PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

CO escape from myoglobin with metadynamics simulations

Matteo Ceccarelli,[1]* Roberto Anedda,[2] Mariano Casu,[2] and Paolo Ruggerone[1,3]

[1] CNR-INFM Slacs and Department of Physics, University of Cagliari, SP Monserrato-Sestu Km 0.700, I-09042 Monserrato, Italy

[2] Department of Chemical Sciences, University of Cagliari, SP Monserrato-Sestu Km 0.700, I-09042 Monserrato, Italy

[3] CNR-INFM Democritos Modeling Center, Via Beirut 2-4, I-34014 Trieste, Italy

ABSTRACT

The relatively small size of myoglobin makes it suitable for the investigation of the ligand escape process in respiratory proteins and, in general, an ideal model system for the study of the more general structure-function paradigm. In this work, we use Molecular Dynamics simulations combined with an accelerated algorithm, the metadynamics, to probe the escape of CO from myoglobin. Our approach permits to quantitatively describe the escape process via the reconstruction of the associated free energy surface. Additionally, hints on the involvement of a larger numbers of residues than hitherto assumed in the gating process are extracted from our data.

Proteins 2008; 71:1231–1236.

Key words: molecular dynamics; free energy; metadynamics; ligand escape.

INTRODUCTION

Since its crystallization and structure determination >50 years ago, myoglobin represents one of the most studied proteins and an ideal model system for the investigation of the structure-function paradigm.[1] The main function of myoglobin is to bind small ligands, such as NO, CO, and O_2, to the heme group,[2] buried inside the protein and protected from the aqueous environment.

The structural characteristic of myoglobin is the presence of four internal cavities in its native state,[3] as depicted in Figure 1. Although Kendrew and coworkers were already aware that xenon can bind myoglobin,[4] only after the crystallization of the protein in presence of 7 atm of xenon the occupation of the internal cavities by Xe was clearly pointed out.[3] These cavities, with a radius larger than 1.2 Å, are lined by hydrophobic residues. Described initially as packing defects, these cavities are nowadays recognized to play an important role for the uptake of ligands.[2] Different experimental techniques, such as time resolved crystallography,[5,6] cryocrystallography,[7] spectroscopy,[8] and kinetic competition experiments with xenon[9–11] have highlighted the ligand trapping in the internal cavities of myoglobin. However, a clear picture of the migration path of the ligands from the buried heme active site to the solvent as well as of the role played by cavities is still missing. The analysis of the myoglobin structure suggests two principal escape paths, one through the distal histidine gate, proposed for the first time by Matthews and Perutz,[12] the other via the temporary occupation of the Xe1 site and then probably through the Xe3 region, the largest internal cavity occupied by xenon and the closest to the external surface. Other authors[11] proposed the first scenario valid near room temperature, while the latter is predominant below 250 K. The first mechanism was also observed in the early molecular dynamics (MD) simulations,[13] and then abandoned in favor of a multiple-pathway escape process.[14]

Recent MD simulations offered important information on the partial diffusion of CO inside the myoglobin and in relation to internal motions.[6,15–17] In particular, by using a single long MD simulation (90 ns) a CO migration path after photodissociation was followed from the heme group to the Xe1 cavity,[17] in qualitative agreement with the picture proposed on the basis of photolysis experiments.[5] The same authors suggested that the presence of CO in a region can perturb the protein environment making possible a transition to neighboring cavities. However, these standard MD simulations do not yield neither a direct evidence of the ligand escape nor, consequently, a quantitative picture of the process because of the very long and competing time scales involved. A recent paper has addressed quantitatively the problem of ligand migra-

Grant sponsor: MIUR; Grant numbers: MIUR-PRIN 2006059902003 and MIUR CYBERSAR.
*Correspondence to: Matteo Ceccarelli, Cittadella Universitaria Monserrato, SP Monserrato-Sestu Km 0.700, Monserrato, Italy 09042. E-mail: matteo.ceccarelli@dsf.unica.it
Received 2 July 2007; Revised 13 August 2007; Accepted 11 September 2007
Published online 27 November 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21817

