PROTEINS: Structure, Function, and Genetics 35:153–162 (1999)

Docking of Flexible Ligands to Flexible Receptors in Solution by Molecular Dynamics Simulation

Massimiliano Mangoni,[1] Danilo Roccatano,[2] and Alfredo Di Nola[1]*
[1]*Department of Chemistry, University of Rome, Rome, Italy*
[2]*Department of Biophysics, University of Groningen, Nijenborgh, The Netherlands*

ABSTRACT **In this paper, a method of simulating the docking of small flexible ligands to flexible receptors in water is reported. The method is based on molecular dynamics simulations and is an extension of an algorithm previously reported by Di Nola et al. (Di Nola et al., Proteins 1994;19:174–182). The method allows a fast exploration of the receptor surface, using a high temperature of the center of mass translational motion, while the ligand internal motions, the solvent, and the receptor are simulated at room temperature. In addition, the method allows a fast center of mass motion of the ligand, even in solution. The dampening effect of the solvent can be overcome by applying different weights to the interactions between system subsets (solvent, receptor, and ligand). Specific ligand–receptor distances have been used to compare the results of the simulations with the crystal structure. The method is applied, as a test system, to the docking of the phosphocholine to the immunoglobulin McPC603. The results show the similarity of structure between the complex in solution and in the crystal. Proteins 1999;35:153–162.** © 1999 Wiley-Liss, Inc.

Key words: docking; molecular dynamics; rational drug design; immunoglobulin

INTRODUCTION

The recognition (docking) process between a ligand and its receptor plays an important role in virtually all biological processes. Understanding the molecular basis of the docking process would open the possibility of designing ligands for any specific receptor.

Experimental work, and computer simulation and analysis, are the main scientific tools for these studies. However, in spite of the rapid spread of computational approaches, stimulated by the availability of high-resolution data on proteins, the basic problem in computational approaches to docking and design, i.e., the determination of the thermodynamically most favorable modes of binding, still remains to be adequately solved.

Among others, the major problems are the following: (1) the type of the target function used to weigh the conformations; (2) the algorithm used for an efficient exploration of the conformational space and for the determination of the minimum value of the target function; (3) the inclusion of flexibility of both ligand and receptor in the calculation; (4) the inclusion of explicit solvent molecules.

In general, simplified solutions to these problems (such as the use of rigid conformations, shape descriptors, simulation in vacuo, etc.) permit rapid scanning of large structural databases, but are incomplete and often fail. On the other hand, more sophisticated methods (based on free energy calculations, use of flexible ligands and receptors, and explicit solvent molecules) are computationally intensive and beyond the capability of modern parallel computers. A good computational approach should include different methods in a sort of hierarchical order: simplified methods to rapidly scan large structural databases and more sophisticated methods that include ligand and receptor flexibility, and explicit solvent treatment, for more detailed information.

The docking algorithms so far proposed can be distinguished by the complexity of the target function and by the extent to which the molecular flexibility is taken into account. It has to be pointed out that none of them includes explicit solvent molecules. The target functions range from surface complementarity,[1,2] to surface area burial,[3] to total molecular mechanics energy,[4] to free energy calculations.[5,6] When only rigid receptors are taken into account, the computational effort can be reduced by pre-computing the potential energy of the receptor so that the ligand energy has only to be evaluated for any specific ligand.[7,8]

Depending upon the extent to which the molecular flexibility is taken into account, a docking algorithm can be categorized into the following classes: (1) both ligand and protein rigid, (2) flexible ligand and rigid protein, and (3) both ligand and protein flexible. Early studies treated ligands and receptors as rigid bodies,[9,10] but more recent works have included the flexibility of ligands.[11–21]

When the structure of the receptor complexed with a ligand is known, it is possible to use this rigid structure for docking studies with other ligands. However, there are several examples where the conformation of the docking pocket deviates significantly from the model structure.[22] Moreover, the inclusion of flexibility is fundamental when

Abbreviations: c.o.m., center of mass; PC, phosphocholine; msd, mean square displacement; rmsd, root mean square deviation; MDD, molecular dynamics docking.

Grant sponsor: European Community; Grant sponsor: Italian Ministero dell'Universitàe della Ricerca Scientifica e Tecnologica.

*Correspondence to: Alfredo Di Nola, Department of Chemistry, University of Rome "La Sapienza", P.O. Box 34 Roma 62, P.le Aldo Moro 5, I-00185, Rome, Italy. E-mail: dinola@degas.chem.uniroma1.it

Received 28 May 1998; Accepted 15 December 1998

the effects of mutations in the receptor are to be studied. Only little effort has been expended to allow for the flexibility of the receptor[16,17,23–25], and most of the docking algorithms in use consider the receptor as a rigid body. To allow for a conformational flexibility of the receptors, Leach [23] used a discrete model of side chain flexibility, and Jones et al.[16] described a genetic algorithm that encodes the torsion of some rotational bonds. Knegtel et al.[24] used an energy-weighted average over a set of crystal structures. Although their method does not explicitly include receptor flexibility, it permits searching over a set of different conformations. Apostolakis et al.[25] performed conjugate gradient minimizations of already randomly generated ligand–protein complexes. Luty et al.[17] used molecular dynamics, with an implicit solvation model, combined with a rigid representation of the bulk of the receptor and a mobile binding site.

Besides the important problem of structural flexibility, another phenomenon, the presence of water molecules, may play an important role in the complex formation. It is well known that structures in the crystal and in solution differ in several important respects, such as radius of gyration, solvent accessible surface, intramolecular hydrogen bonds, and orientation of surface side chains. Moreover, single water molecules may act as intermediates of hydrogen bonds. It should be noted that full flexibility of the receptor requires the inclusion of the solvent in the simulation; in fact, a full flexible protein simulated in vacuo tends to squeeze and maximize the intramolecular contacts.

Di Nola et al.[13] has proposed a method, molecular dynamics docking (MDD), based on a modification of molecular dynamics simulations. Although this method, in principle, allows the simulation of the flexibility of the receptor and explicit solvent molecules, it was only applied to the docking of a flexible ligand, the phosphocholine, onto a rigid receptor, the immunoglobulin McPC603, in vacuo. In this paper the extension of the MDD algorithm to the docking of a flexible ligand onto a flexible receptor in water is reported.

In the present approach, the initial protein structure is taken from the crystal structure of the complex. After equilibration, the structure of the binding pocket deviates from the initial one, so that the present approach permits monitoring local, even significant, rearrangements of the protein. In this respect, it represents a step toward the inclusion of the full flexibility of the receptor. The simulation of large conformational rearrangements of the whole protein structure is beyond present computational capabilities. What is required is a treatment of large-scale modifications in a simplified way, that is, by using the essential dynamics method.[26]

The MDD method consists of a separation of the center of mass motion of the ligand from its internal and rotational motions and a separate coupling to different thermal baths for both types of motion of the ligand and for the motion of the receptor. The temperatures and the time constants of coupling to the baths can be arbitrarily varied. Thus, it is possible to increase the kinetic energy of the center of mass of the ligand without increasing the temperature of the internal motions of the receptor and of the ligand, which allows complete control of the search rate. Moreover, given appropriate values of the temperatures and coupling constants, it is possible to have the ligand and/or the receptor either flexible or rigid.

