THE JOURNAL OF PHYSICAL CHEMISTRY B

ARTICLE

pubs.acs.org/JPCB

Catalytic Mechanism and Roles of Arg197 and Thr183 in the *Staphylococcus aureus* Sortase A Enzyme

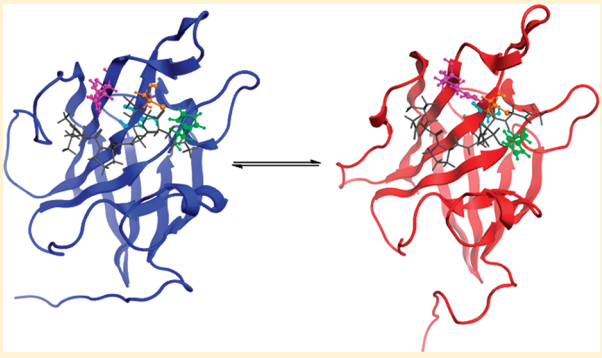
Bo-Xue Tian[†] and Leif A. Eriksson[*,†,‡]

[†]School of Chemistry, National University of Ireland—Galway, Galway, Ireland

[‡]Department of Chemistry, University of Gothenburg, 412 96 Göteborg, Sweden

Supporting Information

ABSTRACT: The sortase A enzyme, which catalyzes the peptidoglycan cell wall anchoring reaction of LPXTG surface proteins, has been proposed to be a universal target for therapeutic agents against Gram-positive bacteria. The catalytic mechanism of the *Staphylococcus aureus* sortase A enzyme has been systematically studied using molecular dynamics simulations, ONIOM(DFT:MM) calculations, and QM/MM charge deletion analysis. The catalytic roles of Arg197 and Thr183 were analyzed. Our calculations show that Arg197 has several important roles in the mechanism. It is crucial for substrate binding, and is capable of reversible shift of its hydrogen bonds between the LP and TG carbonyls of the LPXTG substrate motif, depending on the protonation state of the catalytic Cys184−His120 dyad. Arg197 stabilizes the catalytic dyad in the active ion pair form but at the same time raises the barrier to acylation by approximately 8 kcal/mol. Thr183 is also essential for the catalytic reaction in that it correspondingly lowers the barrier by the same amount via electrostatic interactions. The catalytic mechanism proceeds via proton transfer from His120, followed by nucleophilic attack from the thiolate anion of Cys184. The data thus supports the proposed reverse protonation mechanism, and disproves the hypothesis of the Arg197 generating an oxyanion hole to stabilize the tetrahedral intermediate of the reaction.



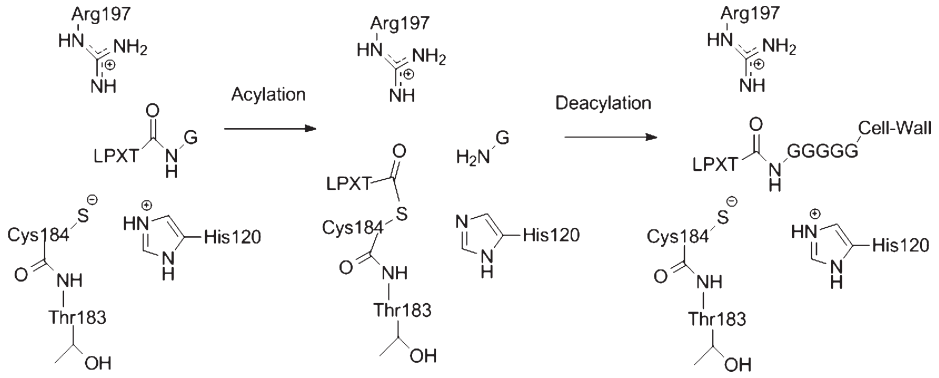
1. INTRODUCTION

The cell wall anchored surface proteins of Gram-positive pathogenic bacteria play vital roles in pathogenicity.[1−3] The sortase A (SrtA) enzyme, which catalyzes the peptidoglycan cell wall anchoring reaction of LPXTG (leucine, proline, X, threonine, and glycine, where X is any amino acid) surface proteins, has been proposed to be a universal target for therapeutic agents against Gram-positive bacteria,[4−6] as inactivation of SrtA reduces the virulence of pathogens significantly.[7,8] The *Staphylococcus aureus* (*S. aureus*) SrtA is currently the most studied SrtA enzyme due to the severity of Methicillin-resistant *S. aureus* (MRSA) infections, and a number of studies have been carried out in order to search for inhibitors against *S. aureus* SrtA.[9,10] However, only a few inhibitors with IC_{50} values in the submicromolar range are to date identified.[11] Knowledge on the enzymatic catalysis can promote the discovery of SrtA inhibitors, and studies on the kinetics,[12,13] mutagenesis,[14−17] and structures[18−22] of the SrtA enzyme have provided important information for understanding the catalytic mechanism. However, no consensus on the catalytic mechanism of SrtA has as yet been reached.

The reaction catalyzed by SrtA is a transpeptidation that undergoes a ping-pong mechanism.[12,13] In the first phase, SrtA cleaves the peptide bond between the threonine and glycine (T−G) residues in the LPXTG motif, forming an LPXT−enzyme intermediate. In the second phase, the second substrate pentaglycine (from the outer cell wall cross bridge) is attached to the LPXT moiety, forming the cell wall anchored LPXT surface protein (Scheme 1). According to mutagenesis studies,[14−17] Cys184, His120, Arg197, and Thr183 are indispensable, as alanine substitution of any of these four residues results in a more than 1000-fold decrease in enzymatic activity. Several catalytic mechanisms were proposed, including a Cys184−His120 ion pair Cys184 thiol activation mechanism[17] which is similar to the enzyme papain (a paradigm cysteine protease[23,24]), a general base mechanism in which the Cys184 thiol is activated by His120 or Arg197[21,25] after the substrate is bound, and a reverse protonation mechanism where a small fraction of the active SrtA is competent for catalysis[12] (see ref 12 for a detailed description of the various mechanisms). According to pK_a measurements, pH dependence experiments, and solvent isotope effect measurements, the Cys184−His120 ion pair mechanism and the general base mechanism have been shown unlikely to occur.[12] The reverse protonation mechanism suggests that Cys184 and His120 exist in the thiolate anion and imidazolium

Received: June 21, 2011
Revised: September 12, 2011
Published: September 27, 2011

Scheme 1. Key Steps of the Cell Wall Anchoring Reaction Catalyzed by SrtA[12]



cation forms before the substrate is bound, and that the ionization state of Cys184 is independent of His120 protonation.[12] Unfortunately, the reverse protonation mechanism is difficult to be proved due to the sparse population of the active form of SrtA.[12] In the recently reported crystal structure of *Streptococcus pyogenes* SrtA[20] and the NMR structure of the *S. aureus* LPXT*−SrtA intermediate,[19] the position of His120 is close to Cys184 (*S. aureus* numbering), whereas Arg197 is far away from the T−G peptide bond. Therefore, it is generally accepted that the carbonyl carbon of the T−G peptide bond in the substrate is attacked by the Cys184 thiolate anion, and the leaving NH group is protonated by the His120 imidazolium cation. Arg197 is considered to be important for substrate binding via hydrogen bonds.[19] The role of Thr183 is currently unknown.

