PROTEINS: Structure, Function, and Genetics 13:223–230 (1992)

Monte Carlo Docking of Oligopeptides to Proteins

Amedeo Caflisch, Peter Niederer, and Max Anliker
Institute of Biomedical Engineering and Medical Informatics, Swiss Federal Institute of Technology, and University of Zürich, CH-8044 Zürich, Switzerland

ABSTRACT **A new two-step procedure has been developed for the docking of flexible oligopeptide chains of unknown conformation to static proteins of known structure. In the first step positions and conformations are sampled and the association energy minimized starting from an approximate preselected docking position. The resulting conformations are further optimized in the second step by a Metropolis Monte Carlo minimization, which optimizes each of these structures. The method has been tested on the HIV-1 aspartic proteinase complex with an inhibitor, whose crystallographic structure is known at 2.3 Å resolution. Furthermore, the application of this method to the docking of the hendecapeptide 58–68 of the influenza A virus matrix protein to the HLA-A2 molecule produced results which are in agreement with experimental observations in identifying side chains critical for T cell recognition and residues responsible of MHC protein binding.** © 1992 Wiley-Liss, Inc.

Key words: force field, global optimization, Metropolis criterion, HIV-1 aspartic proteinase, HLA-A2 MHC protein

INTRODUCTION

Molecular modeling has become a powerful tool in analyzing the structure and function of molecules which play an essential role in living organisms. In fact, the structures of such molecules have in several cases been predicted theoretically and subsequently confirmed experimentally.[1] Recently, by using a structure-based computer-assisted search, a butyrophenone derivative was found which is a selective inhibitor (at high toxic concentrations) of the HIV-1 aspartic proteinase.[2] This documents the potency of molecular data base management, molecular graphics, and computational chemistry.

As for the protein folding simulation, the docking of a small polypeptide chain to a protein on the basis of an energy function presents two major problems: the accuracy of the molecular model or force field and the multiple minima problem.

Wodak and Janin developed an automatic procedure to dock two rigid polypeptide chains described by one interaction center for each amino acid.[3] A system developed by Kuntz et al. automatically fits a relatively flexible ligand, approximated as a small set of large rigid fragments, on a receptor of known geometry represented as a set of overlapping spheres on the basis of molecular shape.[4, 5] Goodsell and Olson[6] recently described a procedure for the automatic docking of small flexible substrates to static proteins by a method which combines Metropolis conformation searching[7] with energy evaluation on the basis of molecular affinity potentials. These docking procedures are based on energy function which does not possess a sufficiently accurate model of hydrogen bonding; their applications can lead to incorrect results if a ligand is held in place tightly by hydrogen bonds and only loosely by steric forces. Furthermore, none of these methods is guaranteed to yield the docked conformation associated with the global minimum of the energy function.

We have developed a two-step procedure for the docking of flexible oligopeptides of unknown conformation into a static protein of known structure. Starting from an approximate preselected docking position of the oligopeptide, the first step (I) generates a set of docked conformations on the basis of an energy function which contains a special intermolecular nonbonded energy term for the elimination of the bad contacts. The force field utilized in the second step (II) considers as interaction centers all the heavy atoms and polar hydrogen atoms of both the ligand molecular and the residues of the binding site of the receptor. Besides the more accurate force field, the essential aspect of the present work consists of the utilization of a Monte Carlo minimization procedure[8] which is based on classical statistical mechanics and overcomes potential energy barriers by random changes and reminimizations. In the original Monte Carlo method a random change of the configuration is accepted if the new structure satisfies the Metropolis criterion, i.e., when $\Delta E \leq 0$, or if $\Delta E > 0$ when $e^{-\Delta E/K_B T} > R$, where R is a random number taken from a uniform distribution over the interval (0, 1), E is the energy of the

Received December 28, 1990; revision accepted June 27, 1991.

Address reprint requests to Dr. Amedeo Caflisch at his new address: Department of Chemistry, Harvard University, 12 Oxford Street, Cambridge, MA 02138.

system, T its temperature, and K_B represents the Boltzmann constant. Even a small internal rotation in a protein can generate bad nonbonded atom contacts associated with very high values of the energy. Accordingly, to avoid sampling high energy conformations a conventional iterative minimization is performed after each random change; the refined structure is then submitted to the Metropolis criterion. Thereby, the Monte Carlo minimization combines the power of the Metropolis Monte Carlo method in global optimization and that of the conjugate gradient algorithm to find local minima. This method is suitable for the study of long-time and global properties of polypeptide chains undergoing large-scale structural changes. Therefore the Monte Carlo minimization procedure can be utilized to simulate the folding of oligopeptides[8] and protein docking.[9]

