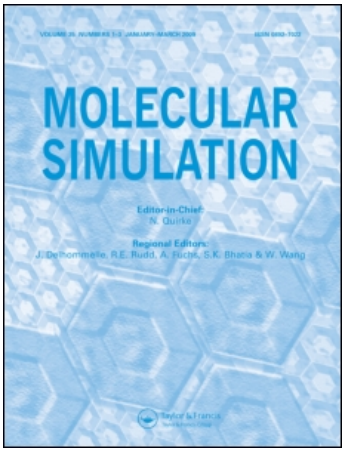

Molecular Simulation

Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Effective Binding Force Calculation in Dimeric Proteins

Luca Maragliano[a]; Mauro Ferrario[a]; Giovanni Ciccotti[b]
[a] INFM National Center on nanoStructure and bioSystems at Surfaces (S 3) and Dipartimento di Fisica, Università degli Studi di Modena e Reggio Emilia, Modena, Italy [b] INFM and Dipartimento di Fisica, Università degli Studi di Roma La Sapienza, Roma, Italy

Online publication date: 26 October 2010

To cite this Article Maragliano, Luca , Ferrario, Mauro and Ciccotti, Giovanni(2004) 'Effective Binding Force Calculation in Dimeric Proteins', Molecular Simulation, 30: 11, 807 — 816
To link to this Article: DOI: 10.1080/0892702042000270205
URL: http://dx.doi.org/10.1080/0892702042000270205

Molecular Simulation, Vol. 30 (11–12), 15 September–15 October 2004, pp. 807–816

Taylor & Francis
Taylor & Francis Group

Effective Binding Force Calculation in Dimeric Proteins

LUCA MARAGLIANO[a], MAURO FERRARIO[a] and GIOVANNI CICCOTTI[b,*]

[a]*INFM National Center on nanoStructure and bioSystems at Surfaces (S^3) and Dipartimento di Fisica, Università degli Studi di Modena e Reggio Emilia, Modena, Italy;* [b]*INFM and Dipartimento di Fisica, Università degli Studi di Roma La Sapienza, Roma, Italy*

(Received September 2003; Accepted January 2004)

We apply the Blue Moon constrained Molecular Dynamics technique to study a particular case of molecular recognition, one of the main issues of modern molecular biology. We investigate the effects of mutation of interface residues on the binding strength of the dimeric protein superoxide dismutase from *Photobacterium leiognathi*. With our technique we produce a specific path describing the separation of the dimers and we calculate the effective mean force involved in the process. We apply the method to two mutants and compare the results with those obtained in an earlier calculation on the native enzyme. The method is sensitive to the mutations and allows us to establish a semi-quantitative hierarchy for the association strengths of the three enzymes.

Keywords: Blue Moon; Dimers; Molecular dynamics; *Photobacterium leiognathi*

INTRODUCTION

A large variety of biochemical processes are regulated by non-covalent interactions between individually folded proteins, usually referred to as protein–protein interactions [1]. They drive fundamental phenomena in biology, such as the formation of hormone–receptor, antibody–antigen or protease–inhibitor complexes. Moreover, they are basic for the formation and stability of quaternary structure in proteins, that is the aggregation of different polypeptide chains in a stable, multimeric complex. The specificity of these interactions is also of great importance. Indeed, if the chemical nature of a ligand in a complex is changed, the obtained structure could have different binding properties.

The ability of two or more molecules to selectively bind to each other forming non-covalent complexes is defined as molecular recognition. In this field, protein–protein recognition in aqueous solution plays a particular role, due to the large molecular surfaces that are desolvated during the association. Direct interactions between the ligands and the removal of water from the contact interface provides the attractive contributions to the free energy of binding that compensate for the loss of translational and rotational entropy in the formation of the stable complex. Hence, the role of the contact interface is crucial. In general, the inter-subunit interface is defined on the base of proximity to the other molecule. A more precise definition relies on the accessibility of the solvent to the residues located at the surface of the polypeptide chain: those residues for which accessibility is lost when going from the isolated structures to the complex are considered part of the inter-subunit interface.

The atomic structure of protein–protein interfaces has been extensively characterized [2,3]. Even if the proteins implicated in these processes are extremely diverse, their recognition sites share common properties. They are generally hydrophobic, and similar in residue composition to the buried regions of the proteins, but they can also resemble the remainder of protein surfaces with a higher percentage of polar residues than protein cores, especially when the complex is formed by pre-folded molecules. Since most of the properties of the complex resides in the contact surface, an efficient tool to characterize the recognition process is to study the effect of interface residue mutations. This approach can elucidate the role of specific residues in the overall stability of the complex [4].

In the context described here, we report on a study by Molecular Dynamics (MD) simulations of the effect of mutations on the association properties of

*Corresponding author. E-mail: giovanni.ciccotti@roma1.infn.it

ISSN 0892-7022 print/ISSN 1029-0435 online © 2004 Taylor & Francis Ltd
DOI: 10.1080/0892702042000270205

a dimeric enzyme, the Cu, Zn superoxide dismutase from *Photobacterium leiognathi* bacterion (PSOD). Cu, Zn superoxide dismutases are a class of homodymeric metallo-enzymes which catalyze the dismutation of superoxide anion into molecular oxygen and hydrogen peroxide, according to the scheme $2O_2^- + 2H^+ \rightarrow H_2O_2 + O_2$, thus representing the first cellular defense against the serious damages produced by oxygen reacting derivatives. Most of Cu, Zn superoxide dismutases are composed of two identical subunits, each one structurally organized in a β-barrel fashion [5], and carrying a Cu^{2+}, Zn^{2+} ion pair. The three-dimensional structure of PSOD has been characterized by Bordo *et al.* [6]. The dimer interface is built by two β-strands and two loops facing each other in the two monomers, involving 18 residues in each subunit. In Fig. 1(a), (b), the three-dimensional structure of PSOD is shown, together with a closer view of the inter-subunit interface.

It has been demonstrated [6] that three distinct clusters of interface residues are crucial in dimer assembly: the first is centered on the side chains of Trp83 and Phe81, including other residues such as Val29 and Pro106, the second on Met41 and Phe96, and the third is formed by the residues located at the top of the interface, Lys25, Tyr26 and Asp85, which are involved in a hydrogen bonded network. In Fig. 1(c), (d), we show the interface structure of the two mutated PSODs studied here. The first presents a single mutation (SM-SOD), the substitution—in each monomer—of Val29 residue in a glicine (Gly). Due to its lack of side chain structure this residue is generally used in interface mutagenesis experiments. Indeed, it is likely to preserve the main chain structure of the molecule, removing all the side chain interactions that can be involved in the dimer assembly. The second mutated enzyme presents a double mutation (DM-SOD), being both Val29 and Met4l substituted—in each monomer—with the negative charged residue glutamic acid (Glu). The insertion of polar residues at interfaces is also a generally used technique to investigate protein association, due to the high degree of perturbation they can produce on the electric field of the molecule. The effect of such mutations are not always easily predictable [4].