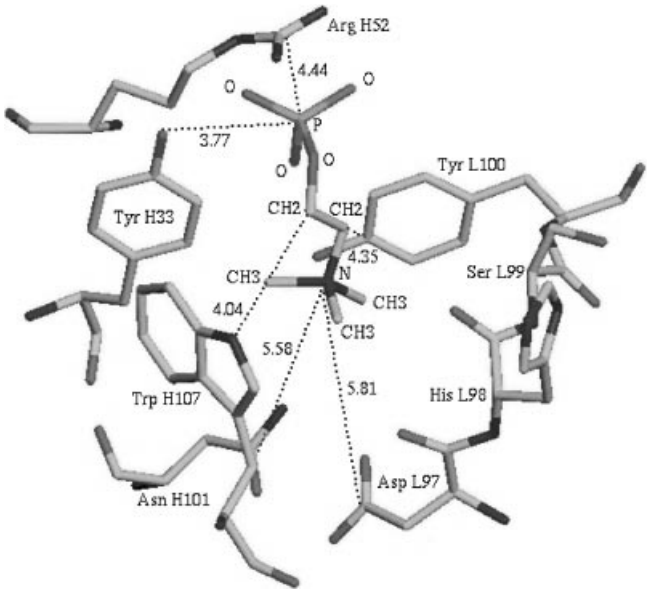

Fig. 1. Main interactions of the PC in the M603 Fab fragment binding site.

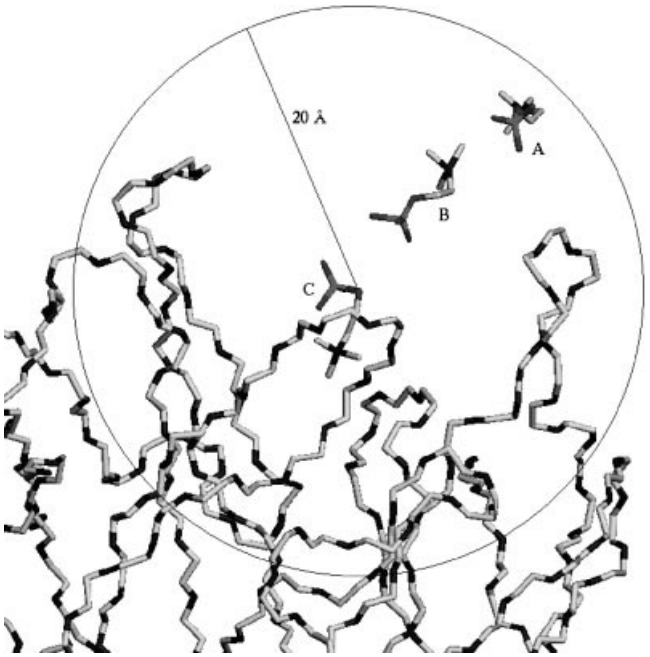

Fig. 2. The simulated system is defined by a sphere of 20 Å around the chain oxygen of the PC in the crystallographic position. C is the position of PC in the crystal complex. A and B are the two different starting positions used in the simulations.

The results showed that, despite the presence of explicit water molecules, the MDD algorithm permits a fast center

TABLE I. Partial Changes for Phosphocholine

Atom	Partial charge (q_e)
CH3	+0.248
N	+0.008
(N)CH2(CH2)	+0.248
(CH2)CH2(O)	+0.000
(CH2)O(P)	−0.300
P	+0.630
(P)O	−0.600
(P)O(H)	−0.528
H	+0.398

TABLE II. Distances Between PC and Protein Side Chains in the Crystal

Ligand	Protein	Distance (Å)
P	C_ζ, Arg H52	4.44
P	OH, Tyr H33	3.77
(N)CH2(CH2)	C_ζ, Tyr L100	4.35
(CH2)CH2(O)	N_ϵ, Trp H107	4.04
N	C_γ, Asp L97	5.81
N	C_γ, Asn H101	5.58

TABLE III. Temperatures and Coupling Constants of the Thermal Baths

System	Temperature (K)	Coupling constant (ps)
Receptor	300	0.002
Ligand (center of mass)	900–1500	0.006
Ligand (internal degrees of freedom)	300	0.004
Solvent	300	0.002

of mass motion of the ligand for the search of the binding site. It was found that the presence of explicit water molecules shields the interactions between the ligand and the receptor, so that the use of different weights for the ligand–receptor or ligand–solvent interactions were necessary to overcome this effect. The average structure of the complex in the binding region is close to the crystal structure. As in the previous study, the docking of the phosphocholine onto the immunoglobulin McPC603 was used as test case.

METHODS

Let us consider a probe molecule (the ligand), that we wish to dock onto a large receptor. Its kinetic energy with respect to the laboratory frame is given by $E_{kin} = \frac{1}{2}\Sigma_i\, m_i\, v_i^2$, where $\mathbf{v}_i$ is the velocity with respect to the laboratory frame. In a coordinate system with origin in the center of mass (c.o.m. frame), the velocities are: $\mathbf{v}_{c,i} = \mathbf{v}_i - \mathbf{V}$, where $\mathbf{V} = \Sigma_i\, m_i \mathbf{v}_i / M$ is the velocity of the c.o.m. in the laboratory frame and M is the total mass. The total momentum $\Sigma_i\, m_i\, \mathbf{v}_{c,i}$ in the c.o.m. frame is zero.

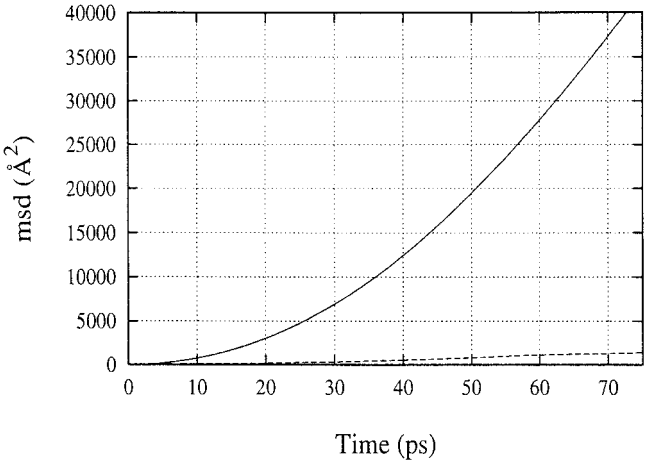

Fig. 3. Mean square displacement (msd) of the PC center of mass during 100 ps simulation in a cubic box with solvent molecules. Solid line: MDD algorithm with temperature of c.o.m. at 1,000 K and temperatures of both water and the internal degrees of freedom of PC at 300 K. Dashed line: temperature of the whole system at 1,000 K.

