Modeling the Protonation States of the Catalytic Aspartates in β-Secretase

Ramkumar Rajamani and Charles H. Reynolds*

Johnson & Johnson Pharmaceutical R&D, Welsh and McKean Roads, PO Box 776, Spring House, Pennsylvania 19477

Received March 4, 2004

β-Secretase (BACE) is a critical enzyme in the production of β-amyloid, a protein that has been implicated as a potential cause of Alzheimer's disease (AD). There are two aspartic acid residues (Asp 32 and Asp 228) present in the catalytic region of BACE that can adopt multiple protonation states. The protonation state and precise location of the protons for these two residues, particularly in the presence of an inhibitor, are subjects of great interest since they have a direct bearing on the mechanism of aspartyl proteases and efforts to model β-secretase. We have carried out full liner-scaling quantum mechanical (QM) calculations that include Poisson−Boltzmann solvation in order to identify the preferred protonation state and proton location in the presence and absence of an inhibitor. These calculations favor the monoprotonated state in the presence of ligand, and di-deprotonated state in the absence of ligand. Further the proton in the monoprotonated state is located on the inner oxygen of Asp 228. These results have implications for the catalytic mechanism of BACE and related aspartyl proteases. They also provide a reference state for the protein in structure-based modeling studies of this therapeutically important target.

Introduction

β-Secretase (BACE) cleaves the amyloid precursor protein (APP) into two fragments, APPsβ and C99. C99 is further cleaved by γ-secretase to form the 40 or 42 residue β-amyloid peptide (Aβ).[1−3] Aβ is the primary constituent of amyloid plaques found in the brain of Alzheimer's disease (AD) victims.[4] Thus, inhibition of β-secretase has been implicated as a promising therapeutic target in the treatment of AD.[5] Tang, Ghosh, and co-workers have reported a number of potent peptidic transition state inhibitors,[6] a common strategy for aspartyl proteases,[7] along with crystal structures for two inhibitors (OM99-2, OM00-3) bound to the enzyme.[6] The first crystal structure revealed an active site that is located in a cleft between the C and N terminal lobes and is capped with a flexible hairpin loop ("flap"). Two catalytically important aspartates (Asp 32, Asp 228) lie in the center of the active site cleft. Mutational studies on HIV-1 protease, a member of the aspartyl protease family, have shown a complete loss in activity with the removal of either carboxyl.[8] It is presumed that Asp 32 and Asp 228 are likewise critical for catalytic activity in BACE. The position of the hydroxyl transition state mimic in the crystal structure indicates a network of hydrogen bonding interactions with the catalytic aspartates. Although the crystal structure provides a wealth of structural information, the protonation state or precise location of the protons on the ionizable groups cannot be determined by X-ray crystallography, at least at the resolution commonly employed. The protonation state is of significant interest since it is a key factor in the mechanism of aspartyl proteases, such as BACE, and may play an important role in modeling the binding affinities of putative inhibitors. The catalytic mechanism of the aspartyl protease family has been a subject of great scientific interest for decades, and has recently assumed greater importance because of the significance of aspartyl protease inhibitors therapeutically.

One of the most extensively studied enzymes in the aspartyl protease family is HIV-1 protease (HIV-1 PR). HIV-1 PR crystallizes as a homodimer with an active site containing two catalytically important aspartates (Asp 25, Asp 25′) that form a symmetric dyad. It has been proposed that the catalytic aspartates have an overall −1 charge, i.e., one aspartate is protonated and the other is deprotonated.[9,10] Further, a bound water molecule is involved in most proposed mechanisms for proteolysis.[11] This arrangement is common to all known aspartyl proteases. A number of mechanisms have been proposed for the proteolytic cleavage of peptides by aspartyl proteases. These mechanisms are based on X-ray crystal structures, kinetics, and general physical organic principles. In general, most mechanisms that have been put forward are based on the general acid−general base theory, and suggest formation of a tetrahedral intermediate as a key step in the cleavage of the peptide bond. One of the critical issues in any proper description of the mechanism is the identification of the protonation states of the catalytic aspartates, and the location of the protons. To this end, many experimental approaches have been employed. These include analysis of crystal structures with bound ligands (natural substrates and transition state analogues),[12−21] pH rate behavior,[9,10] solvent kinetic isotope effects,[22,23] and ^{13}C NMR labeling studies.[24] In addition theoretical/computational methods have also been employed to study this problem. These computational studies have mostly consisted of classical force field calculations with various sampling approaches ranging from simple energy minimization to more complex molecular dynamics or Monte Carlo simulations.[25−37] A few studies have used quantum methods, which would presumably be better able to address this problem that inherently involves bond

* Author to whom correspondence should be addressed. E-mail: CReynol1@prdus.jnj.com. Tel: 215 628 5675. Fax: 215 628 4985.

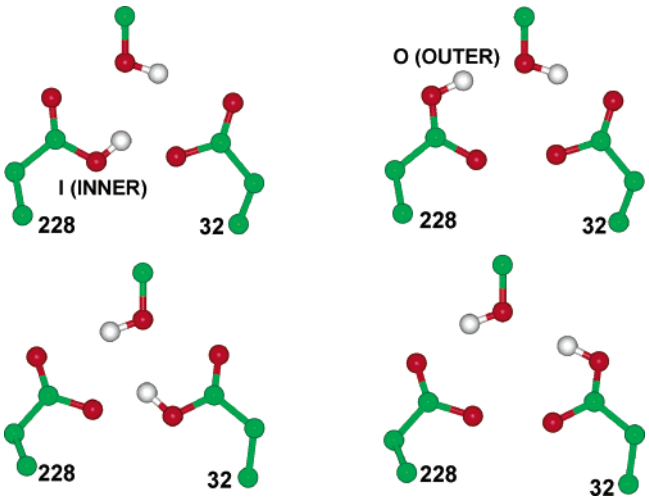
10.1021/jm049817j CCC: $27.50
Published on Web 09/15/2004

breaking/making. However, in order to make these calculations tractable, severe approximations have been made to reduce the size of the molecular system.

