JCTC Journal of Chemical Theory and Computation

Anthrax Lethal Factor Investigated by Molecular Simulations

Rolando Hong, Alessandra Magistrato,* and Paolo Carloni

International School for Advanced Studies (SISSA/ISAS), CNR-INFM-Democritos National Simulation Center, and Italian Institute of Technology (IIT), Trieste, Italy

Abstract: The anthrax disease is caused by the lethal toxin secreted by the bacterium *Bacillus anthracis*. The toxin is a protein aggregate which contains a Zn-based hydrolase called anthrax Lethal Factor (LF). In this work, we investigate the structure of its Michaelis complex with an optimized MAPKK-like substrate using several computational methods including density functional theory, molecular dynamics, and coarse grained techniques. Our calculations suggest that (i) the presence of second-shell ligands is crucial for tuning the structure, energetics, and protonation state of the metal binding site, as found in other Zn-based enzymes; (ii) the nucleophilic agent is a Zn-bound water molecule; (iii) substrate binding to the active site groove is mainly stabilized by van der Waals interactions; (iv) the bonds most likely involved in the substrate hydrolysis are only mildly polarized by the protein scaffold; and (v) part of helix α19, which is present in one solid state structure of LF (PDB: 1JKY), assumes a coiled conformation.

1. Introduction

The anthrax infection caused by the bacterium *Bacillus anthracis* poses a significant threat in biological warfare and terrorism. If ingested or inhaled, the anthrax bacterial spores germinate, resulting in a toxaemia that is usually fatal to the host.[1–4]

Unfortunately, the only way to intervene against anthrax intoxication is to give a generic antibiotic treatment at an early stage of the disease.[5] Thus, there is presently a tremendous effort in investigating the molecular mechanisms responsible for anthrax infection to develop new therapeutic agents.

Most of anthrax's toxic effects are caused by the lethal toxin, a complex consisting of the Protective Antigen (PA) and Lethal Factor (LF) proteins.[6] PA is the membrane-translocating component of the complex; it binds a host cell-surface receptor and translocates LF into the cytosol.[7,10] LF is a cytoplasmatic zinc metalloprotease that cleaves the N-terminal region of selected members of the Mitogen-Activated-Protein-Kinase-Kinase (MAPKK) family;[11] MAP-KKs govern the MAPK signaling pathway, controlling the genomic and physiological response of the cell to its environment.[12] LF alters different cell types, apparently, in an evolutionary conserved manner.[13–16]

X-ray crystallographic studies provided the structural details of LF in the free state (PDB: 1J7N),[17] with a segment of one of its substrates, MAPKK-2 (PDB: 1JKY),[17] with an optimized peptidic substrate (PDB: 1PWV, 1PWW)[18] and with synthetic inhibitors (PDB: 1PWP, 1ZXV).[19,20] [The optimized peptide substrate was built[18] using the consensus residues around the scissible bond based on a peptide library screen, flanked by residues of the actual substrate MAPKK-2.]

Domain **I** of LF binds to PA, while domains **II−IV** create a long groove that holds the MAPKK-2 N-term[18] (Figure 1). Domain **IV** performs the enzymatic catalysis. It features, in the active site, a zinc ion coordinated by two histidines (His686 and His690) and a glutamate (Glu735). The coordination is completed with a water molecule (or an hydroxide group), which is, most probably, the nucleophilic agent in the catalysis. The coordination is that of a distorted tetrahedron. Similar coordination spheres are found in related metalloproteases from the carboxypeptidase and thermolysin families.[21] Clearly the nature and protonation state of the

* Corresponding author phone: +39 040 3787 529; fax: +39 040 3787 528; e-mail: alema@sissa.it. Corresponding author address: SISSA, via Beirut 2-4, 34014 Trieste, Italy.
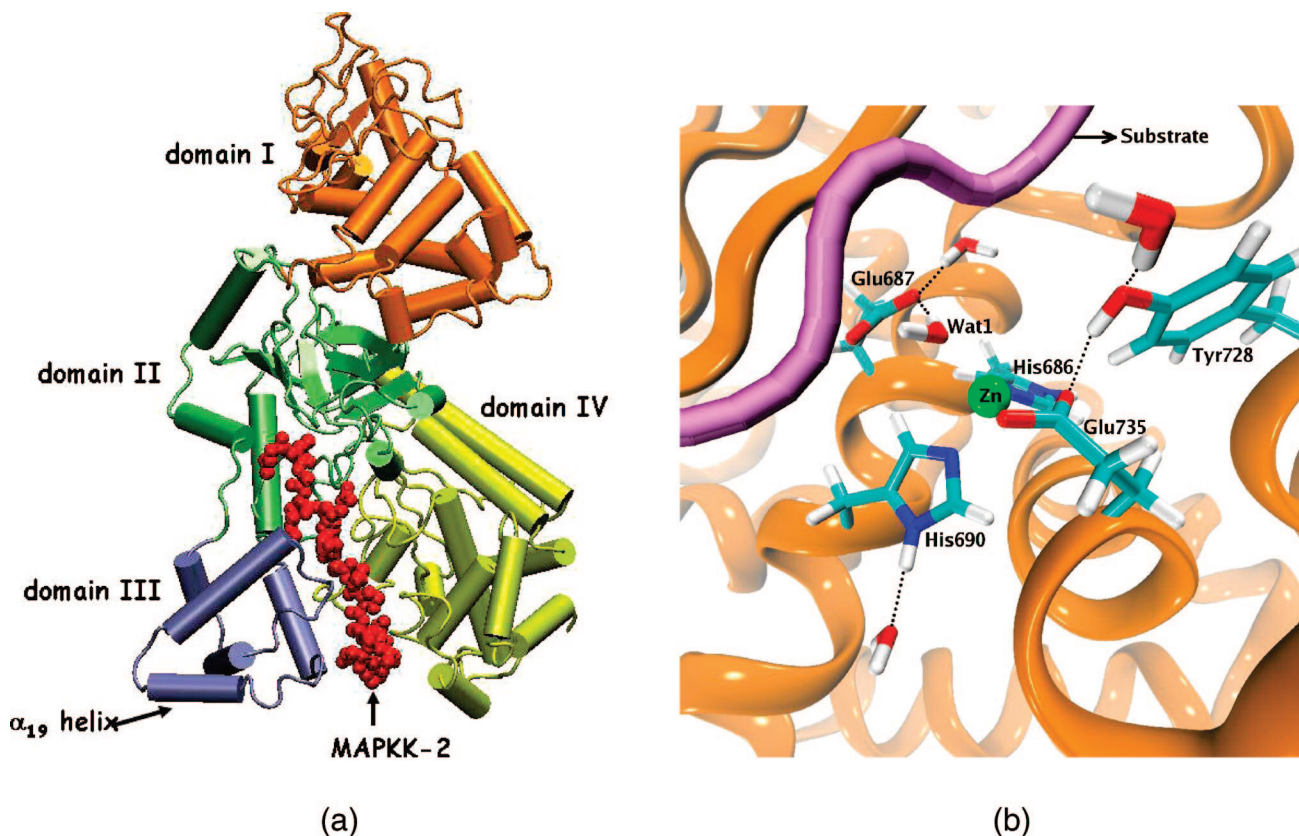
Figure 1. (a) The structure of anthrax Lethal Factor (LF) in complex with its MAPKK-2 substrate (red), as obtained by X-ray crystallography (PDB: 1JKY),[17] includes the following domains: **I** (orange, residues 1−262) is the Protective Antigen (PA) binding domain; **II** (green, residues 263−297, 385−550) is called the Vegetative Insecticidal Protein 2 (VIP2)-like domain because of its similarity with the ADP-ribosyltransferase from *Bacillus cereus* toxin, **III** (blue, residues 303−383) is the helix bundle domain; and **IV** (yellow, residues 552−776) is the catalytic domain. (b) Snapshot from the all-atom MD trajectory featuring the active site (domain **IV**) with the residues invoked to be crucial for the catalytic activity[17,18,25] (shown in licorice representation) and the position of the optimized substrate[18] (shown in purple) used in the simulations.

nucleophilic agent and active site residues involved in the enzymatic process are crucial for the LF proteolytic reaction.

Similarly to what is found in other zinc enzymes,[22] the zinc ligands are stabilized by two outer shells groups,[17,23] including: (i) Glu687, which H-bonds the catalytic water and is believed to play a key role acting as a general base during the reaction.[17] This is consistent with the loss of activity of the LF mutant E687C.[24,25] (ii) Tyr728, which H-bonds Glu735 and whose conservative mutation to phenylalanine (Y728F) impairs the catalytic activity.[23] In addition, Glu739 forms an H-bond with His686 (see PDB: 1JKY).[2] Although no mutagenesis data are available for this residue, this interaction may still play a role because of its closeness to the Zn coordination sphere.

Structural information on the Michaelis complex may help to develop novel peptidomimetic inhibitors with therapeutical properties.

Here we provide a structural model of the complex between LF and an optimized substrate featuring the consensus sequence /VYPYPMEPT/ by using several computational tools. Density Functional Theory (DFT), all-atom Molecular Dynamics (MD) simulations, and mixed DFT-molecular mechanics (DFT/MM) approaches have been employed to study the structural and electronic properties of the Michaelis complex, while calculations based on

Coarse-Grained (CG) models[26−28] and bioinformatics approaches[29−32] have been employed to verify the secondary structure stability of part of the model.