The dimerization process of a molecule can be represented via the separation of the monomers.

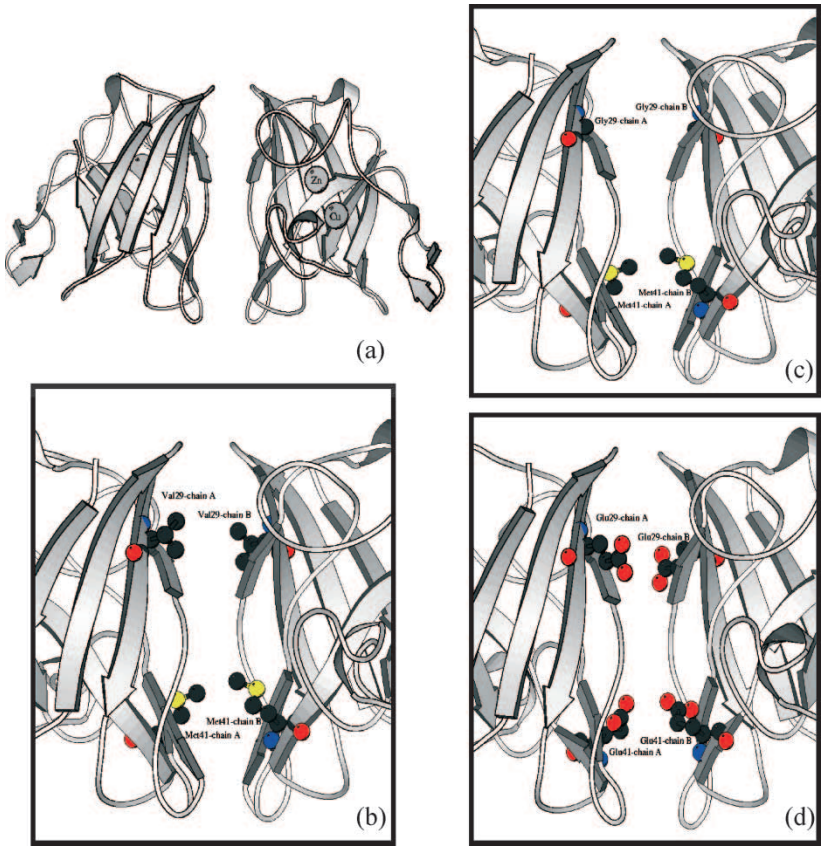

FIGURE 1 Pictorial view of the secondary structure of *P. leiognahi* superoxide dismutase and the two mutants. Arrows represent β-strand structures and thin wires coil segments. (a) Dimeric structure of the native protein. (b) The interface region of the wild-type enzyme, with the mutated residues in ball and stick representation: C atoms are in gray, O atoms in red, S atoms in yellow and N atoms in blue. (c) The interface region of SM-SOD. (d) The interface region of DM-SOD.

Constrained MD simulations can be an efficient tool to build a path that describes the process, along which the relative effective force can be calculated. In the next section, we summarize our procedure to calculate effective forces with constrained MD averages (for the full derivation we refer the reader to Ref. [7]) and we describe the path used to represent the dimerization process for the two mutants of PSOD. In the following sections, we give modeling and computational details of the calculations, and finally we report our results.

THEORY

Blue Moon Formulae

We performed free energy calculations with the constrained MD technique that was derived and first applied to the wild-type PSOD in Ref. [7]. This technique is a generalization of the Blue Moon Ensemble method [8,9] for the case of a multi-dimensional reaction coordinate. We describe the rupture process of the protein in terms of the separation of the monomers by producing a specific representative path and along it we calculate the effective force involved in the process. In principle, the separation could be achieved by constraining, as a reaction coordinate, the relative scalar distance between the monomers while sampling all the other degrees of freedom of the system, including the slow rotational motion of each monomer. However, due to the size and flexibility of the molecule, this requires extremely long trajectories both to equilibrate the system and to collect converged averages. Hence, we adopt a multi-dimensional coordinate space in which both the monomer–monomer separation and their relative orientation can be represented. In this space, we identify a reaction path that separates the monomers at a fixed relative orientation. As a consequence, we do not calculate the full free energy of the association process, but only the contribution coming from the association at a fixed given orientation, the one corresponding to the relative monomer orientations found in the minimized contact configuration.

The vectorial coordinate ξ_α with $\alpha = 1, \ldots, n$, is defined by n functions of the Cartesian coordinates of the atoms in the system $\{\mathbf{r}^N\}$. The free energy variation when going from a state a to a state b can be calculated—independent of the specific path chosen—via the n-dimensional line integral

$$W(\xi_1^b, \ldots, \xi_n^b) - W(\xi_1^a, \ldots, \xi_n^a) = \sum_{\alpha=1}^{n} \int_{\mathscr{C}(\xi_a, \xi_b)} \mathrm{d}\xi'_\alpha \frac{\partial W}{\partial \xi'_\alpha} \tag{1}$$

Using the properties of the delta function [7] it is possible to show that the integrand can be expressed in terms of the multi-dimensional mean force F_{ξ_α} associated to ξ_α,

$$\frac{\partial W}{\partial \xi'_\alpha} = -F_{\xi_\alpha} = \left\langle \frac{\partial \mathscr{H}}{\partial \xi_\alpha} \right\rangle_{\xi'}^{\text{cond.}} \tag{2}$$

where we have introduced the canonical ensemble average conditioned to $\xi_\alpha = \xi'_\alpha$

$$\langle \ldots \rangle_{\xi'}^{\text{cond.}} = \frac{\left\langle \ldots \Pi_{\alpha=1}^{N} \delta(\xi_\alpha(r) - \xi'_\alpha) \right\rangle}{\left\langle \Pi_{\alpha=1}^{N} \delta(\xi_\alpha(r) - \xi'_\alpha) \right\rangle}. \tag{3}$$