Interestingly, studying the structure of a homology model of *Listeria monocytogenes* SrtA with molecular dynamics simulations,[26] we found that altering the protonation status of the Cys−His ion pair leads to two different conformations of the SrtA enzyme. In the inactive form of the SrtA model, Arg197 attains hydrogen bonds with the peptide bond between the LP residues of the LPXTG substrate, whereas, in the active form, Arg197 forms hydrogen bonds with the peptide bond between the TG residues. We suggested that the motion of Arg197 can stabilize the Cys−His ion pair via electrostatic interactions. We also hypothesized that the Arg197 motion may be important for the enzymatic catalysis, since, in the active form, Arg197 could provide an oxyanion hole for stabilizing the tetrahedral intermediate, which has been proposed as a possible role of Arg197.[15,19] To prove our hypothesis, more accurate quantum mechanical calculations on the acylation step were deemed necessary, and are the subject of the current study.

Theoretical chemistry methods, especially the QM/MM method, have been extensively used for studying the catalytic mechanisms of various enzymes.[27−29] The first phases of the reactions catalyzed by the SrtA enzyme and by cysteine proteases are similar. The catalytic mechanism of cysteine proteases has been investigated in a number of theoretical studies.[30−37] QM/MM studies[31−33] on the catalytic pathways of the cysteine proteases have shown that, in the first phase, the protonation of the substrate NH group is most likely to occur prior to, or concerted with, the nucleophilic attack of the carbonyl carbon by the thiolate anion. Alternative pathways were also proposed by DFT-only calculations, and water was proposed as a proton shuttle between the cysteine and histidine of the catalytic dyad.[34]

In this study, we first demonstrate that the previously reported Arg197 motion[26] also occurs in the *S. aureus* SrtA. The QM/MM method is then used to explore the proposed multiple roles of Arg197 in SrtA catalysis, including the influence on the relative stabilities of the inactive and active forms of the Cys184−His120 dyad, substrate binding, as well as the effect on the acylation step. The role of Thr183 in the catalysis, which has hitherto not been explained, is also discussed.

2. COMPUTATIONAL DETAILS

2.1. Preparation of the Michaelis Complex. The first snapshot of the NMR structure of the *S. aureus* LPXT*−SrtA intermediate[19] (residues 62−206, PDB code: 2KID, RCSB Protein Data Bank) was used as the initial model in the current study. We chose this structure instead of the X-ray structure of the *S. aureus* LPXTG−SrtA complex[21] for several reasons. First, in the NMR structure, the Cys184 and His120 residues are close to each other, whereas in the X-ray structure His120 is far away from Cys184. Mutation studies demonstrated that alanine substitution of His120 results in a 10000-fold decrease in enzyme activity, suggesting that His120 is even more important than Cys184 (1000-fold decrease when mutated to alanine). The X-ray structure is unable to explain the role of His120. Second, in the X-ray structure, Cys184 was mutated to alanine, which may result in undesirable conformation changes. It has been shown that the X-ray structure of the *S. aureus* SrtA−LPXTG complex is significantly different from the structures of other SrtA homologues. Moreover, due to the sparse population of the active form of the enzyme, most structures of SrtA resolved are most likely in the inactive form. The NMR structure is suitable for our purpose on studying the catalytic mechanism of SrtA as it mimics the LPXT−SrtA intermediate, the product of the acylation step. Superposing the LPXTG substrate to the LPXT moiety of the LPXT−SrtA intermediate should be more accurate than docking the substrate to the active site of structures without substrates bound. In order to recover the Michaelis complex, an Nme-LPATG-Ace peptide was built, optimized and superposed on the LPAT* moiety of this NMR structure, using the Flexible Alignment module in MOE.[38]

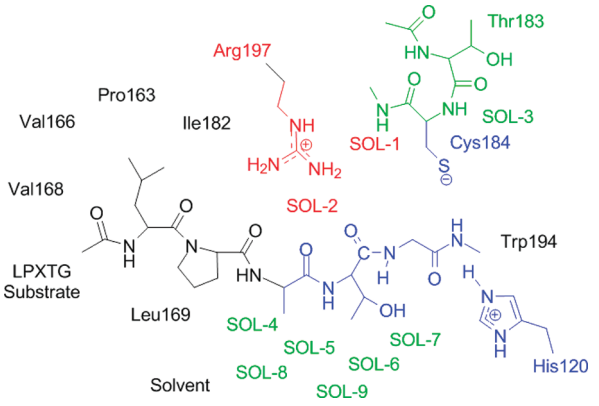
2.2. Molecular Dynamics Simulations. Molecular dynamics (MD) simulations were performed to investigate if the Arg197 motion discovered in our previous study on the *L. monocytogenes* SrtA homology model[26] also applies to the *S. aureus* case. The Amber 03 force field[39] was used for all the MD simulations. The Cys184−His120 dyad was initially set to the active (ion pair) form, while the protonation of the remaining residues was kept the same as in the NMR structure. The obtained structure was