In all known Zn catalytic sites, a solvent species is always a Zn-ligand;[21] however, the nucleophilic agent (e.g., water or OH) can vary in concordance with the Zn environment which can even result in different reaction mechanisms.[33] In agreement with experimental and computational evidences obtained on related zinc metalloproteases,[33b,c] our calculations show that the nucleophilic agent for the LF catalyzed hydrolysis is a water molecule (not an OH group) and that an ionized Glu687 H-bonds the putative catalytic water. However, our study does not give a definitive answer to the LF catalytic mechanism. Substrate binding to the active site is mainly stabilized by van der Waals (vdW) interactions, with Tyr-P1′ and Tyr-P2 constituting the main anchors. The large scale motions of the enzyme do not affect the active site residues, and therefore it is unlikely that they play a mechanical role for the enzymatic reaction. Only a small polarization is exerted by the protein scaffold on the residues involved in the enzymatic reaction, as has been suggested for other proteases.[34] Finally, our results indicate that at least part of helix α19 in domain **III** (Figure 1a) actually assumes a coiled conformation; this is consistent with the unusually large temperature factors reported for this region.[17]

Simulations of Anthrax LF

J. Chem. Theory Comput., Vol. 4, No. 10, 2008 **1747**
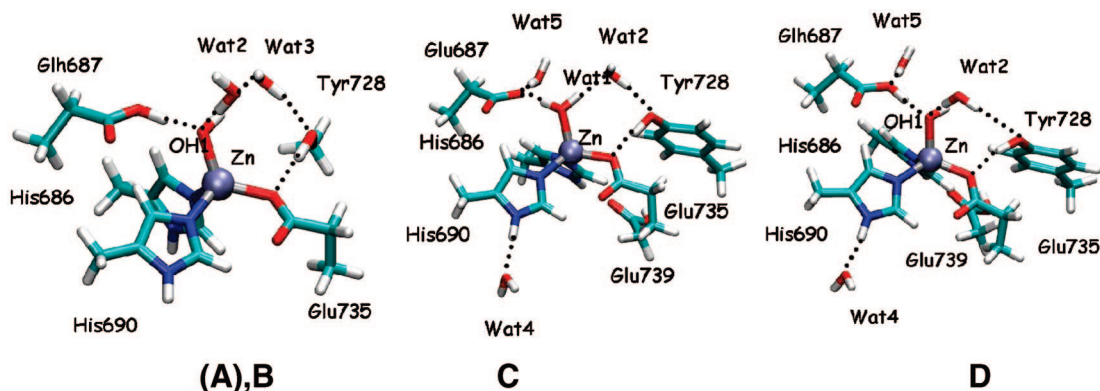

Figure 2. Protomers of LF active site considered in this work: model **A** and **B** (60 atoms) and models **C** and **D** (84 atoms). Model **A** (not shown) was unstable; during geometry optimization it was rapidly transformed into model **B**, bearing a protonated Glu687 (labeled as Glh687 in the figure) and OH1 instead of Wat1. Models **C** and **D** turned out to have the same energetic stability within the accuracy of DFT.

2. Methods

2.1. Construction of the LF Michaelis Complex. The determination of the protonation state of LF active site residues is nontrivial. Here, using the same procedure as that of ref 22b, we perform DFT calculations on a series of models (**A-D**, Figure 2) based on the X-ray structure of the free enzyme (PDB: 1J7N).[17] Models **A** and **B** (60 atoms) included the following: (i) the zinc atom and its ligands: His686 and His690, Glu735 (all cut at $C\beta$ atom, and saturated with hydrogen atoms) and either a water molecule (Wat1, model **A**) or a hydroxide group (OH1, model **B**); (ii) Glu687, either in the ionized state (model **A**) or protonated in $O\epsilon2$ (model **B**); this residue, which forms an H-bond with the nucleophile, was also cut at $C\beta$; (iii) the water molecule Wat2, detected in the X-ray structure, which H-bonds to Wat1; (iv) the water molecule Wat3, which bridges Wat2 and Tyr728 (Figure 2); and (v) Tyr728 (modeled as methanol), which H-bonds to Glu735.

Models **C** and **D** (84 atoms) included the same groups as **A** and **B**, respectively, as well as (i) Glu739 (cut at the $C\gamma$), which H-bonds to His686; (ii) the crystallographic water molecules Wat4 and Wat5, which H-bond to His690 and Glu687, respectively; and (iii) the aromatic ring Tyr728 (cut at $C\beta$), which H-bonds to Wat2 and Glu735.

The DFT calculations were performed using the program CPMD[35] with a plane waves (PW) basis set up to an energy cutoff of 70 Ry. Core/valence interactions were described using norm conserving pseudopotentials of the Martins-Troullier type.[36] Integration of the nonlocal parts of the pseudopotential was obtained via the Kleinman-Bylander scheme[37] for all of the atoms except zinc, for which a Gauss-Hermite numerical integration scheme was used. The gradient corrected Becke exchange functional and the Lee–Yang–Parr correlation functional (BLYP) were used.[38,39] Periodic boundary conditions were applied, and we used orthorhombic cells with edges $a = 16.0$ Å, $b = 19.2$ Å, and $c = 12.8$ Å for models **A** and **B** and $a = b = 17.0$ Å, and $c = 14.5$ Å for models **C** and **D**. Isolated system conditions were applied.[40]

The Michaelis complex was built by reproducing the protonation state resulting from DFT calculations on the X-ray structure of the LF mutant E687C (which is unable to perform catalysis)[18] in complex with an optimized substrate featuring the MAPKK consensus sequence /VYPYPMEPT/ around the scissible bond (PDB: 1PWW).[18] The wild type enzyme was constructed by replacing Cys687 with a glutamic residue and by adding residues 346–367, missing in this X-ray structure, in the same conformation (α-helix) as they were found in the only X-ray structure of LF which provides their positions (PDB: 1JKY). The histidines located outside the LF active site were protonated in $N\pi$, with the exception of His35, His91, His229, His277, His309, and His588, that were protonated in $N\tau$. [The IUPAC-IUB nomenclature was used.[41] $N\pi$ and $N\tau$ correspond to Nd and Ne, respectively, in the PDB atom naming system currently in use.] The two protomers were neutralized by adding 22 potassium counterions immersed in a box of 128.5, 81.3, and 94.0 Å, containing $\sim$26,700 water molecules. The total size of the systems was $\sim$92,600 atoms.

In the MAPKK consensus sequence /VYPYPMEPT/, the conserved reactive proline (i.e., the one placed between two tyrosines) was labeled as P1. Residues on the left side of P1 were labeled as P2, P3... to Pn, while residues on the right side were labeled as P1′, P2′... to Pn′.[18]

The structural characteristics of the Michaelis complex were studied using all-atom MD simulations, coarse-grained methods,[26–28] and bioinformatics tools;[29–32] special attention was given to the structural stability of the reconstructed region α19 (residues 346–367). In addition, we investigated the electrostatic properties of the Michaelis complex using both the Poisson–Boltzmann approach and mixed DFT/MM calculations.

2.2. All-Atom MD Simulations. The AMBER parm98[42] force field was adopted for the substrate, the potassium counterions, and all the enzyme residues with the only exception of the zinc coordination sphere. For the parametrization of the latter we followed the procedure of ref 43 (see the Supporting Information, section 4).

The electrostatic interactions were evaluated using the Particle Mesh Ewald (PME) method.[44] A cutoff of 10 Å was used for the van der Waals interactions and the real part of the electrostatics interactions. The bonds involving hydrogen

1748 *J. Chem. Theory Comput., Vol. 4, No. 10, 2008*

Hong et al.

atoms were kept fixed using the SHAKE algorithm.[45] A time step of 2 fs was used. The initial structures were relaxed by short minimization runs of 2000 steps using the conjugate gradient energy minimization algorithm. 100 ps of MD at constant volume were then performed during which the system was gradually heated to 300 K. Constant temperature (300 K) and pressure (1 atm) production runs were performed by coupling the systems to a Berendsen thermostat and barostat.[46] The NAMD simulation software was used.[47] A trajectory of 50 ns was computed, and the following properties were calculated:

(a) RMSD/RMSF. Root Mean Squared Deviations (RMSD) and Root Mean Squared Fluctuations (RMSF) of the Cα atoms were calculated from the all-atom MD trajectory. The structural stability of the complex during the simulations was monitored by using the RMSD; the RMSF were compared with the temperature factors from the X-ray structure (PDB: 1JKY).[2] The RMSF and the X-ray temperature factors were normalized to compare them (normalized B-values).

(b) Principal Component Analysis (PCA). Large scale motions were calculated as eigenvectors of the covariance matrix of Cα fluctuations, constructed from PCA. The Dynatraj program[48] was used to perform PCA on the last 15 ns of the all-atom MD simulation. For the first three principal components, rigid domains and hinges were identified using the scheme developed by Wriggers and Schulten.[49] Details of this calculation are reported in the Supporting Information, section 6.

(c) Interaction Energies. The interaction energies between LF and the optimized substrate were calculated over the last 15 ns of the all-atom MD simulation. The NAMD Energy plugin (v 1.0) from VMD (v 1.8.6)[50] was used to rerun NAMD[47] on the trajectory to calculate these energies. As the energy values extracted are ultimately dependent on the force field used (AMBER parm98[42]), these calculations are necessarily approximated, and our results are expressed as relative energies (i.e., normalized with respect to the highest interaction energy[51]).

2.3. Hybrid Coarse-Grained/Molecular Mechanics (CG/MM) Simulations. The structural instability of helix α19 (*vide infra*) imposed the employment of a variety of computational techniques to verify our findings obtained from the atomistic simulation.