Hence $W(\xi_1^a, \ldots, \xi_n^a)$ can be seen as a potential of mean force which can be calculated with the aid of Eq. (1) in terms of free energy differences by means of a thermodynamic integration along a specific path. Within the Blue Moon Ensemble method conditional averages, like those in Eq. (2), are calculated via a time average over a constrained trajectory, where the reaction coordinate is fixed at a specific value ξ'. The result for a scalar reaction coordinate was first derived in Ref. [8] and later extended to a vectorial coordinate in Ref. [7]. Indicating with $\langle \ldots \rangle_\xi$ the average over the constrained trajectory, when there is no mixing between molecular constraints and the reaction coordinates, the required expression for Eq. (2) is

$$\frac{\partial W}{\partial \xi'_\alpha} = \frac{\left\langle |\Xi|^{-1/2} [k_B T G_\alpha - \lambda_\alpha] \right\rangle_{\xi'}}{\left\langle |\Xi|^{-1/2} \right\rangle_{\xi'}} \tag{4}$$

where λ_α is the Lagrange multiplier associated with the α component of the constrained multi-dimensional coordinate, $\mathbf{G}$ is a vector defined by

$$G_\alpha = \sum_{i,j=1}^{n} \frac{1}{m_i m_j} \sum_{\mu\sigma\nu} \Xi^{-1}_{\mu\alpha} \frac{\partial \xi_\mu}{\partial \mathbf{r}_i} \frac{\partial^2 \xi_\sigma}{\partial \mathbf{r}_i \partial \mathbf{r}_j} \frac{\partial \xi_\nu}{\partial \mathbf{r}_j} \Xi^{-1}_{\sigma\nu} \qquad \alpha = 1, \ldots, n \tag{5}$$

and the elements of the matrix $\boldsymbol{\Xi}$ are given by

$$\Xi_{\alpha\beta} = \sum_k \frac{1}{m_k} \frac{\partial \xi_\alpha}{\partial \mathbf{r}_k} \frac{\partial \xi_\beta}{\partial \mathbf{r}_k}, \tag{6}$$

with $\alpha, \beta = 1, \ldots, n$. Equation (4) provides us with an explicit expression to compute the effective mean force and, by means of Eq. (1), the reversible work needed to go from state a to state b.

Building the Separation Path

The dimerization process of a molecule can be represented via the separation of the monomers. Constrained MD simulations can be an efficient tool to build a path that describes the process, and to

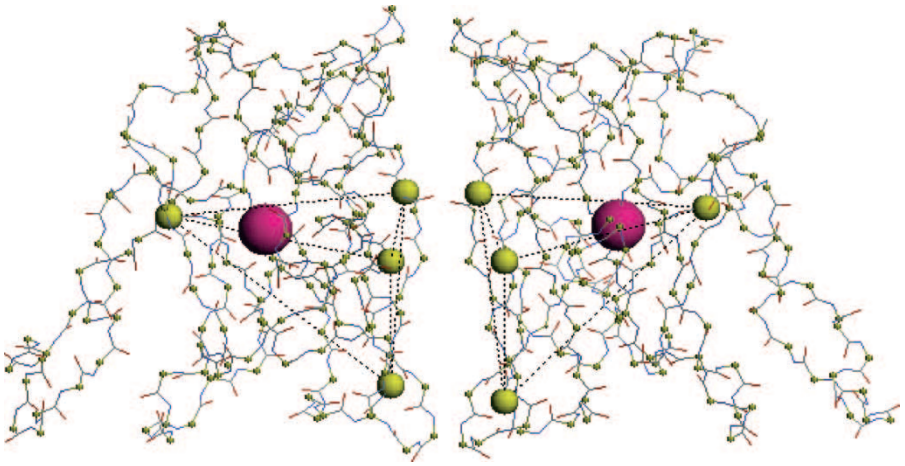

FIGURE 2 Pictorial view of the backbone structure of SM-SOD at atomic level. C_α atoms are depicted as small yellow spheres, all other atoms as rods. Large yellow atoms define the two tetrahedrons described in the text, while the largest spheres represent the centers of mass of the two sets of atoms.

calculate the effective force involved in it. In this section, we describe the method used to represent the dimerization process of the two mutants of PSOD. For a complete description of the separation of the two monomers, we should take into account relative translational and rotational degrees of motion. Due to the large size of the monomers, the latter are expected to evolve on a much longer time scale, so that, in order to sample them correctly, we would have to perform extremely long, time-consuming, simulations. However, the overall three-dimensional structure of the three proteins is nearly identical and, as a consequence, the difference between the translational and rotational entropy for the associated and for the separated monomers will be practically the same for all three proteins. Hence, as we are interested in defining the hierarchy in strength, i.e. in the relative free energy of association, we may at first ignore such contributions and run our calculations at fixed relative orientations.

To this end, the definition of a multi-dimensional coordinate is needed. We can describe the separation and the relative orientation of the two molecules by attaching a reference frame to each monomer. In order to define such a frame we choose a set of four atoms on each monomer,

$$\{\mathbf{r}_i^{(1)}, \quad i = 1,4\} \quad \{\mathbf{r}_i^{(2)}, \quad i = 1,4\}. \tag{7}$$

Given a monomer $I(I = 1,2)$, the four atoms $\{\mathbf{r}_i^{(I)}\}$, $(i = 1,\ldots,4)$ define a tetrahedron which behaves roughly as a rigid body. Since we are studying flexible macromolecules, we will choose atoms of the backbone structure, which is the part of the system with the highest rigidity. Using these atoms we build a multi-dimensional coordinate, $\{\xi\}$, with six components: three are used to describe the monomer relative separation and the remaining three the monomer relative orientation. To define the translational part of $\{\xi\}$, let $\mathbf{C}^{(1)} = (1/4)\sum_i \mathbf{r}_i^{(1)}$ and $\mathbf{C}^{(2)} = (1/4)\sum_i \mathbf{r}_i^{(2)}$ be the centers of mass of each of the two tetrahedrons. In Fig. 2, we show the two tetrahedrons within the structure of the SM-SOD, together with the positions of the centers of mass $\mathbf{C}^{(1)}$ and $\mathbf{C}^{(2)}$. We now define the vector $\mathbf{D}$ as

$$\mathbf{D} = \mathbf{C}^{(1)} - \mathbf{C}^{(2)}, \tag{8}$$

and put $\xi_1 \equiv D_x$, $\xi_2 \equiv D_y$, $\xi_3 \equiv D_z$. As for the rotational part of $\{\xi\}$ we define an intrinsic reference frame on each monomer by introducing the two sets of vectors

$$\mathbf{u}_\beta^{(I)} = \mathbf{C}^{(I)} - \frac{m_{\beta+1}^{(I)}\mathbf{r}_{\beta+1}^{(I)}}{\sum_{i=1}^{4} m_i^{(I)}} \quad \beta = 1,2,3 \quad I = 1,2. \tag{9}$$