MODEL AND METHODS

Force Field

This study used a common approximation[9,10] of the potential energy function, as a sum of terms:

$$E = E_b + E_a + E_t + E_{LJ} + E_C + E_H. \quad (1)$$

The first three terms on the right approximate individually the interactions between covalently bonded atoms, separated by 1, 2, and 3 covalent bonds, while the remaining three represent the interactions between nonbonded atoms, i.e., atoms separated by 3 or more covalent bonds: Lennard–Jones energy, electrostatic energy, and hydrogen bond energy. Polar hydrogen atoms are explicitly considered, while the hydrogens bound to neutral and apolar C, N, and S atoms are considered indirectly by the extended atom approximation. The parameters characterizing the bonded interactions, the Lennard–Jones interactions, and the hydrogen bonding parameters were those used in the molecular mechanics program CHARMM.[11] The distributions of the atomic partial charges are based on the model of Momany et al.[12]

Energy Function for the First Step

Polypeptide chains, which have been coarsely docked by visual assessment with the aid of interactive three-dimensional computer graphics[9] or by algorithms which sample positions and orientations of the ligand in the receptor, present bad intermolecular contacts (BC). Minimizing the potential energy of such structures poses major convergence problems because of the 6–12 Lennard–Jones potential which grows enormously for too closely positioned nonbonded atoms. A special nonbonded interactions term (E_{BC}) was developed to remove bad contacts between the atoms of the ligand polypeptide chain, considered as flexible, and the atoms on the macromolecular receptor chain which are kept fixed in three-dimensional space:

$$E_{BC} = \sum_{i,j}^{N_l \cdot N_r} \begin{cases} k_{BC} & (r_{i,j} - r_0)^4 \\ \dfrac{1}{4\pi\epsilon_0} & \dfrac{q_i q_j}{\epsilon r_{i,j}} \end{cases} \quad \begin{matrix} r_{i,j} < r_0 \\ r_{i,j} \geq r_0 \end{matrix} \quad (2)$$

In this formula i refers to the ligand, j to the receptor, r_0 is usually set equal to 3 Å, k_{BC} to 10 kcal/mol, and the sum runs over all the $N_l \cdot N_r$ pairs where $N_l(N_r)$ represents the number of heavy atoms of the ligand (receptor). The atomic partial charges are the same as for E_C (see above). To describe the nonbonded interactions between ligand and receptor in the first step of our docking procedures, the energy function for the polypeptide ligand [Eq. (1)] has been modified as follows:

$$E_I = E_b + E_a + E_t + E_{LJ} + E_C + E_{BC}. \quad (3)$$

The first five addends approximate the intramolecular potential energy of the oligopeptide ligand. No term for the intramolecular potential energy of the receptor is required since its atoms are kept fixed in space during docking as mentioned above. Although the energy term E_{BC} is not continuous (discontinuity at $r_{i,j} = r_0$ if both atoms i and j possess a partial charge different from zero) the bad contacts were removed from all coarsely docked conformations described in this work by a few iterations of the conjugate gradient algorithm. The polar hydrogen atoms and a hydrogen bonding term are not considered explicitly in this step whose purpose is to perform fast positional and conformational sampling.

Energy Function for the Second Step

In the present study the ligand will be considered flexible while the macromolecular receptor is treated as fixed in space. Accordingly, to describe the nonbonded interactions between ligand and receptor in the second step of our docking procedure, the energy function for one polypeptide chain [Eq. (1)] has been completed as follows:

$$E_{II} = E + E_{12} \quad (4)$$

where

$$E_{12} = E_{LJ,12} + E_{C,12} + E_{H,12} \quad (5)$$

represents the nonbonded interactions between ligand (1) and receptor (2), while E is the potential energy of the ligand.

Docking Procedure and Simulation Parameters

At the beginning an extended or helical conformation of the trial oligopeptide is generated from scratch with an ad hoc developed computer program.[9] This starting structure of the isolated ligand may be refined to reach the next local minimum; the resulting conformation is placed in the putative binding site with the help of an in-house developed

molecular graphics program which also allows one to determine which residues to take into account during computational docking.[9]

Step I of the docking procedure performs fast positional and conformational sampling of the oligopeptide ligand inside the binding groove of the static receptor protein. For this purpose, the coarsely docked structure is subjected to random rigid body translations and rotations (80 cycles of positional sampling). The program allows one to perform the translations along a direction nearly parallel to the long axis of the groove; since both proteins studied in this work possess an elongated binding cleft the translation vector **r** was randomly picked with $|\mathbf{r}| < 0.20$ Å and $|\cos\theta| > 0.8$, where θ represents the angle between **r** and the long axis of the groove. The rotations were performed about the axis of the extended (or helical) structure by choosing the angle in a random fashion. The energy function E_I [Eq. (3)] of each resulting complex is then minimized (40 conjugate gradients iterations) without explicitly considering the polar hydrogen atoms. Step I requires about 10 minutes of CPU time on an IBM 3090 computer for both complexes simulated in this study.