In this respect in the CG/MM approach we treated helix α19 (residues 346–367) and protein or solvent atoms within 12.5 Å from α19 with an all-atom force field (MM region); the rest of the protein was treated with the Go simplified potential[52] (CG region).[26] The effect of the solvent outside the MM region was considered as the sum of stochastic and frictional forces proportional to the mass parm98[42] and Gromos96 43a1[54] force fields were performed; each of them started from a snapshot taken at 3 ns of the all-atom MD trajectory. These simulations were preceded by 1000 steps of energy minimization (using the steepest descend algorithm) followed by a gentle heating of the systems from 0 K to 300 K in 500 ps. From these simulations, the normalized Cα RMSFs were estimated and compared with the normalized temperature factors from the X-ray structure (PDB: 1JKY).[17] To ensure the reproducibility of these results,

additional CG/MM simulations, starting from different initial structures, were computed (see the Supporting Information, section 5). These structures were selected from the equilibrated (last 15 ns) part of the all-atom MD trajectory (*vide infra*).

2.4. Normal Mode Analysis (NMA). NMA was performed with the NOMAD-ref server[55] on the energy minimized structure taken from the last frame of the all-atom MD simulation. In this scheme,[27,28] the protein was represented by a network of beads connected by harmonic springs; only the interactions between beads separated by a distance ≤3 Å were considered.[28] Normalized Cα RMSF were estimated and compared with normalized temperature factors from the X-ray structure (PDB: 1JKY),[2] after normalization of both terms. To ensure the reproducibility of these results, additional NMA calculations, starting from different initial structures, were performed (see the Supporting Information, section 6).

2.5. Bioinformatics. We investigated the propensity for disorder of helix α19 by using several prediction programs: e.g., PredictProtein,[29] PSIPRED,[30] SPRITZ,[31] and HNN[32] (see the Supporting Information, section 7).

2.6. Electrostatics. *(a) Poisson−Boltzmann Calculations.* Electrostatic surface potentials for LF+substrate adduct were calculated by solving the Poisson−Boltzmann equation with the APBS[56] and PDB2PQR[57] programs; the results were visualized using a PYMOL interface. These calculations were made on the energy minimized structure taken from the last frame of the all-atom MD trajectory featuring the protonation state of model C in the active site (see "Construction of the LF Michaelis complex" in the Methods section).

(b) Polarization of the Active Site. The polarization of selected chemical bonds in the active site was investigated using the so-called Bond Ionicity (BI) indexes[58] that can be estimated from DFT/MM[59] calculations. The region treated at the DFT level comprises the Zn, Glu687, Glu735, His686, and His690 (all cut at Cβ), the putative catalytic water (Wat1), and the scissible region of the optimized substrate which is formed by the backbone atoms from Tyr-P2 and Tyr-P1′ (all cut at Cα) and all the atoms from Pro-P1 (see the Supporting Information, Figure S2). The rest of the system was treated with the AMBER parm98[42] force field. We considered active site residues under different environments: e.g., *in vacuo*, and with the influence of the solvent and/or LF electrostatic fields.[60] To construct the models we used 15 equally spaced frames from the last 15 ns of the all-atom MD trajectory.

BI_{AB} of a bond between atoms A and B is defined as

$$BI_{AB} = \frac{d_A}{d_{AB}} \qquad (1)$$

where d_A is the distance between atom *A* and the Boys orbitals[58] along the *AB* bond, and d_{AB} is the length of the bond between *A* and *B*. A value of BI = 0.5 (the Boys orbital is in the middle of the bond) indicates absence of polarization; while values close to 1 or 0 indicate polarization.

Simulations of Anthrax LF

J. Chem. Theory Comput., Vol. 4, No. 10, 2008 **1749**

3. Results and Discussion

The purpose of this work was the characterization of the structural and electrostatic properties of the Michaelis complex formed by LF and an optimized substrate that features the MAPKK consensus sequence /VYPYPMEPT/.[18] To the best of our knowledge this is the first complete structural model of a Zn-bound LF in complex with a scissible substrate. The first step to achieve our goal was determining the correct protonation state in LF active site residues using DFT calculations.

3.1. Protonation State at the Active Site. As mentioned in the Introduction, a critical issue in Zn-based hydrolases is the determination of the protonation states of residues in the active site.[22] In this work we addressed this issue by performing DFT calculations on increasingly complex models of the active site (Figure 2). The smallest models (**A** and **B**) included only the Zn site, while the largest models (**C** and **D**) included additional second-shell ligands of established (e.g., Tyr728)[23] or putative (e.g., Glu739)[2] relevance for the enzymatic reaction (see the Methods section for details). The most likely protomers were defined as those associated with the lowest potential energy and with the lowest RMSD relative to the reference X-ray structure (i.e., LF in the free state, PDB code: 1J7N).[17]

In the smallest models, **A** featured Glu687 in the ionized state and Wat1 as nucleophile, while **B** exhibited Glu687 in its neutral state and the nucleophile was a hydroxide group. During the geometry optimization model **A** was unstable, as Wat1 transferred a proton to $O\epsilon 2@Glu687$, resulting in model **B** (see Figure 2). The latter was instead stable and featured a slightly distorted tetrahedral coordination geometry and establishing the H-bonds $b(O\epsilon 2@Glu687, O@OH1)$ and $b(O\eta@Tyr728, O\epsilon 2@Glu735)$, also putatively present in the X-ray structure (Table 1). The distances Zn-X (X=coordinating atom), decreased by $\sim 0.1-0.3$ Å relative to the X-ray structure (Table 1); the RMSD between model **B** and the X-ray structure was sizable (~ 0.55 Å); both effects were possibly caused by the limited size of the model.

In the largest models, **C** featured Glu687 in the ionized state and Wat1 as nucleophile (like **A**), while **D** exhibited Glu687 in its neutral state and the nucleophile was a hydroxide group (like **B**). Both models (**C** and **D**) were stable. Some general trends in models **C** and **D** with respect to model **B** could be identified: (i) the bonds $b(Zn, N\tau@His686)$ were shorter ($\Delta d=-0.11$ for **C** and $\Delta d=-0.10$ Å for **D**), (ii) one bond $b(Zn, N\tau@His690)$ was longer in **D** ($\Delta d=+0.03$Å) and unaltered in **C**, (iii) the bonds $b(Zn, O@[OH1,Wat1])$ were larger ($\Delta d=+0.08$Å for **C** and $\Delta d=+0.05$ Å for **D**), and (iv) the hydrogen bonds between Glu687 and the nucleophile $b(O\epsilon 2@Glu687, O@[OH1,Wat1])$ were longer ($\Delta d=+0.09$Å for **C** and $\Delta d=+0.01$Å for **D**). On the other hand, the hydrogen bond between Wat5 and Glu687 (not included in the smallest models) was different between **C** and **D** (Table 1).

The RMSD of **C** is smaller than **D**; however, due to the small relative energies of the two protomers ($\Delta E \sim 2$ kcal/ mol), which were not significant with respect to the accuracy of DFT calculations, we could not establish with certainty which is the most likely protonation state of the active site.[61]

Table 1. Comparison between Experimental (X-ray; PDB: 1J7N)[17] and Calculated (DFT) Structural Parameters (Bond Lengths (Å) and Angles (deg)) for the Models of the Free Enzyme Used in These Calculations (**B-D**, Figure 2)[a]

	X-ray	**B**	**C**	**D**
Bond Length (b(A$_i$,B$_j$); in Å)				
b(Zn, O@[OH1,Wat1])	2.1	1.96	2.04	2.01
b(Zn, Nτ@His690)	2.1	2.04	2.04	2.07
b(Zn, Nτ@His686)	2.3	2.20	2.09	2.10
b(Zn, Oϵ2@Glu735)	2.3	2.00	2.01	2.03
b(Oϵ2@Glu687, O@[OH1,Wat1])	3.6	2.57	2.66	2.58
b(Oη@Tyr728, Oϵ2@Glu735)	2.7	2.70	2.94	2.75
b(Nπ@His690, O@Wat4)	2.9		2.97	2.97
b(Nπ@His686, Oϵ1@Glu739)	2.9		2.93	2.88
b(Oϵ2@Glu687, O@Wat5)	2.9		2.85	2.98
Angles (τ(A$_i$, B$_j$, C$_k$), in deg)				
τ(Oϵ1@Glu735, Zn, O@[Wat1, OH1])	96	121	116	113
τ(Nτ@His686, Zn, O@[Wat1,OH1])	107	102	105	102
τ(Nτ@His690, Zn, O@[Wat1,OH1])	128	110	105	105
RMSD (Å)				
		0.55	0.36	0.58

[a] To facilitate the comparison with X-ray data, only the measures involving heavy atoms were used. The RMSD (in Å) with respect to the X-ray structure is also given.

Table 2. Distances between Selected Pairs of Atoms in the LF Active Site, including (i) Zn-Coordination Bonds (*e.g.*, b(Zn, Oϵ2@Glu735)) and (ii) Hydrogen Bonds (*e.g.*, b(H@Wat1, Oϵ2@Glu687)) and Key Geometrical Features (*e.g.*, d(Cδ@Glu735, Cδ@Glu687)) of the Active Site[a]

pairs of atoms	distance (Å)
b(Zn, Oϵ2@Glu735)	2.22 (0.07)
d(Zn, Oϵ2@Glu687)	4.8 (0.2)
b(Zn, Nτ@His686)	2.05 (0.05)
b(Zn, Nτ@His690)	2.15 (0.05)
b(Zn, O@Wat1)	1.98 (0.06)
d(Cδ@Glu735, Cδ@Glu687)	6.9 (0.2)
d(Nτ@His686, Nτ@His690)	2.9 (0.1)
b(H@Wat1, Oϵ2@Glu687)	1.9 (0.2)
b(O'@Tyr-P2, H@Wat1)	1.8 (0.2)
b(Hη@Tyr728, O'@Pro-P1)	1.8 (0.2)
b(Hπ@His686, Oϵ1@Glu739)	1.97(0.2)

[a] The distances were calculated during the last 15 ns of the all-atom MD simulation. Standard deviations are given in parentheses.