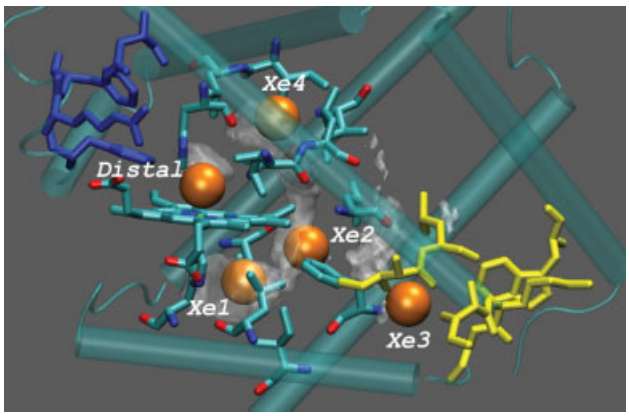

Figure 1

Myoglobin structure with pdbid: 1J52 the distal and the Xenon cavities (in orange). Yellow and blue are the residues of the Xe3 and distal regions, respectively.

tion,[18] and supported the multiple-pathway escape. Despite its computational efficiency—all internal regions of the protein are sampled simultaneously—the method proposed there treats implicitly the presence of the ligand, inferring in some way statically the possible pathway via the evaluation of occupation probability inside the protein.

To possibly compensate "dynamically" this lack of comprehension, we based MD simulations on a non-Markovian algorithm,[19] the so-called metadynamics, which has already demonstrated in several applications to extend successfully MD simulations to biological time scales.[20–22]

On the specific problem of ligand escape, we are aware that also within the framework of metadynamics as well as with any other MD methods, the protein conformational changes at very long times after photolysis, coupled with solvent properties,[23] are impossible to be modelled and prevent any very precise MD description of the process. However, in our work quantitative information on the free energy profile sensed by the ligand is gained and represents an upper and reliable limit of the true energy landscape of the process. Additionally, from our simulations a more subtle involvement of residues in an eventual gating mechanism can be inferred. This might be a good starting point for future studies dedicated to a better characterization of the entire process.

METHODS

We performed all-atom MD simulations with the program Orac[24] using the Amber force field and TIP3 for water.[25–27] The myoglobin (sperm-whale myoglobin: 1BZR at 1.15 Å) was solvated in an orthorombic box of initial side length of 60 Å with 4572 waters in addition to the 206 from the X-ray structure. Following a recent study on water entry in myoglobin[28] no water molecules were introduced in the protein interior. After a slow heating from 10 to 250 K, we performed state-of-the-art MD simulations in the NPT ensemble at 300 K using the SPME (64 grid points and order 5) to treat long-range electrostatic contributions in combination with a multiple-time-steps algorithm.[29] After 9 ns of equilibration in the NPT ensemble we fixed the volume to the average value (length of 59.7343 Å) and we performed additional 600 ps in the NVT ensemble.

Before starting the metadynamics we cut the CO-heme bond, changing the iron coordination from six to five. Contrary to hemoglobin, where the different iron coordination creates a new quaternary structure, in myoglobin the effect is a local distortion of the planar heme moiety. To take into account this relaxation, we adjusted the heme parametrization of Giammona[26] changing the equilibrium values of stretchings and bendings around the iron, while maintaining the same force constants, as reported in Table I. These values were taken from the crystal structure of the heme in the deoxy-myoglobin (pdbid: 1MOA).