The experimental and theoretical studies of the aspartate protonation states have not yielded a consensus on either the protonation state or the precise location of the protons. Using estimates of the pK_a of aspartates and glutamates with a bound nonpeptidomimetic substrate, Yamazaki et al. suggested that the neutral diprotonated state was preferred in the presence of inhibitor.[38] A Monoprotonated form with Asp 32 charged and Asp 228 protonated on the outer oxygen has been proposed by Coates et al. on the basis of high-resolution crystal structures and neutron diffraction studies of endothiapepsin complexes.[19,20] Kinetic isotope studies on BACE using OM99 and statine inhibitors[23] show a preference for the monoprotonated form in the presence of the inhibitors. This work has been used to propose the formation of a short strong hydrogen bond (SSHB) on inhibitor binding. The di-deprotonated form has been reported to be favored in the presence of a reduced amide transition state isostere.[9,39] Using ab initio molecular dynamics simulations Piana et al. suggested the monoprotonated form with a low barrier hydrogen bond (LBHB).[33] Extending the LBHB proposal, Northrop suggested that synchronous intermolecular cyclic proton transfers explain the mechanism of proteolysis rather than a conventional stepwise mechanism.[40] In a more recent computational study of pepstatin, Piana et al. estimated the relative energies of all the protomers of both mono- and diprotonated forms using density functional theory.[34] The inner oxygen of Asp 228 was identified as the preferred proton location for the monoprotonated state in the presence of an inhibitor. However, ab initio molecular dynamics studies and the computed ^{13}C NMR data led the authors to conclude that the diprotonated form is preferred in the presence of the inhibitor. A recent molecular dynamics study suggested the monoprotonated state with the outer oxygen of Asp 32 being protonated.[41] Thus a variety of possible protonation states have been observed experimentally as well as computationally. Given the somewhat contradictory experimental and computational studies outlined above, it is clear that the preferred protonation state and proton locations for aspartyl proteases such as BACE are far from being settled.

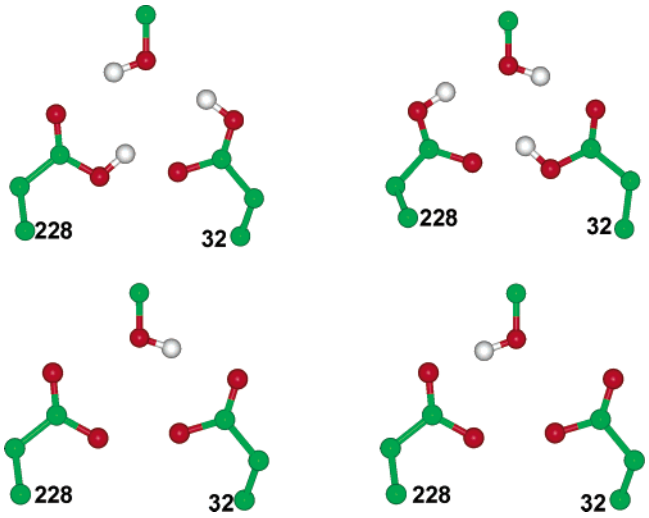
We have undertaken a quantum mechanical study of the protonation state, and preferred locations for the catalytic protons in BACE. The use of a semiempirical quantum method in this study allows us to explicitly include effects that are neglected in classical simulations such as polarization and charge transfer. In addition, a quantum treatment should provide improved electrostatics over fixed point-charges, including solvation effects, and makes it possible to study bond breaking/making processes directly. The semiempirical linear-scaling quantum method[42] employed in this study is fast enough to enable electronic structure calculations for large biomolecules, such as the active site of BACE. Three individual scenarios are probed: where the total charge of the catalytic aspartates is neutral (diprotonated), anionic (monoprotonated), or dianionic (di-deprotonated). In each case all possible positions

Scheme 1. Monoprotonated Tautomers of the Catalytic Aspartates of BACE Considered in This Study[a]



[a] Nomenclature: 228o = proton present on the outer oxygen of Asp 228.

Scheme 2. Proton Locations Considered for the Diprotonated and Di-deprotonated Form[a]



[a] Nomenclature: 32i228o = proton present on the inner oxygen of Asp 32 and outer oxygen of Asp 228. t32 = Substrate hydroxyl points toward Asp 32.

(Schemes 1 and 2) for the protons have been examined using linear-scaling quantum mechanical calculations. The preferred protonation state was determined using propionic acid as a reference acid. Propionic acid was chosen because it has a pK_a similar to that of aspartic acid, and allows us to construct a series of homodesmic reactions.

Computational Details

Linear Scaling Quantum Approach. The DIVCON software package, a linear-scaling quantum mechanical program, was used to perform both single point and energy minimization calculations.[42] The ability to describe large biological macromolecules using quantum mechanics has been made possible by recent developments in reducing the scaling (relative to the number of atoms) of quantum mechanical calculations. DIVCON is based on the divide and conquer (DC) paradigm first described by Yang et al.[43] Other approaches based on density matrix minimization (DMM)[44−47] and localized molecular orbitals (LMO)[48,49] have also been utilized to achieve linear scaling. Briefly, the divide and conquer approach involves

dividing the molecular system of interest into small overlapping subsystems. The subsystems are described in a layered manner containing a core region and surrounding buffers. In general, each core is made up of all atoms of an individual residue, and the inner buffer is constituted by the atoms of all residues that fall within a predefined buffer radius from every atom of the core region except those of the core. All residues that fall within the outer radius cutoff but do not belong to either the core or inner buffer constitute the outer buffer region. The process is repeated until the entire molecular system is subdivided. To speed up the calculations, the density matrices for the local subsystems are first computed and then combined to give the global density matrix. The inclusion of buffer zones primarily addresses truncation effects on the formation of the global density matrix. All calculations were performed using the AM1 Hamiltonian. Solvation was included using the Delphi Poisson−Boltzmann (PB) program[50,51] combined with AM1/CM2 classical charges.[52,53] For a detailed description of the DIVCON program and the divide and conquer linear-scaling quantum approach please refer to Merz et al.[54,55]

Protein Preparation. The BACE crystal structure coordinates with bound OM99-2 was the starting point for these calculations.[6,56] A hydroxyethylene model substrate was constructed by deleting all residues from the crystal structure of OM99-2 except the two residues flanking either side of the transition state mimic. The side chains of the remaining residues were then mutated into alanine, and the ends were capped with methyl groups. All waters, as defined in the crystal structure, were retained. The PPREP module available from Schrodinger[57] was used to define the protonation states of all ionizable residues. Atoms outside a cavity of 17 Å from the ligand were neutralized for the full protein calculations. The positions of the hydrogens were then constructed using the hydrogen atom placement utility in MAESTRO (a commercial modeling package provided by Schrodinger Inc.).[57] All of the heavy atoms were held fixed while the hydrogen atoms of the prepared protein along with the methyl group caps were minimized using the OPLS-AA force field[58] as implemented in Maestro. A gradient convergence criterion of 0.05 was used. This was followed by a refinement protocol where a series of restrained, partial minimizations were performed using progressively weaker restraints on heavy atoms until the average RMSD of the non-hydrogen atoms reached a default value of 0.3 Å. The converged structure was taken as the starting point for construction of all combinations of the proton positions. The complete system consisted of 6040 atoms. The catalytic aspartates and the transition state mimic hydroxyl were minimized using the OPLS-AA parameters, and the resultant structure was used to perform single point calculations using DIVCON. Computation of single point energy (1SCF) for the 6040 atoms using the AM1 Hamiltonian required approximately 20 h of CPU time on an SGI Octane-2 (RS12000). Thus minimization, especially when run in a serial mode, would be quite expensive. In order to make this problem more tractable, a truncated protein system was used for this study.