3.2. Molecular Dynamics of LF Michaelis Complexes. Next, we built two Michaelis complexes using protonation states **C** and **D** for the active site (see the Methods section), and we performed all-atom MD simulations on both Michaelis complexes.

3.2.1. MD of LF Michaelis Complex in the C Protonation State. This complex was stable during the entire simulation (50 ns). In particular, the bond lengths in the coordination sphere had small fluctuations around their average positions during the dynamics (Table 2), and the substrate remained in its binding site for the entire simulation time. The structure of the rest of the protein was also
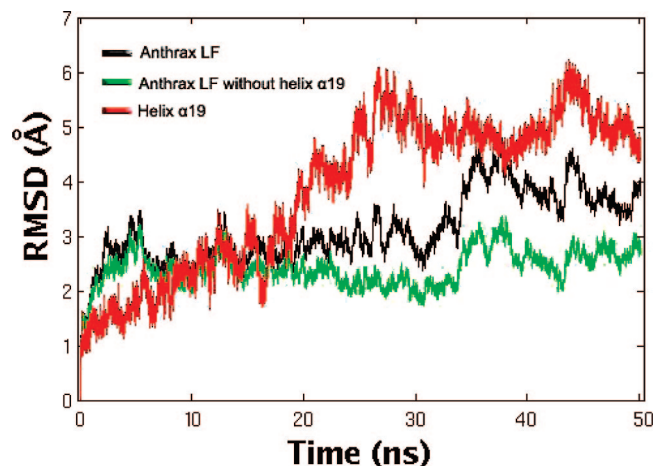
Figure 3. RMSD of LF Cα atoms during the 50 ns all-atom MD simulation. Note the increase in RMSD for helix α19 during the first 25 ns; after this time α19 gradually achieved structural stability.

maintained except for helix α19 (residues 346−367), which became partially unfolded within the first 25 ns of the MD trajectory (Figure 3). This unfolding was localized in the second part of α19 (residues 361−367) and caused a sudden increase in the RMSD of the Cα atoms (Figure 3); the coil-like conformation of the second part of α19 was then maintained until the end of the simulation. [The structure of the complex (LF+substrate) after 50 ns of all-atom MD is available at http://people.sissa.it/∼hong/projects__f iles/50ns_last_fr.pdb.]

To gain further insight into the instability of helix α19, it was necessary to carry out calculations with two types of coarse-grained methods and with disorder prediction servers:

(a) Hybrid Coarse-Grained/Molecular-Mechanics (CG/MM) Simulations. Here, α19 and nearby atoms were treated at the MM level, while the rest of the system was treated at the CG level (see the Methods section). Two 60 ns CG/MM simulations using AMBER parm98[42] and Gromos96 43a1[54] force fields for the MM part were performed. This computationally efficient method allows for the verification of the findings of the all atom MD simulation and for checking the dependence of these results on the employed force field. In both CG/MM simulations, α19 partially unfolds within the first ∼25 ns (see the Supporting Information, Figure S5). The RMSF of the Cα atoms from the two 60 ns trajectories (Figure 4) were larger for α19 than for the rest of the protein. Since the all-atom MD trajectory reaches a fairly constant RMSD only after 25 ns, we have checked the reproducibility of these results by performing CG/MM simulations starting from 15 equally spaced frames taken from the last 15 ns of the all-atom MD simulation (see the Supporting Information, section 5). The results obtained by these simulations, along with those obtained using different force fields, confirm that α19 unfolds in the first 25 ns of simulations independently of the computational technique and force field employed.

(b) Normal Mode Analysis (NMA). Here a CG elastic network of Cα atoms was built based on the energy minimized structure taken from the last frame of the all-atom MD trajectory. As obtained before for CG/MM MD simulations, the calculated RMSF of Cα atoms estimated
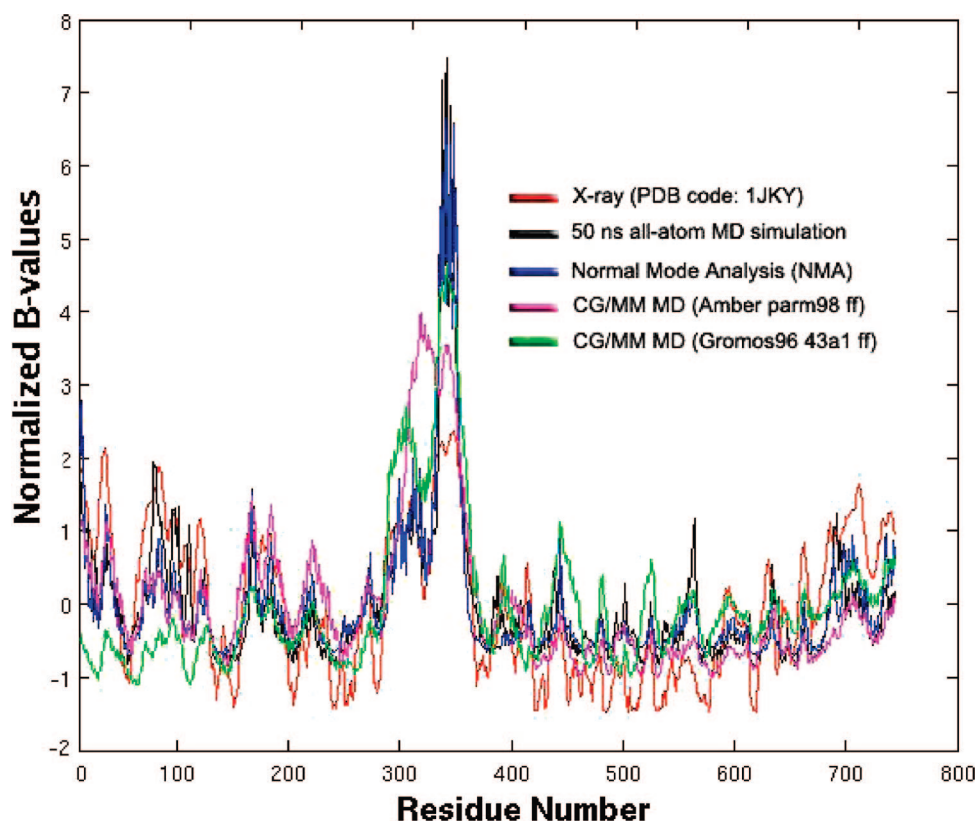

Figure 4. Calculated (RMSF) and experimental (X-ray; (PDB: 1JKY)[17]) normalized B-values for LF. For the all-atom MD and CG/MM simulations, only the last 15 ns were considered.

Simulations of Anthrax LF

J. Chem. Theory Comput., Vol. 4, No. 10, 2008 **1751**

from NMA for the α19 region were larger than those of the rest of the protein (Figure 4). To ensure the reproducibility of these results, NMA calculations were also performed using 15 equally spaced frames taken from the last 15 ns of the all-atom MD simulation (see the Supporting Information, section 6).

The normalized B-values calculated from all-atom MD, NMA, and CG/MM simulations (Figure 4) agree with those reported in the X-ray structure (PDB: 1JKY),[17] except for helix α19: part of the latter (residues 361–367) assumed a coiled conformation in aqueous solution, and its calculated normalized B-values were larger than those of the X-ray structure (Figure 4). This is consistent with the apparent difficulty to determine the solid state structure of α19, which, in fact, has only been resolved in the X-ray structure used here as the starting model for the α19 tract (PDB: 1JKY).[17]

(c) Disorder Prediction. To check the reliability of all-atom MD simulations showing the unfolding of a19 we also performed disorder prediction analysis using bioinformatics tools. The structural predictors PredictProtein,[29] PSIPRED,[30] SPRITZ,[31] and HNN[32] pointed out part of α19 as a disordered region. Particularly, the last segment of this region (residues 361–367) is more likely to be a loop than an α-helix (see Figure S9 in the Supporting Information, section 7).

Finally, we describe the structural and electrostatic features of the active site of the equilibrated Michaelis complex. Only the last 15 ns of all-atom MD simulation were used for this analysis since the RMSD of α19 became stable only at this simulation time.

At the active site, Glu735 acts as a monodentate ligand of Zn and H-bonds to the solvent, interacting on average with ∼0.7 water molecules, as obtained by integrating the radial distribution function of $O\epsilon2$@Glu687 vs O@water (see the Supporting Information, Figure S4a). The Zn is bonded with two histidine residues [b(Zn, $N\tau$@His686)=2.05 ± 0.05 Å and b(Zn, $N\tau$@His690)=2.15 ± 0.05 Å].

The catalytic water molecule, Wat1, H-bonds to Glu687, which is deprotonated; Glu687 is believed to act as a general base during catalysis (i.e., accepting an hydrogen ion from Wat1).[17,62] Besides the H-bond to Wat1, Glu687 also interacts, on average, with ∼1 water molecule from the solvent. On the other hand, Wat1 also H-bonds to Tyr-P2 (Table 2); this interaction could help to orient Wat1 in a proper position to perform the hydrolysis of the substrate.