solvated by TIP3P water molecules in a periodic box with a buffer distance of 8.0 Å. A Cl^- ion was randomly added to satisfy the electroneutrality condition, using the genion module in GROMACS (version 4.0.4) software.[40] The whole system (~20000 atoms) was energy minimized by steepest descent (200 steps) to remove close contacts. A position restrained simulation (2 ns duration, 1.0 fs time step, NPT ensemble, T = 298 K, P = 1 bar) was performed, to enable the water molecules to reach more favorable positions. The leapfrog algorithm was used for the integration of the classical Newton equations of motion. Particle-mesh Ewald (PME) summation[41,42] was used for long-range electrostatics. A 10 Å cutoff was used for both Coulomb and Lennard-Jones interactions. The temperature and pressure were controlled through the Berendsen coupling algorithm,[43] with the time constants 0.1 ps for temperature and 1.0 ps for pressure coupling. All bond lengths were constrained using the LINCS algorithm.[44] In the production run (20 ns duration, 1.0 fs time step, NPT ensemble, T = 298 K, P = 1 bar), the temperature was controlled using the Nose−Hoover thermostat[45] with a time constant of 0.1 ps and the pressure was controlled using the Parrinello−Rahman barostat[46] with a time constant 1.0 ps. Other parameters were similar to the position restrained simulations. Prior to the production run, we performed 1 ns MD simulations using the production run parameters and recorded snapshots at 200 ps, 400 ps, 600 ps, 800 ps, and 1 ns as initial structures for further production runs. After the production run of the active form SrtA−LPXTG complex, we modified the protonation states of the Cys−His dyad to the inactive form and performed another 20 ns MD simulation in order to investigate if the Arg197 side chain can switch to the LP moiety (corresponding to the reverse simulation of our previous *L. monocytogenes* study[26]). 20 ns reversibility tests were also performed to confirm the Arg197 motion (for details, see Supporting Information S6-3). All MD simulations were performed using GROMACS.[40]

2.3. QM/MM (ONIOM) Calculations. Representative MD snapshots were further energy minimized with the Amber 03 force field.[39] In the QM/MM calculations, MM parameters for the active Cys−His catalytic dyad were used. Four main systems were investigated, including the last snapshot of the active form from the 20 ns MD trajectory (called Position1), the last snapshot of the inactive form from the 20 ns MD trajectory (called Position2), the R197A mutation of Position1 (called Position1-R197A), and the T183A mutation of Position1 (called Position1-T183A). Mutations were readily obtained using the MOE program. It should be noted that Position1 and Position2 refer to the preferred positions of Arg197 in the active and inactive SrtA−substrate complexes, whereas active and inactive forms refer to the zwitterionic and neutral forms of the Cys184−His120 catalytic dyad. For the R197A and T183A systems, position restrained simulations were performed prior to the MM optimizations, to equilibrate the surrounding water molecules. The substrate-free systems for Position1, Position2, and Position1-R197A were also studied for comparison of the relative stabilities of the inactive and active forms of the enzyme. Position restrained simulations were performed after the removal of the substrates for the three substrate-free systems.

A two-layer ONIOM (QM/MM) scheme[47] was used for the calculations of the SrtA catalytic mechanism. The standard DFT method B3LYP[48−50] and the two basis sets 6-31G(d) (BS1) and 6-311+G(2d,p) (BS2) were used for treating the QM region. It should be noted that the B3LYP/Amber method has previously been used to study a similar enzyme, papain, and the results are in good accordance with experimental data.[32] Geometries of intermediates and TSs were optimized using the ONIOM(B3LYP/BS1:Amber) method with the ONIOM-ME (mechanical embedding) scheme, followed by frequency calculations at the same level of theory. Atoms far away from the substrate carbonyl group (>15 Å) were kept fixed during the optimizations, and the normal modes of these atoms were excluded from the frequency calculations with the Gaussian 09 keyword *notatoms*. Single-point energy calculations with the ONIOM(B3LYP/BS2:Amber) method were performed on the optimized geometries with the ONIOM-EE (electronic embedding) scheme. Geometries were not optimized with the ONIOM/EE scheme due to the high computational cost and poor convergence. All ONIOM calculations were performed using the Gaussian 09 software.[51]

Scheme 2. QM/MM Partitioning Scheme: Model-1 in Blue, Model-2 Blue + Red, and Model-3 Blue + Red + Green



The QM/MM partitioning is described in Scheme 2. Three different partitioning schemes were used for studying the acylation step at Position1. Model-1 (blue) contains the ATG moiety of the substrate as well as the side chains of Cys184 and His120. Model-2 (blue + red) added Arg197 and two solvent molecules (SOL-1 and SOL-2) to the QM region, and Model-3 further added Thr183, the Thr183-Cys184 backbone, and another seven solvent molecules to the QM part. Model-1 was also used for the calculations on the Position2, Position1-R197A, and Position1-T183A systems.

2.4. Charge Deletion Analysis and Binding Affinity Calculations. Charge deletion analysis was used to investigate the influence of individual residues as well as solvent molecules on the enzymatic catalysis. This approach has been used in a number of theoretical studies.[30,33,52−54] Deletion of atomic charges of a certain residue in the MM region can result in energy changes in the QM region (ΔE_{QM}). If the QM energy changes of a reactant [ΔE_{QM}(R)] and a transition state [ΔE_{QM}(TS)] are different, the reaction barrier is affected. The electrostatic effect of a certain residue on the reaction barrier [$\Delta\Delta E_{QM}$(R→TS)] can be defined as in eq 1. A positive sign of $\Delta\Delta E_{QM}$(R→TS) indicates that the effect of this residue is to lower the barrier height, whereas a negative sign implies that it plays the opposite role. Similarly, the relative stabilities of a reactant and a product (for instance, the relative stabilities of the inactive and active forms of the enzyme) can also be evaluated; see eq 2. All the $\Delta\Delta E_{QM}$ values reported in this study are ONIOM-EE(B3LYP/BS2:Amber) single point energies. To reduce computational cost, every three adjacent residues were combined into a group, and charge deletion analysis was initially used for identifying important

