In-situ Synthesis of a Tacrine-Triazole-Based Inhibitor of Acetylcholinesterase: Configurational Selection Imposed by Steric Interactions

Sanjib Senapati,*[,†,‡] Yuhui Cheng,[†] and J. Andrew McCammon[†,§,||]

Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093-0365, Department of Biotechnology, Indian Institute of Technology Madras, Chennai 600036, India, Howard Hughes Medical Institute, University of California, San Diego, La Jolla, California 92093-0365, and Department of Pharmacology, University of California, San Diego, La Jolla, California 92093-0365

Received November 10, 2005

Recently, researchers have used acetylcholinesterase (AChE) as a reaction vessel to synthesize its own inhibitors. Thus, **1** (*syn*-TZ2PA6), a femtomolar AChE inhibitor, which is formed in a 1:1 mixture with its anti-isomer by solution phase reaction from **3** (TZ2) and **4** (PA6), can be synthesized exclusively inside the AChE gorge. Our computational approach based on quantum mechanical/molecular mechanical (QM/MM) calculations, molecular dynamics (MD), and targeted molecular dynamics (TMD) studies answers why **1** is the sole product in the AChE environment. Ab initio QM/MM results show that the reaction in the AChE gorge occurs when **3**/azide and **4**/acetylene are extended in a parallel orientation. An MD simulation started from the final structure of QM/MM calculations keeps the azide's and acetylene's parallel orientations intact for 10 ns of simulation time. A TMD simulation applied on an antiparallel azide−acetylene conformation flips the acetylene easily to bring it to a postion that is parallel to azide. A second set of QM/MM calculations performed on this flipped structure generates a similar minimum-energy path as obtained previously. Even a TMD simulation carried out on a parallel azide−acetylene conformation could not deform their parallel arrangement. All of these results, thus, imply that inside the AChE gorge, the azide group of **3** and the acetylene group of **4** always remain parallel, with the consequence that **1** is the only product. The architecture of the gorge plays an important role in this selective formation of **1**.

Introduction

Acetylcholinesterase terminates the impulse signaling at cholinergic synapses by rapid hydrolysis of the neurotransmitter, acetylcholine (ACh)[1]. Symptomatic treatment of certain neurological diseases can be achieved by controlled inhibition of AChE. Acetylcholinesterase inhibitors are, thus, of therapeutic importance.[2−4] The active site of AChE is located at the base of a long and narrow 20 Å gorge. It consists of an esteratic subsite containing the catalytic machinery, and an anionic subsite responsible for binding the quaternary trimethylammonium tailgroup of ACh.[5] The catalytic functional unit of AChE is the catalytic triad consisting of Ser 203, His 447, and Glu 334. The structure of the gorge also reveals a constricted region halfway up the gorge formed by the side chains of Tyr124, Phe297, Tyr337, and Phe338.[6,7] An allosteric site remote from the active center and located at the rim of the gorge[8,9] also plays an important role in ligand binding. This site called the peripheral anionic site (PAS) is composed of surface residues Tyr72, Trp286, and His287.

The molecule 3,8-diamino-6-phenyl-5-[6-[1-[2-[(1,2,3,4-tetrahydro-9-acridinyl)amino]ethyl]-1*H*-1,2,3-triazol-5-yl]hexyl]-phenanthridinium (**1**) is an ultra high-affinity inhibitor of acetylcholinesterase (dissociation constant, K_d is 77 fM to 410 fM depending on the AChE source)[6,10−12]. This is the highest affinity reversible organic inhibitor of AChE known, with association rate constants close to the diffusion limit (rate constant: 1.7×10^{10} M^{-1} min^{-1} in mouse AChE).[6] The corresponding anti-isomer 3,8-diamino-6-phenyl-5-[6-[1-[2-[(1,2,3,4-tetrahydro-9-acridinyl)amino]ethyl]-1*H*-1,2,3-triazol-4-yl]hexyl]-phenanthridinium (**2**) is also a respectable but weaker AChE inhibitor. These two triazole regioisomers differ in the nitrogen substitution positions on the 1,2,3-triazole and can be derived by thermal reaction between a tacrine azide moiety, including a 2-carbon linker **3** and a phenylphenanthridinium alkyne moiety including a chain of six carbons **4** (Figure 1). This thermal reaction at room temperature proceeds very slowly and provides a mixture of syn and anti-isomers ≈1:1. The crystal structures of **1** and **2** complexed with mouse AChE (mAChE) have recently become available.[6] These two crystalline complexes differ mainly in the PAS conformations, where the aromatic plane of the *syn*-phenanthridinium orients 90° from its position in the anti complex, and the Trp286 side chain in the syn complex dislodges from the PAS surface and swings into the solvent. In both the complexes, however, the tacrine moiety of the inhibitor is placed nicely at the active center of the gorge with the triazole moiety occupying the constricted region.

Molecule **1** can also be produced in situ from precursor fragments by click chemistry[6,10−12]. The click chemistry approach uses the molecular structure of a target enzyme as a scaffold to synthesize its own inhibitors. The method relies on the simultaneous binding of two ligands, which contain reacting substrate functionalities, to adjacent sites on the protein; the proximity of the ligands is then likely to accelerate the reaction that connects them. Thus, the above 1,3-dipolar cycloaddition reaction between azide and acetylene, which proceeds very slowly by thermal reaction, is observed to occur in an accelerated rate when carried out in the presence of AChE, and most interestingly, the enzyme catalyzes the reaction in a highly selective fashion with the syn isomer as the only product. Here,

* To whom correspondence should be addressed. Phone: +91-44-2257-4122. Fax: +91-44-2257-4102. E-mail: sanjibs@iitm.ac.in.
† Department of Chemistry and Biochemistry, University of California, San Diego.
‡ Department of Biotechnology, Indian Institute of Technology, Madras.
§ Howard Hughes Medical Institute, University of California, San Diego.
|| Department of Pharmacology, University of California, San Diego.

10.1021/jm051132b CCC: $33.50
Published on Web 09/15/2006

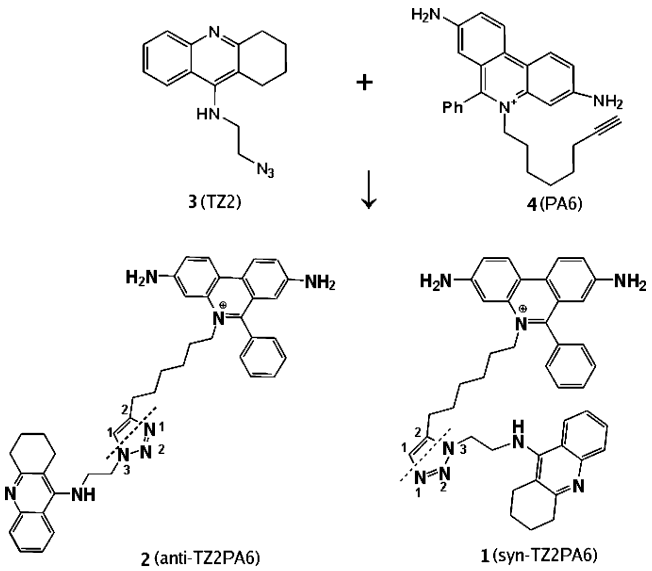

Figure 1. Structures of **1** and **2** formed by the 1,3-dipolar cycloaddition reaction. The tacrine, triazole, and phenylphenanthridinium moieties in **1** and **2** are shown from bottom to top, respectively.

we demonstrate the applicability of a computational approach, the combination of ab initio quantum mechanical/molecular mechanical calculations, standard molecular dynamics, and targeted molecular dynamics studies to answer why **1** is the sole product in the mAChE environment.