Tyr728 forms an H-bond with the reactive carbonyl group of proline substrate (O'@Pro-P1, Table 2) for ∼35% of the simulation time, orienting the substrate for the nucleophilic attack. This H-bond interaction might play a role for the catalysis, providing a plausible, yet speculative, explanation on why the Y728F mutant is not catalytically active.[23] Also Tyr728 and O'@Pro-P1 are exposed towards the solvent, and they H-bond, on average, to ∼2 and 1 water molecules, respectively.

The motions of Tyr728 (of the enzyme) and Tyr-P1' (of the substrate) are correlated (see the Supporting Information, Figure S1b). The aromatic rings of these two residues lay at a distance smaller than 5 Å for approximately 53% of the simulation time, being that Tyr-P1' is accommodated in the
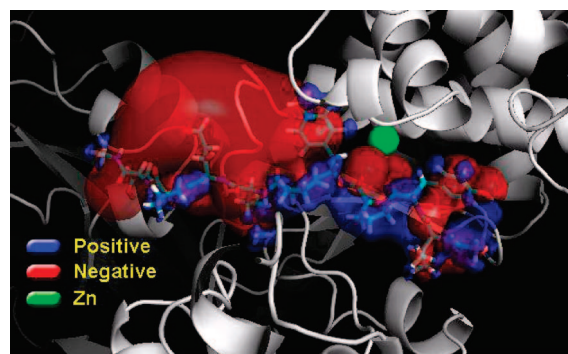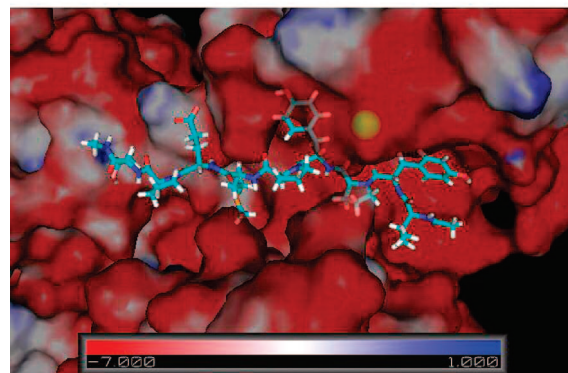
Figure 5. Electrostatic isosurfaces on the Michaelis complex. (**a**) Optimized MAPKK2-like substrate. Note the complementarity of the bilobular (created by O'@Pro-P1 and O'@Tyr-P1') around the Zn ion, while the larger negative patch (form by Glu-P4') is outside the negative groove of LF. (**b**) LF active site groove. Note the tyrosine on the left side of the Zn ion (Tyr-P1') inserted in the hydrophobic pocket (S1') of the enzyme.

hydrophobic pocket (S1') of the enzyme which is partially formed by Tyr728 (Figure 5b). The hydroxylic moiety of Tyr-P1' also forms an H-bond with the carbonyl group in the Val675 backbone ($H\eta$@Tyr-P1'...O'@Val675=1.8 ± 0.1 Å), and this interaction is maintained for ∼80% of the simulation time. The remaining 20% of the simulation time, Tyr-P1' H-bonds with Glu739 ($H\eta$@Tyr-P1'...$O\epsilon2$@Glu739=1.9 ± 0.2 Å) in S1'. At the entrance of S1' the backbone of Tyr-P1' also forms two stable (∼90% of the simulation time) H-bonds with backbone atoms in LF; the first one, with Lys656 (O'@Tyr-P1'... H'@Lys656=1.9 ± 0.1 Å) and the second with Gly657 (H'@Tyr-P1'... O'@Gly657=2.0 ± 0.2 Å).

Poisson–Boltzmann calculations suggest that the substrate fits in the groove of the enzyme forming complementary electrostatic interactions (Figure 5). The Zn ion interacts between a bilobular negative patch on the substrate (formed by O'@Pro-P1 and O'@Tyr-P1').

To dissect the single contributions of the substrate residues in binding to LF, we performed a calculation of substrate/ LF interaction energies based on the AMBER parm98[42] force field. Such calculations are necessarily approximate, and they are used here only for qualitative comparisons; therefore,

Table 3. LF Positional Selectivity[18] and Relative Interaction Energies (Electrostatics, van der Waals (vdW), and Total) between Individual Residues of the Substrate and LF, Calculated for the Last 15 ns of the All-Atom MD Simulation (Model **C**)[a]

	Val	Tyr	Pro	Tyr	Pro	Met	Glu
	P3	P2	P1	P1′	P2′	P3′	P4′
selectivity	1.5	**3.1**	-	**3.0**	1.9	1.3	1.6
total energy	(−0.26 ± 0.05)	**(−0.45 ± 0.09)**	(−0.1 ± 0.1)	**(−1.0 ± 0.1)**	(−0.48 ± 0.05)	(−0.43 ± 0.07)	(0 ± 0.5)
electrostatic energy	(−0.02 ± 0.02)	(−0.07 ± 0.07)	(0.1 ± 0.1)	**(−0.5 ± 0.1)**	(−0.24 ± 0.05)	(−0.09 ± 0.05)	(0.2 ± 0.6)
vdW energy	(−0.21 ± 0.05)	**(−0.36 ± 0.02)**	(−0.19 ± 0.05)	(−0.43 ± 0.05)	(−0.21 ± 0.02)	(−0.33 ± 0.05)	(−0.16 ± 0.05)

[a] Because these values are based on the AMBER parm98[42] force field, they are used here only for qualitative comparisons. The energy values were normalized to the largest absolute value as in ref 51.

Table 4. Comparison between Bond Ionicity Indexes (BIs)[58,59] in the Reactive Bonds of LF under Different Electrostatic Conditions Showing the Effect of the Protein and Solvent Electrostatic Properties on the Polarization of the Active Site[a]

	charged	no charge in protein	no charge in solvent	no charge
$d(N_{pep}\text{-}BO^1)/d(N_{pep}\text{-}C_{pep})$	0.37(0.02)	0.38(0.02)	0.37(0.02)	0.37(0.02)
$d(N_{pep}\text{-}BO^2)/d(N_{pep}\text{-}C_{pep})$	0.35(0.02)	0.36(0.02)	0.35(0.02)	0.36(0.02)
$d(O_{pep}\text{-}BO^1_{lone})$	0.32(0.01)	0.31(0.01)	0.32(0.01)	0.32(0.01)
$d(O_{pep}\text{-}BO^2_{lone})$	0.31(0.01)	0.31(0.01)	0.30(0.01)	0.30(0.01)
$d(O_{pep}\text{-}BO^1_{C=O})/d(O_{pep}\text{-}C_{pep})$	0.39(0.01)	0.40(0.01)	0.40(0.01)	0.40(0.01)
$d(O_{pep}\text{-}BO^2_{C=O})/d(O_{pep}\text{-}C_{pep})$	0.39(0.01)	0.39(0.03)	0.40(0.01)	0.40(0.01)
$d(O_{Wat}\text{-}BO^1_{lone})$	0.28(0.04)	0.28(0.05)	0.28(0.02)	0.27(0.02)
$d(O_{Wat}\text{-}BO^2_{lone})$	0.43(0.06)	0.47(0.05)	0.43(0.04)	0.45(0.04)
$d(O_{Wat}\text{-}BO_{O-H1})/d(O_{Wat}\text{-}H_{Wat})$	0.50(0.01)	0.49(0.01)	0.50(0.01)	0.49(0.01)
$d(O_{Wat}\text{-}BO_{O-H2})/d(O_{Wat}\text{-}H_{Wat})$	0.51(0.01)	0.51(0.01)	0.51(0.01)	0.51(0.01)

[a] Four conditions were evaluated: (i) "**charged**", calculation assigning charges to all atoms in the system; (ii) "**no charge in protein**", calculation assigning charges equal to zero in all atoms of the protein; (iii) "**no charge in solvent**", calculation assigning charges equal to zero to the atoms of all the water molecules; and (iv) "**no charge**", calculation assigning charges equal to zero to all atoms in the system. Atomic charges of the MM atoms were assigned using the AMBER parm98 force field,[42] while active site atoms were treated at the DFT level.[59] DFT/MM calculations were performed on 15 equally spaced frames taken from the last (equilibrated) 15 ns of the all-atom MD trajectory. Standard deviations are given in parentheses.

only normalized values[51] of the interaction energies are reported (Table 3).

The substrate's residues with the largest LF positional-dependent selectivity,[18] namely Tyr-P2, Tyr-P1′, and Pro-P2′, featured the strongest interaction energies with LF. Indeed, Tyr-P1′ is the main anchor for substrate binding forming significant vdW interactions (mainly provided by Try-P1′ aromatic ring and Tyr728, His686, and Leu677 side chains) as well as H-bonding and electrostatic interactions, mostly with Val675, Glu739, Lys656, and Gly657. Tyr-P2 plays also a significant role for the binding, forming vdW interactions with His690, Tyr659, Leu658, Pro661, and Ala734 (Table 3).