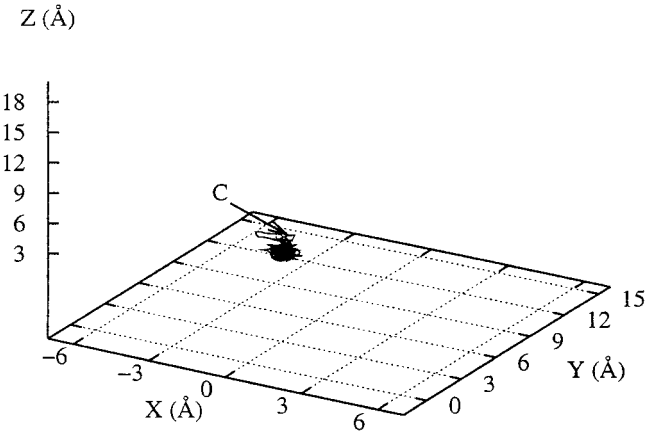

Fig. 4. Trajectory of the center of mass of PC, starting from the crystallographic position. C is the position of PC c.o.m. in the crystal complex.

The kinetic energy in the laboratory frame can be expressed as:

$$E_{kin} = \tfrac{1}{2}\sum_i m_i v_i^2 = \tfrac{1}{2}\sum_i m_i v_{c,i}^2 + \tfrac{1}{2} MV^2.$$

Let us call these two energies internal energy and c.o.m. energy. For an isolated molecule these two energies are completely decoupled, as the motion of the c.o.m. can be changed only by external forces. When the ligand interacts with the receptor, the intermolecular forces affect both internal and c.o.m. energies of the ligand. As a consequence, there is an energy flow between these two pools of energy and, of course, between the ligand and the receptor. In practice, when a ligand approaches a rigid or flexible receptor, its c.o.m. kinetic (or translational) energy is converted into internal energy of both ligand and receptor and it gets trapped.

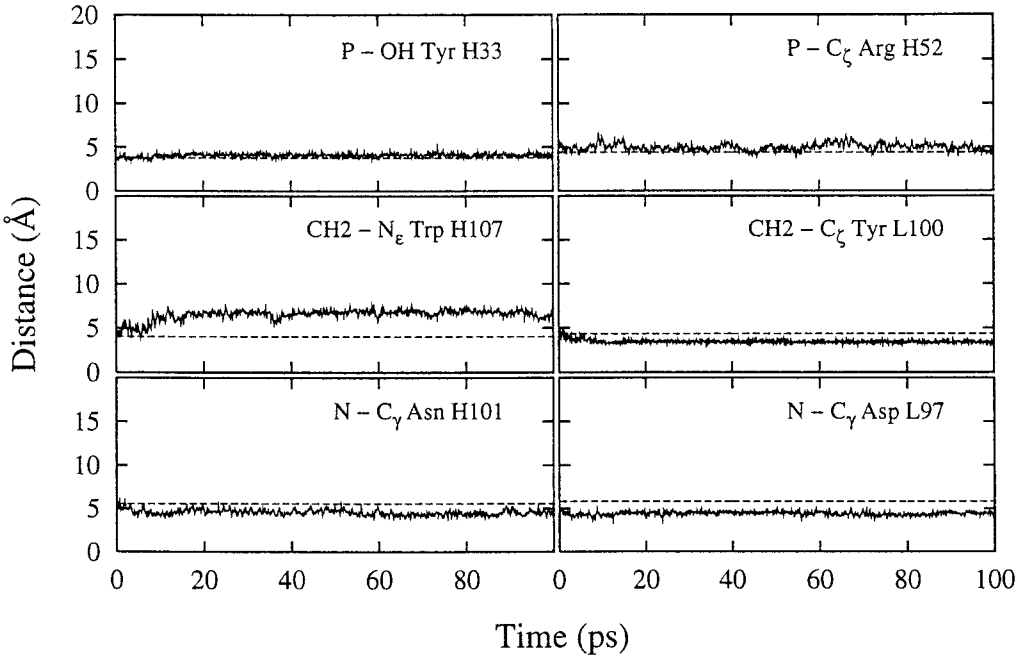

Fig. 5. Trajectories of the six ligand–receptor distances reported in table II (solid line). The starting position of PC coincides with its crystal structure. The dashed lines correspond to the values in the crystal structure.

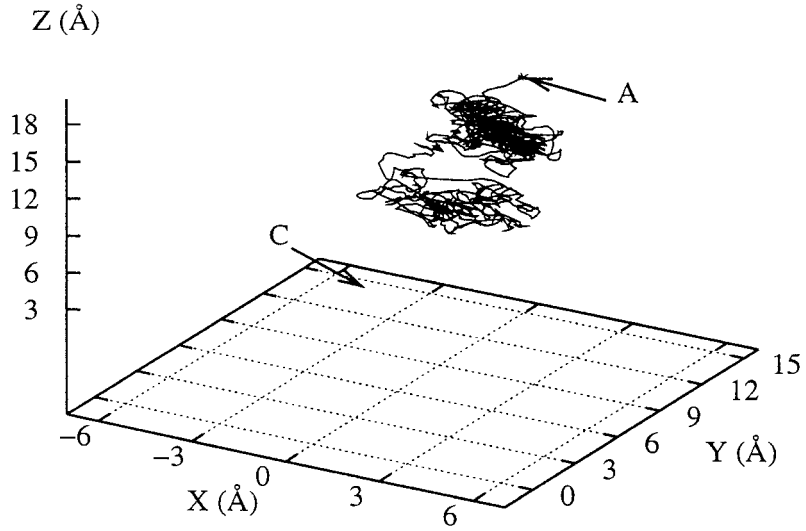

Fig. 6. Trajectory of the center of mass of PC during the 180 ps simulation at 300 K, starting from position A of Fig. 2 and using only one thermal bath. C is the position of PC c.o.m. in the crystal complex.

The MDD algorithm, as developed by Di Nola et al.,[13] couples the internal and c.o.m. kinetic energies of the ligand to two different Berendsen's thermal baths[27] with different time constants. The translational temperature of the ligand was chosen in the range 900–1,500 K to allow a fast search of the receptor surface. The internal degrees of freedom were kept at room temperature. Small values of the time constants were chosen to control the energy flow between the translational and internal degrees of freedom.

In the present study, the flexibility of the receptor and explicit water molecules have been added, so that two additional baths were introduced to keep the temperatures of the receptor and the solvent at room temperature.

Computational Procedure

We have applied the method to the binding of the phosphocholine (PC) to the immunoglobulin McPC603. The binding depends on both van der Waals and electrostatic interactions. The ligand binds into a hydrophobic pocket, with its choline group interacting favorably with an aspartic acid (Asp L97) and an asparagine (Asn H101) at the base of the pocket, while the phosphate end forms

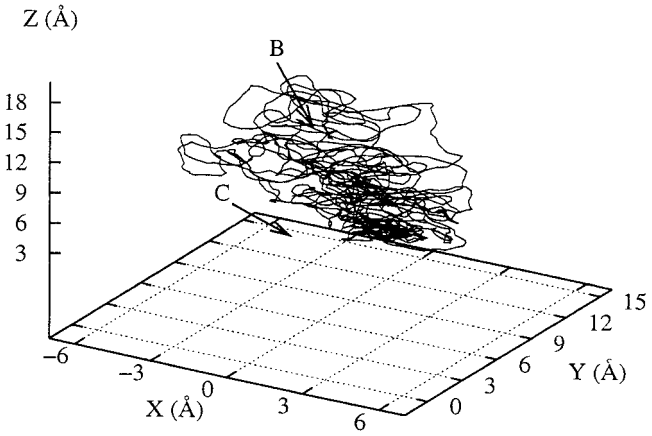

Fig. 7. Trajectory of the center of mass of PC during the 180 ps simulation, starting from position B and using the MDD algorithm. C is the position of PC c.o.m. in the crystal complex.