QM/MM approaches have been demonstrated to be effective in the accurate description of chemical bond formation and breaking in enzymatic reactions.[13,14] In these approaches, a small active part directly participating in the making and breaking of bonds is treated by quantum mechanical methods, whereas the rest of the enzyme containing a large number of atoms is described by molecular mechanical methods. Thus, the combined QM and MM approach takes advantage of the applicability and accuracy of the ab initio QM methods for chemical reactions and the computational efficiency of the MM calculations. The QM/MM boundary that very often bisects some covalent bonds is treated by a pseudobond approach,[15] where a one-free-valence atom with an effective core potential is constructed to replace the boundary atom of the environment part and to form a pseudobond with the boundary atom of the active part.

MD simulations have proven to be a valuable tool to investigate the dynamic behavior of stable macromolecules at finite temperatures.[16−18] However, a very large and complex conformational change that sometimes occurs in an enzyme structure, is unlikely to be observed in the time scale of an ordinary MD simulation (nanoseconds). TMD[19] is a method based on MD simulation and in principle is able to simulate more elaborate conformational changes in proteins. In the TMD method, the reaction coordinate is defined by a single mass-weighted root-mean-square target distance between a known initial structure and a fixed final (target) structure. By gradually reducing the constrained target distance to zero, the system is driven from the reactant to product state without explicitly defining the reaction pathway. TMD using a classical force field has been applied previously to investigate conformational transitions of insulin, ras-p21, chymotrypsin, aspartate transcarbamylase, and other molecules.[19−28]

We start this work by performing ab initio QM/MM calculations on **1**-mAChE and **2**-mAChE complexes (step 1). The structures of these complexes were taken from the Brookhaven Protein Data Bank. The results show that the cycloaddition in the syn complex occurs when the azide and acetylene are extended in a parallel orientation, whereas they would have to lie antiparallel to form the anti triazole complex. An MD simulation (step 2) started from the final structure of the QM/MM calculations on the syn complex keeps the stable parallel orientation of azide and acetylene intact for a further 10 ns of simulation time. A combination of MD and TMD simulations (step 3) applied on the final structure from QM/MM calculations on the anti complex flips the orientation of acetylene and brings it parallel to that of azide. Note that the constraint due to TMD was applied only on the phenanthridinium moiety and Trp286 side chain, keeping the alkane chain (to which acetylene is attached) completely free. Finally, in step 4, another QM/MM calculation was performed on the resultant structure from the TMD simulation to obtain and compare the reaction energy barrier with the previously obtained barrier from step 1. A very consistent barrier height between step 1 and step 4 shows the robustness of our approach.

Basic Models and Preparation of the Systems

We started our study by performing density functional theory QM/MM calculations on the crystal structures of the **1**−mAChE and **2**−mAChE complexes. The crystal structures were obtained from the Protein Data Bank (pdb code: 1Q83 for syn complex[6] and 1Q84 for anti complex[6]) and were completed in the following way before carrying out QM/MM calculations on them. The coordinates for missing residues Pro259, Gly260, Gly261, Ala262, Gly263, Ala542, and Thr543 in both of the structures and for the missing side chains of Ser495, Lys496 in 1Q83 and of Glu1, Arg493, Ser495, and Lys496 in 1Q84 were modeled using the InsightII graphics package.[29] For both the 1Q83 and 1Q84 models, hydrogens for heavy atoms were added by Leap in the Amber 7.0 package,[30] whereas hydrogens for crystal conserved water molecules were added and optimized by WHATIF.[31] Added hydrogens were energy minimized for 100 steps using the steepest descent algorithm. For histidine residues, calculations of the local electrostatic environment and the effective pK_a with WHATIF indicated that residues Hid381, Hid387, Hid405, and Hid447 should be protonated on N_δ, whereas the others are protonated at the epsilon nitrogen position. Whether Hid or Hie is chosen is determined by the local hydrogen bonding network. A set of partial atomic charges of the ligand isomers was obtained via quantum electronic structure calculations and is included in the Supporting Information. Using the Gaussian 98 program[32] with the 6-31G* basis set, we performed a Hartree−Fock geometry optimization procedure. The atom-centered RESP charges were determined via fits to the electrostatic potentials obtained from the calculated wave functions. The missing interaction parameters in the ligands were generated using antechamber tools in Amber.[33]

After relaxing the added atoms using the Amber 7.0 package in the gas phase, each structure was solvated in a cubic box of water of 90 Å in length. Eight sodium ions (as the ligand has charge +1) were randomly placed in the solvent to neutralize the −9 charge of the protein. A set of minimization and thermalization runs of the starting structures was performed to remove the initial strain. Both of the systems were then equilibrated by short MD simulations of 100 ps in the NPT ensemble followed by a 200 ps run in the NVT ensemble. All minimization and MD steps were performed using the Amber 7.0 package with Amber force fields.[34] A particle-mesh Ewald summation[35] with a 8 Å short-range cutoff was used to treat long-range electrostatics. SHAKE was used to constrain bond lengths between heavy atoms and hydrogens. The time step used

was 2 fs, and the system pressure was restrained to 1 atm, with a coupling time of 1.0 ps. The simulations were performed on eight processors of a Xeon 2.8 GHz linux cluster. The resulting structures were used to obtain the starting model for the QM/MM calculations by deleting the water molecules beyond 25 Å from the triazole ring of the ligand. The total number of atoms in the QM/MM models were then found to be 10407 in 1Q83 and 10328 in 1Q84.

The enzyme−substrate systems were partitioned into a QM subsystem and a MM subsystem, and a proper treatment of the boundary atoms was made by a pseudobond approach.[15] Following this approach, a boundary atom B in the MM subsystem was replaced by a pseudoatom B_{pb}, which has one free valance but a parametrized effective core potential, such that $A - B_{pb}$ pseudobond mimics the covalent A−B bond at the QM−MM boundary (A stands for a boundary atom in the QM subsystem). With this pseudoatom, the QM subsystem now forms a closed-shell active part, which was treated quantum mechanically, and the rest of the atoms were represented by a molecular mechanical force field. The QM subsystem consists of the triazole ring, one methyl group on each side of the ring, and two pseudoatoms (replacing one methyl carbon each) for a total of 14 atoms. The total energy of the QM/MM system can be written as follows.

$$E_{\text{Total}} = E_{\text{qm}}(QM) + E_{\text{mm}}(MM) + E_{\text{qm/mm}}(QM/MM) \quad (1)$$

where $E_{\text{qm}}(QM)$ is the quantum mechanical energy of the QM subsystem, and $E_{\text{mm}}(MM)$ is the standard molecular mechanical interactions exclusively involving atoms in the MM subsystem. The QM/MM interaction between the QM and MM subsystems consists of three terms.

$$E_{\text{qm/mm}}(QM/MM) = E_{\text{electrostatic}}(QM/MM) + E_{\text{vdW}}(QM/MM) + E_{\text{MM-bonded}}(QM/MM) \quad (2)$$

where $E_{\text{MM-bonded}}(QM/MM)$ refers to the MM bond, angle, and dihedral energy terms, which involve terms with at least one atom from each subsystem. $E_{\text{MM-bonded}}(QM/MM)$, $E_{\text{vdW}}(QM/MM)$, and $E_{\text{mm}}(MM)$ are calculated with an MM force field, and $E_{\text{qm}}(QM) + E_{\text{electrostatic}}(QM/MM)$ is calculated as the eigenvalue of an effective Hamiltonian. The effective Hamiltonian describes the electrons of the triazole ring, two methyl groups, and the valence electrons from the two boundary pseudoatoms. These electrons move in the potential generated from the nuclei of the atoms in the QM subsystem, the point charge of the MM atoms, and the effective core potential of the boundary atoms.