Although in the present work we do not perform any reactivity study, the polarization of the active site residues induced by the electrostatic properties of the protein environment may provide some information on the Michaelis complex, which is an important species in the catalytic cycle of the enzyme. Consequently, we decided to consider the polarization effect of different electrostatic environments (e.g., *in vacuo*, and with the influence of the solvent and/or LF electrostatic fields) on the putative reactive bonds of the Michaelis complex.[63] Our results (Table 4) show that the environment does not have a marked role in polarizing the active site residues, with the only exception of the nucleophilic water, which trivially, upon coordination to Zn, undergoes a polarization of the lone pairs. The polarization of the bonds involved in the LF-mediated substrate cleavage can be quantitatively compared with homologous reactive

Table 5. Comparison between Bond Ionicity Indexes (BIs)[58,59] in the Reactive Bonds of Three Prototypical Proteases: Furin, HIV-1 PR, and LF[a]

	furin	HIV-1 PR	anthrax LF
$d(N_{pep}\text{-}BO^1)/d(N_{pep}\text{-}C_{pep})$	0.33(0.02)	0.31(0.03)	0.37(0.02)
$d(N_{pep}\text{-}BO^2)/d(N_{pep}\text{-}C_{pep})$	0.38(0.01)	0.38(0.02)	0.35(0.02)
$d(O_{pep}\text{-}BO^1_{lone})$	0.33(0.01)	0.34(0.01)	0.32(0.01)
$d(O_{pep}\text{-}BO^2_{lone})$	0.31(0.01)	0.32(0.01)	0.31(0.01)
$d(O_{pep}\text{-}BO^1_{C=O})/d(O_{pep}\text{-}C_{pep})$	0.37(0.01)	0.38(0.01)	0.39(0.01)
$d(O_{pep}\text{-}BO^2_{C=O})/d(O_{pep}\text{-}C_{pep})$	0.38(0.01)	0.39(0.01)	0.39(0.01)
$d(O_{Hyd}\text{-}BO^1_{lone})$	0.31(0.01)		
$d(O_{Hyd}\text{-}BO^2_{lone})$	0.31(0.01)		
$d(O_{Hyd}\text{-}BO_{O-H})/d(O_{Hyd}\text{-}H_{Hyd})$	0.50(0.01)		
$d(O_{Hyd}\text{-}BO_{C-O})/d(O_{Hyd}\text{-}C)$	0.39(0.01)		
$d(O_{Wat}\text{-}BO^1_{lone})$		0.32(0.31)	0.28(0.04)
$d(O_{Wat}\text{-}BO^2_{lone})$		0.33(0.01)	0.43(0.06)
$d(O_{Wat}\text{-}BO_{O-H1})/d(O_{Wat}\text{-}H_{Wat})$		0.47(0.02)	0.50(0.01)
$d(O_{Wat}\text{-}BO_{O-H2})/d(O_{Wat}\text{-}H_{Wat})$		0.53(0.02)	0.51(0.01)

[a] The location of the Boys Orbitals (BOs) for LF is given in Figures S2 and S3 of the Supporting Information. The calculations were performed on 15 equally spaced frames taken from the last (equilibrated) 15 ns of the all-atom MD trajectory. Standard deviations are given in parentheses. X_{pep} refers to the atom type X from the peptidic bond which undergoes nucleophilic attack. X_{Wat} refers to the atom type X from the nucleophilic water molecule. X_{Hyd} refers to the atom type X from the nucleophilic hydroxilic group (belonging to the side chain of Ser).

bonds of two other proteases, for which calculations have been carried out with an identical setup. These are the aspartyl protease from human immunodeficiency virus of type 1 (HIV-1 PR),[63,64] which is believed to use a water molecule for the hydrolysis,[64] and the serine protease furin,[63] which uses the hydroxylic group from a serine residue as a

Simulations of Anthrax LF

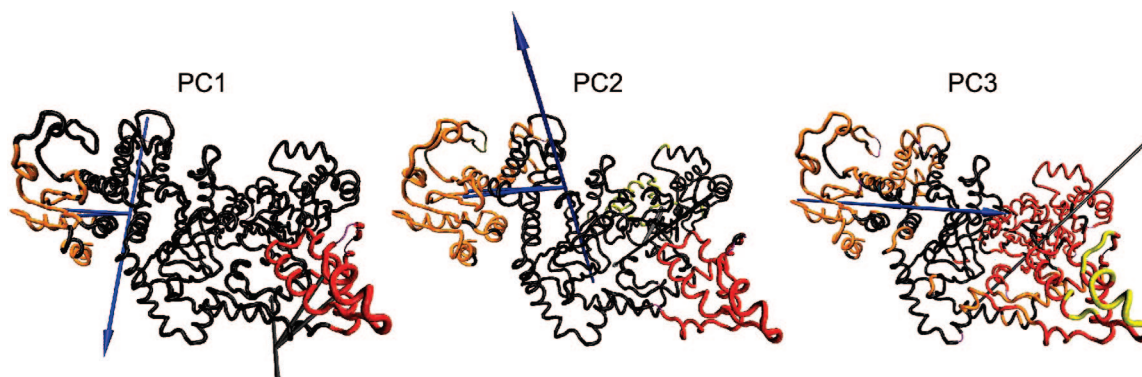*J. Chem. Theory Comput., Vol. 4, No. 10, 2008* **1753**

Figure 6. The first three Principal Components (PCs) as calculated with Dynatraj[48] from the last 15 ns of the all-atom MD simulation. Rigid domains, as calculated with the procedure of Wriggers and Schulten,[49] are depicted in different colors; details of these calculations can be found in the Supporting Information, section 6. The arrows indicate the effective rotation axis between the two adjacent rigid domains. The direction of the arrow represents the sense (e.g., clockwise) used to find the rotations axis between two rigid domains, and it is ultimately dependent on the choice of the "reference" rigid domain (the one kept steady during the calculation of the rotational angle). Note that PC3 was able to capture a movement in the α19 region (PC3, yellow), pointing to the higher flexibility of this region (see Figure 4).

nucleophile. As a measure of polarization we used the Bond Ionicity indexes[58] (see the section "Electrostatics" in the Methods section).

Based on the values of the BIs of $C_{pep}=O_{pep}$ and N_{pep}-C_{pep} bonds in the substrate (Table 5 and Figure S3), we conclude that substrate's reactive bonds in LF are less polarized than those in HIV-1 PR and furin. In addition, the water (Wat1) O−H bonds were also less polarized in LF than in HIV-1 PR: (i) in LF, BI_{O-H1}@Wat1(LF) = 0.50 ± 0.01, BI_{O-H2}@Wat1(LF) = 0.51 ± 0.01 and (ii) in HIV-1 PR, BI_{O-H1}@Wat1(HIV-1 PR) = 0.47 ± 0.02 and BI_{O-H2}@Wat1(HIV-1 PR) = 0.53 ± 0.02. However, the lone pairs on the oxygen atom of the catalytic water (represented as $d(O_{Wat}-BO^{1[2]}_{lone})$ in Table 5) were more asymmetric in LF than in HIV-1 PR simply due to the coordination of the water to Zn.

Thus, although our analysis provides no information on the polarization of the transition state or of other species of the catalytic cycle (which were not investigated in this study), we do suggest here that small polarization effects on the substrate are induced by the LF scaffold, while the presence of zinc has a more critical role for the polarization of the reactants (Tables 4 and 5).

The large scale motions may play a role for substrate recognition and/or for enzymatic catalysis.[34,63−66] We explore this issue by performing PCA[48] on the last (equilibrated) 15 ns of the all-atom MD trajectory. In the first three Principal Components (PCs), which account for ∼54% of the overall motion, we identified large motions involving domain **III** (i.e., the domain that includes helix α19, see Figure 1a). In particular, PC3 was able to capture a relatively independent movement of α19 with respect to domain **III** (Figure 6). The observed large scale motions of LF did not affect the active site (see Table 2), similar to what has been found in another protease studied with the same computational setup, the serine protease furin.[63] Therefore, we conclude that it is unlikely that the large scale motions of LF could play a mechanical role for the enzymatic activity, although we cannot exclude that they may indirectly affect catalysis by long-range electrostatics.

3.2.2. MD of LF Michaelis Complex in D Protonation State. This protonation state turned out to be already unstable in the first 0.5 ns of simulation: the H-bond network was disrupted because of a rotation of Glu687 along the Cγ-Cδ bond, allowing the entrance of additional water in the active site and the departure of the substrate from the active site. This complex was therefore discarded.

4. Conclusions

We characterized the structural properties of the Michaelis complex formed by anthrax LF and an optimized substrate using a variety of computational tools. Our findings can be summarized as follows:

(i) The second shell ligands affect the structural properties of the Zn active site as has been observed in other Zn-based enzymes.[22,24,67] Our calculations confirm that second shell ligands have an influence on the protonation state of the active site geometry, stabilizing the hydrogen bond network within the active site residues.

(ii) The nucleophilic agent is a Zn-bound water molecule (not an OH group) H-bonded to Glu687.

(iii) The calculated substrate per residue interaction energies with LF correlate with the experimentally derived LF positional selectivity.[18] In particular, Tyr-P1′ and Tyr-P2 constitute the main substrate anchors to the LF active site. In fact, Tyr-P1′ has the largest interaction energy with LF, with similar contributions from electrostatics (featured mainly by a H-bond established between Hη@Tyr-P1′ and the backbone of the enzyme (O′@Val675)) and vdW interactions (with Tyr-P1′ interacting with the side chains of Tyr728, His686, and Leu677). In contrast, Tyr-P2 exhibits large vdW interactions principally with the Zn-bound His690. These results may provide a rationale to explain the selectivity of LF for substrates with tyrosines in the vicinity of the reactive proline.[18]

(iv) The LF scaffold induces small polarization effects on the substrate. A larger polarization is observed for the lone pairs of the nucleophilic agent Wat1 (see (ii)); the obvious cause for this effect is the Wat1 coordination to Zn.