hydrogen bonds with tyrosine (Tyr L100) and arginine (Arg H52) groups, as shown in Fig. 1 (obtained with the software program MOIL-View[28]). The coordinates of the crystal structure of the complex between M603 Fab fragment and the PC were taken from Segal et al.[29]

The simulation system was obtained by putting the binding site of M603 Fab fragment and the PC (in position A or B of Fig. 2) in a cubic box of 50 Å length, filled with water. The PC and the water were subjected to 100 ps molecular dynamics equilibration at a temperature of 300 K, while the protein was kept rigid. From this box a sphere with 20 Å radius was cut around the position of the chain oxygen of the PC in the crystal complex (Fig. 2). All the atoms included in an inner sphere of radius15 Å were free. The protein and water atoms included between the two spheres were position restrained. The total number of atoms in this system was 2874, 1204 of which were in the inner sphere. In this way, 648 water molecules were included in the system, 417 position restrained in the boundary region and 231 in the inner sphere. There were 112 protein residues included in the simulation, 47 of which were free to move. All the residues included in the system belong to the five hypervariable loops of the antibody binding site, as shown in Fig. 2.

The programs for the molecular dynamics simulations were taken from the GROMOS87 library[30] and were adapted to perform separate scaling of the temperatures. The SHAKE algorithm[31] was used to keep bond lengths rigid. The parameters of the simulation were chosen as follows: a cut-off radius of 16 Å for the ligand–protein interactions, and a cut-off radius of 9 Å for the remaining interactions; the time step was 2 fs and the dielectric permittivity was 1. The charges in the phosphocholine, taken from Di Nola et al.,[13] are reported in Table I. The charges of the immunoglobulin were taken from the GROMOS87 package. SPC[32] model was used for water. Only polar hydrogens were explicitly treated. For non-polar hydrogens, the united atoms representation was adopted.

Due to the flexibility of the protein, the results could not be evaluated by the root mean square deviation (rmsd) of the positions of the PC in the crystal and in the simulation. We have therefore monitored the main interactions between the PC and the protein, reported in Table II and shown in Fig. 1.

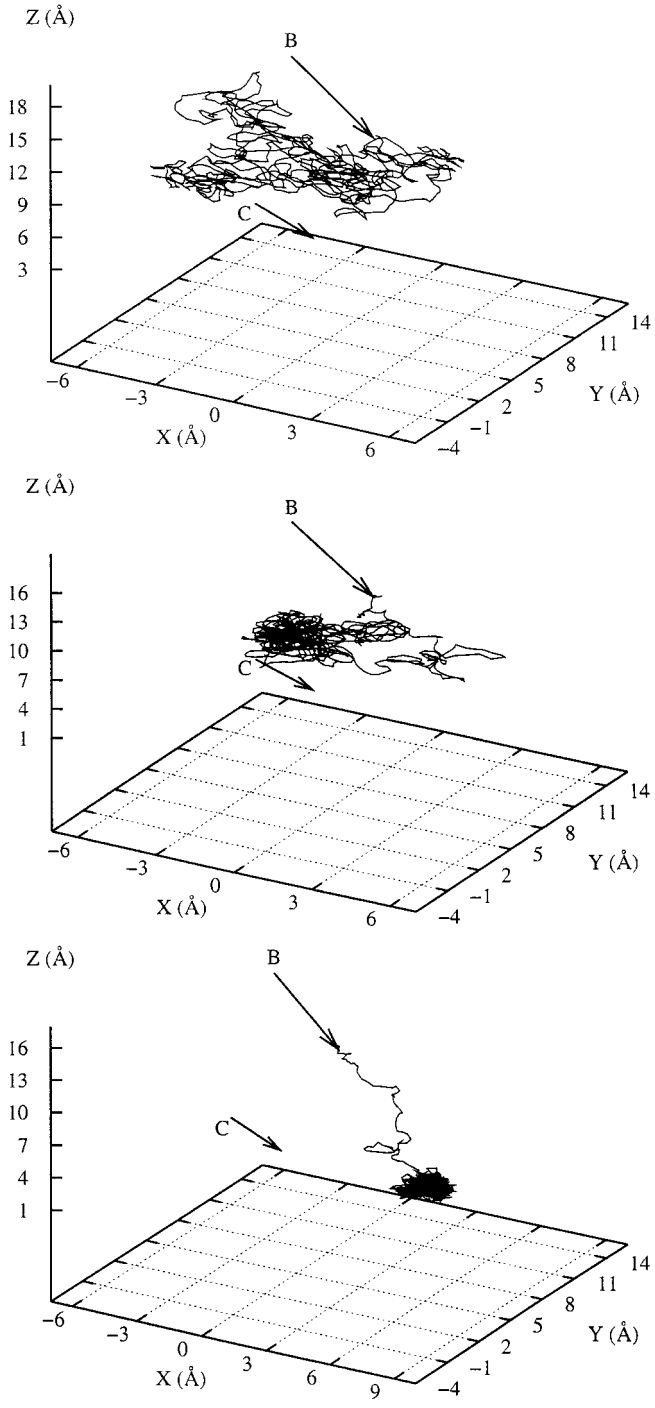

Fig. 8. Trajectories of the center of mass of PC during the 180 ps simulation starting from position B. Protein rigid and solvent and ligand at T = 1,000 K. **A:** Full interactions; **B:** simulation with half the weight of the ligand–solvent interactions; **C:** simulation with double the weight of the ligand–receptor interactions.

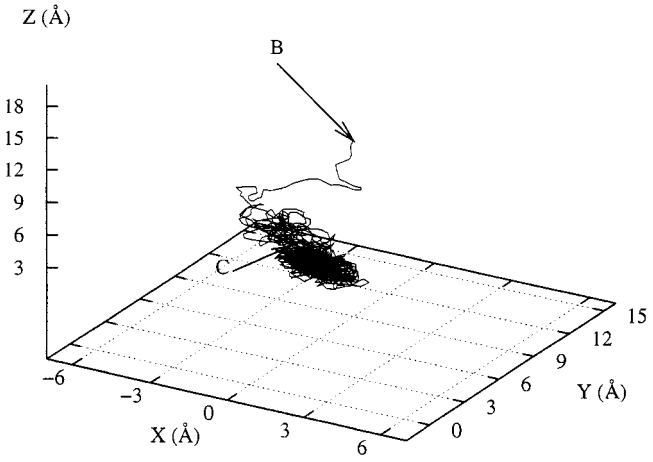

Fig. 9. Trajectory of the center of mass of PC during the 180 ps simulation, starting from position B. MDD simulation with half the weight of the ligand–solvent interactions. C is the position of PC c.o.m. in the crystal complex.

During the simulations, the translational temperature of the ligand was kept in the range 900–1,500 K to allow a fast exploration of the receptor surface, while the internal degrees of freedom were kept at 300 K. Small values of the time constants were chosen to control the energy flow between the translational and internal degrees of freedom. This was also necessary to prevent the ligand, with its fast motions, from inducing unrealistic local modifications of the receptor structure. The values of the temperatures and time constants used for the coupling to the four thermal baths are reported in Table III.