The truncated system was constructed by deleting all residues that fell outside a 15 Å cutoff of any atom of the ligand (Figure 1a,b). The dangling terminals were capped with methyl groups. All of the deleted residues were already neutralized by the protein preparation described above. Crystallographic waters that were present within 6 Å of the ligand were retained, and the rest were deleted. A total of 4 waters were retained (total number of atoms = 1477). In order to assess the effect of truncation on the computed relative energies, the monoprotonated state was studied using both the full enzyme and the truncated protein structure. The relative energies of different monoprotonated states for the complete and the truncated enzyme were found to be very close, i.e., the average relative error over the four states is 0.77 kcal/mol. On the basis of this comparison, the truncated version was employed for all further studies.

QM Minimization. The system was minimized at the AM1 level using the LBGFS algorithm in DIVCON. DIVCON was

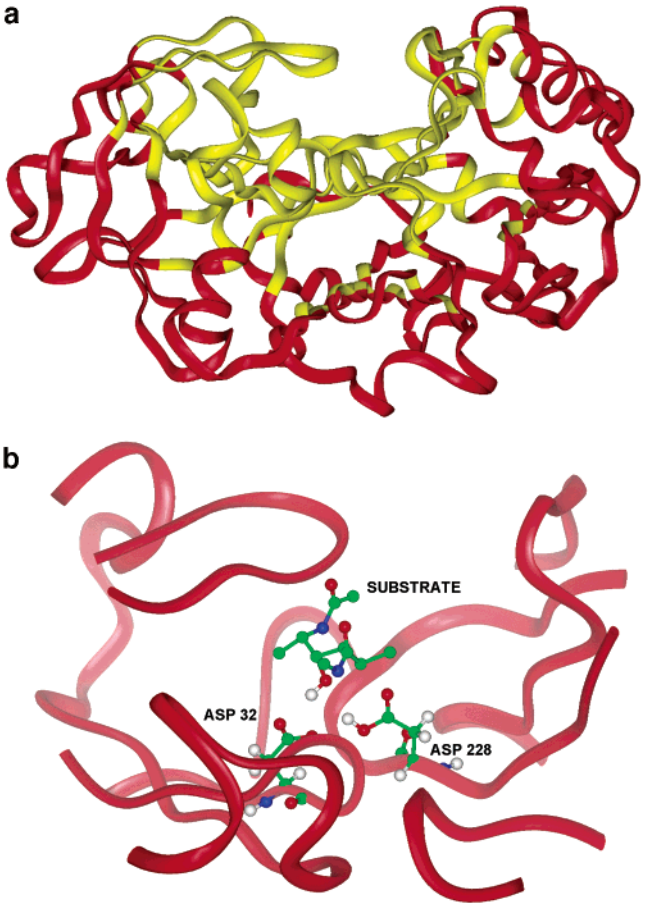


Figure 1. (a) A ribbon representation of the crystal structure of BACE. The region in yellow defines the 15 Å cutoff used to truncate the protein in this study. (b) A ribbon representation of the region close to the active site. The catalytic aspartates and the hydroxyethylene based model inhibitor are represented using ball and stick drawings. Of the represented catalytic aspartates the deprotonated form is Asp 32 and the protonated form is Asp 228.

used to selectively optimize the residues Asp 32, Asp 228, and hydroxyl of the substrate, while all remaining protein atoms were held fixed. The recommended cutoffs of 4.5 Å and 2 Å for the dual buffer were utilized in this study (details provided in the linear scaling quantum approach section). A total of 95 subsystems were defined inclusive of the waters and the bound ligand. Optimization criteria with the maximum component of gradient GTEST = 0.5 kcal/(mol Å) was utilized during minimization.[55] A total of 75−125 steps were required to obtain converged minimized structures, depending on the system considered. The minimized structures were used as input for single point calculations that included a Poisson−Boltzmann solvation term. The PB calculations had an inside dielectric of 1 and outside dielectric of 80. CM2 charges[52,53] were used in the PB calculations. Work published by Schroeder et al.[59] showed that AM1 is capable of reproducing proton transfer energies, such as those being evaluated in this study. Their work supports using AM1 with PB solvation to evaluate the protonation states in BACE.

Reaction Scheme. In order to determine the preferred protonation state, the three most probable protonation states were compared energetically. These are monoprotonated (charge = −1), diprotonated (charge = 0), and di-deprotonated (charge = −2). The three states differ in the net charge and hence cannot be reliably compared directly. To facilitate comparison, a series of homodesmic reactions were constructed using propionic acid as a reference compound.[60] Propionic acid was chosen because of its size and pK_a similarities to aspartic acid (aspartic acid pK_a = 3.8−4.4, propionic acid = 4.8). One representative structure (only the lowest energy conformer)

Scheme 3. Reaction Scheme Used for Comparison of Protonation States[a]

Di-Deprotonated	→	Mono-protonated	→	Di-protonated
+		+		+
$2(CH_3CH_2COOH)$		$CH_3CH_2COO^-$		$2(CH_3CH_2COO^-)$
		+		
		CH_3CH_2COOH		
I		II		III

[a] I represents di-deprotonated state, II represents monoprotonated state, and III represents diprotonated state.

Table 1. Relative Energies (kcal/mol) for the Monoprotonated, Diprotonated, and Di-deprotonated States

	substrate bound		free enzyme	
	gas phase	solvated (aq)	gas phase	solvated (aq)
monoprotonated				
228i	0	0	0	0
228o	16.7	13.9	21.9	9.8
32i	23.4	30.5	21.2	30.4
32o	31.7	40.7	32.0	33.3
diprotonated				
228i32o	0	0	0	0
228o32i	18.5	16.9	20.7	10.6
di-deprotonated				
t228	7.5	0.5	4.8	0
t32	0	0	0	13.2

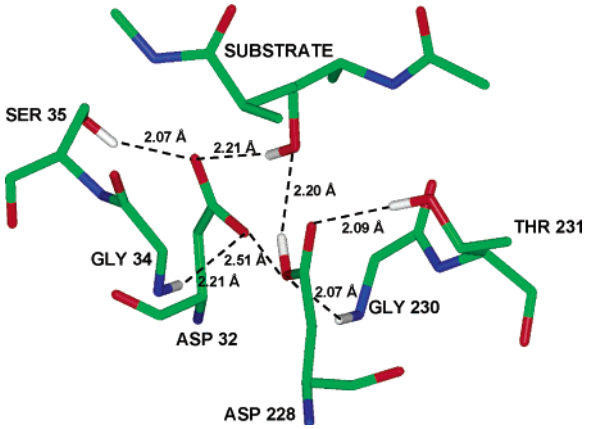


Figure 2. A snapshot of key residues in the active site of the monoprotonated form (228i) taken from the AM1 optimized geometry.