The vectors $\{\mathbf{u}^{(I)}\}$ are then orthonormalized according to the Graham–Smith procedure, obtaining the set of vectors

$$\begin{aligned} \mathbf{e}_1^{(I)} &= \frac{\mathbf{u}_1^{(I)}}{\sqrt{\mathbf{u}_1^{(I)}\cdot\mathbf{u}_1^{(I)}}} \\ \mathbf{e}_\gamma^{(I)} &= \frac{\mathbf{u}_\gamma^{(I)} - \sum_{\beta=1}^{\gamma-1} \mathbf{e}_\beta^{(I)}\left(\mathbf{e}_\beta^{(I)}\cdot\mathbf{e}_\gamma^{(I)}\right)}{\sqrt{\mathbf{u}_\gamma^{(I)}\cdot\mathbf{u}_\gamma^{(I)} - \sum_{\beta=1}^{\gamma-1} \mathbf{e}_\beta^{(I)}\cdot\mathbf{e}_\gamma^{(I)}}} \end{aligned} \tag{10}$$

with $\gamma = 2,3$ and $I = 1,2$. This gives an orthonormal basis of vectors $\{\mathbf{e}^{(I)}\}$ on each molecule. The monomer relative orientation can be defined by means of the Euler angles, whose expression in terms of the basis vectors are [10]:

$$\begin{aligned} \theta &= \cos^{-1}\left(\mathbf{e}_3^{(1)}\cdot\mathbf{e}_3^{(2)}\right) \\ \psi &= \cos^{-1}\left(\frac{\mathbf{e}_1^{(2)}\cdot\mathbf{e}_3^{(1)}\times\mathbf{e}_3^{(2)}}{\sin\theta}\right) \\ \phi &= \cos^{-1}\left(\frac{\mathbf{e}_1^{(1)}\cdot\mathbf{e}_3^{(1)}\times\mathbf{e}_3^{(2)}}{\sin\theta}\right) \end{aligned} \tag{11}$$

Finally, we take for the six components of the full reaction coordinate:

$$(\xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6) = (D_x, D_y, D_z, \theta, \phi, \psi). \qquad (12)$$

The vectorial coordinate $\{\xi\}$ in Eq. (12) is by construction an explicit function of the coordinates in Eq. (7). This enables us to compute analytically the derivatives in the definition of both the vector $\mathbf{G}$ in Eq. (5) and the factor Ξ defined in Eq. (6), whose explicit dependence on the atomic coordinates is needed in Eq. (4).

In order to get rid of the rotational degrees of freedom (components ξ_α of $\{\xi\}$, with $\alpha = 4, 5, 6$), we have to constrain them to fixed values. Indeed, this is not enough to completely fix the way the two monomers face each other unless we fix the values of ξ_2 and ξ_3, considering the variations of distance along the component ξ_1. This amounts to using a straight line path $\mathscr{C}(\{\xi^a\}, \{\xi^b\})$ to explore the multi-dimensional space of the reaction coordinate. With this choice, the line integral of Eq. (1), which gives the free energy difference between two states of the system, is reduced to a scalar integration. It is important to note here that in this way we do not compute the full free energy variation along the separation of the molecules, but only the contribution between the two states at a given relative orientation. One state roughly corresponds to the minimum at contact distance and the other, at large monomer separation, to the region of flat dependence for the free energy where monomers can be assumed to be isolated. In order to calculate the absolute dissociation energies one should also compute and add the, mainly entropic, different contributions to the free energy coming from allowing overall rotations and orthogonal translations in the two situations, one with associated monomers and the other with, independently, solvated monomers.

Although, the theory has been developed by considering only the constraints $\{\xi\}$, our molecular model contains also "molecular" bond constraints. Recently, the Blue Moon formalism has been fully extended to a system with molecular constraints [11]. The resulting formula for the effective force involves an unbiasing factor whose calculation is very complicated and much too time-consuming for a system with a large number of molecular constraints, as it would involve the calculation of determinants for very large matrices. For the present calculation, we avoid the coupling between the two sets of constraints by replacing the molecular constraints which involve the atoms that enter the definition of $\{\xi\}$ with harmonic springs. This does not change the connectivity in the system, but—as we will see later—it will force us to use a time step smaller than that which could be used with constraints, roughly doubling the computational cost of each mean force calculation.

COMPUTATIONAL METHODS

We have studied by MD two dimeric proteins in water, which are mutated forms of superoxide dismutase from *P. leiognathi*. The atomic coordinates of the starting configuration for the SM-SOD protein were taken from the 2.2 Å resolution crystal structure of reference [12]. In order to prepare the starting configuration for the DM-SOD protein, the Val29 and Met41 residues were substituted by glutamic acids in the crystal structure coordinate file of the wild-type enzyme [6], maintaining, where possible, the original positions for the common atoms in the sidechains. After changing the sidechains, the atomic coordinates of the residues have been checked to remove overlapping contacts with atoms of the other residues and with water. The sidechains of the substituted residues easily accommodate in the dimer interface. The non-polar hydrogen of the proteins have not been explicitly taken into account but, within a united atoms approach, they have been represented by modifying the interaction parameters of the atoms to which they are linked. Polar hydrogens were explicitly added to the atomic coordinates files of the two proteins. Finally, the number of "atoms" of the protein models are 2688 for SM-SOD and 2700 for DM-SOD. All simulations were performed with the DL-protein code [13] modified to include the subroutines performing the Blue Moon mean force calculations. We used the GROMOS force field [14] with the set of parameters denoted by "37c", and water molecules represented by the SPC/E [15].

As explained in the previous section, the force field was modified by substituting a few distance constraints by harmonic springs, i.e. the one used to connect the atoms entering the definition of the reaction coordinate to neighboring atoms in the proteins. Each protein dimer was put in a parallelepiped shaped box (96, 60 and 60 Å) which was filled with water molecules, by using an equilibrated water configuration as a building block. No counter ions were added, but, when needed, the systems were neutralized by means of a uniform charge background. Water molecules whose atoms were less than 1.8 Å from any protein atom were removed. The final number of water molecules in the systems were 9241 for SM-SOD and 10344 for DM-SOD simulations. Apart from the atoms defining the multidimensional coordinates ξ, all covalent bonds in the proteins and in water were represented by constraints that keep bond distances fixed to the proper chemical values. The constraint terms in the equation of motion were integrated implementing the SHAKE algorithm [16]. The four atoms used

for building the multi-dimensional coordinate (i.e. those in the set of Eq. (7)) were, in both monomers, the C_α atoms of residues Gly40, Val107, Ala121 and Asn130.