With each prepared enzyme−substrate system, an iterative optimization procedure[36] was applied at the B3LYP/6-31G* level. The reaction coordinate for step 1 QM/MM calculations was chosen to be the simultaneous bond breaking of the two C−N bonds in the triazole ring (Figure 1). Similarly, the reaction coordinate for step 4 QM/MM calculations was chosen to be the simultaneous bond formation of the two new C−N bonds between acetylene and azide. This choice of reaction coordinate was based on the reaction mechanism of 1,3-dipolar cycloadditions of azides and acetylenes as established from the pioneering work of Huisgen.[37] An iterative restrained minimization was then repeatedly applied to different points along the reaction coordinate, resulting in an optimal path for the reaction in the enzymatic environment and its associated potential energy surface. Stationary points obtained along the minimum energy paths and Hessian matrixes for degrees of freedom involving atoms in the QM subsystem were calculated leading to determination of the corresponding vibrational frequencies. Given that the determined minimum energy path is smooth and continuous, the energy maximum on the path with one and only one imaginary frequency is the transition state (TS), whereas the energy minima along the path with no imaginary frequencies are characterized as the reactant or the intermediate. For the reactant, transition state, and intermediate state we further carried out single-point high-level (MP2 and B3LYP) QM/MM calculations with a larger 6-31+G^* basis set. Throughout the calculations, the pseudobonds were treated with the 3-21G basis set and its corresponding effective core potential parameters.[36] The calculations were carried out using modified versions of the Gaussian 98[32] and the TINKER programs.[38] The MM model used is the AMBER 95 all-atom force field[34] and the TIP3P model for water.[39] For the QM subsystems, criteria used for geometry optimizations follow Gaussian 98 defaults. For the MM subsystem, the convergence criterion used is to have the root-mean-square energy gradient be less than 0.1 kcal mol^{-1} Å^{-1}. In the MM minimization, only atoms within 20 Å of the triazole ring were allowed to move. No cutoff for nonbonding interactions was used in the QM/MM calculations and the MM minimizations.

We also carried out MD simulations on the final structures obtained from the QM/MM calculations in step 1. Each QM/MM resultant structure was first solvated in a cubic box of water of 90 Å in length before carrying out simulations following a similar procedure as outlined above. The temperature of the simulations was maintained at 300 K.

A number of TMD simulations were also performed on the final structures of QM/MM calculations in step 1. TMD is a method to induce a conformational transition relatively quickly by applying a time-dependent, purely geometrical restraint. The TMD simulations were performed by introducing an additional constraint force in a normal MD simulation. It incorporates an energy restraint based on the mass-weighted root-mean-square (RMS) distance of a set of atoms in the current conformation with respect to a reference target conformation. The functional form of the RMS restraint energy can be written as follows.

$$E_{\text{RMS}} = 0.5*k_{\text{TMD}}*M*[D_{\text{RMS}}(X(t), X^{\text{target}}) - d_{\text{o}}]^2 \quad (3)$$

where k_{TMD} is the force constant, M is the number of constrained atoms, D_{RMS} represents the relative RMS distance for a selected set of atoms between the instantaneous conformation $X(t)$ and the reference X^{target}, and d_{o} is an offset constant in Å. With the decreasing of d_{o} (RMSD offset) as a function of the simulation time, the conformational change is driven from the initial to the final targeted conformation.[28] In this study, the conformational change is driven by applying RMSD restraints to all heavy atoms of the phenanthridinium moiety of the ligand and the Trp286 side chain of the enzyme.

Results and Discussion

We first present the results of QM/MM calculations carried out on the **1**−mAChE and **2**−mAChE complexes. Figure 2 displays the calculated minimum-energy paths along the reaction coordinate for the bond-breaking processes in the ligand. Note that here we followed a back-calculation procedure in which starting from **1** or **2** in the enzyme environment, we broke its C−N bonds to obtain **3** and **4**. Thus, in these calculations, essentially starting from the product state (P), we arrive at the intermediate state (E.I) through the transition state ($E + I \rightleftharpoons E.I \rightarrow P$; E, enzyme; I, inhibitor). The intermediate state is the one where **3** and **4** find themselves in close proximity inside

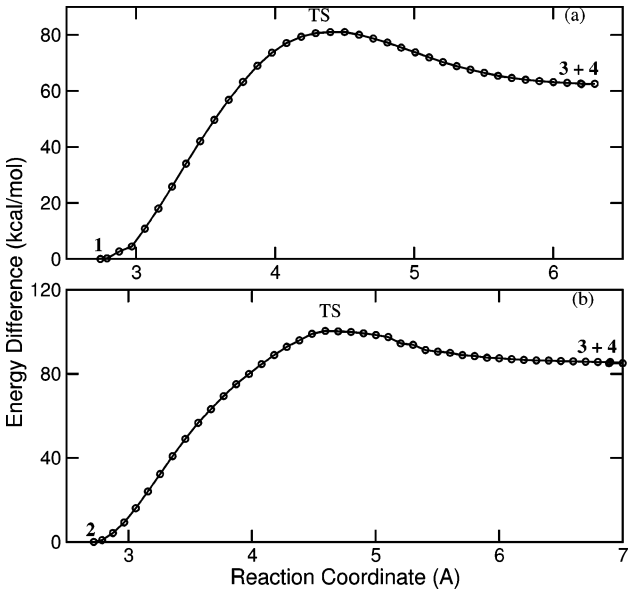

Figure 2. Determined minimum-energy path along the reaction coordinate for the (a) syn complex and (b) anti complex. The reaction coordinate is the simultaneous bond breaking of the two C−N bonds in the triazole ring of **1** and **2** to get back to **3** + **4** (see Figure 1). Here, we follow a back-calculation procedure by which starting from the product state (**1** or **2**), we reach the intermediate state (**3** + **4**). The activation energies ($E_{TS} - E_{3+4}$) are obtained at various high-level (B3LYP and MP2) QM/MM calculations and are tabulated in Table 1.

the protein gorge from which an accelerated reaction takes place in the framework of click chemistry. Figure 3 represents the geometries of these intermediates in both systems. This Figure clearly implies that the cycloaddition in the syn complex occurs when the azide and acetylene are extended in a parallel orientation, whereas they would have to lie antiparallel in the formation of the anti triazole complex. Now let us look back at Figure 2. The determined minimum-energy path for the syn complex (Figure 2a) is very smooth as the geometry changes along the reaction coordinate. The potential energy barriers (for the forward reaction) are computed to be 18.06 and 14.41 kcal/mol for the syn and anti complex,respectively, at the B3LYP/6-31G* level. A more accurate estimate of these barriers can be obtained using a larger 6-31+G* basis set. The calculated values are 13.35 and 7.94 kcal/mol for the syn and anti complexes, respectively, at the level of B3LYP/6-31+G*, and 9.08 and 4.34 kcal/mol for the syn and anti complexes, respectively, at the MP2/6-31+G* level (see Table 1). Although a comparison of these barrier heights suggests that a reaction of the acetylene and azide moieties should be relativeley favorable in the anti geometry, we will show below that the protein environment strongly disfavors this geometry and forces the reaction to occur only through a parallel acetylene-azide geometry. These values of the barrier heights can also be compared to the experimental activation barrier estimated from the value of k_2 ($E + I \rightleftharpoons E.I \overset{k_2}{\rightarrow} P$). According to the enzymatic reaction rate theory, the relationship between the free energy of activation and the rate constant can be written as follows.[40]

$$k(T) = (k_B T/h)(c^o)^{1-n} \exp[-\Delta G_o^{\#}(T)/RT] \qquad (4)$$

where $\Delta G_o^{\#}$ is the standard-state molar free energy of activation. If we assume that the entropic contribution in $G_o^{\#}$ is negligible, the experimental value of the activation barrier comes out to be 11.27 kcal/mol. This value corroborates our computed energy barriers of 13.35 and 9.08 kcal/mol for the syn complex obtained at the B3LYP and MP2 levels, respectively, with a larger 6-31+G* basis set. These observations are, therefore, indicative of a good environment for the parallel azide and acetylene inside the enzyme active site.