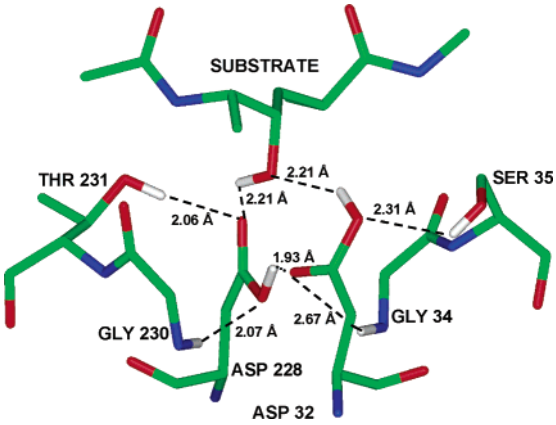


Figure 3. A snapshot of key residues in the active site of the diprotonated form (228i32o) taken from the AM1 optimized geometry.

was used for each of the three protonation states. A representation of the comparison scheme is presented in Scheme 3. The calculated relative energies for the model reactions were used to identify the preferred protonation state.

Results

Monoprotonated Aspartates. The side chains of the aspartic dyad and the hydroxyl of the bound substrate were minimized for all possible tautomers (Scheme 1) of the monoprotonated state. The computed relative energies (Table 1) indicate that protonation of the inner oxygen of Asp 228 is preferred over all other proton locations. Placement of the proton on either oxygen of Asp 32 was found to be higher in energy, with the outer oxygen of Asp 32 being most unfavorable. The optimized geometry for 228i and selected interactions between the catalytic aspartates and other active site residues are highlighted in Figure 2. For 228i, the interatomic distance between the proton located on the inner oxygen of Asp 228 and the inner oxygen of Asp 32 (inner O^{32}—O^{228} distance of 2.51 Å) was 1.91 Å.

The calculations described above were carried out with the model inhibitor bound to the protein active site. It is certainly possible that the ligand influences the preferred proton location, as well as overall protonation state. The role of the ligand in influencing proton location was examined by repeating these calculations for the apo-enzyme (i.e., without the ligand). The relative energies in Table 1 show that presence or absence of ligand has no effect on proton location. This is true in the gas phase and with aqueous solvation (Table 1). Any solvation effects may be mitigated by the extensive hydrogen bond network that surrounds the catalytic aspartates in the active site.

Diprotonated Aspartates. Two specific combinations were explored for the diprotonated state (Scheme 2). Of the two combinations, the lower energy conformer had the inner oxygen of Asp 228 and the outer oxygen of Asp 32 protonated. A snapshot of the minimized geometry is shown in Figure 3. Inspection of the structure revealed a strong network of hydrogen bonds between the substrate hydroxyl and the catalytic aspartates. A similar network has been reported for other aspartyl proteases.[34] Hydrogen bond distances for O^{32i}—O^{228i} of 2.7 Å and angles for O^{228i}—H···O^{32i} of 133.0° are measured between the inner oxygens of Asp 32 and Asp 228. The substrate hydroxyl interacts with Asp 228 with an O^{228o}—O^{OH} distance of 2.9 Å and O^{OH}—H···O^{228o} angle

Table 2. Dihedral Angles (deg) between the Oxygens of the Catalytic Aspartates

state	dihedral measure (deg) (O^{32o}–O^{32i}–O^{228i}–O^{228o})			
monoprotonated	228i	228o	32i	32o
	−27	−30	−46	−56
diprotonated	228i32o			
	−56.0			
di-deprotonated	t228	t32		
	−56.0	−46.0		

of 127.0°. This is compared to a O^{32o}–O^{OH} distance of 3.1 Å and O^{OH}–H···O^{32o} angle of 146.0° for Asp 32. The dihedral measured between the aspartyl oxygens of −56.0° (Table 2) shows a considerable deviation from coplanarity and is close to that observed in the crystal structure (−75°).[6]

The proton positions of the catalytic aspartates are supported by interactions between the hydroxyl of Ser 35 and the outer oxygen of Asp 32 $d(O^{35}–O^{32o})$ of 2.98 Å, and between the hydroxyl of Thr 231 and the outer oxygen of Asp 228 $d(O^{228o}–O^{231})$ of 3.00 Å. It has been proposed that additional interactions with the backbone amides of Gly 34 and Gly 230 also contribute to enforcing a coplanar arrangement of the catalytic aspartates. Hydrogen bonding interactions with $d(NH^{231}\cdots O^{228i}) = 2.92$ Å and $d(NH^{34}\cdots O^{32i}) = 3.5$ are observed in the computed structures.

Inclusion of solvation had no effect on the location of the protons. Hence the preferred combination of proton locations is predicted to be 228i32o in the gas phase as well as solution. Again the influence of the ligand on the location of the protons was examined by repeating the calculations without the ligand. As was true for the monoprotonated state, the ligand has no effect on the preferred proton locations.

Di-deprotonated Aspartates. Since there are no protons on the catalytic aspartates, the conformation of the hydroxyl and resulting direction of the hydrogen bond between the hydroxyl and charged aspartates defines the two structures sampled for the di-deprotonated state (t228 and t32). This is shown in Scheme 2. In the gas phase, the conformer where the hydroxyl hydrogen is pointed toward Asp 32 was found to be most favorable. Geometry optimization of the di-deprotonated states was performed from the initial construct derived by deleting a proton from the energy minimized structure of monoprotonated 228i. This starting structure moves spontaneously out of the coplanar arrangement with respect to the catalytic aspartates, and in general agreement with the crystal structure of the BACE/OM99 complex. This compares with the more twisted aspartate conformations that were observed for other protonation states, and the experimental structure. For example the computed dihedral angles (O^{32o}–O^{32i}–O^{228i}–O^{228o}) for the monoprotonated state are −46.4° and −56.0° for the 32i and 32o structures, respectively. Interestingly, inspection of the structures shows that very little change occurs in the position and orientation of Asp 32, but there is a large change in Asp 228. In the published crystal structure there is very little room between the flap and Asp 32, while Asp 228 is more exposed to solvent. The different spatial arrangement and environment for the two catalytic aspartates may explain the observed selective motion of Asp 228.