In step II the energy function E_{II} [see Eq. (4)] is optimized by 400 cycles of Monte Carlo minimization[8] for all the conformations obtained by step I whose energy E_I are < 1500 kcal/mol. This number of cycles was chosen to balance a sufficient amount of conformational sampling with a reasonable computational time; it allows us to perform ~10 random torsions about each variable dihedral angle for oligopeptide lengths of ~10 residues. The computational time is proportional to the number of intermolecular nonbonded atom pairs which is about the same for both complexes simulated in this study, i.e., 32,256 and 35,570 for the HIV-1 proteinase-inhibitor and HLA-A2-antigen complex, respectively. The E_I cutoff value was set to 1500 kcal/mol since we found that submitting conformations associated with a higher value of E_I to step II did not produce low minima of E_{II} in a reasonable computation time. Each cycle begins with a random change of a randomly picked variable dihedral angle of the oligopeptide ligand followed by 100 iterations of the conjugate gradient algorithm applied to the energy function E_{II} whose gradient is computed analytically. To accept the new conformation the Metropolis criterion is applied at a temperature of 310 K. As suggested by Li and Scheraga[8] only one variable dihedral angle pro cycle is varied in order to avoid sampling the energetically unfavorable regions more frequently. Step II was developed to perform an exhaustive conformational sampling on the basis of an accurate force field; it needs about 2.5 hr of CPU time on an IBM 3090 for each complex, i.e., between 25 and 100 hr for each simulation because step I generates between 10 and 40 complexes with $E_I < 1500$ kcal/mol. Furthermore, step II is independent of step I and can be used for the optimization of a known complex or a mutated oligopeptide inside a known complex. Also, step II is more powerful than a conventional minimization algorithm since it allows us to surmount intervening potential barriers in moving through several discrete local minima.

RESULTS AND DISCUSSION

Docking of a Heptapeptide Inhibitor to the HIV-1 Aspartic Proteinase

We decided to test the two-step docking procedure on the HIV-1 aspartic proteinase complex with the heptapeptide Gly-Thr-Ile-Met-Ψ[CH_2-NH]-Met-Gln-Arg since the crystal structure of the complex between this proteinase and the inhibitor *N*-acetyl-Thr-Ile-Nle-Ψ[CH_2-NH]-Nle-Gln-Arg.amide exists at 2.3 Å resolution[13] (Brookhaven Protein Data Bank[14] listing 4HVP). We were obliged to utilize the original methionine side chains and to substitute the *N*-acetyl terminus with a glycine residue because our force field has been parameterized for the 20 standard amino acids only. (The *N*-acetyl terminus of the inhibitor is partially exposed to the surrounding solvent in the complex. Accordingly, substitution of *N*-acetyl with glycine should not prevent inhibitor binding.) The crystallographic binding cleft contains a water molecule[13, 15] which was considered in the simulation as static.

A completely extended conformation of the heptapeptide was coarsely docked by an in-house developed molecular graphics program[9] at three different positions with the Met-4 and Met-5 side chains approximately oriented in the direction opposite to that of the crystallographic result. Starting C_α rms deviations from the crystallographic conformation measured 2.4, 2.9, and 3.8 Å respectively. Each of the first two starting complexes was submitted to the two-step procedure twice (with different random-number generators), while the third complex was submitted three times. The acceptance ratio (accepted/discarded structures), according to the Metropolis criterion at a T value of 310 K, was about 27%. The results are listed in Table I.