RESULTS

The effect of the MDD algorithm in vacuo, as reported in the previous paper,[13] was twofold: it allowed a fast motion of the ligand, and it avoided the trapping in local minima. To evaluate how much the presence of explicit solvent molecules slows down the translational motion of the ligand, we performed two molecular dynamics simulations of the PC in water, with and without the MDD algorithm, respectively. In the first one, the whole simulation system was kept at a temperature of 1,000 K with only one thermal bath. In the second one, the MDD algorithm was used to keep the c.o.m. temperature at 1,000 K, while the rest of the system was kept at 300 K. The mean square displacement (msd) of the PC in these two simulations are shown in Fig. 3. The diffusion coefficient of the PC c.o.m. in the first simulation, evaluated using the slope of the msd curve,[33] was $3.06 \times 10^{-4} cm^2/s$ and the fitting with a function $y = at^k$ (with y representing the msd and t the time) gave the result $k = 1.11$. Fig. 3 shows that the mean square displacement is not proportional to the time when the MDD algorithm is applied. The fitting with a function $y = at^k$ gave the value $k = 1.96$. The slope of the curve at t = 50 ps divided by 6 (in order to make a comparison with the diffusion coefficient of the previous case) was $1.27 \times 10^{-2} cm^2/s$, showing that the MDD algorithm increases the molecule translational diffusion by at least one order of magnitude.

To evaluate if the MDD algorithm induces unrealistic changes in the internal structure of the ligand, we have calculated the N–P distance in two different simulations of the PC in water: the first with no application of the MDD algorithm and at a temperature of 300 K, the second with the MDD algorithm and with a temperature of the c.o.m. at 1,000 K, while the solvent and internal degrees of freedom of the PC were kept at 300 K. The average distances in the two simulations were 4.65 and 4.71 Å, respectively, showing that the MDD algorithm does not significantly affect the internal structure of the ligand.

To have a reference simulation, we performed 100 ps of MD simulation of the ligand–receptor complex at T = 300 K, without using the MDD algorithm. The starting position of the PC corresponded to the crystal structure of the complex. The trajectory of the PC c.o.m. is shown in Fig. 4. Fig. 5 shows the trajectories of the six distances reported in Table II. It can be noted that the six contacts are basically conserved. The main deviation is observed for the CH2 – N_ϵ of Trp H107 and corresponds to a weak hydrophobic contact, as reported by Novotny et al.,[34] so that a deviation of the structure in water from the structure in the crystal is not surprising. The shorter values of the distances involving the choline interactions could indicate that our force field somewhat emphasizes these interactions. The average rmsd of these six distances from the distances in the crystal was 1.4 Å.

A second reference simulation was performed. The MDD algorithm was not applied in this simulation, and one thermal bath was used at T = 300 K. The initial position of the PC was 14 Å, far from the crystal structure position, and the length of the simulation was 180 ps. The trajectory of the PC c.o.m. is shown in Fig. 6. The docked position of the crystal structure is indicated in the figure (C). The figure clearly shows that within the simulation time the ligand explores two restricted zones, but it does not get into the docking site, giving at t = 180 ps the best rms deviation ~5 Å from the crystal position.

The last simulation was repeated using the MDD algorithm, with a temperature of the c.o.m. motion of the ligand of 1,500 K and with a different initial position. The trajectory of the c.o.m. of the ligand (Fig. 7) shows that the ligand explores a larger surface of the protein in the same time period as that of the previous simulation; it gets closer to the correct position, but it does not fit correctly to the binding site. Analogous results were obtained with simulations up to 300 ps and with c.o.m. temperatures up to 2,400 K. A close inspection of the binding site during the simulation showed that three water molecules, close to the pocket, prevented the correct docking of the PC. It can be hypothesized that a much longer simulation would be necessary to remove the water molecules from the pocket.

To circumvent this effect, we have performed two types of simulations. The first used a weight of the ligand–water interactions halved with respect to the remaining interactions; the second used a weight of the ligand–protein interactions doubled with respect to the remaining interac-

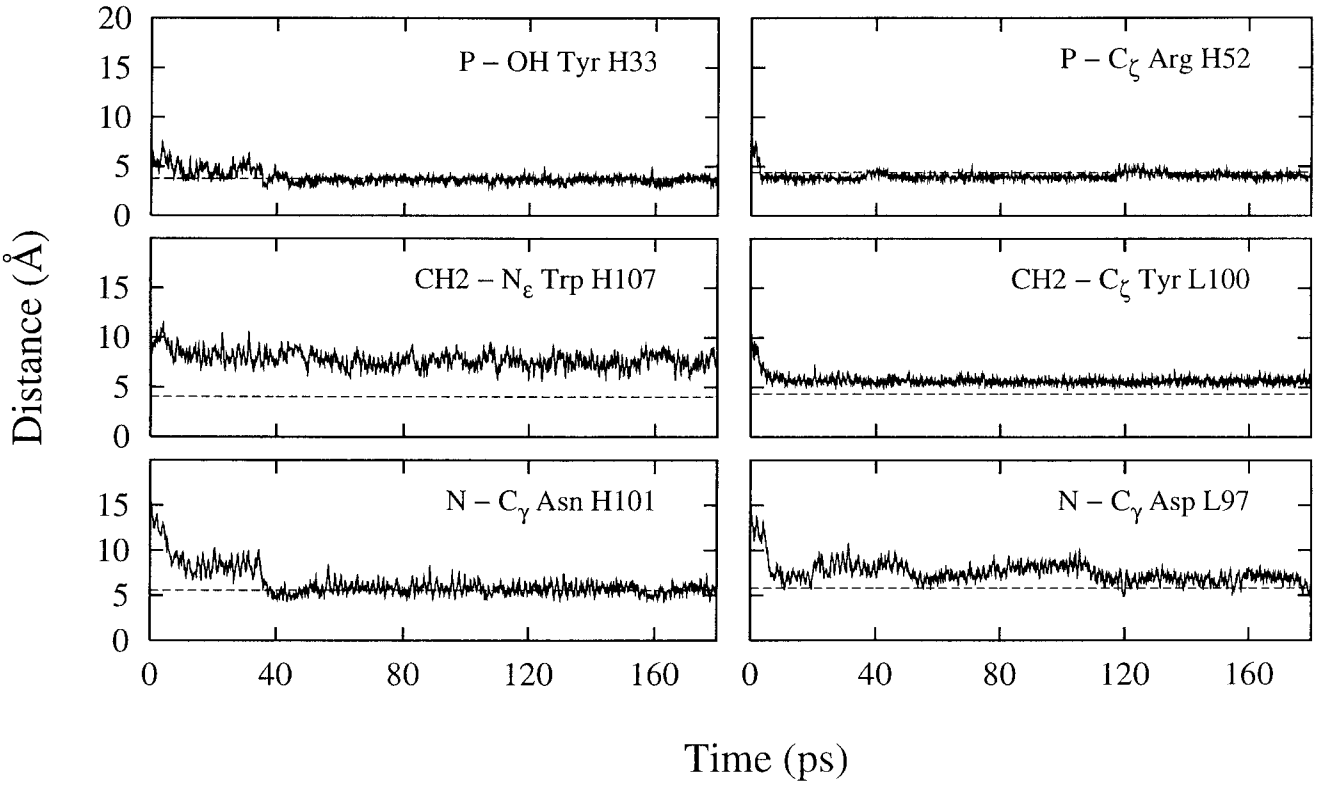

Fig. 10. Trajectories of the six ligand–receptor distances reported in Table II (solid line). MDD simulation with half weight of the ligand–solvent interactions. The dashed lines correspond to the values in the crystal structure.