Solvation attenuates the relative energies of the two hydroxyl rotamers significantly. On inclusion of PB water the two rotamers become energetically equivalent. A snapshot of the active site of the minimized structure for t32 is shown in Figure 4. The large difference in the relative energies for the gas and solution phase estimates may be understood from the positions of the catalytic Asp's (Figure 5). In the t228 case, a bifurcated hydrogen bond is formed between Asp 228 and the hydroxyls of the ligand and Ser 231. However, in the case of t32, the charges on the two aspartyls are fully solvated through hydrogen bonding interactions thus providing a preference in the gas phase calculations.

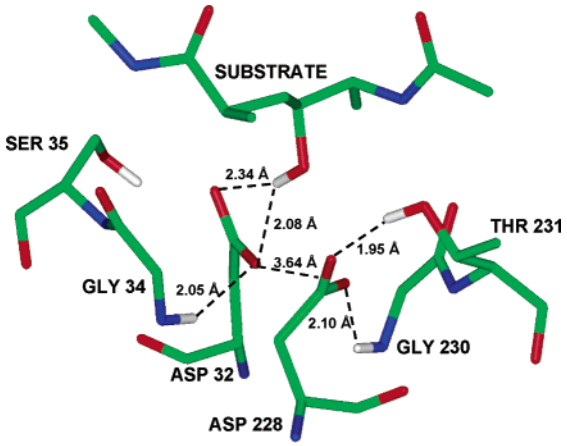


Figure 4. A snapshot of key residues in the active site of the di-deprotonated form (t32) taken from the AM1 optimized geometry.

Relative Stability. The relative stabilities of the three protonation states were compared using propionic acid as the reference state. A net charge of −2 was maintained for all three steps in the homodesmic reactions. Only the minimium energy conformations for the individual states were compared. The relative energies of the three possible protonation states are given in Table 3. Clearly the monoprotonated form is preferred over the di-deprotonated form with a difference in relative energy of ∼30 kcal/mol. By comparison, the diprotonated state is destabilized only by 3.5 kcal/mol. Thus the calculations suggest that the monoprotonated state (charge = −1) is preferred for BACE. This is consistent with most previous literature on the catalytic mechanism of aspartyl proteases. The influence of the ligand on protonation state was probed by comparing the preferences in the ligand free case. The di-deprotonated state (charge = −2) was the most stable protonation state when the ligand was absent. The monoprotonated state is 13.2 kcal/mol higher in energy, and the diprotonated form is 21.1 kcal/mol higher in energy. Thus we predict that the apo-protein (with no ligand bound) should be di-deprotonated (charge = −2), but the aspartates should adopt the monoprotonated state (charge = −1) upon ligand binding. This has interesting implications for the basic mechanism of

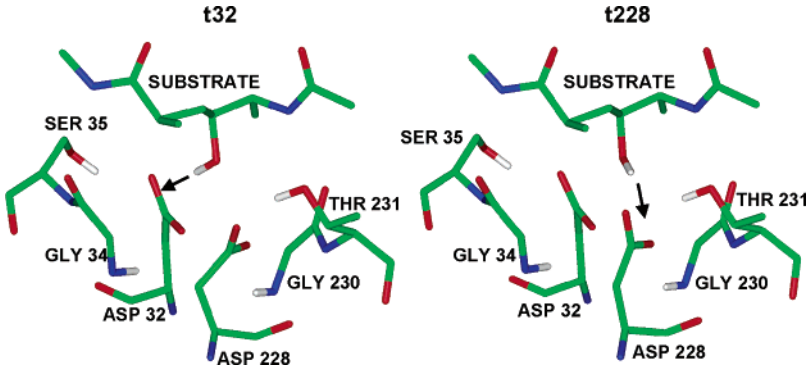


Figure 5. Comparison of the active site residue positioning of di-deprotonated form t32 and t228 taken from AM1 minimized geometries. The figure shows a large deviation in Asp 228 orientation depending upon the direction of the hydroxyl hydrogen.

Scheme 4. Schematic Representation of the General Acid−Base Mechanism for Proteolysis by Aspartyl Proteases

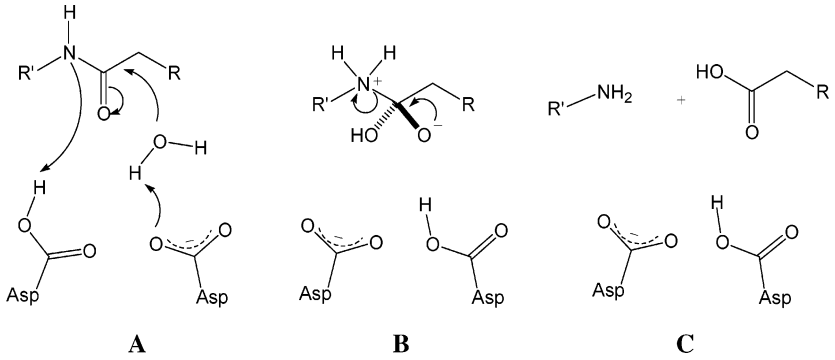


Table 3. Relative Energies in kcal/mol for the Reaction Scheme Outlined in Scheme 3

state	substrate bound	free enzyme
I	30.7	0.0
II	0.0	13.2
III	3.5	21.1

aspartyl proteases, and for efforts to model protein−ligand binding affinities in BACE.[61]

Discussion

Protonation State. The homodesmic reactions indicate that the most stable protonation state, at neutral pH and in the presence of a ligand, is the monoprotonated (charge = −1) state. This is consistent with several of the experimental studies for BACE and other aspartyl proteases. It should be noted, however, that the diprotonated (charge = 0) state is only computed to be 3.5 kcal/mol higher in energy, and might be accessible at lower pH. This is a significant consideration since the enzyme is most active under acidic conditions. It also means that care should be exercised when comparing the calculations with experimental studies to take into account the pH of the experiment.

One of the most interesting conclusions from this work is the prediction that the protonation state changes depending on the presence of a ligand. When the inhibitor is removed, the most stable protonation state is the di-deprotonated (charge = −2) state. Other calculations have shown a ligand dependence on the protonation state. For example, an ab initio molecular dynamics study of the HIV-1 protease and pepstatin complex by Piana et al.[34] concluded that the diprotonated form was preferred when the ligand is bound and the monoprotonated state is preferred when the ligand is absent.[33] While their work is consistent with ours in predicting a change in protonation state upon binding, the protonation states predicted by the ab initio molecular dynamics study do not agree with our results. This may be because the ab initio calculations used a severely approximated protein structure (i.e., only included the catalytic aspartates (25, 25′) and glycine (26, 26′)) and neglected solvent). There is precedence for the di-deprotonated state in the apo-enzyme. Wang and Kollman proposed the di-deprotonated state for HIV protease in the absence of any ligand[35] on the basis of an extensive molecular dynamics simulation. However, this is a difficult problem to address using a classical valence force field since bonds are being made and broken, and there are large changes in the total charge of the system.