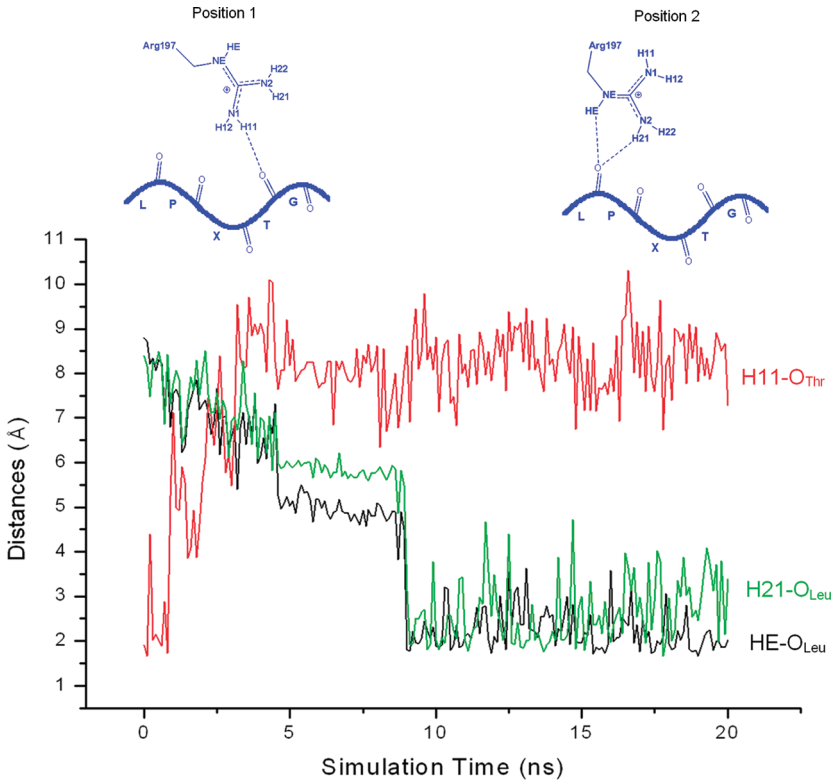


Figure 1. Hydrogen bonds between Arg197 and substrate in the 20 ns MD simulation of the inactive SrtA−LPXTG complex, starting from the simulated active complex. Position1 and Position2 refer to the preferred positions of Arg197 in the active and inactive SrtA−substrate complexes.

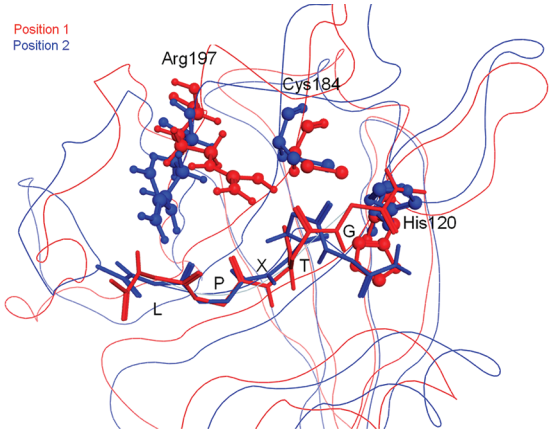


Figure 2. Superposed structures with Arg197 in Position1 (in red) and Position2 (in blue).

groups ($\Delta\Delta E_{QM}$ > 1 kcal/mol). On the basis of this initial screening, the individual residues in the important groups, as well as all active site residues (not frozen in the ONIOM calculations), were subsequently analyzed (Supporting Information S1).

$$\Delta\Delta E_{QM}(R\rightarrow TS) = \Delta E_{QM}(TS) - \Delta E_{QM}(R) \quad (1)$$

$$\Delta\Delta E_{QM}(R\rightarrow P) = \Delta E_{QM}(P) - \Delta E_{QM}(R) \quad (2)$$

The binding affinity of the substrate was evaluated using the MM/GBVI approach[55] in the MOE software, which includes implicit solvent effects. The average value of binding affinities from 10 snapshots (with all solvent molecules removed) in the last 1 ns MD trajectory was evaluated (for the Position1-R197A system, only the last snapshot of the position restrained simulation was used).

3. RESULTS AND DISCUSSION

3.1. Arg197 Motion. In our previous study on the *L. monocytogenes* SrtA homology model,[26] Arg 197 was found capable to change its hydrogen donor interactions within 30 ns MD simulations, from the LP backbone carbonyl groups of the LPXTG substrate in the inactive form to the TG backbone carbonyls in the active form, and that this motion was fully reversible. To explore the Arg197 motion in *S. aureus* SrtA, we performed a 20 ns MD simulation on the active form, followed by another 20 ns MD simulation with the Cys−His dyad switched to the inactive form.

As expected, Arg197 forms hydrogen bonds to the carbonyl oxygen of the TG peptide bond in the active form, and swings over to form hydrogen bonds to the LP part in the hydrophobic pocket of the active site within the 20 ns MD simulation on the inactive form (five individual 20 ns simulations were performed and showed consistent results). We have shown previously that 20−30 ns simulation is sufficient for this small system, and that extending the simulation time to 100 ns does not substantially influence the results for the Arg197 motion.[26] The hydrogen bonds between Arg197 and the substrate in Position1 and Position2 are illustrated in Figure 1, and a superposed view is shown in Figure 2. In Position1, the side chain of Arg197 forms one direct hydrogen bond with the carbonyl group of the TG peptide bond (H11···O$_{Thr}$, Figure 1). In the simulation of the inactive form of the SrtA−LPXTG complex, the H11···O$_{Thr}$

Table 1. Proton States of the Cys−His Dyad under Different Conditions

entry	system	relative energies[a] active	relative energies[a] inactive
1	Position1 without LPXTG	0.0	27.4 ± 3.0[b]
2	Position1-R197A without LPXTG	0.0[c]	−10.8
3	Position2 without LPXTG	0.0[c]	−8.2
4	Position1 with LPXTG	0.0	2.5
5	Position1-R197A with LPXTG	0.0	−15.2
6	Position2 with LPXTG	0.0	−14.1

[a] The energy of the inactive form relative to the active form is used for each case, and the energies for different systems are not comparable, as the number of atoms used is different. All energies are ONIOM-EE(B3LYP/BS2:Amber) single point energies on Model-3 partitioning (Scheme 2), in kcal/mol. Relative energies of Model-1 are summarized in Supporting Information S2. [b] Average energy of five snapshots, with error bar. [c] The N−H distance in His120 in the active form is constrained at 1.05 Å as optimization of the active form spontaneously leads to the inactive form.

hydrogen bond is broken spontaneously, while two new hydrogen bonds are formed (H21···O_{Leu} and HE···O_{Leu}, Figure 1). The repositioning of Arg197 depending on the protonation state of the His−Cys dyad may be essential for the SrtA catalysis, for one or more reasons: (1) the Arg197 motion may alter the relative stabilities of the inactive and active forms of the Cys184−His120 catalytic dyad; (2) this motion may increase the substrate binding affinity; (3) Arg197 may form an oxyanion hole to stabilize the transition state or tetrahedral intermediate in the acylation step. To elucidate the role of Arg197, different systems were therefore investigated by means of QM/MM (ONIOM) calculations, as outlined above.

3.2. Relative Stabilities of the Inactive and Active Forms. The relative stability of the inactive and active forms is an important issue in the SrtA enzyme, as pK_a measurements[12,25] have shown that the majority of the enzyme exists in the inactive form. In this respect, SrtA seems to differ from the cysteine proteases such as papain. However, the bell-shaped pH dependencies of kinetic parameters k_{cat}/K_m for SrtA imply that one residue in the Cys−His dyad must be protonated and the other deprotonated for activity.[12] The reverse protonation mechanism was proposed,[12] i.e., that a small fraction of SrtA with the Cys−His dyad in the active form catalyzes the reaction.