The heptapeptide structure associated with the lowest E_{II} energy (D1) has the lowest value of the C_α rms deviation, C_α DME (distance matrix error), and C_β DME from the native structure; furthermore, D1 possesses the same intermolecular hydrogen bonds and nearly the same position and side chains orientation (with the exception of Gln-6) as the native inhibitor inside the HIV-1 proteinase binding site. The three conformations, D1, D2, and D3, which are closest to the native structure are similar (small values of the C_α rms deviation and C_α DME, cf. Table II). This implies that they have partly converged toward a common conformation similar (C_α rms deviation within 2.0 Å) to the native one (Fig. 1). The

TABLE I. Results of the Two-Step Procedure Applied to the HIV-1 Proteinase–Inhibitor Complex

Structure	Starting C_α rms (Å)	E_{II} (kcal/mol)	C_α rms* (Å)	C_α DME† (Å)	C_β DME (Å)	C_α radius of gyration‡ (Å)
	0.00§	−129.46	0.51	0.21	0.28	6.16
D1**	2.39	−128.58	0.72	0.37	0.39	6.27
D2	2.39	−122.55	1.93	0.82	0.64	6.47
D3	3.80	−118.39	1.60	0.44	1.10	6.20
D4	3.80	−115.42	4.03	1.12	0.89	5.70
D5	3.80	−111.30	3.11	0.86	1.11	6.44
D6	2.90	−110.29	3.58	1.76	2.21	6.27
D7	2.90	−107.39	2.81	2.02	1.07	5.33
GS††	5.35	−127.47	1.74	0.76	1.44	5.81

*In this work the rms difference has been calculated without optimally translating and rotating the computed oligopeptide structure onto the crystallographic conformation.

†The rms distance matrix error (DME) determines the degree of similarity in the pattern of intramolecular contacts between the calculated heptapeptide structure and the crystallographic conformation. The C_α DME and C_β DME from the native conformation of the extended starting structure measured 1.88 and 1.50 Å, respectively.

‡The C_α radius of gyration of the extended starting structure measured 7.30 Å, while the one of the native inhibitor measures 6.10 Å.

§The first line contains the results of a conjugate gradient refinement of the complex to an rms energy gradient of 0.08 kcal/(mol Å) starting from the crystallographic conformation.

**The two-step procedure was applied to the three starting conformations with different random-number generators.

††Lowest energy structure obtained by the two-step procedure with the grid search sampling during step I.

TABLE II. Differences Between Computed Structures*

	C_α rms							C_α DME					
	D2	D3	D4	D5	D6	D7		D2	D3	D4	D5	D6	D7
D1	2.07	1.69	4.24	3.26	3.43	2.92	D1	0.64	0.45	1.23	0.66	1.62	2.12
D2		2.35	4.07	3.57	4.14	3.31	D2		0.94	1.61	0.57	1.80	2.41
D3			3.84	2.64	2.88	3.20	D3			1.13	0.85	1.56	2.02
D4				2.60	4.81	4.26	D4				1.35	1.31	1.12
D5					3.43	3.33	D5					1.41	2.11
D6						2.89	D6						1.57

*The seven simulations have been listed in the same order as in Table I. All values are given in Å.

main differences are located in the Met-5 side chain in D3 and the Arg-7 side chain in D2 and D3.

The following three requirements must be satisfied by an ideal docking procedure: (1) random starting conformations, positions, and orientations of the ligand inside the receptor binding site should converge to the same minimized complex, (2) the computed lowest energy conformation of the complex should be similar to the native, experimentally observed structure, and (3) low lying local energy minima, i.e., metastable conformations should be determined.

These three goals have been reached in the test case, albeit the first only partially. Four of seven simulations did not converge to the minimum associated with the native, experimentally determined structure. We suppose that this is mainly due to the relatively small number of positional sampling cycles of step I, which had been chosen with the aim to balance a sufficient number of conformations with the computational time needed in step II.

Recently, we modified the sampling strategy in step I by replacing the random rigid body movements of the Monte Carlo method with a grid search in the same degrees of freedom. The axis of translation was automatically determined by moving spheres of different radii along several directions starting from the center of mass of the binding site residues. The number of contacts between each sphere and the atoms of the binding site residues are calculated and that axis is selected along which the number of clashes is minimal. The totally extended structure of the heptapeptide was positioned at 5.35 Å C_α rms deviation from the native conformation with half of the chain lying outside the binding site. This very poor starting position would not enter the binding site in a reasonable computing time as a result of random Monte Carlo sampling. Stepsize for translation was 0.3 Å, and for rotation about the chain axis, 20°. The search was performed over 9 Å and full rotation, i.e., $[(9.0/0.3) \times (360/20)] = 540$ conformations. The lowest energy complex obtained by a two-step procedure consisting of grid search (step I) followed by Monte Carlo optimization (step