Location of the Proton. The coplanar arrangement of the carboxyl oxygens in the catalytic Asp's has been used as a measure for the preferred proton location. In our study this effect is greatest on protonating the inner oxygen of Asp 228. The coplanarity between the catalytic aspartates was measured through the dihedral formed between the oxygens of the dyad. For the preferred conformation a dihedral of ∼−27° was measured compared to ∼−75° obtained from the crystal structure. If structural similarity between the optimized geometry and the crystal geometry was used to predict the proton location, the Asp 32o form would be the most preferred as the measured dihedral of ∼−56° compares closest to that observed for the crystal structure.

Mechanistic Significance. The identification of the monoprotonated form as the preferred protonation state in the presence of a transition state inhibitor supports the chemical mechanism where one catalytic aspartate

acts as a general acid and the other acts as a general base. Upon forming the tetrahedral intermediate, the proton locations are reversed (Scheme 4). If we assume that the transition state mimic binds in the same manner as the tetrahedral intermediate, then Asp 228 would be protonated and Asp 32 deprotonated. Previous studies using neutron diffraction studies on endothiapepsin complexed with a hydroxyethylene inhibitor provided evidence for the monoprotonated state.[19] Subsequent collapse of the tetrahedral intermediate and regeneration of the native enzyme state completes the mechanism. Our calculations provide further evidence that the monoprotonated state is most likely the catalytically active state, and the state of the enzyme in the tetrahedral intermediate. They also indicate that the di-deprotonated state (charge = −2) is most stable in the apo-enzyme.

Conclusions

This is the first quantum mechanical modeling study of the protonation state and proton locations for the catalytic aspartates in any aspartyl protease, where the protein active site was included explicitly. These calculations give the monoprotonated form with the proton located on the inner oxygen of Asp 228 as the most stable protonation state in the presence of bound ligand. This is likely also the state present on the formation of the tetrahedral intermediate during proteolysis, and is consistent with proposed catalytic mechanisms for other aspartyl proteases. Given the small energy difference between the monoprotonated and diprotonated states, the preferred state might be expected to change depending on the nature of the ligand and the pH. At low pH the diprotonated (charge = 0) state would become more energetically accessible. When the inhibitor is absent, our calculations find that the di-deprotonated (−2) state is preferred by a wide margin. Thus the protonation state is predicted to change upon ligand binding, with one of the catalytic aspartates (Asp 228) abstracting a proton from the surrounding aqueous environment. This may represent a key step in proteolysis.

These results have significant implications for drug discovery, since they provide an assessment of the protonation state in BACE that can be used to model prospective inhibitors. As such this work provides a starting point for structure-based drug design efforts aimed at designing new, more potent, ligands for this therapeutically important enzyme target. They also highlight the expanding role of quantum methods in biology and medicinal chemistry. The ability to study protein catalytic sites using quantum mechanics provides for inclusion of all electronic effects such as polarization and charge transfer explicitly, and should be a fundamentally more rigorous description of the enzyme environment. Further, quantum methods allow straightforward calculation of reactions that involve bond making and breaking, something that is difficult with classical valence force-field models.

Acknowledgment. R.R. thanks the Johnson and Johnson Corporate Office of Science and Technology for his Excellence in Science Postdoctoral Fellowship. We thank Ellen Baxter, Douglas Brenneman, Kelly Conway, Allen Reitz, Brett Tounge, Jian Li, and the Spring House CNS Team for their helpful comments and insights. We also thank Kaushik Raha and Lance Westerhoff of The Pennsylvania State University and QuantumBio Inc. for their assistance.