Thus far we have two facts: (i) the click chemistry method generates only **1** in the mAChE active site, starting from the precursor fragments **3** and **4**, and (ii) molecule **1** is produced when the azide and acetylene groups of **3** and **4** align in parallel. Following these facts, we carried out an MD simulation on each of the intermediate states (i.e., in Figure 3a and b), to test whether the enzyme and inhibitor will undergo conformational rearrangements to keep/bring azide and acetylene in parallel geometry. We call these MD simulations MD-I and MD-II for the syn and anti intermediate state, respectively. Figure 4 shows the average structure of the intermediate state obtained from MD-I (this simulation was carried out for 10 ns after solvating the structure represented by Figure 3a in a box of water). In a close-up view, only the structures of the ligand fragments along with a few active site residues are displayed. The average structure from the simulation was obtained by superposing the enzyme−inhibitor conformations onto that in the starting structure and taking the mean of the atom positions at every 10 ps interval during the 10 ns run. This Figure clearly shows that **3** and **4** with parallel azide and acetylene find a very good environment in the active site gorge of AChE and remain stable for a long simulation time.

The ligand structures in simulation MD-II (starting structure: Figure 3b), however, undergo some conformational rearrangements. The average structure shows that the C−C vector of the acetylene group (Figure 3b) rotates anticlockwise by about 70°. During the 10 ns simulation, we also found that the phenanthridinium moiety and the protein residue Trp286 undergo much larger fluctuations than those observed in the MD-I simulation. These results suggest that a conformational transition could take place in the protein−ligand structure over a longer simulation time. We then introduced TMD in our study to induce the conformational transition relatively quickly. A series of TMD simulations were initiated from the resultant structure of MD-II. These TMD simulations necessarily perform a best-fit of the phenanthridinium moiety and Trp286 of the simulation structure to the phenanthridinium moiety and Trp286 of the syn intermediate state (the target structure). Remember that these are the PAS sites where the **1**-mAChE and **2**-mAChE X-ray structures mainly differ. It is also worth mentioning that during the TMD simulations, the alkane chain of **4** to which acetylene is attached was kept completely free. Also note that the TMD simulations in this series differ from each other only in the force constant, k_{TMD}.

The first TMD simulation we carried out had a force constant $k_{TMD} = 0.5$ kcal mol^{-1} Å^{-2}. This force was found to be too small to induce any conformational transition within a long enough simulation time, say 1 ns. Note that a conformational transition even at a smaller value of force constant is feasible, but that takes a longer time. With a successive increment of 0.5 kcal mol^{-1} Å^{-2}, we picked up the lowest force constant, $k_{TMD} = 2.0$ kcal mol^{-1} Å^{-2} that could induce the desired transition. At this value of the force constant, the phenanthridinium moiety and Trp286 reach the target structure in just 1 ns of simulation time. More importantly, the flexible alkane chain in **4** reorganizes itself (note that no TMD was applied on it) in a manner that brings the acetylene group parallel to the azide group of **3**. We call this structure CONFIG$_{TMD}$. We continued this simulation for another 10 ns and noticed that the acetylene and azide groups maintain their parallel alignment

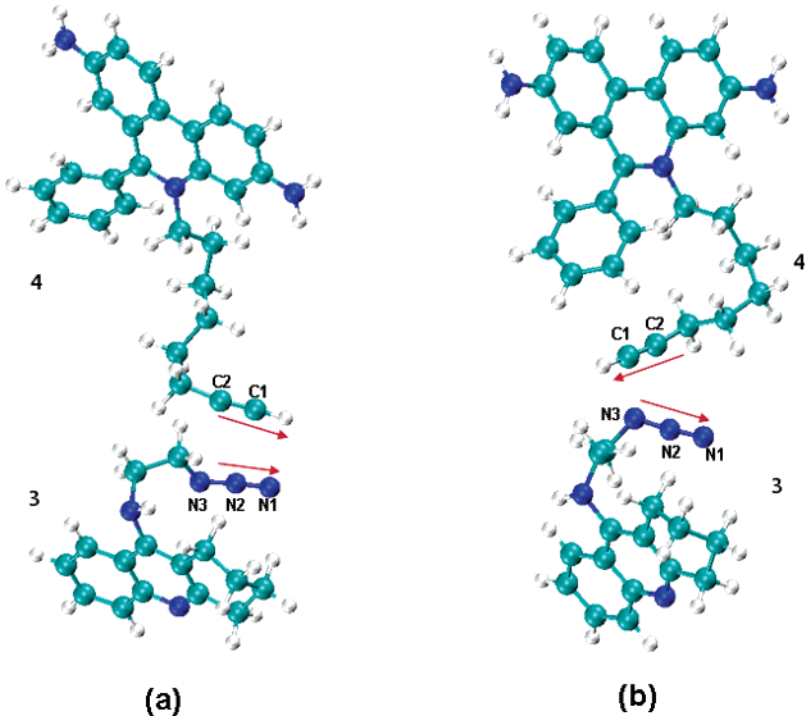

Figure 3. Geometries of the intermediate states obtained from QM/MM calculations on the (a) syn complex and (b) the anti complex. The protein residues are omitted for clarity.

Table 1. Calculated QM/MM Potential Energies for the 1,3-Dipolar Cycloaddition Reaction in mAChE[a]

	MP2/6-31+G*/MM	B3LYP/6-31+G*/MM	B3LYP/6-31G*/MM
	1−mAChE Complex		
transition state	66.44	70.84	77.59
intermediate state	57.36	57.49	59.53
$\Delta H^{\#}$	9.08	13.35	18.06
	2−mAChE Complex		
transition state	91.83	92.52	96.87
intermediate state	87.49	84.58	82.46
$\Delta H^{\#}$	4.34	7.94	14.41

[a] All energies are in units of kcal/mol.

throughout. We also noticed much smaller fluctuations in the PAS residues. The average structure obtained from the last 10 ns TMD run resembles Figure 4 very closely.

To further check the stability of the transformed structure, we performed an ordinary MD simulation starting from the structure, $CONFIG_{TMD}$. In this simulation all of the TMD restraints were taken off, letting every atom including the phenanthridinium moiety and Trp286 go free. The dynamics of the molecules were then observed for 10 ns. Figure 5 shows the average structure of the system obtained in this MD simulation. The average structure as well as the visual inspection of the dynamics of the molecules again indicate that the acetylene and azide groups always tend to remain parallel to each other in the AChE environment.