In all simulations periodic boundary conditions were used [17]. van der Waals interactions were truncated at a cut-off distance of 10 Å, while electrostatic interactions were fully calculated by using the Smooth Particle Mesh Ewald method [18]. The cut-off for the real space sum was again 10 Å with the Ewald convergence parameter set to $0.35\,\text{Å}^{-1}$ for SM-SOD and to $0.36\,\text{Å}^{-1}$ for DM-SOD. Sixth order cubic splines were used for interpolation, with grids of 90 × 64 × 64 points for both proteins. The equations of motion, within the extended dynamics approach for simulation in the NVT and NPT ensembles [19], were integrated with the Velocity Verlet scheme [20,21] with a time step of 0.5 fs. The choice of such a small value was dictated by the need to properly integrate the fast motion connected with the harmonic springs, which were introduced in the model of the protein. Coordinates of all atoms in the system were saved every 100 fs for successive data analysis.

RESULTS AND DISCUSSION

In calculating the effective force involved in the separation of the monomers in solution, the first step was to build a specific path $\mathscr{C}(\{\xi^a\},\{\xi^b\})$ that describes this separation. As we explained earlier, although our path is defined in a multi-dimensional coordinate space, we are interested only in the variation of the ξ_1 variable, keeping the other ξ_α, with $\alpha = 2,\ldots,6$, constrained at fixed values. As a consequence, the definition of these values is very important in the determination of the binding force. In order to find physically meaningful values for the components $\xi_2,\ldots,\xi_6$, we require that their values correspond to those for which the dimer is at the mechanical energy minimum. In the following we describe the procedure we used to build the separation path $\mathscr{C}(\{\xi^a\},\{\xi^b\})$ for the two monomers. Starting from the atomic structures of the two monomers in the associated dimer configurations, we first equilibrated the proteins in water for about 200 ps. Then, we slowly quenched each of the two systems until atomic velocities vanished at zero temperature, while monitoring the behavior of the Euler angles defined in Eq. (11). When the energy minimum was reached the values of the angles stopped oscillating and we could calculate the corresponding values for the coordinate $\{\xi\}$. We took them as our initial values, denoted by $\{\xi^{(0)}\} = \{\xi_1^{(0)},\xi_2^{(0)},\xi_3^{(0)},\xi_4^{(0)},\xi_5^{(0)},\xi_6^{(0)}\}$. In Fig. 3, we plot the values of the Euler angles during the quenching procedure of DM-SOD.

Having defined the initial point $\{\xi^{(0)}\}$, we still have to build the values $\{\xi_1^{(i)},\xi_\gamma^{(0)},\gamma = 2,\ldots,6\}$ that will identify the other points along the separation path. In order to achieve this we first separate the monomers in water treating them as rigid bodies, by freezing all intermolecular degrees of freedom. Starting from the initial $\{\xi^{(0)}\}$, we integrate the protein dynamics with atomic positions frozen relative to the monomer center of mass and with a fixed relative group

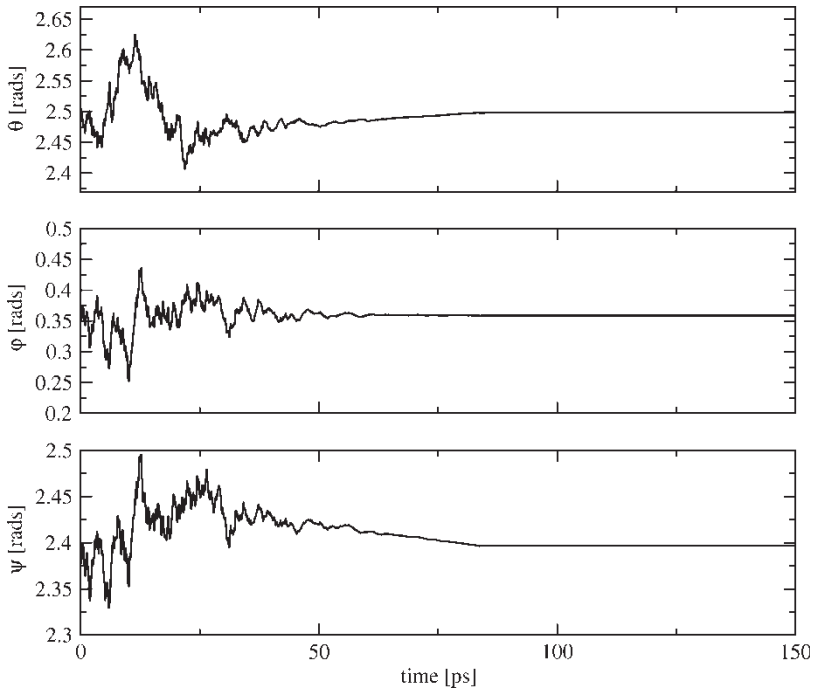

FIGURE 3 Euler's angles calculated as in Eq. (11) during the quenching procedure of DM-SOD.

velocity of 10^{-3} Å/fs along the direction defined by ξ_1. While doing this, we continue to calculate the forces on water molecules and integrate their equations of motion at a fixed temperature by using a Nosè–Hoover thermostat [22]. This allows water to diffuse following the translational motion of the monomers avoiding undesired overlap, which could severely disrupt the structure of the protein. At the same time the solvent can smoothly fill up the region that is being emptied at the interface. Starting from a configuration labeled with the value of the reaction coordinate $\{\xi_1^{(i)}, \xi_\gamma^{(0)}\}$, when a distance increment of 1 Å has been reached, we store the final configuration which now corresponds to the values $\{\xi_1^{(i+1)}, \xi_\gamma^{(0)}\}$. Iterating this procedure, we can produce a number of configurations with the monomers at various distances along ξ_1 and at a fixed relative orientation. The values of ξ_1 in the generated configurations for the two proteins are 17.88 Å, 18.88 Å,. . ., 23.88 Å, 24.88 Å for SM-SOD and 16.78 Å, 17.78 Å,. . ., 25.78 Å, 26.78 Å for DM-SOD.

Starting from each of the configurations obtained with the procedure outlined above, we fully equilibrate the system, first by integrating only the dynamics of water molecules while keeping positions of the protein atoms frozen. By increasing the temperature up to 500 K for a few tens of picoseconds we allow water to hydrate properly the monomer–monomer interface, and then, after returning to the standard temperature, we resume the full atomic motion for the protein while applying Blue Moon constraints to maintain the chosen values of the multi-dimensional coordinate $\{\xi_1^{(i)}, \xi_\gamma^{(0)}\}$ fixed. In this last equilibration step, we let the entire system evolve for at least 100 ps in the NVT ensemble at $T = 300$ K, so as to reach a satisfactory equilibrium state at fixed monomer separation. After repeating this procedure for each of the constraint values $\{\xi_1^{(i)}, \xi_\gamma^{(0)}\}$, the calculation of the effective force defined in Eq. (4) is started.