tions. The reference simulation at T = 300 K, with a single thermal bath, was repeated in both cases and was unsuccessful for lengths up to 300 ps. To have further reference simulations, three additional simulations of 180 ps length, with a single thermal bath at T = 1,000 K for the ligand and the solvent, were performed. In these cases, the protein backbone has been taken as rigid, because at this temperature, unrealistic folding can obtain. The three simulations, (without any scaling of the interactions, with the ligand–water interactions halved, and with the ligand–protein interactions doubled), are reported in Fig. 8. The figure shows that the ligand does not dock the protein (Fig. 8A) or it binds the protein in the wrong place (Fig. 8B,C).

The trajectory of the c.o.m. of the ligand, with the MDD algorithm, and with a weight of the ligand–water interactions halved with respect to the remaining interactions, is reported in Fig. 9 and shows that the ligand fits the pocket. The parameter setting of the four thermal baths were as follows: center of mass T = 1,500 K, time constant 6 fs; internal degrees of the ligand T = 300 K, time constant 4 fs; solvent T = 300 K, time constant 2 fs; protein T = 300 K, time constant 2 fs. The trajectories of the six distances characterizing the docking are shown in Fig. 10. There is a good agreement with the trajectories of the reference simulation reported in Fig. 5.

In particular, we found the same deviation from the crystallographic structure for the distances CH2 – N_ϵ of Trp H107 and N – C_γ of Asn H101. The rmsd of these six distances from those of the reference simulation, evaluated over the time range 130–180 ps, was 1.34 Å, and the rmsd from the distances in the crystal was 1.15 Å. Fig. 11 shows a snapshot of the binding site at t = 123 ps (obtained with the software program QUANTA 97 from MSI[35]). In the figure, the water molecules nearest to the pocket are also shown. Fig. 12 shows the deviation between the crystal structure of the protein and the snapshot at t = 130 ps. Deviations up to 3.2 Å for the backbone, and up to 5.0 Å for the side chains can be observed. Deviations of such amplitude from the crystal structure are hard to detect with other docking algorithms that allow limited conformational flexibility.[16,17,23–25]

Figs. 13, 14, and 15 show the results obtained with a weight of the ligand–protein interactions doubled with respect to the remaining interactions, and with the same parameter setting as the previously cited simulation. The results show that the ligand enters into the pocket and reaches a position close to the crystallographic one. The rmsd from the reference simulation was 0.78 Å, evaluated over the time range 130–180 ps, and the rmsd from the crystal structure was 0.83 Å. The results show that in the present case there is a better agreement with the crystal structure than with the reference simulation reported in Fig. 5.

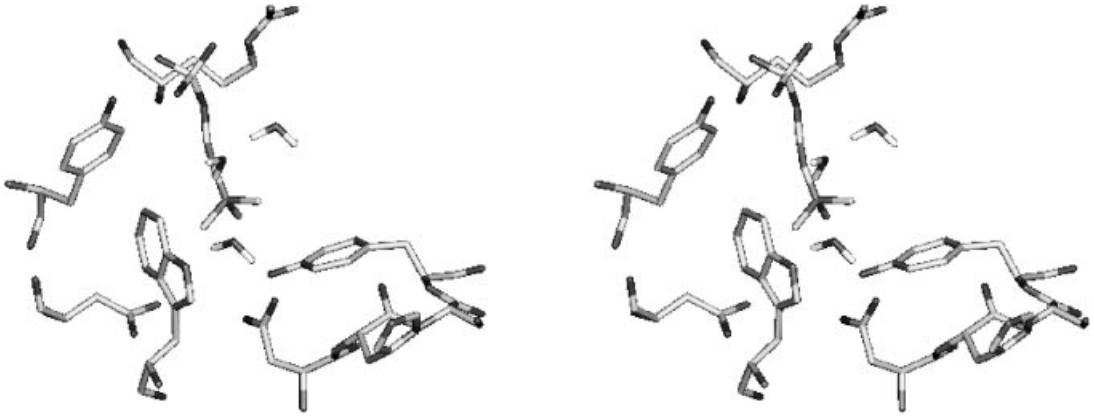

Fig. 11. Stereoscopic view of the binding site after 123 ps simulation, using MDD with half the weight of the ligand–solvent interactions.

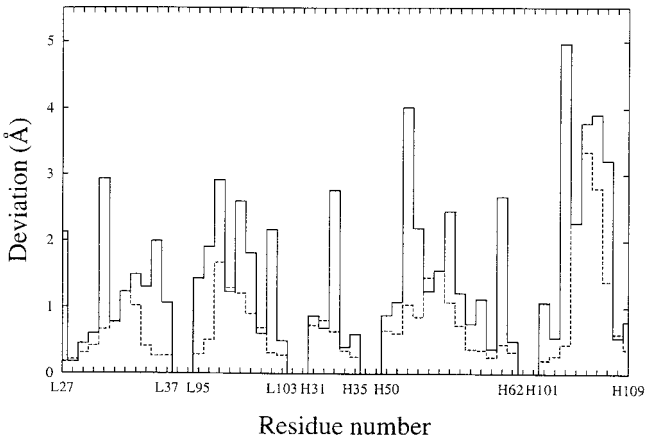

Fig. 12. Deviation between the crystal structure and the structure at t = 130 ps of the simulation, with half the weight of the ligand–solvent interaction for the backbone (dashed line) and for the side chains (solid line).

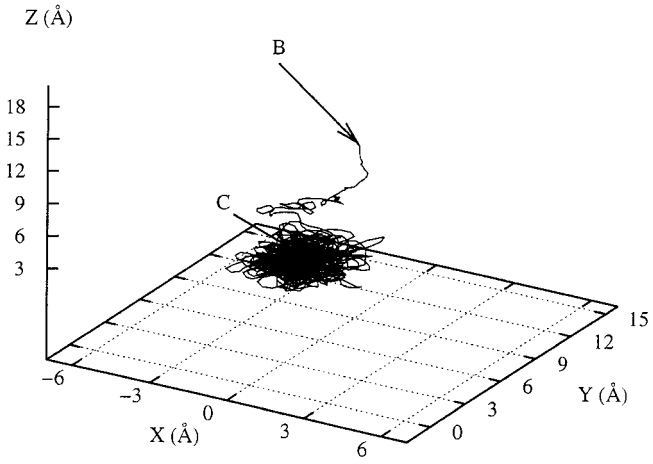

Fig. 13. Trajectory of the center of mass of PC during the 180 ps simulation, starting from position B. MDD simulation with double the weight of the ligand–receptor interactions. C is the position of PC c.o.m. in the crystal complex.