To test how good the transformed structure $CONFIG_{TMD}$ is, we carried out QM/MM calculations on the structure represented by Figure 5. This is the structure obtained by transforming the anti-intermediate state to the syn intermediate state by a combination of TMD and MD simulations. The QM/MM calculations are performed by a procedure similar to that described above. The reaction coordinate was chosen to be the simultaneous bond formation of the two new C−N bonds between acetylene and azide (C1−N1 and C2−N3) to form *syn*-triazole (see Figure 3). Figure 6 displays the calculated

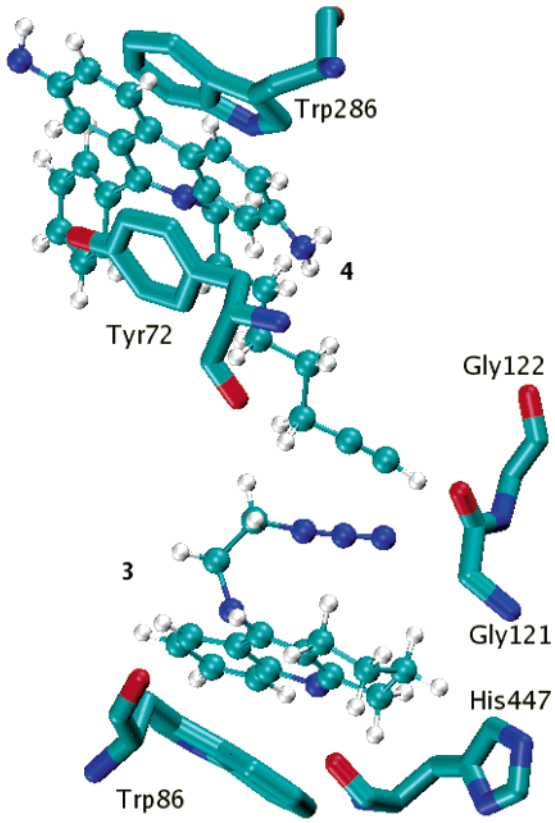

Figure 4. Average structure of the syn intermediate state obtained from a 10 ns MD simulation. In a close-up view, only the structures of the ligand fragments (ball-and-stick) along with a few active site residues (licorice) are displayed. The color scheme: cyan, C; blue, N; red, O; and white, H.

minimum-energy path along the reaction coordinate for the bond formation between **3** and **4**. Note the difference between this Figure and Figure 2a. Here, starting from the intermediate state (closely residing **3** and **4** inside the protein gorge), we reach

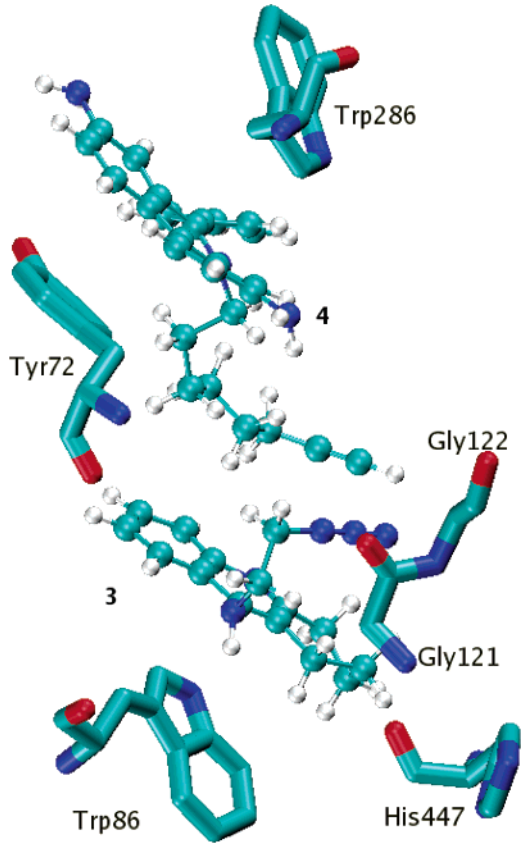

Figure 5. Average structure of the anti-intermediate state obtained from combined TMD and MD simulations. The target structure for the TMD simulation constitutes the PAS residues of the syn intermediate state. The color scheme is similar to that in Figure 4.

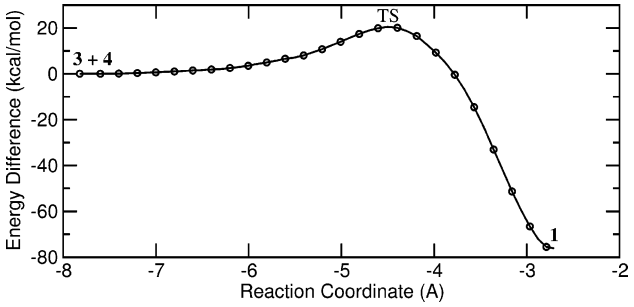

Figure 6. Determined minimum-energy path along the reaction coordinate for the bond formation between **3** and **4** to obtain **1**.

the product state through a transition state (step 2 in $E + I \rightleftharpoons E.I \rightarrow P$). The product is **1**. The minimum-energy path is seen to be very smooth with a well-defined transition and product state. The calculated potential energy barrier at the B3LYP/6-31G* QM/MM level is 20.4 kcal/mol compared to the value of 18.06 kcal/mol from Figure 2a. Thus, a very consistent barrier height and a smooth minimum-energy path similar to that in Figure 2a strengthen our finding that inside the AChE active site, the azide group of **3** and the acetylene group of **4** always try to go parallel. It also indicates that our chosen reaction coordinate is well defined beyond the transition state.

Another series of TMD simulations initiated from the resultant structure of MD-I has also been performed, where an applied RMSD restraint best fits the phenanthridinium moiety and Trp286 of the simulation structure to the phenanthridinium moiety and Trp286 of the anti-intermediate state, represented by Figure 3b. These simulations again differ form each other in the force constant, k_{TMD}. As before, we keep the alkane chain

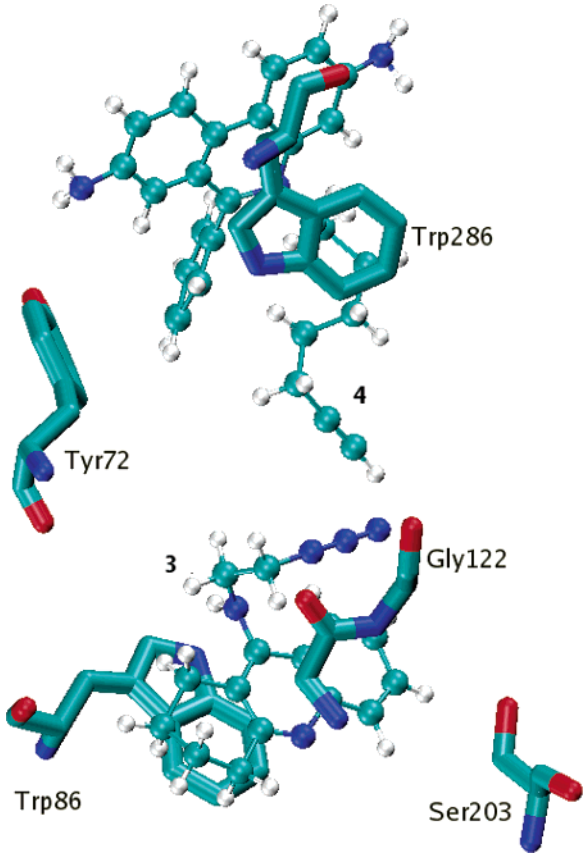

Figure 7. Average structure of the syn intermediate state obtained from a 10 ns TMD simulation. The target structure constitutes the PAS residues of the anti-intermediate state. The color scheme is similar to that in Figure 4.