For each configuration we now accumulate the values of the force $F_{\xi_1} = -\partial W/\partial \xi_1$, using Eq. (4), until convergence is reached, by monitoring the running average calculated as a function of the progressing time τ (Figs. 4 and 5).

$$\bar{F}_{\xi_1}(\tau) = \frac{\int_0^\tau F_{\xi_1}(t')|\Xi|^{-1/2}(t')\,dt'}{\int_0^\tau |\Xi|^{-1/2}(t')\,dt'}. \tag{13}$$

In Figs. 6 and 7, we report the results for the effective force and the mean force potential for two mutants: the free energy barrier for dissociation turns out to be 16.43 ± 1.30 kcal/mol for SM-SOD and 26.05 ± 1.60 kcal/mol for DM-SOD.

In a previous calculation [7], a dissociation barrier of 39.4 ± 2.26 kcal/mol was obtained for the wild-type enzyme. In Fig. 8, we show the behavior of the potential of mean force for the wild-type enzyme and for the two mutants. By comparing the barrier heights for the three proteins, we can say that the most stable dimer corresponds to the wild-type protein, since both mutations induce a decrease in the free energy barrier for dissociation. The SM-SOD case has the lowest value. These results confirm that

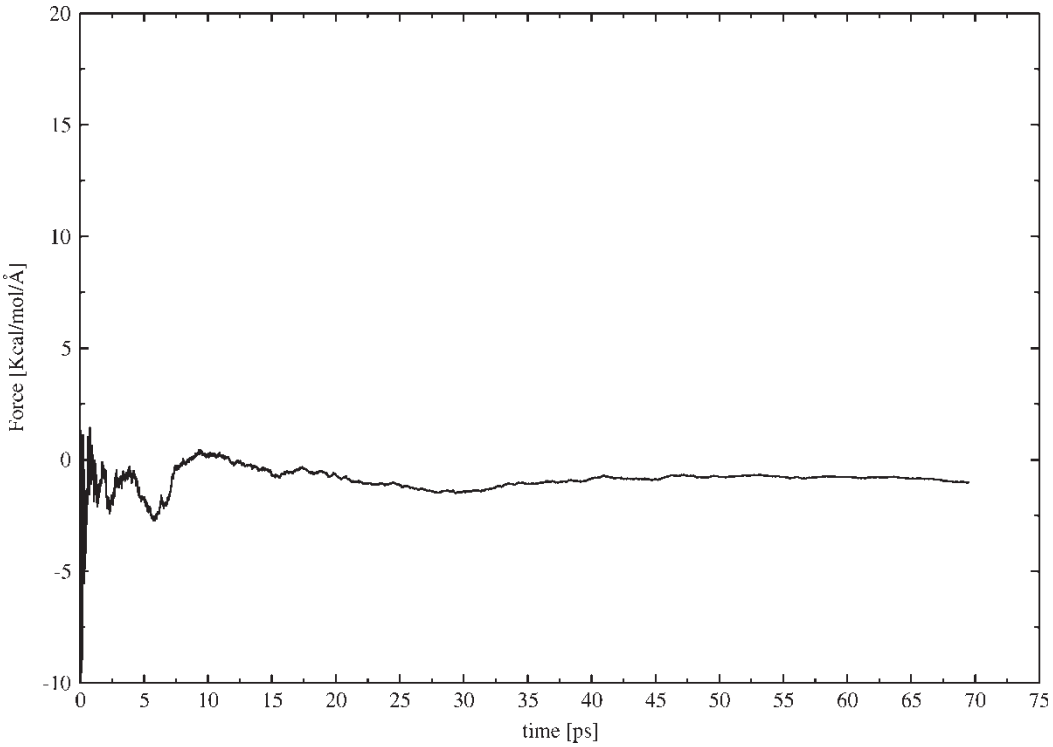

FIGURE 4 Running average of the effective force for SM-SOD at $\xi_1 = 17.88$ Å.

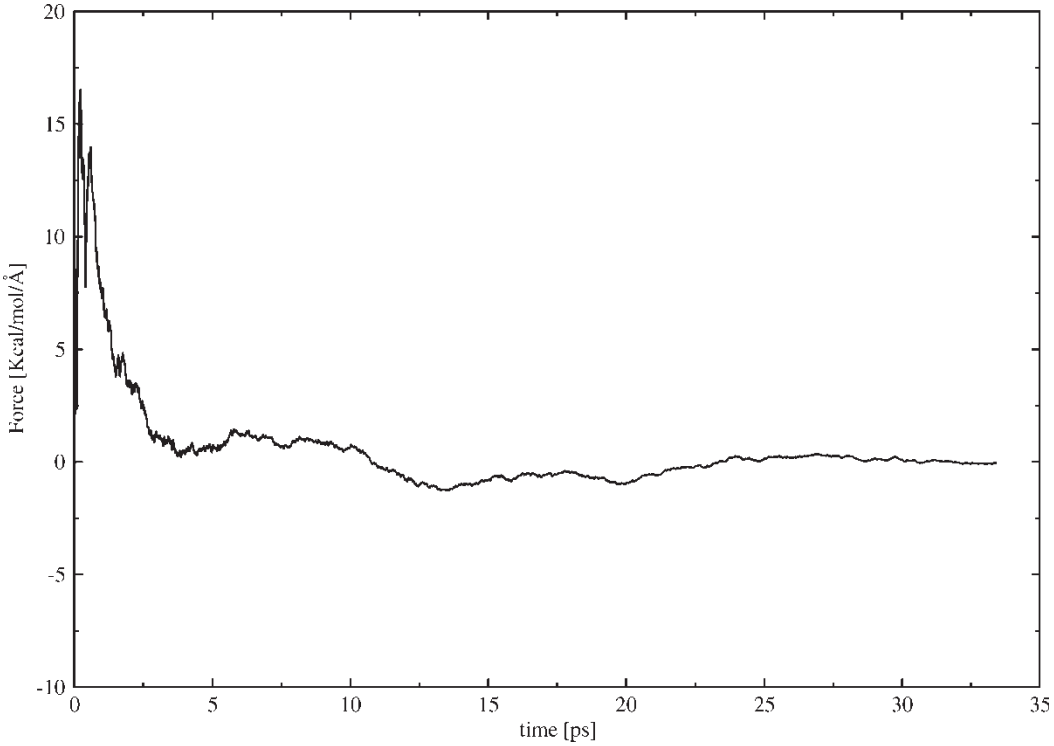

FIGURE 5 Running average of the effective force for DM-SOD at $\xi_1 = 16.78$ Å.

the mutated residues are among those crucial for the recognition process, as stated in previous structural studies [23]. With regards to the Val29 → Gly mutation, it has been observed that it induces a variation of the functional properties of the enzyme, increasing the catalytic rate of a factor 2 [12]. This conclusion, together with our numerical result, suggests that the residue Val29 plays an important role in both functional and association properties of the enzyme.

