., van den Broek, L., Fleuren, G.J., Kenemans, P., Kast, W.M., and Melief, C.J. 1997. Tumor eradication by wild-type p53-specific cytotoxic T lymphocytes. *J. Exp. Med.* **186**: 695–704.
- Zacharias, M. and Springer, S. 2004. Conformational flexibility of the MHC class I α 1- α 2 domain in peptide bound and free states: A molecular dynamics simulation study. *Biophys. J.* **87**: 2203–2214.
- Zhao, Y., Pinilla, C., Valmori, D., Martin, R., and Simon, R. 2003. Application of support vector machines for T-cell epitopes prediction. *Bioinformatics* **19**: 1978–1984.