We made use of the metadynamics algorithm to accelerate the CO diffusion on the MD time scale. This algorithm performs a sort of dynamical Umbrella Sampling in a high dimensional space for systems where the path spanned by the slow reaction variables is a priori unknown. The evolution of these selected coordinates is accelerated adding at given time interval Δt additional (energy) penalty terms, a bias that prevent the system from visiting conformations already sampled.[19] In simple models, in the limit of a long metadynamics run, the sum of the penalty terms tends to compensate the underlying FES in the reduced space permitting a reconstruction of an approximate FES.[30]

The main advantage of metadynamics with respect to umbrella sampling is that an a-priori-defined range of variation of the reaction coordinates is not required, letting the system evolve toward the lowest transition state, thus obtaining in this particular case the minimum escape path. This prevents the sampling of uninteresting regions and, in principle, it allows the introduction of a

Table I

Heme Structural Parameters

	Ligand bound (coordination 6)	Ligand unbound (coordination 5)
Fe—Ne (His93)	2.08 Å	2.32 Å
N—Fe—N	90°	98°
N—Fe—N	180°	164°

Differences between heme structural parameters for the ligand bound (taken from Ref. 26) and unbound (taken from the crystal structure of deoxy myoglobin, pdbid: 1MOA).

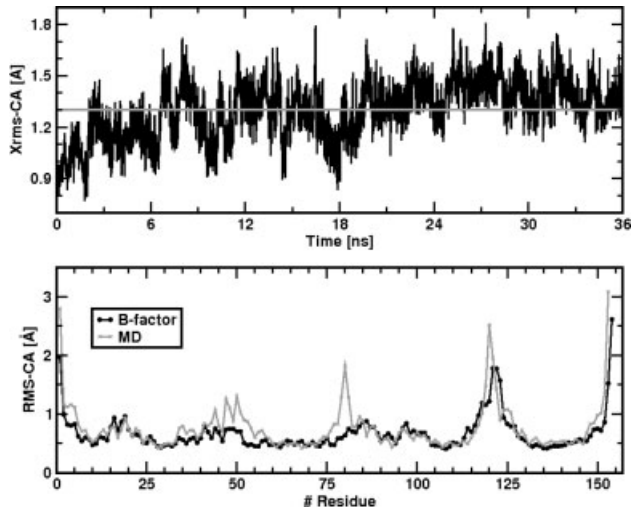

Figure 2

Top: Root mean square deviations of Cα atoms from the X-ray structure as a function of time (black) with the average value (gray). Bottom: fluctuations of Cα atoms (gray) compared with the experimental B-factor (black).

high number of reaction coordinates. However, the use of a time-dependent biasing potential is in some way a nonequilibrium procedure with respect to the slow modes not explicitly included in the selected coordinates. In this case, the choice of the parameters controlling the bias potentials (i.e., deposition time step, height and scale factor) is crucial for an equilibration of the system each time a new term is added. The efficient sampling of non-explicit slow modes within metadynamics can be tackled in different ways, either improving the sampling with the replica exchange method,[31] or correcting the reconstructed FES with a subsequent refining umbrella sampling.[32]

One of the main problems of metadynamics is the choice of the slow reaction variables to accelerate. To describe the CO diffusion process with the metadynamics we chose as coordinates the position of the ligand in a Cartesian system centred at the heme Fe atom. Using the metadynamics with these variables, CO visited also the high energy (or the low probability) regions in the protein interior and eventually induced local rearrangements of the protein, such as dihedral transitions that may be central in hopping from one cavity to the other, as recently proposed.[17] The use of the Cartesian coordinates allowed reconstructing the FES as a three-dimensional map of the occupied regions. The FES can easily be represented using a 3D plot of the isosurface at constant energy.

A scale factor of 0.3 Å for the Cartesian coordinates of the CO molecule was used to reach a resolution in space of 0.5 Å. To account for the relaxation rates on the nanosecond time scale, we added a Gaussian penalty term with a low frequency (4 ps) and a height of only 0.8 k_BT. These metadynamics parameters were selected to have on the reconstructed FES an error of 1.0 k_BT;[30] smaller than the error associated with the empirical force fields adopted in simulations.