As for the choice of the reaction coordinate, this is crucial when the free energy barrier between a stable

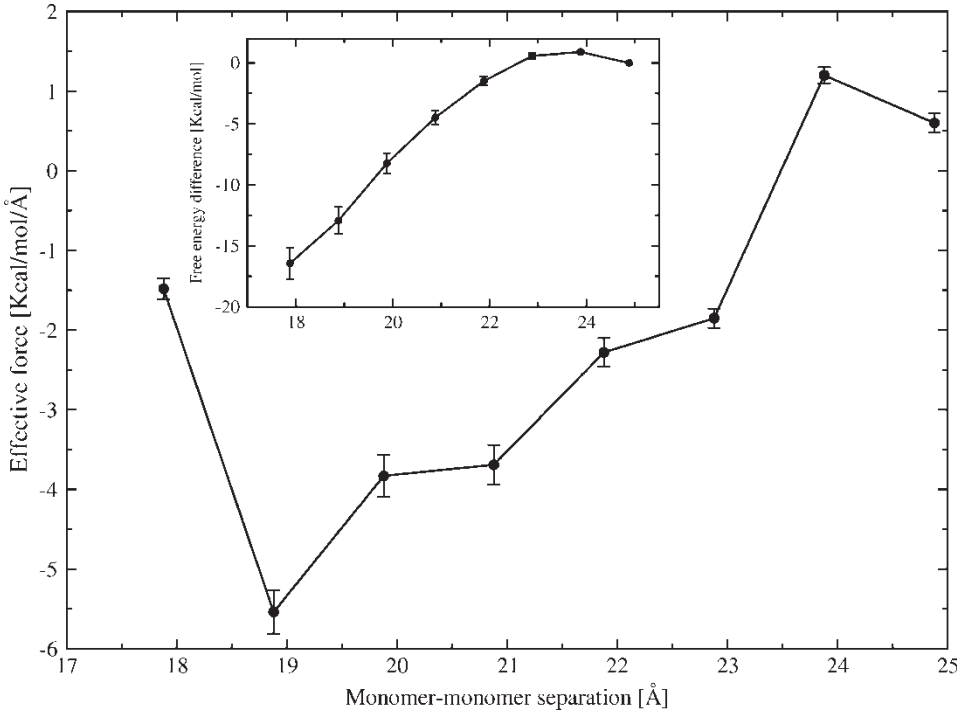

FIGURE 6 Effective force as a function of the monomer–monomer separation in SM-SOD. Inset shows the calculated mean force potential.

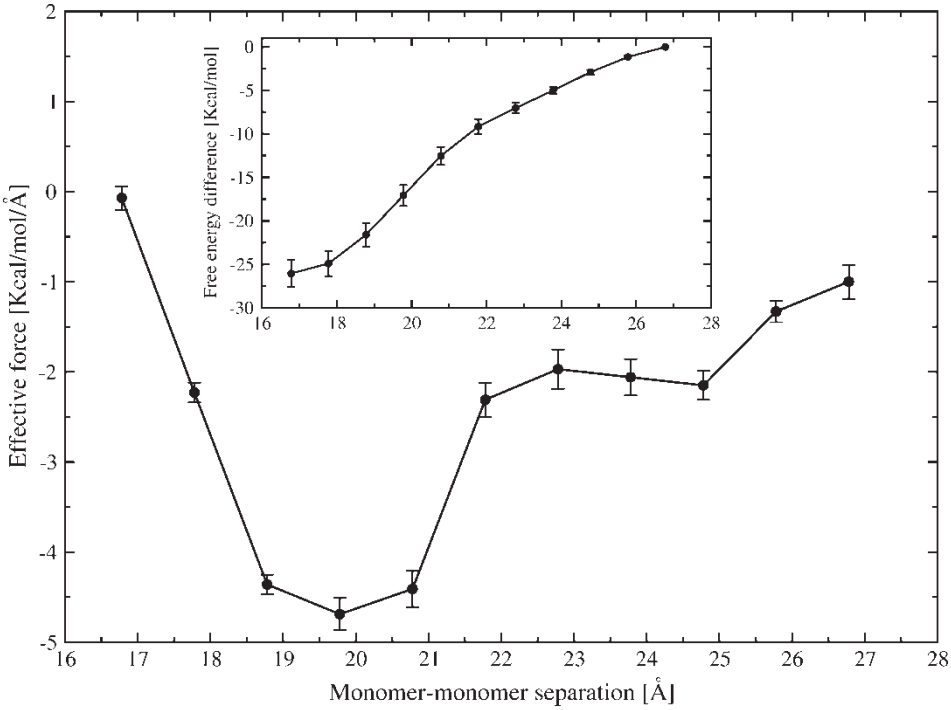

FIGURE 7 Effective force as a function of the monomer–monomer separation in DM-SOD. Inset shows the calculated mean force potential.

and a transition state has to be calculated. In our calculation, we are interested in the difference between two stable states, which is independent of the particular path of integration chosen. However, both the initial and final state have the same fixed relative orientation, so that they differ from the "true" associated and fully dissociated states for the dimers. Indeed, the missing contributions could be separately evaluated with two separate calculations for the free energy difference between our "restricted" starting and final points and the "true" ones. They could be well approximated by a normal mode analysis at constant distance and by using the entropic contribution for free rigid body rotations of the associated dimers and of the isolated monomers. However, thanks to the high homology of the mutants with the wild-type enzyme and the fact that three-dimensional structures are nearly identical, we expect such missing free energy contributions to be almost identical as well. Therefore, the values

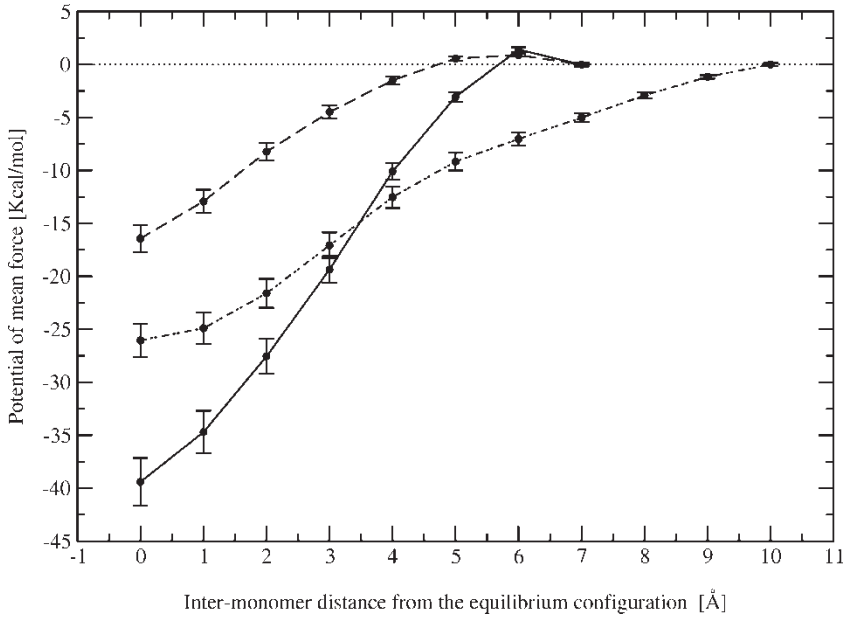

FIGURE 8 Mean force potential curves for the wild-type (full line), SM (dashed line) and DM-SOD (dashed-dotted line).

we calculated along the "restricted dissociation" path contain the information which is most relevant for the comparison of the affinities and provide a well defined answer to the problem of defining the hierarchy in the relative strength of association for the three proteins under investigation.