Recently, Mladenovic et al.[30] demonstrated that the enzyme environment is capable of altering the relative stabilities of the neutral versus zwitterionic forms of the Cys−His dyad in cathepsin (one of the cysteine proteases[23]) using QM/MM molecular dynamics simulations, which explained the abnormal pK_a values of the Cys−His dyad measured in cysteine proteases.[56−58] The question we posed was whether the Arg197 motion is capable of altering the relative stabilities of the inactive form versus the active form in SrtA. To address this issue, we performed systematic QM/MM calculations on the transformations between the inactive and active forms in Position1, Position1-R197A, and Position2.

Strictly speaking, the substrate may bind to the inactive or active SrtA enzymes, corresponding to the general base mechanism and the reverse protonation mechanism, respectively. We therefore studied the relative stabilities of the inactive and active forms with and without the substrate bound. The protonation states in Position1 (with and without substrate bound) were initially explored using five different simulation snapshots (the last snapshot from five individual MD simulations). In Position1, the active form was found to be more stable than the inactive form (Table 1, entry 1 and 4) by 24−30 kcal/mol. Given the small difference between the five simulation snapshots (Table 1, entry 1), only one starting structure was used in the Position2 and Position1-R197A systems. From Table 1, it is clear that the inactive form becomes more stable than the active form in either Position1-R197A (Position1 with Arg197 mutated to alanine) or Position2 (Arg197 located far away from the Cys−His dyad). In addition, the active form becomes less stable with the substrate bound (Table 1). The rationale for this is that the substrate consists of hydrophobic residues which prevents solvent molecules from entering the enzyme pocket and that the Cys−His ion pair is less favored in the resulting nonpolar environment. It should be mentioned that the His120 proton forms a stable hydrogen bond (or salt bridge) with the Cys184 thiolate group in all the MD simulations, and we did not observe any water molecules (which could be proton shuttles as described in ref 34) in the middle of the Cys−His ion pair.

Charge deletion analysis of Position1 (with substrate bound) showed that Arg197, Lys134, Lys198, and SOL-1 are essential for the stabilization of the active Cys−His dyad, whereas Thr183, Thr121, Asp186, and SOL-2 play an opposite role (Table 2).

Table 2. Charge Deletion Analysis for Position1 (with LPXTG): Influence of the Environment on the Relative Stabilities of the Inactive and Active forms

influence of residues	$\Delta\Delta E_{QM}$ (active → inactive)[b]	influence of nearby solvent molecules[a]	$\Delta\Delta E_{QM}$ (active → inactive)[b]
Arg197	23.0	SOL-1	15.8
Thr183	−10.0	SOL-2	−5.1
Thr121	−3.1	SOL-3	−0.3
Lys134	5.9	SOL-4	0.0
Asp185	1.9	SOL-5	−0.2
Asp186	−6.0	SOL-6	0.5
Trp194	−0.2	SOL-7	−0.9
Lys198	5.5	SOL-8	0.8
Thr in substrate	−3.7[c]	SOL-9	1.9

[a] Positions of the solvent molecules as described in Scheme 2. [b] As defined in eq 2, in kcal/mol. [c] Charge deletion analysis on Model-1 with the substrate in the MM region.

Table 3. Substrate Binding Affinity under Different Conditions

entry	system	MM/GBVI binding affinity[a]
1	Position1 active	−55.8 ± 0.8
2	Position1 inactive	−54.9 ± 0.6
3	Position1-R197A active	−49.6 ± 0.9
4	Position1-R197A inactive	−50.9 ± 1.2
5	Position2 active	−55.0 ± 0.5
6	Position2 inactive	−54.0 ± 0.8

[a] In kcal/mol.

dx.doi.org/10.1021/jp2058113 |J. Phys. Chem. B 2011, 115, 13003–13011

Table 4. Properties of the TS in Position1 for Different ONIOM Schemes

ONIOM model	barrier height[a]	imaginary[b] frequency	$S_{Cys184}\cdots C_{Thr}$ distance in TS[c]	$HE_{His120}\cdots N_{Gly}$ distance in TS[c]
Model-1	19.4	−730.4	2.880	1.412
Model-2	14.8	−700.2	2.907	1.573
Model-3	17.5	−536.8	3.350	1.190

[a] In kcal/mol. [b] In i cm^{-1}. [c] In Å.

Nonpolar residues such as Trp194 have a minor influence on the relative stabilities of the inactive and active forms.

Besides the influence on the relative stability of the Cys−His dyad, the Arg197 motion may also affect the binding affinity of the substrate (Table 3). As expected, the hydrogen bonding between Arg197 and LPXTG is important for substrate binding, as the R197A mutation results in a 4−6 kcal/mol decrease in binding affinity (Table 3, entry 1 vs entry 3 and entry 2 vs entry 4). However, the binding affinity is not significantly affected by the Arg197 motion (entry 1 vs entry 5 and entry 2 vs entry 6), implying that the substrate does not selectively bind to the active form of the enzyme. The ∼5 kcal/mol reduced binding affinity of the substrate also correlates well with the 1000-fold reduction in activity noted for the R197A mutant.[17]

We note that the general base mechanism remains unlikely to occur, since with the LPXTG bound to the inactive enzyme, the proton transfer from Cys184 to His120 becomes more endothermic (Table 1, entries 3 and 6), unless the Arg197 side chain moves to the TG region. According to our MD simulations, however, the Arg197 side chain stays in the LP region if the Cys−His dyad is in the inactive form. We also performed MD simulations of the active form SrtA without the substrate bound. Unlike the SrtA−substrate complex system, no preferred position of the Arg197 side chain could be identified in the substrate-free system, probably because it becomes more flexible without the substrate to bind to (simulations started from Position1 and Position2, Supporting Information S3). With the substrate bound, Arg197 orients the side chain hydrogen bonds to the TG moiety of the substrate (Position1), which is the preferred position thereof. Taken together, our calculations favor the reverse protonation mechanism;[12] i.e., the LPXTG substrate directly binds to the active form of the enzyme and undergoes acylation, even though the majority of the enzyme is in the inactive form. This is also supported by the fact that the active form is more stable (Arg197 in Position1)/less unstable (Arg197 in Position2), and thus more populated, when the substrate is not bound (Table 1).