RESULTS AND DISCUSSION

Starting with the relaxed structure of myoglobin and after cutting the CO-heme bond, we have performed metadynamics simulations biasing the position of the CO inside the protein. Metadynamics simulations last totally 50 ns. Interestingly, no water molecules diffuse inside the Mb structure, and the displacement of the whole protein with respect to the X-ray structure, using a Cα best-fitting procedure, is on average 1.3 Å in the first 36 ns, as shown in Figure 2 (upper panel). Here (lower panel) we compare also the calculated fluctuations with the experimental B-factor. The agreement is remarkably good, apart two regions, near residue 50 and 80 respectively, where simulations assign larger fluctuations with respect to the reference X-ray system. The amino acids of these two regions are in contact with the solvent, respectively near the distal pocket and the Xe3 cavity. Note that these regions represent the two possible escape zone for ligands proposed in literature.

A first escape is observed after 36 ns of metadynamics. This, however, does not guarantee that the reconstructed FES is well converged. This time has to be intended as a meta time, a measure of how fast we fill the free energy minima, without any correlation with the real escape time. Thus, a series of successive metadynamics runs have been carried out restarting with the ligand inside the protein and maintaining all the previous history-

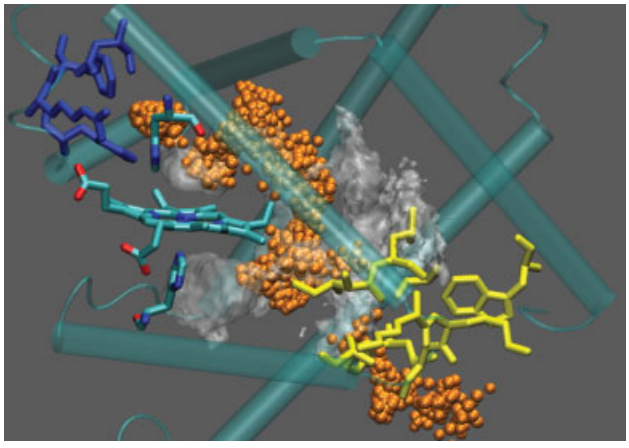

Figure 3

CO escape path from the distal pocket extracted from the standard MD simulation and compared with the isosurface at 13 k_BT.

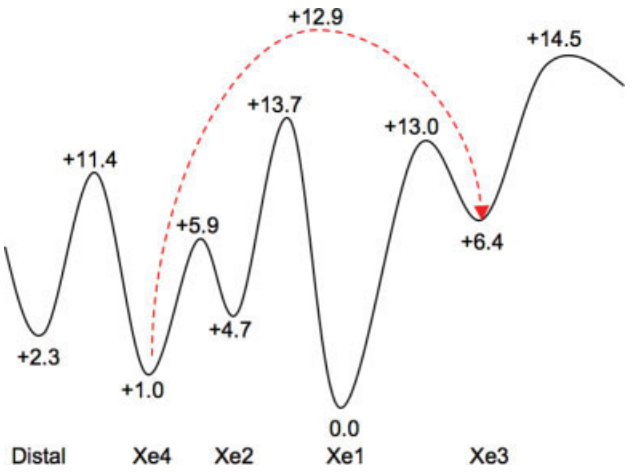

Figure 4

Scheme of minima and transition states between minima with energies in k_BT. The last barrier (14.5 k_BT) separates the interior of the myoglobin from the solution. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

dependent penalty terms. With the CO initially in Xe4 and in the distal cavity, new escapes occur after 12 ns and 2 ns, respectively. All three events (the original one and these latter) take place through the region of Xe3, the largest Xe cavity and also the closest to the external surface. The decreasing meta time of the three escape events, 36-12-2 ns, indicates that the procedure is reaching the convergence, and the penalty terms tend to compensate exactly the underlying FES. However, to confirm the reached diffusion limit of our metadynamics runs we have performed a standard MD simulation locating the CO molecule in the distal pocket and maintaining the penalty terms we added during the previous 50 ns of metadynamics. CO, pushed by the penalty terms, escapes in only 1 ns still through Xe3, proving the diffusive regime of CO and the accurate and reliable reproduction of the underlying FES. In Figure 3, the positions assumed by CO after departure from the distal cavity and on the way to reach the solvent through the Xe3 cavity during this short MD simulation are represented in orange.