CONCLUSIONS

In this work, we have applied the multi-dimensional coordinate generalization of the Blue Moon method to investigate a case of molecular recognition, that is the difference in binding strength of two mutated versions of the dimeric enzyme superoxide dismutase from *P. leiognathi*, with respect to the native protein. The mutants were obtained through the substitution of specific amino acids located at the inter-subunit interface of the molecule. The technique we developed enabled us to give an estimate of the order of magnitude for the free energy dissociation barrier of the molecules using a separation path at fixed relative monomer–monomer orientation. Our results showed that the method is able to differentiate between the behavior of the three systems studied, and allowed us to conclude that the native enzyme is the most strongly bound dimer, while the single mutant is the less bound. We have explored mean forces that would tend to take the system off the chosen path and found that they are significantly different from zero only in the immediate vicinity of the minima, reflecting a slight difference between the chosen energy minima that define the fixed relative orientation and the collection of states significant for the corresponding free energy minima. The procedure we have followed does not permit us to estimate the absolute free energy for dimer dissociation, but still gives a well defined hierarchy for the effect of mutations since the neglected contribution coming from allowing full relative motion both at contact and at large monomer separation is, to a very good approximation, the same for all the three enzymes.

Acknowledgements

We are grateful to Alessandro Desideri, Mattia Falconi and Alessandro Sergi for stimulating collaboration on this subject over the years and Ray Kapral for a careful reading of the manuscript.

References

[1] Janin, J. (1996) "Protein–protein recognition", *Prog. Biophys. Mol. Biol.* **64**, 145.
[2] Chakrabarti, P. and Janin, J. (2002) "Dissecting protein–protein recognition sites", *PROTEINS: Struct. Funct. Genet.* **47**, 334.
[3] LoConte, L., Chothia, C. and Janin, J. (1999) "The atomic structure of protein–protein recognition sites", *J. Mol. Biol.* **285**, 2177.
[4] Stitesand, W. (1997) "Protein–protein interactions: interface structure, binding thermodynamics and mutational analysis", *Chem. Rev.* **97**, 1233.
[5] Creighton, T. (1993) *Proteins. Structure and Molecular Properties* (W.H. Freeman and Company, New York).
[6] Bordo, D., Matak, D., Djinovic-Carugo, K., Rosano, C., Pesce, A., Bolognesi, M., Stroppolo, M., Falconi, M., Battistoni, A. and Desideri, A. (1999) "Evolutionary constraints for dimer formation in prokaryotic Cu, Zn superoxide dismutase", *J. Mol. Biol.* **285**, 283.
[7] Sergi, A., Ciccotti, G., Falconi, M., Desideri, A. and Ferrario, M. (2002) "Effective binding force calculation in a dimeric protein by molecular dynamics simulation", *J. Chem. Phys.* **116**, 6329.
[8] Carter, E.A., Ciccotti, G., Hynes, J.T. and Kapral, R. (1989) "A new ensemble for molecular dynamics simulation of rare events", *Chem. Phys. Lett.* **156**, 472.
[9] Sprik, M. and Ciccotti, G. (1998) "Free energy from constrained molecular dynamics", *J. Chem. Phys.* **109**, 7737.
[10] Goldstein, H. (1980) *Classical Mechanics* (Addison-Wesley, Reading, MA.
[11] Coluzza, I., Sprik, M. and Ciccotti, G. (2003) "Constrained reaction coordinate dynamics for systems with constraints", *Mol. Phys.* **101**, 2885–2894.
[12] Stroppolo, M., Pesce, A., Falconi, M., O'Neill, P., Bolognesi, M. and Desideri, A. (2000) "Single mutation at the subunit interface confers extra efficiency to Cu, Zn superoxide dismutase", *FEBS Lett.* **483**, 17.
[13] Melchionna, S. and Cozzini, S. (2000) *DL-protein User Manual*, Tech. Rep., http:/www.sissa.it/cm/DLPROTEIN
[14] Gunsteren, W.V and Berendsen, H.J.C. (1987) *Gromos Manual*, Tech. Rep., University of Groningen, Groningen.
[15] Berendsen, H., Postma, J., Gusteren, W. and Hermans, J. (1983) "The missing term in effective pair potential", *J. Phys. Chem.* **91**, 6269.
[16] Ciccotti, G. and Ryckaert, J. (1986) "Molecular dynamics simulations of rigid molecules", *Comp. Phys. Rep.* **4**, 345.
[17] Frenkel, D. and Smit, B. (2002) *Understanding Molecular Simulations* (Academic Press, New York).
[18] Esmann, U., Perera, L., Berkowitz, M., Darden, T., Lee, H. and Pedersen, L. (1995) "A smooth particle mesh Ewald method", *J. Chem. Phys.* **105**, 8577.
[19] Martyna, G.J., Tuckerman, M.E., Tobias, D.J. and Klein, M.L. (1996) "Explicit reversible integrators for extended systems dynamics", *Mol. Phys.* **87**, 1117.
[20] Andersen, H. (1983) "Rattle: a "velocity" version of the shake algorithm for molecular dynamics calculation", *J. Comput. Phys.* **52**, 24.
[21] Tuckerman, M.E., Berne, B.J. and Martyna, G.J. (1992) "Reversible multiple time scale molecular dynamics", *J. Chem. Phys.* **97**, 1990.
[22] Nosè, S. (1991) "Constant temperature molecular-dynamics methods", *Prog. Theor. Phys. Suppl.* **1**, 103.
[23] Stroppolo, M., Pesce, A., D'Orazio, M., O'Neill, P., Bordo, D., Rosano, C., Milani, M., Battistoni, A., Bolognesi, M. and Desideri, A. (2001) "Single mutations at the subunit interface modulate copper reactivity in *Photobacterium leiognathi* Cu, Zn superoxide dismutase", *J. Mol. Biol.* **308**, 555.