3.3. Acylation. Another possible role of Arg197 is to form an oxyanion hole (Position1) that will stabilize the tetrahedral intermediate in the acylation step. However, since we have demonstrated that the Cys−His ion pair, which is the reactant of the acylation step, is significantly stabilized by Arg197, the barrier height of acylation may in fact increase due to the stabilization of the reactant. Therefore, evaluation of the change in barrier height for the acylation step caused by the motion of Arg197 is of interest.

The acylation of cysteine proteases has previously been modeled in a number of theoretical studies.[31−33] On the basis of AM1 calculations, Arad et al.[31] suggested that in the enzyme papain a proton transfer from histidine (PT-His) occurs prior to, or concerted with, the nucleophilic attack by the thiolate anion

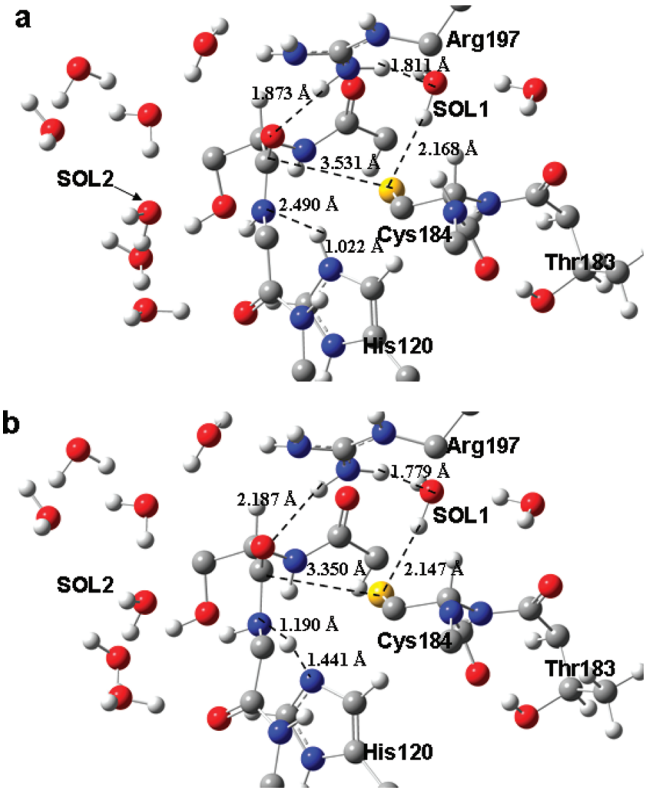


Figure 3. Model-3 geometries (QM region) of (a) reactant and (b) TS in the acylation step in Position1.

(NA-Cys). The papain system was further studied by Harrison et al.[32] with the AM1/MM and B3LYP/MM methods, and the concerted pathway without the involvement of a tetrahedral intermediate (TI) was suggested. Recently, Ma et al.[33] determined the two-dimensional free energy surfaces of the acylation of human cathepsin by using AM1/MM molecular dynamics simulations. The PT-His and NA-Cys reactions were found highly coupled, and a TI was formed along the NA-Cys pathway. Although these impressive studies provided insights into the catalytic mechanism of cysteine proteases, some limitations are obvious. First, the use of semiempirical methods may not be sufficient for the QM part. Second, no imaginary frequency of the TS was reported. Moreover, alternative QM/MM partitioning schemes were not investigated.

The standard QM/MM method ONIOM(B3LYP:Amber), which has been extensively used for studying enzyme reactions,[29] was chosen, and three different QM/MM partitioning schemes were used to model the acylation step in the active form of SrtA in Position1. The deacylation step was not modeled. The second substrate pentaglycine is similar to glycine (the product released in the acylation step) in terms of the structure. It can therefore be expected that the deacylation is highly similar to the reverse process of the current acylation.

Initial scans of two reaction coordinates, i.e., NA-Cys ($S_{Cys184}\cdots C_{Thr}$ distance) and PT-His ($HE_{His120}\cdots N_{Gly}$ distance), were performed separately. In the NA-Cys process, the ONIOM-EE (B3LYP/BS2:Amber) energy increases continuously, whereas, in the PT-His process, the energy starts decreasing when the $HE_{His120}\cdots N_{Gly}$ distance becomes less than 1.60 Å. Interestingly, the $S_{Cys184}\cdots C_{Thr}$ distance is significantly shortened (although not forming a $S_{Cys184}-C_{Thr}$ bond) when

Table 5. Charge Deletion Analysis of the Acylation Step in Position1

influence of residues		influence of nearby solvent molecules	
	$\Delta\Delta E_{QM}(R\rightarrow TS)^a$		$\Delta\Delta E_{QM}(R\rightarrow TS)^a$
Arg197	−18.0	SOL-1	−2.8
Thr183	3.6	SOL-2	1.5
Thr121	−3.4	SOL-3	−1.0
Lys134	−3.9	SOL-4	0.2
Asp185	−0.1	SOL-5	−0.2
Asp186	3.0	SOL-6	−0.2
Trp194	−0.3	SOL-7	−0.5
Lys198	−2.7	SOL-8	0.7
		SOL-9	1.2

[a] As defined in eq 1, Model-1 was used, in kcal/mol.

the $HE_{His120}\cdots N_{Gly}$ distance becomes less than 1.60 Å (Supporting Information S4). These results suggest that the PT-His process is the major route to reach the TS, and NA-Cys may occur concertedly with or immediately after PT-His. Our results are consistent with the experimental fact that H120A mutation causes a larger reduction in enzyme activity than C184A (∼10000-fold decrease for H120A and ∼1000-fold decrease for C184A, respectively),[17] as well as with previous theoretical studies on cysteine proteases.[31−33]

The TS for Position1 was then optimized; the results are summarized in Table 4. The obtained barrier heights are in reasonable accordance with the experimental value ∼20 kcal/mol[12,13] (calculated using the Eyring equation). Checking the normal modes of the imaginary frequency, we found that the TS corresponds to the PT-His reaction coordinate. Interestingly, an IRC-like optimization of the TS (the $HE_{His120}\cdots N_{Gly}$ distance of TS was shortened, followed optimization of the product) gave the final product rather than the protonated substrate or TI, probably because the barrier heights for NA-Cys and the cleavage of the peptide bond are very small (or barrierless) once the proton transfer is completed.