Figure 4 pictorially summarizes the reconstructed FES showing the energies of different minima, labelled with the Xenon cavities notation, as well as the transitions between them. Note that, starting from the lowest minimum (Xe1) to the highest (Xe3), CO occupies all Xe cavities before exiting through Xe3. The energy rank of the cavities is Xe1 < Xe4 < Distal < Xe2 < Xe3, with Xe1 and Xe4 differing only by 1 k_BT. In Figure 5, panels A–C, we report the isosurfaces spanned by CO at different free energies, from 4 k_BT to 13 k_BT, together with the residues lining the Xenon cavities: two large regions are recognizable, the first constituted by Xe1, Xe2, Xe4, and the distal cavity, the second by Xe3. These two large regions are connected by two different paths [red arrows in Fig. 5 (A–C)], through either Xe1 or Xe4. The two transition states are almost at the same energy, 13 k_BT, with the energetically more demanding transition between Xe2 and Xe1, 13.7 k_BT. Finally, the energy cost to exit from Xe1 to the solvent is 14.5 k_BT, which, according to transition state theory, corresponds to an escape time of a few microsecond (considering a prefactor of 1 ps^{-1}). This value agrees with the experimentally determined time scale of ligand geminate rebinding, i.e. $\approx$ 1 μs.[11]

Although our simulations have identified a single escape path, at this stage we cannot conclude that the escape through Xe3 is the only existing route. Indeed, we pick out the escape through Xe3 as the energetically most favorable, provided that no low frequency conformational rearrangements of the protein are taken into account. This would be the case at low temperature, as suggested by experiments.[11] The other escape path proposed in the past involves the distal pocket as exiting region.[9] It should be pointed out that the calculated fluctuations of the distal pocket and the Xe3 cavity (the blue and yellow residues in Fig. 3) deviate from the experimental B-factor

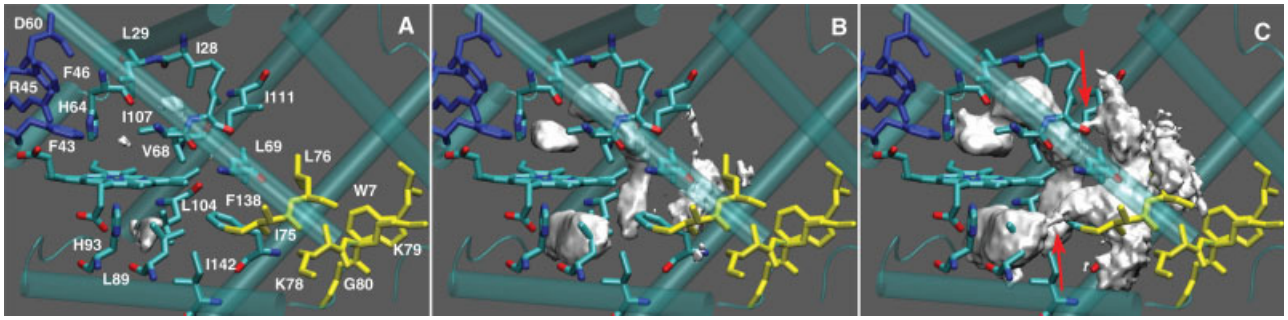

Figure 5

Isosurfaces spanned by CO at increasing free energies, respectively 4 k_BT (A), 10 k_BT (B) and 13 k_BT (C). In blue are highlighted residues of the histidine gate (F43, R45, F46, D60), while in yellow residues lining the Xe3 cavity (W7, I75, L76, K78, K79, G80).