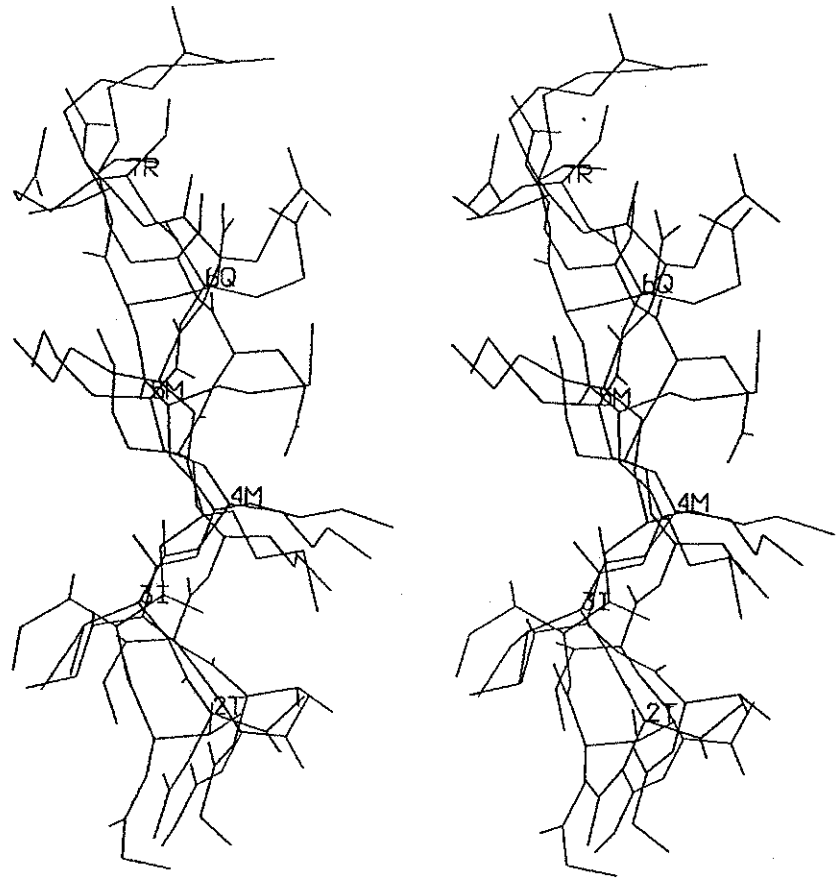

Fig. 1. Stereo picture of the D1, D2, and D3 structures superimposed, without optimally positioning, to the native, experimentally determined conformation of *N*-acetyl-Thr-Ile-Nle-Ψ[CH_2-NH]-Nle-Gln-Arg.amide. C_α atoms of the native structure are labeled.

II) is similar to the native structure (cf. the last line of Table I), except that the Thr-2 and Ile-3 side chain positions are interchanged with respect to the crystallographic result.

Docking of a Hendecapeptide Antigen to the HLA-A2 MHC Protein

Based on the above discussed validation of both model and methods we decided to study the docking of the epitope corresponding to the hendecapeptide 58–68 from the influenza A virus matrix protein (MP58)[16, 17] to the HLA-A2 class I MHC protein.[18, 19] Protein coordinates were kindly provided by Dr. Don C. Wiley.

An α-helical conformation was generated as a starting structure for the foreign antigen since we believe with the majority of researchers in this field[20–24] that the MHC protein binding groove imposes this stable secondary structure on the oligopeptide. This structure was initially constructed by Monte Carlo minimization of an α-helical conformation by changing only the variable side chain torsion angles. The resulting structure of MP58 was coarsely docked parallel to the long axis of the groove, i.e., parallel to the axis of the binding site helices, in two different positions for each of the two possible orientations. The four coarsely docked structures were then submitted to the two-step procedure. In this process, 32 residues that point toward the binding groove of HLA-A2[19] interacted with the hendecapeptide. The acceptance ratio, according to the Metropolis criterion at a T value of 310 K, was about 25%. The calculated lowest energy complex is shown in Figure 2.

The minimum values of E_{II} obtained by the two-step procedure applied to the four coarsely docked conformations are −137.18 and −135.35 kcal/mol (up orientation of the helical axis, as in Fig. 2) and −126.80 and −126.69 kcal/mol (down orientation). The values suggest that the peptide docks with the up orientation, in which the electrostatic interactions are favored and more hydrogen bonds can form.[9]

To assess the validity of this result one has to compare it with the functional data since no structure of an MHC protein complex with a recognizable