This can be ascribed to the different weight used for the ligand–receptor interactions. At the end of the simulation there is also an increment of the CH2 – C_ζ of Tyr L100 distance, which is attributable to a rotation of the tyrosine side chain that let the aromatic ring out of the docking site. The presence of the hydroxyl group also makes it favorably interact with the water molecules outside the docking site. This is consistent with data showing a relatively smaller contribution to the binding of the tyrosine L100 hydrophobic contacts with PC choline group and aliphatic PC chain.[34]

Figs. 4, 9, and 13 show that in, simulations using the MDD algorithm, the volume explored, when the ligand is docked into the binding site, is significantly larger than the volume explored in a MD simulation without the MDD algorithm. Analogous results were obtained starting at different positions. Two out of three simulations were successful. In the unsuccessful simulation, a rotation of the tyrosine L100 aromatic ring brought the hydroxyl group into a favorable position for the interaction with the aspartic acid L97 at the bottom of the binding site, thereby denying access to the PC at the crystal position. Finally, we have performed MDD simulations starting from the structure of the complex. The rmsd of the six monitored distances (not shown) are comparable to those of the corresponding full simulations, of Figs. 10 and 14, respectively.

To summarize, we have used four thermal baths: for the protein, the solvent, the internal motions of the ligand, and the c.o.m. motion of the ligand. The temperature of the c.o.m. motion of the ligand was set at T = 1,500 K, and the remaining baths were at T = 300 K. The time constants of the coupling to the baths were chosen to be close to the time step of the integrator (2–6 fs), in order to control the heat flow between different degrees of freedom. To speed up the search, the ligand–solvent interactions must be halved or the ligand–protein interactions doubled. The CPU time required for a 100 ps simulation on a Silicon Graphics Power Challenger R10000 was 6.5 hr.

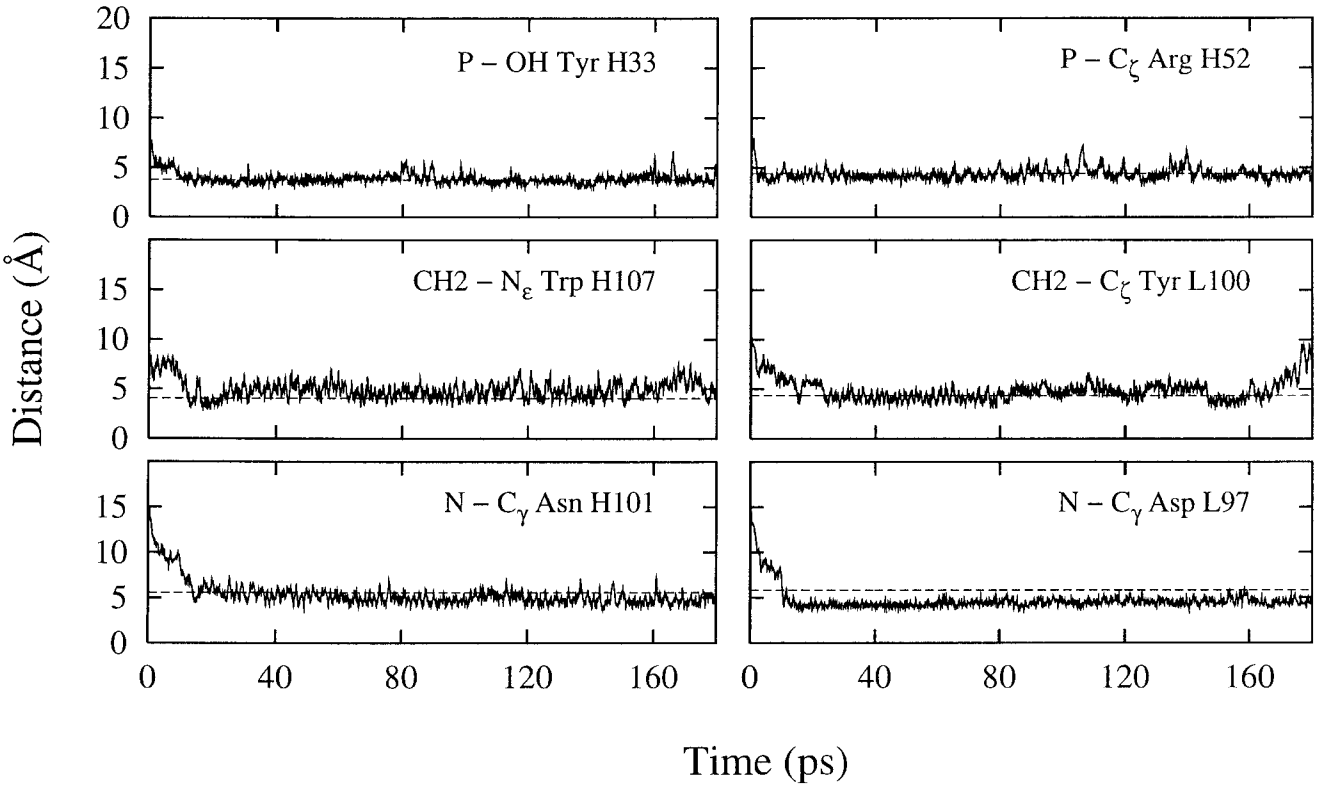

Fig. 14. Trajectories of the six ligand–receptor distances reported in Table II (solid line). MDD simulation with double weight of the ligand–receptor interactions. The dashed lines correspond to the values in the crystal structure.

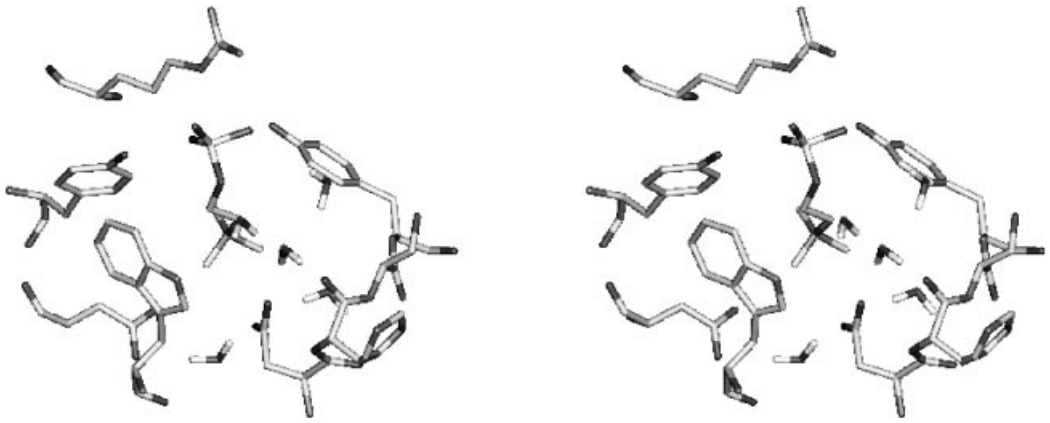

Fig. 15. Stereoscopic view of the binding site after 150 ps simulation, using MDD with double the weight of the ligand–receptor interactions.