(see Fig. 2). Following these indications, we analyzed the region around the distal pocket and we identified the residue Arg-45 as a very stable gating structure separating the solvent from the distal pocket (the blue residues in Fig. 3). Arg-45 is stabilized by the interaction with two charged groups, Asp-60 and the heme prosthetic group. We performed metadynamics simulations to estimate the energy cost of the gate opening. As variables we used two distances, Arg-45 to Asp-60 (CZ-CG) and Arg-45 to the prosthetic group (CZ-CGD), while on CO coordinates no penalty term is added. The opening process (the above distances increase from 4 to 10 Å) requires an energy of 16 k_BT, but the distal histidine essentially maintains its initial position, suggesting that the conformational rearrangement involves probably more residues than simply Arg-45 and Asp-60. Future simulations will focus on this more specific process.

CONCLUSIONS

We make use of an algorithm to accelerate MD simulations, the metadynamics, to investigate quantitatively and explicitly the escape path of CO in myoglobin. Biasing the CO position inside the protein, our first result suggests that CO escapes through the Xe3 region, the largest internal cavity of myoglobin, with an energy barrier of 14.5 k_BT, and without any large scale motion associated but local fluctuations of single amino acids. Moreover, the time associated to the overcoming of this barrier, following the transition state theory, agrees well with the geminate rebinding at low temperature, where many large fluctuations are hindered. On the other hand, the region of the distal histidine shows large fluctuations with respect to the experimental B-factor, making it a potential alternative escape path as suggested in literature. A first attempt to the opening of this gate gave an energy of 16 k_BT, higher than the escape through Xe3. This indicates that displacements of Arg-45 and Asp-60 alone cannot account properly for the eventual gating process, and that this latter should involve a larger number of residues than hitherto assumed. The resulting process appears more involved with a probably complicated temperature dependence because of its enhanced collective character. The issue requires future investigations and computational simulations.

ACKNOWLEDGMENT

Simulations were performed at the CASPUR computer center, Roma, IT, and Cosmolab Center (CA), IT.