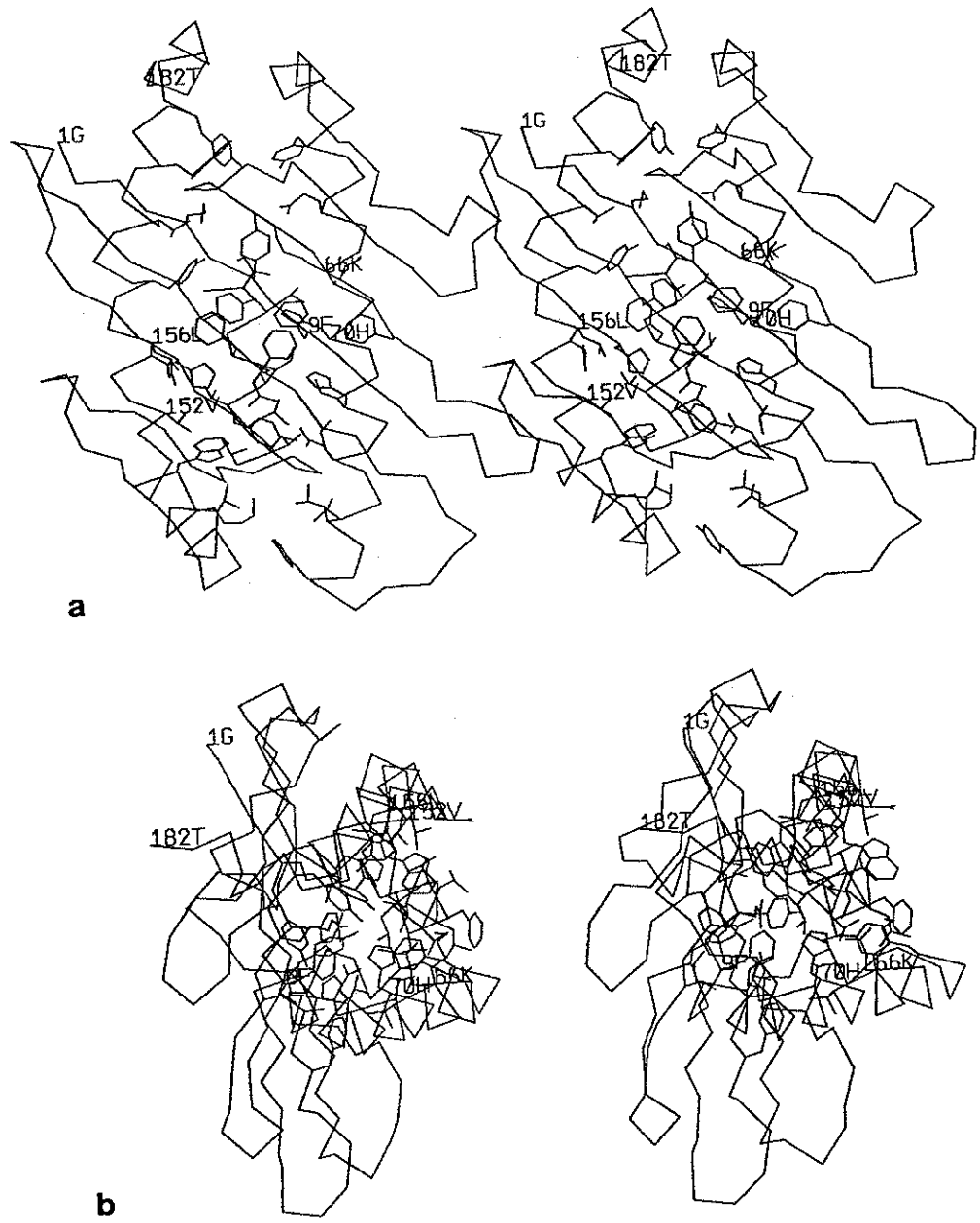

Fig. 2. Stereo pictures of the calculated lowest energy complex between MP58 and HLA-A2. (a) C_α atoms representation containing the side chains of MP58 and the ones of HLA-A2 which point toward the binding groove. (b) Same complex; the axis of the helices are perpendicular to the plane of the page.

foreign antigen has been cocrystallized. The effect of mutations and variations of HLA-A2 on the recognition capabilities of the influenza A matrix peptides 55–73 and 56–68 by CTL has been investigated.[25] On the basis of the associated experimental results the authors[25] stated that the principal interactions in the binding groove are a consequence of close contacts between the peptide antigen and residues Lys-66, Val-152, and Leu-156 on the α-helices of the HLA-A2 molecule and that different amino acids like Phe-9 (on the β-pleated sheet of the α_1-domain) and the closely positioned His-70 (on the α_1-helix) modulate the peptide interactions so that some T cell clones react and some do not. Furthermore, in the same study the authors[25] postulated that the lack of effect on CTL peptide recognition by mutation at residues his-74, Asp-77, and Thr-80 implies, at least with some CTL clones, that they are not necessarily involved in peptide binding although they point into the groove. To experimentally determine how the influenza A virus matrix peptide 57–68 fits into the antigen binding site of HLA-A2 a