of **4** completely free. In this series, the lowest force constant that can induce a conformational transition was found to be 4.0 kcal mol^{-1} Å^{-2}. Figure 7 shows the average structure of this system obtained from a 10 ns simulation run. Two important points that need to be noted from these TMD simulations are the following. Although the phenanthridinium moiety and Trp286 reach the target structure, the flexible alkane chain in **4** does not respond to the applied restraints with the acetylene group still remaining parallel to the azide group of **3**. We also needed a force that was two-times larger to induce the conformational transition in this system compared to that in the previous series of TMD simulations. These observations imply that the intermediate state of the syn complex with parallel acetylene and azide groups lies lower in the potential energy surface and that the anti-intermediate state can overcome the barrier to transform to the syn structure more easily. We also found that no matter what the conformations of the PAS residues are, the acetylene and the azide inside the AChE gorge always tend to remain parallel to each other. And if they remain parallel, according to our QM/MM results, the only possibility is the formation of *syn*-triazole. This explains why **1** is the sole product in an AChE environment.

We then tried to find the possible reasons why the acetylene and azide groups prefer to remain parallel in the active site. The structure of the AChE gorge shows a constricted region, ≈5−8 Å from the active site base and formed by the side chains of Tyr124, Phe297, Tyr337, and Phe338. The lengths of the alkane chains in **3** and **4** are such that the acetylene and azide groups fall in this region of constriction. Figure 8 shows a topography of this region along with the favorable locations of **3** and **4** in the gorge. The azide group of **3** finds a stable

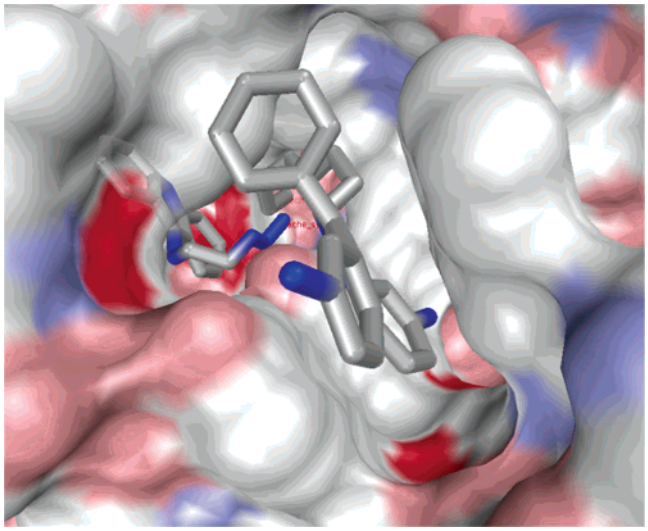

Figure 8. Topography of the constricted region in the active site gorge of AChE. The favorable positions of **3** and **4** around this region are also shown.

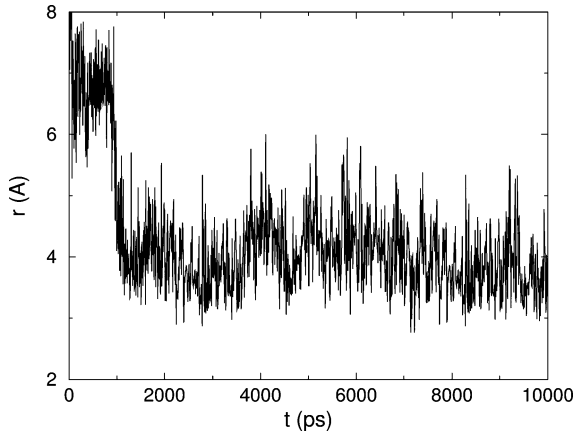

Figure 9. Distance between the acetylene carbon of **4** and Gly122−N as a function of time from the TMD simulation of the anti-intermediate state.

arrangement in this region by protruding inward to form hydrogen bonds (H bonds) with Tyr124 and Tyr337. The hydroxyl groups of these protein residues maintain stable hydrogen bonding with the N2 and N3 atoms of azide. A water-bridged hydrogen bond was also found to persist between the azide−N3 and the Gly−121 backbone.

The acetylene group of **4** also gains stability in the constricted region when it protrudes inward. In this arrangement, it maintains a number of H bonds with the adjacent residues, apart from gaining an electrostatic stabilization due to **3**−azide, Tyr337, and Phe338. As an example, Figure 9 explores the possibility of the formation of H bonds between the acetylene and its adjacent residues Gly121 and Gly122. The H-bond interaction, ≡C···H−N is modeled on the basis of the partial charges of the atoms, where the acetylene carbon serves as an acceptor, and Gly−N serves as the donor. Note that there exists other H bonds, for example, ≡C−H···O_{Gly} in this region (O_{Gly} acceptor, acetylene carbon donor). Figure 9, however, plots the distance between the acetylene carbon and Gly122−N as a function of time by analyzing the trajectories of the TMD simulation that had been carried out on the anti-intermediate state. As is clear from this Figure, except for the initial 1 ns time, which was required to flip the acetylene orientation from antiparallel to parallel, the acetylene maintains (an intermittent) H bond with Gly122. The time evolution of the distances between the acetylene carbon and Gly−O has a similar pattern and, therefore, is not included in the Figure. A number of water-mediated H bonds with the adjacent residues in the parallel orientation of acetylene also help it to acquire stability in this region. Moreover, the aromatic rings of Tyr 124, Phe297, Tyr337, and Phe 338 in the constricted region are placed in such an orientation that they induce a $\pi-\pi$ stacking stabilization to the π-systems of acetylene and azide.

Now let us think about the consequences of this parallel orientation. **3** and **4** react to produce **1**. The triazole moiety of this isomer inserts into the Tyr124−Tyr337 parallel sandwich and provides a greater surface complementarity with the gorge walls, leading to a tightly bound complex (the dissociation constant of the **1**−mAChE complex is about 20 times less than that of the anti complex.[6]) The *anti*-triazole, however, adopts an elongated flat shape and cannot find such a surface complementarity with the gorge walls. Thus, the architecture of the AChE active site gorge along with the linker distances of **3** and **4** favor the formation of **1** in the AChE environment.

Summary and Conclusions

Recently, researchers have used AChE as a reaction vessel to synthesize its own inhibitor. Thus, **1**, an ultra high-affinity inhibitor of AChE, which is formed in an ≈1:1 mixture with its anti-isomer by solution phase reaction from **3** and **4**, can be synthesized exclusively (with no formation of the anti-isomer) and with an accelerated rate inside the protein gorge. In this work, we demonstrate the applicability of a computational approach to answer why **1** is the sole product in the AChE environment.

Ab initio QM/MM calculations on the X-ray structures of **1**−mAChE and **2**−mAChE complexes (step 1) show that the cycloaddition reaction in the syn complex occurs when the azide of **3** and the acetylene of **4** are extended in a parallel orientation, whereas they would have to lie antiparallel for the formation of the anti complex. An MD simulation (step 2) started from the final structure of the QM/MM calculations on the syn complex keeps the parallel orientation of the azide and acetylene intact for 10 ns of simulation time. A combination of MD and TMD simulations (step 3) applied to the final structure from QM/MM calculations on the anti complex flip the orientation of acetylene and brings it parallel to azide. A second set of QM/MM calculations (step 4) performed on this flipped structure generates a minimum-energy path and barrier height similar to those obtained in step 1. Even a TMD simulation carried out on a parallel azide−acetylene conformation could not deform the parallel arrangement of acetylene and azide. All of these results, thus, imply that inside the AChE active site, the azide of **3** and the acetylene of **4** always remain parallel, with the consequence that **1** is the only product. The architecture of the AChE active site gorge is found to play an important role in this selective formation of **1**.