References

(1) Vassar, R.; Citron, M. Aβ-generating enzymes: recent advances in β- and γ-secretase research. *Neuron* **2000**, *27*, 419−422.
(2) Selkoe, D. Alzheimer's disease: genes, proteins and therapy. *Physiol. Rev.* **2001**, *81*, 741−766.
(3) Roggo, A. Inhibition of BACE, a promising approach to Alzheimer. *Curr. Top. Med. Chem.* **2002**, *2*, 359−370.
(4) Hardy, J.; Allsop, D. Amyloid deposition as the central event in the aetiology of Alzheimer's disease. *Trends Pharmacol.* **1991**, *12*, 383−388.
(5) Citron, M. Emerging Alzheimer's disease therapies: inhibition of β-Secretase. *Neurobiol. Aging* **2002**, *23*, 1017−1022.
(6) Hong, L.; Koelsch, G.; Lin, X.; Wu, S.; Terzyan, S.; Ghosh, A. K.; Zhang, X. C.; Tang, J. Structure of the Protease Domain of Memapsin 2 (β-Secretase) Complexed with Inhibitor. *Science* **2000**, *290*, 150−153.
(7) Leung, D.; Abbenante, G.; Fairlie, D. P. Protease Inhibitors: Current Status and Future Prospects. *J. Med. Chem.* **2000**, *43*, 305−341.
(8) Hussain, I.; Powell, D.; Howlett, D. R.; Tew, D. G.; Meek, T. D. et al. Identification of a Novel Aspartic Protease (Asp 2) as -Secretase. *Mol. Cell. Neurosci.* **1999**, *14*, 419−427.
(9) Hyland, L. J.; Tomaszek, J. T. A.; Meek, T. D. Human immunodeficiency virus-1 protease: II. Use of pH rate studies and solvent kinetic isotope effects to elucidate details of chemical mechanism. *Biochemistry* **1991**, *30*, 8454−8463.
(10) Ido, E.; Han, H.; Kezdy, F. J.; Tang, J. Kinetic studies of human immunodefeciency virus type I protease and its active-site hydrogen bond mutant A28S. *J. Biol. Chem.* **1991**, *266*, 24349−24366.
(11) Wlodawer, A.; Miller, M.; Jaskolski, M.; Sathyanarayana, B. K.; Baldwin, E.; Weber, I. T.; Selk, L. M.; Clawson, L.; Schneider, J.; Kent, S. B. H. Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. *Science* **1989**, *245*, 616−621.
(12) Bott, R.; Subramanian, E.; Davies, D. R. Three-dimensional structure of the complex of the *Rhizopus chinensis* carboxyl protienase and pepstatin at 2.5-Å resolution. *Biochemistry* **1982**, *21*, 6956−6962.
(13) James, M. N. G.; Sielecki, A. R.; Salituro, F.; Rich, D. H.; Hofmann, T. Conformational flexibility in the active site of aspartyl proteinases revealed by a pepstatin fragment binding to penicillopepsin. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 6137−6142.
(14) Foundling, S. I.; Cooper, J.; Watson, F. E.; Cleasby, A.; Pearl, L. H.; Sibanda, B. L.; Hemmings, A.; Wood, S. P.; Blundell, T. L.; Valler, T. L.; Kay, J.; Boger, J.; Dunn, B. M.; Leckie, B. J.; Jones, D. M.; Atrash, B.; Hallett, A.; Szelke, M. High resolution X-ray analysis of renin inhibitor-aspartic proteinase complexes. *Nature* **1987**, *327*.
(15) Suguna, K.; Padlan, E. A.; Smith, C. W.; Carlson, W. D.; Davies, D. Binding of a reduced peptide inhibitor to the aspartic proteinase from Rhizopus chinensis: Implications for a mechanism of action. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 7009−7013.
(16) Sali, A.; Veerapandian, B.; Cooper, J.; Foundling, S. I.; Hoover, D. J.; et al. High resolution X-ray diffraction study of the complex between endothiapepsin and an oligopeptide inhibitor: The analysis of inhibitor binding and description of the rigid body shift in enzyme. *EMBO J.* **1989**, *8*, 2179−2188.
(17) Jaskolski, M.; Tomasselli, A. G.; Sawyer, T. K.; Staples, D. G.; Heinrikson, R. L.; et al. Structure at 2.5-Å Resolution of Chemically Synthesized Human Immunodeficiency Virus Type 1 Protease Complexed with a Hydroxyethylene-Based Inhibitor. *Biochemistry* **1991**, *30*, 1600−1609.
(18) Veerapandian, B.; Cooper, J. B.; Sali, A.; Blundell, T. L.; Rosati, R. L.; et al. Direct observation by X-ray analysis of tetrahedral "intermediate" of aspartic proteinases. *Protein Sci.* **1992**, *1*, 322−328.
(19) Coates, L.; Erskine, P. T.; Wood, S. P.; Myles, D. A. A.; Cooper, J. A Neutron Laue Diffraction Study of Endothiapepsin: Implications for the Aspartic Proteinase Mechanism. *Biochemistry* **2001**, *40*, 13149−13157.
(20) Coates, L.; Erskine, P. T.; Crump, M. P.; Wood, S. P.; Cooper, J. B. Five Atomic Resolution Structures of Endothiapepsin Inhibitor Complexes: Implications for the Aspartic Proteinase Mechanism. *J. Mol. Biol.* **2002**, *318*, 1405−1415.
(21) Erskine, P. T.; Coates, L.; Mall, S.; Gill, R. S.; Wood, S. P.; Myles, D. A. A.; Cooper, J. B. Atomic resolution analysis of the catalytic site of an aspartic proteinase and an unexpected mode of binding by short peptides. *Protein Sci.* **2003**, *12*, 1741−1749.