TABLE III. Close Contacts Between MP58 and HLA-A2 Residues for the Lowest Energy Complex*

MP58	G58	I59	L60	G61	F62	V63	F64	T65	L66	T67	V68
HLA-A2	Y159	K66	K66		Y99	A69		W147	T73	K146	K146
		H70	A69		H114	H70		V152	D77		W147
		Y99	H70		L156	T73		Q155	R97		A150
		Y159			Y159			L156	W147		

*The following criterion defines a close contact: one or more heavy atoms of residue X on molecule 1 must be located at a distance smaller than 4.2 Å from one or more atoms of residue Y on molecule 2. The lack of close contacts for Phe-64 and the close contacts between Phe-62, Val-63, and the HLA-A2 binding site residues agree with experimental results (see text).

series of peptide analogues was synthetized, each containing a point mutation, and then tested for their ability to sensitize target cells and to compete with the natural sequence for recognition by CTL.[17] Thereby, it was possible to identify MP58 side chains critical for T cell recognition (Leu-60, Phe-64) and residues responsible of MHC protein binding (Gly-61, Phe-62, and Val-63). Table III lists the close contacts between MP58 and HLA-A2 for the lowest energy complex obtained by the two-step docking procedure. This computed structure (Fig. 2 and Table III) indicates that the docked conformation of MP58 into HLA-A2 is consistent with the one that can be derived from the analysis of the experimental results.

CONCLUSIONS

A new two-step procedure for docking a flexible oligopeptide to a static protein on the basis of an accurate force field has been developed and tested on the HIV-1 aspartic proteinase complex with the Gly-Thr-Ile-Met-Ψ[CH_2-NH]-Met-Gln-Arg heptapeptide. The computed lowest energy conformation corresponds to the native, experimentally observed structure; furthermore, among the seven simulations performed, the three docked conformations associated with the lowest energy values are similar (C_α rms deviation within 2.0 Å) to each other and to the native structure. The lowest energy conformation generated by a two-step procedure with grid search sampling in step I is similar to the crystallographic result in spite of the fact that the simulation was started from an initial position of the totally extended oligopeptide translated more than 5 Å away from the experimentally determined binding site.

An improved understanding of the details of inhibitor–enzyme interactions will be helpful in designing new HIV-1 aspartic proteinase inhibitors with improved binding characteristics. Such theoretical studies could become of fundamental importance in the development of new AIDS therapeutics.

Furthermore, we applied the two-step docking procedure to a typical recognition problem of cellular immunology. The agreement between theoretical and experimental findings for the HLA-A2 complex with the hendecapeptide 58–68 of the influenza A virus matrix protein represents a further validation of our method. Accordingly, we conclude that the two-step docking procedure developed here can be useful to predict the effects of mutagenesis of either the restriction element or the peptide ligand on the association energy of the complex. Since it is unlikely that many different peptide MHC molecule complexes will be cocrystallized, the interplay between functional data and computer simulations will become essential as a guide for the synthesis of analogous peptides that bind MHC molecules with higher affinity. In particular, it may provide help in the prediction of new epitopes to develop new vaccines or inhibiting drugs.

ACKNOWLEDGMENTS

We thank Dr. Hein Osenberg for computer assistance, and Dr. Marc Brandis, Olivier Forni, and Marco Sassi who assisted in the development of the molecular graphics program. Furthermore, we wish to thank Drs. Leslie Bisset, Walter Fierz, Hans Hengartner, and Rolf Zinkernagel for helpful discussions.