CONCLUSIONS

In the present study, we have shown that molecular dynamics simulations can be successfully used in docking problems involving small flexible ligands and flexible receptors in water. This is the first time that the flexibility of the receptor and the presence of explicit water molecules have been taken into account. We showed that the MDD algorithm makes the search of the binding site sufficiently fast, even in the presence of the solvent and of the full flexibility of the binding pocket. For an efficient search it was necessary to use different weights for the ligand–water or ligand–receptor interactions. In both cases, the results agreed with the crystal structure of the complex and with our reference simulation. The main advantage of the MDD algorithm with respect to other algorithms is that it allows one to introduce "naturally" the flexibility of the receptor and the presence of water molecules. Moreover, it avoids the trapping into local minima. On the other

hand, it requires a relatively long computational time compared to other methods. As a consequence, it should be preferred when the presence of explicit solvent molecules cannot be neglected. In other cases, it can be used after rapid scanning of different ligands with simplified methods.

Another characteristic of the MDD algorithm is that, once the ligand has reached the docking site, the algorithm allows a sampling of the local conformational space, i.e., it allows local structural transitions of both ligand and receptor. Due to this characteristic, it can be used when the structural effects of mutations on both ligand and receptor have to be studied. In this last case, the simulation can directly start at the known structure of the complex. In conclusion, the MDD algorithm can be successfully used in addition to simplified methods that allow for rapid scanning of structural databases, and it is useful for obtaining detailed information on the binding mechanism.

ACKNOWLEDGMENTS

Support is acknowledged from the European Community (TMR program "Protein folding" ERBFMRXRX), and from Italian Ministero dell'Università e della Ricerca Scientifica e Tecnologica MURST (National project "Structural Biology"). M. M. acknowledges support from CASPUR-Italy.

REFERENCES

1. Katchalski-Katzir E, Shariv I, Eisenstein M, Friesem A, Aflalo C, Vakser IA. Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. Proc Natl Acad Sci USA 1992;89:2195–21999.
2. Walls PH, Sternberg MJE. New algorithm to model protein–protein recognition based on surface complementarity: application to antibody–antigen docking. J Mol Biol 1992;228:277–297.
3. Wang H. Grid-Search Molecular accessible surface algorithm for solving the protein docking problem. J Comput Chem 1991;12:746–750.
4. Shoichet BK, Kuntz ID. Protein Docking and Complementarity. J Mol Biol 1991;221:327–346.
5. Horton N, Lewis M. Calculation of the free energy of association for protein complexes. Protein Sci 1992;1:169–181.
6. Krystek S, Stouch T, Novotny J. Affinity and specificity of serine endopeptidase–protein inhibitor interactions: empirical free energy calculations based on x-ray crystallographic structures. J Mol Biol 1993;234:661–679.
7. Goodford PJ. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. J Med Chem 1985;28:849–875.
8. Goodsell DS, Olson AJ. Automated docking of substrates to proteins by simulated annealing. Proteins 1990;8:195–202.
9. Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE. A geometric approach to macromolecule–ligand interactions. J Mol Biol 1982;161:268–288.
10. Des Jarlais RL, Sheridan RP, Seibel GL, Dixon JS, Kuntz ID, Venkataraghavan R. Using shape complementarity as an initial screen in designing ligands for a receptor binding site of known three-dimensional structure. J Med Chem 1988;31:722–729.
11. Goodsell DS, Garrett M, Olson AJ. Automated docking of flexible ligands: applications of AutoDock. J Mol Recognit 1996;9:1–5.
12. Leach AR, Kuntz ID. Conformational analysis of flexible ligands in macromolecular receptor sites. J Comput Chem 1992;13:730–748.
13. Di Nola A, Roccatano D, Berendsen HJC. Molecular dynamics simulation of the docking of substrates to proteins. Proteins 1994;19:174–182.
14. Mizutani MY, Tomioka N, Itai A. Rational automatic search method for stable docking models of protein and ligand. J Mol Biol 1994;243:310–326.
15. Clark KP, Ajay. Flexible ligand docking without parameter adjustment across four ligand receptor complexes. J Comput Chem 1995;16:1210–1226.
16. Jones G, Willet P, Glen RC. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. J Mol Biol 1995;245:43–53.
17. Luty BA, Wasserman ZR, Stouten PFW, Hodge CN, Zacharias M, McCammon JA. A molecular mechanics/grid method for evaluation of ligand–receptor interactions. J Comput Chem 1995;16:454–464.
18. Oshiro CM, Kuntz ID, Dixon JS. Flexible ligand docking using a genetic algorithm, J Comput Aided Mol Des 1995;9:113–130.
19. Gulukota K, Vajda S, DeLisi C. Peptide docking using dynamic programming. J Comput Chem 1996;17:418–428.
20. Rarey M, Kramer B, Lengauer T, Klebe G. A fast flexible docking method using an incremental construction algorithm. J Mol Biol 1996;261:470–489.
21. Welch W, Ruppert J, Jain AN. Hammerhead: fast, fully automated docking of flexible ligands to protein binding sites. Chem Biol 1996;3:449–462.
22. Koshland Jr DE. The joy and vicissitudes of protein science. Protein Sci 1993;2:1364–1368.
23. Leach AR. Ligand docking to proteins with discrete side-chain flexibility. J Mol Biol 1994;235:345–356.
24. Knegtel RMA, Kuntz ID, Oshiro CM. Molecular docking to ensembles of protein structures. J Mol Biol 1997;266:424–440.
25. Apostolakis J, Plückthun A, Caflisch A. Docking small ligands in flexible binding sites. J Comput Chem 1998;19:21–37.
26. Amadei A, Linssen ABM, Berendsen HJC. Essential dynamics of proteins. Proteins 1993;17:412–425..
27. Berendsen HJC, Postma JPM, van Gunsteren WF, Di Nola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys 1984;81:3684–3690.
28. Simmerling C, Elber R, Zhang J. MOIL-View— a program for visualization of structure and dynamics of biomolecules and STO—a program for computing stochastic paths, In: Pullman A. et al., editors. Modeling of biomolecular structure and mechanisms. Dordrecht, Netherlands: Kluwer; 1995. p 241–265.
29. Segal DM, Padlan EA, Cohen GH, Rudikoff S, Potter M, Davies DR. The three-dimensional structure of a phosphorylcholine-binding mouse immunoglobulin Fab and the nature of the antigen binding site. Proc Natl Acad Sci USA 1974;71:4298–4302.
30. van Gunsteren WF, Berendsen HJC. GROMOS: Groningen molecular simulation package manual. Nijenborgh, Netherlands: BIOMOS; 1987.
31. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comp Phys 1977;23:327–341.
32. Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J. Interaction models for water in relation to protein hydration. In: Pullman, B, editor. Intermolecular forces. Dordrecht, Netherlands: Reidel; 1981. p 331–42.
33. Allen MP, Tildesley DJ. Computer simulation of liquids. Oxford: Clarendon Press; 1987. 385 p.
34. Novotny J, Bruccoleri RE, Saul FA. On the attribution of binding energy in antigen–antibody complexes McPC 603, D1.3, and HyHEL-5. Biochemistry 1989;28:4735–4749.
35. QUANTA 97 Basic Operations. San Diego: Molecular Simulations Inc; 1997.