We previously have reported an induced fit in mAChE upon binding **1** using a combined molecular dynamics and ligand−protein docking study.[41] Using a fragment-based docking approach,[42] we found that although the fragment **3** of **1** can be accommodated nicely at the bottom of the active center gorge (with its azide group protruding inward in the constricted region), the fragment **4** cannot be fit owing to strong steric clashes with the PAS residues. The fragment **4** of **2**, however, can fit nicely with no steric clashes but in a nonparallel orientation of its acetylene group with respect to the orientation

of previously docked **3**−azide. Recall that **3** has a similar conformation in both complexes (Figure 3).

On the basis of these results and the results from the X-ray structure, the entire mechanism of this reaction can be predicted as follows. The moiety **3** with its azide group protruding inward in the constricted region first binds to the active center of AChE (**3** has a higher affinity than **4**; $K_d \approx 1.8 \times 10^{-8}$ M for **3** and $\approx 1.1 \times 10^{-6}$ M for **4**[10]). The moiety **4** then enters into the gorge with a nonparallel orientation of the acetylene group (similar to **4** in Figure 3b). Acetylene, in the constricted region, immediately flips to a parallel orientation to acquire better stabilization. A parallel orientation of **3**−azide and **4**−acetylene then led the exclusive formation of **1**. The corkscrew-like topology of this isomer shortens the linker distance of **4** by about 1.5 Å, as is found in the X-ray structure.[6] As a result of the reduced linking distance, the phenanthridinium moiety is pulled deeper into the gorge, where it sterically clashes with the Trp286 indole at the PAS surface. This causes the phenanthridinium moiety and Trp286 side chain to be dislodged from their original positions and attain a conformation similar to the one found in the syn complex (induced fit in mAChE[41]).

As part of the entire mechanism of the cycloaddition reaction in AChE, in this work, we successfully show that because of the unique structure of the protein gorge and the distinctive linker lengths of the ligand fragments, the acetylene and azide groups always align parallel in the constricted region. Following the above prediction, one may now start calculations with the anti-intermediate state, represented by Figure 3b, and directly apply TMD on its acetylene group to bring it parallel to **1**−azide. An ab initio QM/MM calculation then can be carried out for the formation of C−N bonds between acetylene and azide to produce *syn*-triazole. An ordinary MD simulation then can be performed on this resultant QM/MM structure, to see if the subsequent induced fit changes around the triazole ring, and the PAS region can eventually lead to a structure similar to that of the **1**−mAChE complex. Work along these directions is in progress. The future study also intends to determine the free energy change associated with the dislodgement of the phenanthridinium moiety and Trp286 side chain during the reaction using ab initio QM/MM free energy perturbation methods.[36]

It is worth mentioning that the intermediate states generated by the QM/MM calculations (Figure 3) are rather arbitrary and may not correspond to what exactly happens in nature. However, this is a unique technique that can shred some light on such important intermediate states.[15,36] Interestingly, these intermediate states demonstrated considerable stability during the subsequent classical MD studies, indicating that the generated states (Figure 3) correspond to at least some local minima on the potential surface.

It is also worth mentioning here that we observed a reduced flexibility in AChE upon binding **1** in the MD simulation study of the **1**−mAChE complex.[41] A similar decrease in flexibility upon ligand binding has been reported for several other proteins. Examples include FK506 binding protein FKBP-12[43], acyl-coenzyme A binding protein,[44] intestinal fatty acid binding protein,[45] and 2′GMP binding protein ribonuclease T1[46]. Zidek et al.,[47] however, found an increased protein (MUP-I) backbone flexibility upon binding 2-*sec*-butyl-4,5-dihydrothiazole, a small ligand. Increased protein conformational entropy is indicated to be the dominant factor associated with the free energy change of ligand binding in this study. We, however, do not speculate on a major contribution from AChE conformational entropy upon binding **1**. AChE has a long 20 Å gorge, and **1** is a relatively large ligand (molecular mass of **1** is 661.871 compared to 141.231 for 2-*sec*-butyl-4,5-dihydrothiazole). These factors may help the energetically favorable hydrophobic effect to outweigh the entropic term in our case. Note that the hydrophobic effect is proportional to the protein/ligand surface areas buried in the interface.[47]

Acknowledgment. S.S. gratefully acknowledges Professor Hartmuth C. Kolb for providing a topographic picture of the enzyme (Figure 8). We thank Dr. Zoran Radic, Professor Hartmuth C. Kolb, Professor Palmer Taylor, and Professor Barry Sharpless for helpful discussions. This work was supported in part by NIH, NSF, the Howard Hughes Medical Institute, the National Biomedical Computation Resource, the NSF Center for Theoretical Biological Physics, the W. M. Keck Foundation, and Accelrys, Inc.

Supporting Information Available: Complete References 30 and 32, and the partial atomic charges on the ligand isomers. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Kandel, E. R.; Schwartz, J. H.; Jessell, T. M. *Principles of Neural Science*, 3rd ed.; Appleton & Lange: Norwalk, CT, 1991.
(2) Barnard, E. A. In *The Peripheral Nervous System*; Hubbard, J. I., Ed.; Plenum Publishers: New York, 1974; pp 201−206.
(3) Dunnett, S. B.; Fibiger, H. C. Role of forebrain cholinergic systems in learning and memory: relevance to the cognitive deficits of aging and Alzheimer's dementia. *Prog. Brain Res.* **1993**, *98*, 413−420.
(4) *Current Trends in the Development of Acetylcholinesterase Inhibitors as Anti-Alzheimer Agents Current Pharmaceutial Design*; Camps, P., Munoz-Torrero, D. Eds.; Bentham Science Publishers: 2004; Vol. 10, pp ii−iii.
(5) Quinn, D. M. Acetylcholinesterase: enzyme structure, reaction dynamics, and virtual transition states. *Chem. Rev.* **1987**, *87*, 955−979.
(6) Bourne, Y.; Kolb, H. C.; Radic, Z.; Sharpless, K. B.; Taylor, P.; Marchot, P. Freeze-frame inhibitor captures acetylcholinesterase in a unique conformation. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 1449−1454.
(7) Sugimoto, H.; Yamanishi, Y.; Limura, Y.; Kawakami, Y. Donepezil Hydrochloride (E2020) and other acetylcholinesterase inhibitors. *Curr. Med. Chem.* **2000**, *7*, 303−339.
(8) Radic, Z.; Reiner, E.; Taylor, P. Role of the peripheral anionic site on acetylcholinesterase: inhibition by substrates and coumarin derivatives. *Mol. Pharmacol.* **1991**, *39*, 98−104.
(9) Bourne, Y.; Taylor, P.; Radic, Z.; Marchot, P. Structural insights into ligand interactions at the acetylcholinesterase peripheral anionic site. *EMBO J.* **2003**, *22*, 1−12.
(10) Lewis, W. G.; Green, L. G.; Grynszpan, F.; Radic, Z.; Carlier, P. R.; Taylor, P.; Finn, M. G.; Sharpless, K. B. Click chemistry in situ: Acetylcholinesterase as a reaction vessel for the selective assembly of a femtomolar inhibitor from an array of building blocks. *Angew. Chem., Int. Ed.* **2002**, *41*, 1053−1057.
(11) Manetsch, R.; Krasinski, Q.; Radic, Z.; Raushel, J.; Taylor, P.; Sharpless, K. B.; Kolb, H. C. In situ click chemistry: Enzyme inhibitors made to their own specifications. *J. Am. Chem. Soc.* **2004**, *126*, 12809−12818.
(12) Krasinski, Q.; Radic, Z.; Manetsch, R.; Raushel, J.; Taylor, P.; Sharpless, K. B.; Kolb, H. C. In situ selection of lead compounds by click chemistry: Target-guided optimization of acetylcholinesterase inhibitors. *J. Am. Chem. Soc.* **2005**, *127*, 6686−6692.
(13) Zhang, Y.; Kua, J.; McCammon, J. A. Role of the Catalytic Triad and Oxyanion Hole in Acetylcholinesterase Catalysis: An ab initio QM/MM Study. *J. Am. Chem. Soc.* **2002**, *124*, 10572−10577.
(14) Cheng, Y.; Zhang, Y.; McCammon, J. A. How does the cAMP-dependent protein kinase catalyze the phosphorylation reaction: An ab initio QM/MM study. *J. Am. Chem. Soc.* **2005**, *127*, 1553−1562.
(15) Zhang, Y.; Lee, T.; Yang, W. A pseudobond approach to combining quantum mechanical and molecular mechanical methods. *J. Chem. Phys.* **1999**, *110*, 46−54.
(16) Piana, S.; Carloni, P.; Perrinello, M. Role of conformational fluctuations in the enzymatic reaction of HIV-1 protease. *J. Mol. Biol.* **2002**, *319*, 567−583.
(17) Senapati, S.; Berkowitz, M. L. Molecular dynamics simulation studies of polyether and perfluoropolyether surfactant based reverse micelles in supercritical carbon dioxide. *J. Phys. Chem. B* **2003**, *107*, 12906−12916.