(v) Large-scale motions do not affect the structure of the LF active site; it is therefore unlikely that these motions could play a mechanical role during the first step of the catalysis.

(vi) Part of helix α19 (residues 361−367) assumes a coil-like conformation in aqueous solution. This behavior was qualitatively predicted by several bioinformatics tools and confirmed by several computational approaches.

Supporting Information Available: Calculations on LF correlated of motions; parametrization of the active site residues, results polarization; and hydration of the active site, additional simulations on hybrid Coarse-Grained/Molecular Mechanics (CG/MM) simulations, Principal Component Analysis (PCA), Normal Mode Analysis (NMA), and additional structural bioinformatics predictions. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Mock, M.; Mignot, T. Anthrax toxins and the host: a story of intimacy. *Cell. Microbiol.* **2003**, *5*, 15–23.

(2) Guidi-Rotani, C. The alveolar macrophage: The Trojan horse of *Bacillus anthracis*. *Trends. Microbiol.* **2002**, *10*, 405–409.

(3) Dixon, T. C.; Fadl, A. A.; Koheler, T. M.; Swanson, J. A.; Hanna, P. C. Early *Bacillus anthracis*-macrophage interactions: intracellular survival and escape. *Cell. Microbiol.* **2000**, *2*, 453–463.

(4) Friedlander, A. M. Anthrax: clinical features, pathogenesis, and potential biological warfare threat. *Curr. Clin. Top. Infect. Dis.* **2000**, *20*, 335–349.

(5) Shoop, W. L.; Xiong, Y.; Woods, A.; Guo, J.; Pivnichny, J. V.; Felcetto, T.; Michael, B. F.; Bansal, A.; Cummings, R. T.; Cinnungam, B. R.; Friedlander, A. M.; Douglas, C. M.; Patel, S. B.; Wisniewski, D.; Scapin, G.; Spaolowe, S. P.; Zaller, D. M.; Chapman, K. T.; Scolnick, E. M.; Schmatz, D. M.; Bartizal, K.; MacCoss, M.; Hermes, J. D. Anthrax lethal factor inhibition. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 7958–7963.

(6) Pezard, C.; Berche, P.; Mock, M. Contribution of individual toxin components to virulence of *Bacillus anthracis*. *Infect. Immun.* **1991**, *59*, 3472–3477.

(7) Petosa, C.; Collier, R. J.; Klimpel, K. R.; Leppla, S. H.; Liddington, R. C. Crystal structure of the anthrax toxin protective antigen. *Nature* **1997**, *385*, 833–838.

(8) Montecucco, C.; Tonello, F.; Zanotti, G. Stop the killer: how to inhibit the anthrax lethal factor metalloprotease. *Trends. Biochem. Sci.* **2004**, *29*, 282–285.

(9) Abrami, L.; Reig, N.; van der Goot, F. G. Anthrax toxin: the long and winding road that leads to the kill. *Trends. Microbiol.* **2005**, *13*, 72–78.

(10) Bann, J. C.; Hultgreen, S. J. Structural Biology: Anthrax hijacks host receptor. *Nature* **2004**, *430*, 843–844.

(11) Duesbery, N. S.; Webb, C. P.; Leppla, S. H.; Gordon, V. M.; Klimpel, K. R.; Copeland, T. D.; Ahn, N. G.; Oskarsonn, M. K.; Fukasawa, K.; Paul, K. D.; Vande Woude, G. F. Proteolytic inactivation of MAP-Kinase-Kinase by anthrax lethal factor. *Science* **1998**, *280*, 734–737.

(12) Weston, C. R.; Lambright, D. G.; Davis, R. J. MAP Kinase signalling specificity. *Science* **2002**, *296*, 2345–2347.

(13) Friedlander, A. M. Macrophages are sensitive to anthrax lethal toxin through an acid-dependent process. *J. Biol. Chem.* **1986**, *261* (16), 7123–7126.

(14) Kirby, J. E. Anthrax lethal toxin induces human endothelial cell apoptosis. *Infect. Immun.* **2004**, *72* (1), 430–439.

(15) Milne, J. C.; Furlong, D.; Hanna, P. C.; Well, J. S.; Collier, R. J. Anthrax protective antigen forms oligomers during intoxication of mammalian cells. *J. Biol. Chem.* **1994**, *269*, 20607–20612.

(16) Guichard, A.; Park, J. M.; Cruz-Moreno, B.; Karin, M.; Bier, E. Anthrax lethal factor and edema factor act on conserved targets in Drosophila. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 3244–3249.

(17) Pannifer, A. D.; Wong, T. Y.; Schwarzenbacher, R.; Renatus, M.; Petosa, C.; Collier, R. J.; Bienkowska, J.; Lacy, D. B.; Park, S.; Leppla, S. H.; Hanna, P.; Liddington, R. C. Crystal structure of the anthrax lethal factor. *Nature* **2001**, *414*, 229–233.

(18) Turk, B. J.; Wong, T. Y.; Schwarzenbacher, R.; Jarrell, E. T.; Leppla, S. H.; Collier, R. J.; Liddington, R. C.; Cantley, L. C. The structural basis for substrate and inhibitor selectivity of the anthrax lethal factor. *Nat. Struct. Mol. Biol.* **2004**, *11*, 60–66.

(19) Panchal, R. G.; Hermone, A. R.; Nguyen, T. L.; Wong, T. Y.; Schwarzenbacher, R.; Schmidt, J.; Lane, D.; McGrath, C.; Turk, B. E.; Burnett, J.; Aman, M. J.; Little, S.; Sausville, E. A.; Zaharevitz, D. W.; Cantley, L. C.; Liddington, R. C.; Gussio, R.; Bavari, S. Identification of small molecular inhibitors of anthrax lethal factor. *Nat. Struct. Mol. Biol.* **2004**, *11*, 67–72.

(20) Forino, M.; Johnson, S.; Wong, T. Y.; Rozanov, D.; Savinov, A. Y.; Li, W.; Fattorusso, R.; Becattini, B.; Orry, A. J.; Abagyan, R. A.; Smith, J. W.; Alibek, K.; Liddington, R. C.; Strongin, A. Y.; Pellecchia, M. Efficient synthetic inhibitors of anthrax lethal factor. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102* (27), 9499–9504.

(21) Auld, D. S. Zinc coordination sphere in biochemical zinc sites. *Biometals* **2001**, *14*, 271–313.

(22) (a) Magistrato, A.; DeGrado, W. F.; Laio, A.; Rothlisberger, U.; VandeVondele, J.; Klein, M. L. Characterization of the dizinc analogue of the synthetic diiron protein DF1 using *ab initio* and hybrid quantum/classical molecular dynamics simulations. *J. Phys. Chem. B* **2003**, *107*, 4182–4188. (b) Dal Peraro, M.; Vila, A. J.; Carloni, P. Structural determinants and hydrogen-bond network of the mononuclear zinc(II)-beta-lactamase active site. *J. Biol. Inorg. Chem.* **2002**, *7*, 704–712. (c) Gervasio, F. L.; Schettino, V.; Mangani, S.; Krack, M.; Carloni, P.; Parrinello, M. Influence of outer-shell metal ligands on the structural and electronic properties of horse liver alcohol dehydrogenase zinc active site. *J. Phys. Chem. B* **2003**, *107*, 6886–6892.

(23) Tonello, F.; Naletto, L.; Romanello, V.; Dal Molin, F.; Montecucco, C. Tyrosine-728 and glutamic acid 735 are essential for the metalloproteolitic activity of the lethal factor of *Bacillus anthracis*. *Biochem. Biophys. Res. Commun.* **2004**, *131*, 496–502.

Simulations of Anthrax LF

J. Chem. Theory Comput., Vol. 4, No. 10, 2008 **1755**

(24) Klimpel, K. R.; Arora, N.; Lepla, S. H. Anthrax toxin lethal factor contains a zinc metalloprotease consensus sequence which is required for lethal toxin activity. *Mol. Microbiol.* **1994**, *13* (6), 1093–1100.

(25) Hammond, S. E.; Hanna, P. C. Lethal factor active-site mutations affect catalytic activity in vitro. *Infect. Immun.* **1998**, *66*, 2374–2378.

(26) Neri, M.; Anselmi, C.; Cascella, M.; Maritan, A.; Carloni, P. Coarse-Grained model of proteins incorporating atomistic detail of the active site. *Phys. Rev. Lett.* **2005**, *95*, 218102.

(27) Tozzini, V. Coarse grained models for proteins. *Curr. Opin. Struct. Biol.* **2005**, *15*, 144–150.

(28) Tirion, M. M. Large amplitude elastic motions in proteins from a single-parameter atomic analysis. *Phys. Rev. Lett.* **1996**, *77*, 1905–1908.

(29) Rost, B.; Yachdav, G.; Liu, J. The PredictProtein server. *Nucleic Acids Res.* **2004**, *32* (Web Server Issue), W321–W326.

(30) Bryson, K.; McGuffin, L. J.; Marsden, R. L.; Ward, J. J.; Sodhi, J. S.; Jones, D. T. Protein structure prediction servers at University College London. *Nucleic Acids Res.* **2005**, *33*, W36–38.

(31) Vullo, A.; Bortolami, O.; Pollastri, G.; Tosatto, S. C. E. Spritz: a server for the prediction of intrinsically disordered regions in protein sequences using kernel machines. *Nucleic Acids Res.* **2006**, *34*, W164–168.

(32) Guermeur, Y.; Geourjon, C.; Gallinari, P.; Deleage, G. Improved performance in protein secondary structure prediction by inhomogeneous score combination. *Bioinformatics* **1999**, *15* (5), 413–421.