REFERENCES

1. Frauenfelder H, McMahon BH, Austin RH, Chu K, Groves JT. The role of structure, energy landscape, dynamics, and allostery in the enzymatic function of myoglobin. Proc Natl Acad Sci USA 2001;98: 2370–2374.
2. Brunori M, Bourgeois D, Vallone B. The structural dynamics of myoglobin. J Struct Biol 2004;147:223–234.
3. Tilton RF, Kuntz ID, Jr, Petzko GA. Cavities in proteins: structure of a metmyoglobin-xenon complex solved to 1.9 Å. Biochem 1984; 23:2849–2857.
4. Schoenborn BP, Watson HC, Kendrew JC. Binding of xenon to sperm whale myoglobin. Nature 1965;207:28–30.
5. Srajer V, Ren Z, Teng T-Y, Schmidt M, Ursby T, Bourgeois D, Pradervand C, Schildkamp W, Wulff M, Moffat K. Protein conformational relaxation and ligand migration in myoglobin: a nanosecond to millisecond molecular movie from time-resolved Laue X-ray diffraction. Biochem 2001;40:13802–13815.
6. Hummer G, Schotte F, Anfinrud PA. Unveiling functional protein motions with picosecond X-ray crystallography and molecular dynamics simulations. Proc Natl Acad Sci USA 2004;101:15330–15334.
7. Brunori M, Vallone B, Cutruzolla F, Travaglino-Allocatelli C, Berendzen J, Chu K, Sweet RM, Schlichting I. The role of cavities in protein dynamics: crystal structure of a photolytic intermediate of a mutant myoglobin. Proc Natl Acad Sci USA 2000;97:2058–2063.
8. Nienhaus K, Lamb DC, Deng P, Nienhaus GU. The effect of ligand dynamics on heme electronic transition band III in myoglobin. Biophys J 2002;82:1059–1067.
9. Scott EE, Gibson QH, Olson JS. Mapping the pathways for O2 entry into and exit from myoglobin. J Biol Chem 2001;276:5177–5188.
10. Nienhaus K, Deng P, Kriegl JM, Nienhaus GU. Structural dynamics of myoglobin: effect of internal cavities on ligand migration and binding. Biochem 2003;42:9647–9658.
11. Tetreau C, Blouquit Y, Novikov E, Quiniou E, Lavalette D. Competition with xenon elicits ligand migration and escape pathways in myoglobin. Biophys J 2004;86:435–447.
12. Perutz MF, Matthews FS. An X-ray study of azide methaemoglobin. J Mol Biol 1966;21:199–202.
13. Case DA, Karplus M. Dynamics of ligand binding to heme proteins. J Mol Biol 1979;132:343–368.
14. Elber R, Karplus M. Enhanced sampling in molecular dynamics: use of the time-dependent Hartree approximation for a simulation of carbon monoxide diffusion through myoglobin. J Am Chem Soc 1990;112:9161–9175.
15. Vitkup D, Petsko GA, Karplus M. A comparison between molecular dynamics and X-ray results for dissociated CO in myoglobin. Nat Struct Biol 1997;4:202–208.
16. Nutt DR, Meuwly M. CO migration in native and mutant myoglobin: atomistic simulations for the understanding of protein function. Proc Natl Acad Sci USA 2004;101:5998–6002.
17. Bossa C, Anselmi M, Roccatano D, Amadei A, Vallone B, Brunori M, Di Nola A. Extended molecular dynamics simulation of the carbon monoxide migration in sperm whale myoglobin. Biophys J 2004;86:3855–3862.
18. Cohen J, Arkhipov A, Braun R, Schulten K. Imaging the migration pathways for O_2, CO, NO, and Xe inside myoglobin. Biophys J 2006;91:1844–1857.
19. Laio A, Parrinello M. Escaping free energy minima. Proc Natl Acad Sci USA 2002;20:12562–12566.
20. Ceccarelli M, Danelon C, Laio A, Parrinello M. Microscopic mechanism of antibiotics translocation through a porin. Biophys J 2004; 87:58–64.
21. Branduardi D, Gervasio FL, Cavalli A, Recanatini M, Parrinello M. The role of the peripheral anionic site and cation-π interactions in the ligand penetration of the human ache gorge. J Am Chem Soc 2005;127:9147–9155.
22. Gervasio FL, Parrinello M, Ceccarelli M, Klein ML. Exploring the gating mechanism in the ClC Chloride Channel via metadynamics. J Mol Biol 2006;361:390–398.
23. Fenimore PW, Frauenfelder H, McMahon BH, Parak FG. Slaving: solvent fluctuations dominate protein dynamics and functions. Proc Natl Acad Sci USA 2002;99:16047–16051.

24. Procacci P, Darden T, Paci E, Marchi M. Orac: a molecular dynamics program to simulate complex molecular systems with realistic electrostatic interactions. J Comput Chem 1997;18:1848–1862.
25. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A second generation force field for the simulation of proteins, nucleic acids and organic molecules. J Am Chem Soc 1995;117:5179–5197.
26. Giammona DA. 1984. Ph.D. Thesis, University of California Davis, California.
27. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79:926–935.
28. Goldbeck RA, Bhaskaran S, Ortega C, Mendoza JL, Olson JS, Soman J, Kliger D, Esquerra RM. Water and ligand entry in myoglobin: assessing the speed and extent of heme pocket hydration after CO photodissociation. Proc Natl Acad Sci USA 2006;103:1254–1259.
29. Marchi M, Procacci P. Coordinates scaling and multiple time step algorithms for simulation of solvated proteins in the NPT ensemble. J Chem Phys 1998;109:5194–5202.
30. Laio A, Rodriguez-Fortea A, Gervasio FL, Ceccarelli M, Parrinello M. Assessing the accuracy of metadynamics. J Phys Chem B 2005; 109:6714–6721.
31. Bussi G, Gervasio FL, Laio A, Parrinello M. Free-energy landscape for β-hairpin folding from combined parallel tempering and metadynamics. J Am Chem Soc 2006;128:13435–13441.
32. Babin V, Roland C, Darden TA, Sagui C. The free energy landscape of small peptides as obtained from metadynamics with umbrella sampling corrections. J Chem Phys 2006;125:204909.