(18) Kua, J.; Zhang, Y.; McCammon, J. A. Studying enzyme binding specificity in acetylcholinesterase using a combined molecular dynamics and multiple docking approach. *J. Am. Chem. Soc.* **2002**, *124*, 8260−8267.
(19) Schlitter, J.; Engels, M.; Kruger, P. Targeted molecular dynamics: A new approach for searching pathways of conformational transitions. *J. Mol. Graphics* **1994**, *12*, 84−89.
(20) Engels, M.; Jacoby, E.; Kruger, P.; Schlitter, J.; Wollmer, A. The T → R structural transition of insulin; pathways identified by targeted energy minimization. *Protein Eng.* **1992**, *5*, 669−677.
(21) Jacoby, E.; Kruger, P.; Schlitter, J.; Wollmer, A. Simulation of a complex protein structural change: The T ↔ R transition in the insulin hexamer. *Protein Eng.* **1995**, *9*, 113−125.
(22) Wroblowski, B.; Diaz, J. F.; Schlitter, J.; Engelborghs, Y. Modeling pathways of α−chymotrypsin and deactivation. *Protein Eng.* **1997**, *10*, 1163−1174.
(23) Diaz, J. F.; Schlitter, J.; Engelborghs, Y. Calculation of pathways for the conformational transition between GTP and GDP bound states of the Ha-ras-p21 protein. *Proteins: Struct., Funct., Genet.* **1997**, *28*, 434−451.
(24) Ma, J.; Karplus, M. Molecular switch in signal transduction: Reaction paths of the conformational changes in ras p21. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 11905−11910.
(25) Roche, O.; Field, M. J. Simulations of the T ↔ R conformational transition in aspartate transcarbamylase. *Protein Eng.* **1999**, *12*, 285−295.
(26) Fererra, P.; Apostolakis, J.; Calfisch, A. Computer simulations of protein folding by targeted molecular dynamics. *Proteins: Struct., Funct., Genet.* **2000**, *39*, 252−260.
(27) Schlitter, J.; Swegat, W.; Mulders, T. Distance-type reaction coordinates for modeling activated processes. *J. Mol. Model.* **2001**, *7*, 171−177.
(28) Yang, L.; Beard, W. A.; Wilson, S. H.; Roux, B.; Broyde, S.; Schlick, T. Local deformations revealed by dynamics simulations of DNA polymerase β with DNA mismatches at the primer terminus. *J. Mol. Biol.* **2002**, *321*, 459−478.
(29) InsightII, Accelrys, Inc., 9685 Scranton Road, San Diego, CA, 1996.
(30) Case, D. A. et al. *Amber 7.0*, University of California, San Francisco, CA, 2002.
(31) Vriend, G. WHAT IF: A molecular modeling and drug design program. *J. Mol. Graphics* **1990**, *8*, 52−56.
(32) Frisch, M. J. et al. *Gaussian 98*, revision A.6; Gaussian, Inc.: Pittsburgh, PA, 1998.
(33) Wang, J.; Cieplak, P.; Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **2000**, *21*, 1049−1074.
(34) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(35) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(36) Zhang, Y.; Liu, H.; Yang, W. Free energy calculation on enzyme reactions with an efficient iterative procedure to determine minimum energy paths on a combined ab initio QM/MM potential energy surface. *J. Chem. Phys.* **2000**, *112*, 3483−3492.
(37) Huisgen, R. Kinetics and mechanism of 1,3-dipolar cycloadditions. *Angew. Chem., Int. Ed. Engl.* **1963**, *2*, 633−645.
(38) Ponder, J. W. *TINKER, Software Tools for Molecular Design*, version 3.6, 1998.
(39) Jorgensen, W. L.; Chandrsekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−933.
(40) Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. How enzymes work: Analysis by modern rate theory and computer simulations. *Science* **2004**, *303*, 186−195.
(41) Senapati, S.; Bui, J.; McCammon, J. A. Induced fit in mouse acetylcholinesterase upon binding a femtomolar inhibitor: A molecular dynamics study. *J. Med. Chem.* **2005**, *48*, 8155−8162.
(42) Lin, J.; Perryman, A. L.; Schames, J. R.; McCammon, J. A. Computational drug design accommodating receptor flexibility: The relaxed complex scheme. *J. Am. Chem. Soc.* **2002**, *124*, 5632−5633.
(43) Cheng, J.; Lepre, C. A.; Moore, J. M. ^{15}N NMR relaxation studies of the FK506 binding protein: Dynamic effects of ligand binding and implications for calcineurin recognition. *Biochemistry* **1994**, *33*, 4093−4100.
(44) Rischel, C.; Madsen, J. C.; Andersen, K. V.; Poulsen, F. M. Comparison of backbone dynamics of *apo*- and *holo*- acyl-coenzyme A binding protein using ^{15}NMR relaxation measurements. *Biochemistry* **1994**, *33*, 13997−14002.
(45) Hodsdon, M. E.; Cistola, D. P. Ligand binding alters the backbone mobility of intestinal fatty acid-binding protein as monitored by ^{15}NMR relaxation and ^{1}H exchange. *Biochemistry* **1997**, *36*, 2278−2290.
(46) Fushman, D.; Weisemann, R.; Thuring, H.; Ruterjans, H. Backbone dynamics of ribonuclease T1 and its complex with 2′GMP studied by two-dimensional heteronuclear NMR spectroscopy. *J. Biomol. NMR* **1994**, *4*, 61−78.
(47) Zidek, L.; Novotny, M. V.; Stone, M. J. Increased protein backbone conformational entropy upon hydrophobic ligand binding. *Nat. Struct. Biol.* **1999**, *6*, 1118−1121.

JM051132B