(22) Cho, Y.; Rebholz, K. L.; Northrop, D. B. Solvent Isotope Effects on the Onset of Inhibition of Porcine Pepsin by Pepstatin. *Biochemistry* **1994**, *33*, 9637−9642.
(23) Toulokhonova, L.; Metzler, W. J.; Witmer, M. R.; Copeland, R. A.; Marcinkeviciene, J. Kinetic Studies on β-Site Amyloid Precursor Protein-cleaving Enzyme (BACE). *J. Biol. Chem.* **2003**, *278*, 4582−4589.
(24) Smith, R.; Brereton, I. M.; Chai, R. Y.; Kent, S. B. H. Ionization states of the catalytic residues in HIV-1 protease. *Nat. Struct. Biol.* **1996**, *3*, 946−950.
(25) Beveridge, A. J.; Heywood, G. C. A Quantum Mechanical Study of the Active Site of Aspartic Proteinases. *Biochemistry* **1993**, *32*, 3325−3333.
(26) Harte, W. E., Jr.; Beveridge, D. L. Prediction of the Protonation State of the Active Site Aspartyl Residues in HIV-1 Protease-Inhibitor Complexed via Molecular Dynamics Simulation. *J. Am. Chem. Soc.* **1993**, *115*, 3883−3886.
(27) Cheng, X.; Tropsha, A. Relative Binding Free Energies of Peptide Inhibitors of HIV-1 Protease: The Influence of the Active Site Protonation State. *J. Med. Chem.* **1995**, *38*, 42−48.
(28) Goldblum, A.; Rayan, A.; Fliess, A.; Glick, M. Extending Crystallographic Information with Semiempirical Quantum Mechanics and Molecular Mechanincs: A Case of Aspartic Proteinases. *J. Chem. Inf. Comput. Sci.* **1993**, *33*, 270−274.
(29) Lee, H.; Darden, T. A.; Pedersen, L. G. An ab Initio Quantum Mechanical Model for the Catalytic Mechanism of HIV-1 Protease. *J. Am. Chem. Soc.* **1996**, *118*, 3946−3950.
(30) Liu, H.; Muller-Plathe, F.; van Gunsteren, W. F. A Combined Quantum/Classical Molecular Dynamics Study of the Catalytic Mechanism of HIV Protease. *J. Mol. Biol.* **1996**, *261*, 454−469.
(31) van Gunsteren, W. F.; Liu, H.; Muller-Plathe, F. The elucidation of enzymatic reaction mechanisms by computer simulation: Human Immunodeficiency Virus protease catalysis. *J. Mol. Struct.* **1998**, *432*, 9−14.
(32) Venturini, A.; Lopez-Ortiz, F.; Alvarez, J. M.; Gonzalez, J. Theoretical Proposal of a Catalytic Mechanism for the HIV-1 Protease Involving an Enzyme-Bound Tetrahedral Intermediate. *J. Am. Chem. Soc.* **1998**, *120*, 1110−1111.
(33) Piana, S.; Carloni, P. Conformational Flexibility of the Catalytic Asp Dyad in HIV-1 Protease: an Ab Initio Study on the Free Enzyme. *Proteins: Struct., Funct., Genet.* **2000**, *39*, 26−36.
(34) Piana, S.; Sebastiani, D.; Carloni, P.; Parrinello, M. Ab Initio Molecular Dynamics-Based Assignment of the Protonation State of Pepstatin A/HIV-1 Protease Clevage Site. *J. Am. Chem. Soc.* **2001**, *123*, 8730−8737.
(35) Wang, W.; Kollman, P. A. Free Energy Calculations on Dimer Stability of the HIV Protease using Molecular Dynamics and a Continuum Solvent Model. *J. Mol. Biol.* **2000**, *303*, 567−582.
(36) Trylska, J.; Bala, P.; Geller, M.; Grochowski, P. Molecular Dynamics Simulations of the First Steps of the Reaction Catalyzed by HIV-1 Protease. *Biophys. J.* **2002**, *83*, 794−807.
(37) Silva, A. M.; Cachau, R. E.; Sham, H. L.; Erickson, J. W. Inhibition and Catalytic Mechanism of HIV-1 aspartic protease. *J. Mol. Biol.* **1996**, *255*, 321−346.
(38) Yamazaki, T.; Nicholson, L. K.; Torchia, D. A.; Wingfield, P.; Stahl, S. J.; Kaufman, J. D.; Eyermann, C. J.; Hedge, C. N.; Lam, P. Y. S.; Ru, Y.; Jadhav, P. K.; Chang, C.; Webers, P.C. NMR and X-ray Evidence That the HIV Protease Catalytic Aspartyl Groups Are Protonated in the Complex Formed by the Protease and a Non-Peptide Cyclic Urea-Based Inhibitor. *J. Am. Chem. Soc.* **1994**, *116*, 10791−10792.
(39) Hyland, L. J.; Tomaszek, J. T. A.; Roberts, G. D.; Carr, S. A.; Magaard, V. W.; Bryan, H. L.; Fakhoury, S. A.; Moore, M. L.; Minnich, M. D.; Culp, J. S.; DesJarlais, R. L.; Meek, T. D. Humn Immunodeficiency Virus-1 Protease. 1. Initial Velocity Studies and Kinetic Characterization of Reaction Intermediates by ^{18}O Isotope Exchange. *Biochemistry* **1991**, *30*, 8441−8453.
(40) Northrop, D. Follow the Protons: A Low-Barrier Hydrogen Bond Unifies the Mechanisms of Aspartic Proteases. *Acc. Chem. Res.* **2001**, *34*, 790−797.
(41) Park, H.; Lee, S. Determination of the Active Site Protonation State of β-Secretase from Molecular Dynamics Simulation and Docking Experiment: Implications for Structure-Based Inhibitor Design. *J. Am. Chem. Soc.* **2003**, *125*, 16416−16422.
(42) Dixon, S. L.; van der Vaart, A.; Gogonea, V.; Vincent, J. J.; Brothers, E. N.; Suarez, D.; Westeroff, L. M.; Merz, K. M., Jr. *DivCon*; QuantumBio Inc.: University Park, PA, 1999.
(43) Yang, W.; Lee, T. S. A Density-Matrix Form of The Divide and Conquer Approach for Electronic Structure Calculations of Large Molecules. *J. Chem. Phys.* **1995**, *103*, 5674−5678.
(44) Li, X. P.; Nunes, R. W.; Vanderbilt, D. Density-matrix electronic-structure method with linear system-size scaling. *Phys. Rev. B* **1993**, *47*, 10891.
(45) Nunes, R. W.; Vanderbilt, D. Generalization of the density-matrix method to a nonorthogonal basis. *Phys. Rev. B* **1994**, *50*, 17611.
(46) Daniels, A. D.; Millam, J. M.; Sucseria, G. E. Semiempirical methods with conjugate gradient density matrix search to replace diagonalization for molecular systems containing thousands of atoms. *J. Chem. Phys.* **1997**, *107*, 425.
(47) Millam, J. M.; Sucseria, G. E. Linear scaling conjugate gradient density matrix search as an alternative to diagonalization for first principles electronic structure calculations. *J. Chem. Phys.* **1997**, *106*, 5569.
(48) Stewart, J. J. P. Application of localized molecular orbitals to the solution of semiempirical self-consistent field equations. *Int. J. Quantum Chem.* **1996**, *58*, 133.
(49) Stewart, J. J. P. Calculation of the geometry of a small protein using semiempirical methods. *J. Mol. Struct. (THEOCHEM)* **1997**, *401*, 195.
(50) Honig, B.; Nicholls, A. Classical Electrostatics in Biology and Chemistry. *Science* **1995**, *268*, 1144−1149.
(51) Sharp, K.; Honig, B. Electrostatic Interactions in Macromolecules: Theory and Applications. *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 301−332.
(52) Li, J.; Zhu, T.; Cramer, C. J.; Thrular, D. G. *J. Phys. Chem.* **1998**, *102*, 1820−1831.
(53) Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Comput. Aided Mol. Des.* **1995**, *9*, 87−110.
(54) Gogonea, V.; Merz, K. M., Jr. Fully quantum mechanical description pf proteins in solution. Combining Linear Scaling Quantum Mechanical Methodologies with the Poisson−Boltzmann Equation. *J. Phys. Chem. A* **1999**, *103*, 5171−5188.
(55) van der Vaart, A.; Suarez, D.; Merz, K. M., Jr. Critical assessment of the performance of the semiempirical divide and conquer method for single point calculations and geometry optimizations of large chemical systems. *J. Chem. Phys.* **2000**, *113*, 10512−10523.
(56) Yon, J.; Cleasby, A.; Bruinzeel, W. D.; Masure, S. L. J.; Tickle, I.; Sharff, A. Cloning, sequence and crystal structure of human synthetic BACE catalytic domain and use of atomic structure coordinates for drug screening and drug design. In *PCT Int. Appl.* (Astex Technology Ltd., U.K.; Janssen Pharmaceutica NV); Wo, 2003; p 272.
(57) *First Discovery*, version 2.0; Schrodinger, Inc: Portland, OR.
(58) Jorgensen, W. L.; Tirado-Rives, J. The OPLS [Optimized Potentials for Liquid Simulations] Potential Functions for Proteins, Energy Minimizations for Crystals of Cyclic Peptides and Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.
(59) Schroeder, S.; Daggett, V.; Kollman, P. A. A comparison of the AM1 and PM3 semiempirical models for evaluating model compounds relevant to catalysis by serine proteases. *J. Am. Chem. Soc.* **1991**, *113*, 8922−8925.
(60) Diez y Riega, H.; Rincon, L.; Almeida, R. Semi-empirical studies of substituent effects on the ionization of bicyclooctane carboxylic acids and quinuclidines. *J. Phys. Org. Chem.* **2003**, *16*, 107.
(61) Tounge, B. A.; Reynolds, C. H. Calculation of the binding affinity of β-secretase inhibitors using the linear interaction energy method. *J. Med. Chem.* **2003**, *46*, 2074−2082.

JM049817J