The geometries of the reactant and TS of Position1 in the acylation step are shown in Figure 3 using Model-3 (∼100 QM atoms). In the reactant (R), Arg197 forms a direct hydrogen bond with the carbonyl group of the TG peptide bond ($H_{Arg}\cdots O_{Thr}$) and an indirect hydrogen bond with the thiolate sulfur through SOL-1. In the TS, the $HE_{His120}\cdots N_{Gly}$ distance decreases from 2.490 to 1.190 Å, whereas the $S_{Cys184}\cdots C_{Thr}$ distance is nearly unchanged (3.531 Å in R and 3.350 Å in TS). Interestingly, the $H_{Arg}\cdots O_{Thr}$ hydrogen bond becomes weaker in the TS (1.873 Å in R and 2.187 Å in TS), whereas the indirect hydrogen bond between Arg197 and the thiolate anion becomes stronger (Figure 3). These geometrical changes show that Arg197 stabilizes R more than TS, which is further confirmed by charge deletion analysis (based on the smaller Model-1; Table 5).

According to the charge deletion analysis, Thr183 might be able to lower the barrier height of acylation (Table 5). To further investigate the role of Arg197 and Thr183 on the acylation step, we modeled this for Position1-R197A, Position1-T183A, and Position2 (Table 6). The changes of barrier heights are consistent with charge deletion analysis; i.e., the R197A mutation significantly lowers the acylation barrier (from 19.4 to 11.7 kcal/mol), whereas the T183A mutation works in the opposite manner (from 19.4 to 28.8 kcal/mol). The acylation barrier in Position2 is lower than that of Position1, implying

Table 6. Comparison of Acylation Barriers at Different States

system	acylation barrier[a]
Position1	19.4
Position2	16.8
Position1-R197A	11.7
Position1-T183A	28.2

[a] Model-1 was used, in kcal/mol.

that the Arg197 motion activates the Cys−His dyad but at the same time raises the acylation barrier. Experimentally, both R197A and T183A mutations result in ∼1000-fold decrease in enzyme activity,[14−16] and the roles of these two residues are still somewhat controversial.[19] Our calculations show that Arg197 has dual roles; i.e., it is needed both for binding the substrate and for activating the Cys−His dyad. However, Arg197 itself gives rise to an increased barrier for the acylation step. To compensate for this, the presence of Thr183 affects the catalysis by lowering the acylation barrier by roughly the same amount via electrostatic interactions.

The reaction energy for the system in Position1 is with the larger Model-3 endothermic by 11.9 kcal/mol (5.9 kcal/mol for the smaller Model-1); Table S5 in the Supporting Information. The barrier for the second half of the reaction, assuming the reaction profile to be symmetric, is thus only 5.6 kcal/mol (13.5 kcal/mol for the smaller model), which enables a rapid formation of the final product once the pentaglycine substrate from the peptidoglycan cell wall is in position to covalently anchor the LPXT sequence. These energies are consistent with the fact that the acylation step is the rate-limiting step in the overall transpeptidation process.

4. CONCLUSIONS

The catalytic mechanism of the *Staphylococcus aureus* sortase A enzyme has been systematically studied. Arg197 is capable of reversibly switching its hydrogen donor interactions from the LP backbone carbonyl groups of the LPXTG substrate in the inactive form to the TG backbone carbonyls in the active form, within 20 ns MD simulations. QM/MM calculations and charge deletion analysis show that Arg197 stabilizes the thiolate anion in the Cys184−His120 catalytic dyad (the precursor of the acylation step), which is essential for SrtA catalysis. Arg197 is important for substrate binding, but the Arg197 motion does not affect the binding affinity. On the basis of the MD simulations and QM/MM calculations of the relative stabilities of the inactive and active Cys−His dyad, we suggest that the general base mechanism is unlikely to occur. The reverse protonation mechanism proposed by Frankel et al.[12] is the most likely mechanism for SrtA. By analyzing the normal modes of the imaginary frequency for the TS of the acylation step, we conclude that this corresponds to proton transfer from His120 (PT-His), rather than nucleophilic attack by the thiolate anion of Cys184 (NA-Cys). IRC-like optimization of TS yields the final product, due to the barrier heights for NA-Cys and the cleavage of the peptide bond being very small (or lacking) when the PT-His step is completed. Interestingly, with the presence of Arg197, the barrier height of acylation increases, which disproves the previous hypothesis that Arg197 may stabilize the TS by forming an oxyanion hole. To compensate for this, Thr183 significantly reduces the acylation barrier via electrostatic interactions, which explains the unknown role of Thr183 in SrtA catalysis. All results are consistent with available experimental data, and provide a

detailed explanation for the observed features of the sortase A enzymatic mechanism.

■ ASSOCIATED CONTENT

Supporting Information. Charge deletion analysis details, Model-1 calculations on the protonation states of the Cys−His dyad under different conditions, MD simulations of the substrate-free systems, scan of the PT-His reaction coordinate, ONIOM energies, rmsd values of MD simulations on the SrtA−LPXTG complex, Cartesian coordinates of the QM region in the Model-3 ONIOM calculations on the acylation step at Position1, and complete citation for ref 51. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: leif.eriksson@chem.gu.se.

■ ACKNOWLEDGMENT

The National University of Ireland—Galway is gratefully acknowledged for financial support. The SFI/HEA Irish Centre for High-End Computing (ICHEC) is acknowledged for the provision of computational facilities and support.