REFERENCES

1. Scheraga, H. A. Protein structure and function, from a colloidal to a molecular view. Carlsberg Res. Commun. 49:1–55, 1984.
2. DesJarlais, R. L., Seibel, G. L., Kuntz, I. D., Furth, P. S., Alvarez, J. C., Ortiz de Montellano, P. R., DeCamp, D. L., Babe, L. M., Craik, C. S. Structure-based design of nonpeptide inhibitors specific for the human immunodeficiency virus 1 protease. Proc. Natl. Acad. Sci. U.S.A. 87: 6644–6648, 1990.
3. Wodak, S. J., Janin, J. Computer analysis of protein-protein interaction. J. Mol. Biol. 124:323–342, 1978.
4. Kuntz, I. D., Blaney, J. M., Oatley, S. J., Langridge, R., Ferrin, T. E. A geometric approach to macromolecule-ligand interactions. J. Mol. Biol. 161:269–288, 1982.
5. DesJarlais, R. L., Sheridan, R. P., Dixon, J. S., Kuntz, I. D., Venkataraghavan, R. Docking flexible ligands to macromolecular receptors by molecular shape. J. Med. Chem. 29:2149–2153, 1986.
6. Goodsell, D. S., Olson, A. J. Automated docking of substrates to proteins by simulated annealing. Proteins: Struct. Funct. Genet. 8:195–202, 1990.
7. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., Teller, E. Equation of state calculations by fast computing machines. J. Chem. Phys. 21:1087–1092, 1953.
8. Li, Z., Scheraga, H. A. Monte Carlo-minimization approach to the multiple-minima problem in protein folding. Proc. Natl. Acad. Sci. U.S.A. 84:6611–6615, 1987.
9. Caflisch, A. Monte Carlo-docking of polypeptide chains; a general model with applications to oligopeptides bonded to

class I MHC proteins. PhD thesis, ETH Zürich, 1991. ETH diss 9400.
10. Burkert, U., Allinger, N. L. "Molecular Mechanics," Vol. 177. Washington, DC: ACS Monograph, 1982.
11. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. J. Comput. Chem. 4:187–217, 1983.
12. Momany, F. A., McGuire, R. F., Burgess, A. W., Scheraga, H. A. Energy parameters in polypeptides. VII. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen bond interactions, and intrinsic torsional potentials for the naturally occuring amino acids. J. Phys. Chem. 79:2361–2381, 1975.
13. Miller, M., Schneider, J., Sathyanarayana, N. K., Toth, M. V., Marshall, G. R., Clawson, L., Selk, L., Kent, S. B. H., Wlodawer, A. Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 Å resolution. Science, 246:1149–1152, 1989.
14. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., Tasumi, M. The Protein Data Bank: A computer-based archival file for macromolecular structures. J. Mol. Biol. 112:535–542, 1977.
15. Erickson, J., Neidhart, D. J., VanDrie, J., Kempf, D. J., Wang, X. C., Norbeck, D. W., Plattner, J. J., Rittenhouse, J. W., Turon, M., Wideburg, N., Kohlbrenner, W. E., Simmer, R., Helfrich, R., Paul, D. A., Knigge, M. Design, activity, and 2.8 Å crystal structure of a C_2 symmetric inhibitor complexed to HIV-1 protease. Science 249:527–533, 1990.
16. Gotch, F., Rothbard, J., Howland, K., Townsend, A., McMichael, A. Cytotoxic T lymphocytes recognize a fragment of influenza virus matrix protein in association with HLA-A2. Nature (London) 326:881–882, 1987.
17. Gotch, F., McMichael, A., Rothbard, J. Recognition of influenza A matrix protein by HLA-A2-restricted cytotoxic T lymphocytes. J. Exp. Med. 168:2045–2057, 1988.
18. Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L., Wiley, D .C. Structure of the human class I histocompatibility antigen, HLA-A2. Nature (London) 329:506–512, 1987.
19. Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L., Wiley, D. C. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. Nature (London) 329:512–518, 1987.
20. DeLisi, C., Berzofsky, J. A. T-cell antigenic sites tend to be amphipathic structures. Proc. Natl. Acad. Sci. U.S.A. 82: 7048–7052, 1985.
21. Spouge, J. L., Guy, H. R., Cornettte, J. L., Margalit, H., Cease, K., Berzofsky, J. A., DeLisi, C. Strong conformational propensities enhance T cell antigenicity. J. Immunol. 138(1):204–212, 1987.
22. Reyes, V. E., Chin, L. T., Humphreys, R. E. Selection of class I MHC-restricted peptides with the strip-of-helix hydrophobicity algorithm. Mol. Immunol. 25(9):867–871, 1988.
23. Allen, P. M., Matsueda, G. R., Evans, R. J., Dunbar, J. B., Jr., Marshall, G. R., Unanue, E. R. Identification of the T-cell contact and Ia contact residues of a T-cell antigenic epitope. Nature (London) 327:713–715, 1987.
24. Rothbard, J. B., Taylor, W. R. A sequence pattern common to T cell epitopes. EMBO J 7(1):93–100, 1988.
25. McMichael, A. J., Gotch, F. M., Santos-Aguado, J., Strominger, J. L. Effect of mutations and variations of HLA-A2 on recognition of a virus peptide epitope by cytotoxic T lymphocytes. Proc. Natl. Acad. Sci. U.S.A. 85: 9194–9198, 1988.