(33) (a) Makinen, M. W.; Kuo, L. C.; Dymowski, J. J.; Jaffer, S. Catalytic role of the metal ion of carboxypeptidase A in ester hydrolysis. *J. Biol. Chem.* **1978**, *254* (5), 356–366. (b) Christianson, D. W.; Lipscomb, W. N. Carboxypeptidase A. *Acc. Chem. Res.* **1989**, *22*, 62–69. (c) Pelmenschikov, V.; Blomberg, M. R. A.; Siegbahn, P. E. M. A theoretical study of the mechanism for peptide hydrolysis by thermolysin. *J. Biol. Inorg. Chem.* **2002**, *7*, 284–298. (d) Cross, J. B.; Vreven, T.; Meroueh, S. O.; Mobashery, S.; Schlegel, H. B. Computational investigation of irreversible inactivation of the zinc-dependent protease carboxypeptidase A. *J. Phys. Chem. B* **2005**, *109*, 4761–4769.

(34) Piana, S.; Carloni, P.; Parrinello, M. Role of conformational fluctuations in the enzymatic reaction of hiv-1 protease. *J. Mol. Biol.* **2002**, *319*, 567–583.

(35) *CPMD*; Copyright IBM Corp 1990–2006, Copyright MPI für Festkörperforschung Stuttgart 1997–2001.

(36) Trouiller, N.; Martins, J. L. Efficient pseudopotentials for plane-wave calculation. *Phys. Rev. B* **1991**, *43*, 1993–2006.

(37) Keinman, L.; Bylander, D. M. Efficacious Form for Model Pseudopotentials. *Phys. Rev. Lett.* **1982**, *48*, 1425–1428.

(38) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behaviour. *Phys. Rev. A* **1998**, *38*, 3098–3100.

(39) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789.

(40) Barnett, R. N.; Landman, U. Born-Oppenheimer molecular-dynamics simulations of finite systems: Structure and dynamics of (H2O)2. *Phys. Rev. B* **1993**, *48*, 2081–2097.

(41) (a) IUPAC-IUB. Abbreviations and symbols for description of conformation of polypeptide chains. Tentative rules Biochemistry **1970**, *9* (18), 3471–3479. (b) IUPAC-IUB. Nomenclature and symbolism for amino acid and peptides. Recommendations 1983 Pure Appl. Chem. **1984**, *56* (5), 595–624.

(42) (a) Cheatham, T. E., III; Cieplak, P.; Kollman, P. A. A modified version of the Cornell *et al.* force field with improved sugar pucker phases and helical repeat. *J. Biomol. Struct. Dyn.* **1999**, *16*, 845–862. (b) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E., III; DeBolt, S.; Ferguson, D.; Seibel, G. L.; Kollman, P. A. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics an1d free energy calculations to simulate the structural and energetic properties of molecules. *Comput. Phys. Commun.* **1995**, *91*, 1–41.

(43) (a) Suárez, D.; Díaz, N.; Merz, K. M., Jr. Molecular dynamics simulations of the mononuclear zinc β-lactamase from *Bacillus cereus* complexed with Benzylpenicillin and a quantum chemical study of the reaction mechanism. *J. Am. Chem. Soc.* **2001**, *123*, 9867–9879. (b) Suarez, D.; Diaz, N.; Merz, K. M., Jr. Molecular dynamics simulation of the dinuclear zinc-β-lactamase from *Bacteroides fragilis* complexed with imipenem. *J. Comput. Chem.* **2002**, *28*, 1587–1600. (c) Dal Peraro, M.; Villa, A. J.; Carloni, P. Substrate binding to Mononuclear metallo-β-lactamase from *Bacillus cereus*. *Proteins* **2004**, *54*, 412–423.

(44) Cheatham, T. E., III; Miller, J. L.; Fox, T.; Darden, T. A.; Kollman, P. A. Molecular dynamics simulations on solvated biomolecular systems: the Particle Mesh Ewald method leads to stable trajectories of DNA, RNA, and proteins. *J. Am. Chem. Soc.* **1995**, *117*, 4193–4194.

(45) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.

(46) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(47) (a) Laxmikant, K.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. NAMD2: greater scalability for parallel molecular dynamics. *J. Comp. Phys.* **1999**, *151*, 283–312. (b) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781–1802.

(48) Barret, C. P.; Hall, B. A.; Noble, E. M. Dynamite: a simple way to gain insight into protein motions. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *D60*, 2280–2287.

(49) Wriggers, W.; Schulten, K. Protein domain movements: Detection of rigid domains and visualization of hinges in comparisons of atomic coordinates. *Proteins* **1997**, *29*, 1–14.

(50) Humphrey, W.; Dalke, A.; Schulten, K. VMD-Visual Molecular Dynamics. *J Mol. Graphics* **1996**, *14*, 33–38.

(51) Guidoni, L.; Torre, V.; Carloni, P. Water and potassium dynamics inside the KcsA K+ channel. *FEBS Lett.* **2000**, *477*, 37–42.

(52) Noguti, T.; Go, N. Collective variable description of small-amplitude conformational fluctuations in a globular protein. *Nature* **1982**, *296*, 776–278.

(53) Doi, M. *Introduction to Polymer Physics*; Oxford University Press: Oxford, 1996.

(54) van Gunsteren, W. F.; Daura, X.; Mark, A. E. GROMOS force field In *Encyclopedia of Computational Chemistry*; John Wiley & Sons: New York, 1998; 1211p.

(55) Lindahl, E.; Azuara, C.; Koehl, P.; Delarue, M. NOMAD-Ref: visualization, deformation and refinement of macromolecular structures based on all-atom Normal Mode Analysis. *Nucleic Acids. Res.* **2006**, *34*, W52−56.

(56) Bank, R.; Holst, M. A new paradigm for parallel adaptive meshing. *SIAM J. Sci. Comput.* **2000**, *22*, 1411–1443.

(57) Dolinsky, T.; Nielsen, J.; McCammon, A.; Baker, N. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatic calculations. *Nucleic Acid. Res.* **2004**, *32*, W665−W667.

(58) (a) Silvestrelli, P. L.; Marzari, N.; Vanderbilt, D.; Parrinello, M. Maximally-localized Wannier functions for disordered systems: application to amorphous silicon. *Solid State Commun.* **1998**, *107*, 7–11. (b) Berghold, G.; Mundy, C. J.; Romero, A. H.; Hutter, J.; Parrinello, M. General and efficient algorithms for obtaining maximally localized Wannier functions. *Phys. Rev. B* **2000**, *61*, 10040–10048. (c) Magistrato, A.; Robertazzi, A.; Carloni, P. Nitrogen fixation by a molybdenum catalyst mimicking the function of the nitrogenase enzyme: A critical evaluation of DFT and solvent effects. *J. Chem. Theory Comput.* **2007**, *3*, 1708–1720.

(59) Laio, A.; VandeVondele, J.; Rothlisberger, U. A Hamiltonian electrostatic coupling scheme for hybrid Car-Parrinello molecular dynamics simulations. *J. Chem. Phys.* **2002**, *116*, 6941–6947.

(60) (a) Vargiu, A. V.; Ruggerone, P.; Magistrato, A.; Carloni, P. Anthramycin-DNA binding explored by molecular simulations. *J. Phys. Chem. B* **2006**, *110*, 24687–24695. (b) Spiegel, K.; Magistrato, A. Modeling anticancer drug-DNA interactions via mixed QM/MM molecular dynamics simulations. *Org. Biomol. Chem.* **2006**, *4*, 2507–2517.

(61) Simona, F.; Magistrato, A.; Vera, D. M. A.; Garau, G.; Vila, A. J.; Carloni, P. Protonation state and substrate binding to B2 metallo-beta-lactamase CphA from. *Aeromonas hydrofila. Proteins* **2007**, *69*, 595–605.

(62) Jedrzejas, M. The structure and function of novel proteins of *Bacillus anthracis* and other spore forming bacteria: Development of novel prophylactic and therapeutic agents. *Crit. Rev. Biochem. Mol. Biol.* **2002**, *37*, 339–373.

(63) Carnevale, V.; Raugei, S.; Micheletti, C.; Carloni, P. Large-scale motions and electrostatic properties of furin and HIV-1 protease. *J. Phys. Chem. A* **2007**, *111*, 12327–12332.

(64) Piana, S.; Bucher, D.; Carloni, P.; Rothlisberger, U. Reaction Mechanism of HIV protease by hybrid Car-Parrinello/Classical MD simulations. *J. Phys. Chem. B* **2004**, *108*, 11139–11149.

(65) Henzler-Wildman, K. A.; Thai, V.; Lei, M.; Ott, M.; Wolf-Watz, M.; Fenn, T.; Pozharski, E.; Wilson, M. A.; Petsko, G. A.; Karplus, M.; Hübner, C. G.; Kern, D. Intrinsic motions along an enzymatic reaction trajectory. *Nature* **2007**, *450*, 838–844.

(66) Perryman, A. L.; Lin, J. H.; McCammon, J. A. HIV-1 Protease Molecular Dynamics of a Wild-Type and of the V82F/I84V Mutant: Possible Contributions to Drug Resistance and a Potential New Target Site for Drugs. *Protein Sci.* **2004**, *13*, 1108–1123.

(67) Christianson, D. W.; Cox, D. Catalysis by metal-activated hydroxide in zinc and manganese metalloenzymes. *Annu. Rev. Biochem.* **1999**, *68*, 33–57.