■ REFERENCES

(1) Paterson, G. K.; Mitchell, T. J. *Trends Microbiol.* **2004**, *12*, 89–95.
(2) Marraffini, L. A.; DeDent, A. C.; Schneewind, O. *Microbiol. Mol. Biol. Rev.* **2006**, *70*, 192–221.
(3) Navarre, W. W.; Schneewind, O. *Microbiol. Mol. Biol. Rev.* **1999**, *63*, 174–229.
(4) Mazmanian, S. K.; Liu, G.; Hung, T. T.; Schneewind, O. *Science* **1999**, *285*, 760–763.
(5) Ton-That, H.; Liu, G.; Mazmanian, S. K.; Faull, K. F.; Schneewind, O. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 12424–12429.
(6) Cossart, P.; Jonquieres, R. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5013–5015.
(7) Mazmanian, S. K.; Liu, G.; Jensen, E. R.; Lenoy, E.; Schneewind, O. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5510–5515.
(8) Bierne, H.; Mazmanian, S. K.; Trost, M.; Pucciarelli, M. G.; Liu, G.; Dehoux, P.; European Listeria Genome Consortium; Jänsch, L.; Garcia-del Portillo, F.; Schneewind, O.; Cossart, P. *Mol. Microbiol.* **2002**, *43*, 869−881.
(9) Suree, N.; Jung, M. E.; Clubb, R. T. *Mini-Rev. Med. Chem.* **2007**, *7*, 991–1000.
(10) Maresso, A. W.; Schneewind, O. *Pharmacol. Rev.* **2008**, *60*, 128–141.
(11) Suree, N.; Yi, S. W.; Thieu, W.; Marohn, M.; Damoiseaux, R.; Chan, A.; Jung, M. E.; Clubb, R. T. *Bioorg. Med. Chem.* **2009**, *17*, 7174–7185.
(12) Frankel, B. A.; Kruger, R. G.; Robinson, D. E.; Kelleher, N. L.; McCafferty, D. G. *Biochemistry* **2005**, *44*, 11188–11200.
(13) Huang, X. Y.; Aulabaugh, A.; Ding, W. D.; Kapoor, B.; Alksne, L.; Tabei, K.; Ellestad, G. *Biochemistry* **2003**, *42*, 11307–11315.
(14) Bentley, M. L.; Lamb, E. C.; McCafferty, D. G. *J. Biol. Chem.* **2008**, *283*, 14762–14771.
(15) Frankel, B. A.; Tong, Y.; Bentley, M. L.; Fitzgerald, M. C.; McCafferty, D. G. *Biochemistry* **2007**, *46*, 7269–7278.
(16) Marraffini, L. A.; Ton-That, H.; Zong, Y. N.; Narayana, S. V. L.; Schneewind, O. *J. Biol. Chem.* **2004**, *279*, 37763–37770.
(17) Ton-That, H.; Mazmanian, S. K.; Alksne, L.; Schneewind, O. *J. Biol. Chem.* **2002**, *277*, 7447–7452.
(18) Weiner, E. M.; Robson, S.; Marohn, M.; Clubb, R. T. *J. Biol. Chem.* **2010**, *285*, 23431–23441.
(19) Suree, N.; Liew, C. K.; Villareal, V. A.; Thieu, W.; Fadeev, E. A.; Clemens, J. J.; Jung, M. E.; Clubb, R. T. *J. Biol. Chem.* **2009**, *284*, 24465–24477.
(20) Race, P. R.; Bentley, M. L.; Melvin, J. A.; Crow, A.; Hughes, R. K.; Smith, W. D.; Sessions, R. B.; Kehoe, M. A.; McCafferty, D. G.; Banfield, M. J. *J. Biol. Chem.* **2009**, *284*, 6924–6933.
(21) Zong, Y. N.; Bice, T. W.; Ton-That, H.; Schneewind, O.; Narayana, S. V. L. *J. Biol. Chem.* **2004**, *279*, 31383–31389.
(22) Ilangovan, U.; Ton-That, H.; Iwahara, J.; Schneewind, O.; Clubb, R. T. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 6056–6061.
(23) Rawlings, N. D.; Barrett, A. J. In *Proteolytic Enzymes: Serine and Cysteine Peptidases*; Academic Press Inc: San Diego, CA, 1994; Vol. 244, pp 461−486.
(24) Otto, H. H.; Schirmeister, T. *Chem. Rev.* **1997**, *97*, 133–171.
(25) Connolly, K. M.; Smith, B. T.; Pilpa, R.; Ilangovan, U.; Jung, M. E.; Clubb, R. T. *J. Biol. Chem.* **2003**, *278*, 34061–34065.
(26) Tian, B. X.; Eriksson, L. A. *Proteins: Struct., Funct., Bioinf.* **2011**, *79*, 1564–1572.
(27) Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. *Science* **2004**, *303*, 186–195.
(28) Senn, H. M.; Thiel, W. *Curr. Opin. Chem. Biol.* **2007**, *11*, 182–187.
(29) Shaik, S.; Cohen, S.; Wang, Y.; Chen, H.; Kumar, D.; Thiel, W. *Chem. Rev.* **2010**, *110*, 949–1017.
(30) Mladenovic, M.; Fink, R. F.; Thiel, W.; Schirmeister, T.; Engels, B. *J. Am. Chem. Soc.* **2008**, *130*, 8696–8705.
(31) Arad, D.; Langridge, R.; Kollman, P. A. *J. Am. Chem. Soc.* **1990**, *112*, 491–502.
(32) Harrison, M. J.; Burton, N. A.; Hillier, I. H. *J. Am. Chem. Soc.* **1997**, *119*, 12285–12291.
(33) Ma, S.; Devi-Kesavan, L. S.; Gao, J. *J. Am. Chem. Soc.* **2007**, *129*, 13633–13645.
(34) Miscione, G. P.; Calvaresi, M.; Bottoni, A. *J. Phys. Chem. B* **2010**, *114*, 4637–4645.
(35) Han, W. G.; Tajkhorshid, E.; Suhai, S. *J. Biomol. Struct. Dyn.* **1999**, *16*, 1019–1032.
(36) Byun, K.; Gao, J. L. *J. Mol. Graphics* **2000**, *18*, 50–55.
(37) Strajbl, M.; Florian, J.; Warshel, A. *J. Phys. Chem. B* **2001**, *105*, 4471–4484.
(38) *MOE 2010.10*; Chemical Computing Group Inc.: Montreal, Quebec, Canada; http://www.chemcomp.com.
(39) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G. M.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; et al. *J. Comput. Chem.* **2003**, *24*, 1999–2012 (for the complete citation, see Supporting Information S8).
(40) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
(41) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
(42) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
(43) Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
(44) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
(45) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695–1697.
(46) Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
(47) Vreven, T.; Byun, K. S.; Komaromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comput.* **2006**, *2*, 815–826.
(48) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
(49) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
(50) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
(51) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.;

Petersson, G. et al. *Gaussian 09*, revision B.01; Gaussian, Inc.: Wallingford, CT, 2009 (for the complete citation, see Supporting Information S8).

(52) Bash, P. A.; Field, M. J.; Davenport, R. C.; Petsko, G. A.; Ringe, D.; Karplus, M. *Biochemistry* **1991**, *30*, 5826–5832.

(53) Dinner, A. R.; Blackburn, G. M.; Karplus, M. *Nature* **2001**, *413*, 752–755.

(54) Tuttle, T.; Thiel, W. *J. Phys. Chem. B* **2007**, *111*, 7665–7674.

(55) Labute, P. *J. Comput. Chem.* **2008**, *29*, 1693–1698.

(56) Lewis, S. D.; Johnson, F. A.; Shafer, J. A. *Biochemistry* **1981**, *20*, 48–51.

(57) Johnson, F. A.; Lewis, S. D.; Shafer, J. A. *Biochemistry* **1981**, *20*, 44–48.

(58) Lewis, S. D.; Johnson, F. A.; Shafer, J. A. *Biochemistry* **1976**, *15*, 5009–5017.