JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Automated Docking with Protein Flexibility in the Design of Femtomolar "Click Chemistry" Inhibitors of Acetylcholinesterase

Garrett M. Morris,[†] Luke G. Green,[‡] Zoran Radić,[§] Palmer Taylor,[§] K. Barry Sharpless,[‡] Arthur J. Olson,[∥] and Flavio Grynszpan*[,⊥]

[†]Crysalin, Ltd., Cherwell Innovation Center, 77 Heyford Park, Upper Heyford, Oxfordshire, OX25 5HD, U.K.

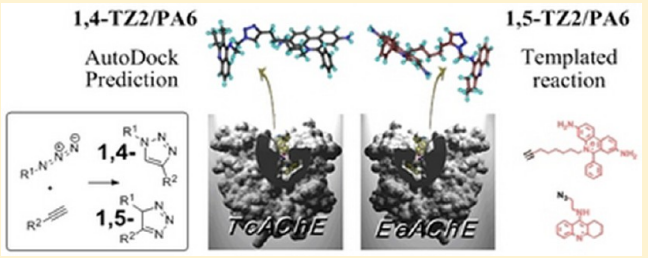
[∥]Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037-1000, United States

[⊥]Department of Biological Chemistry, Ariel University, Ariel 40700, Israel

[‡]Department of Chemistry, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037-1000, United States

[§]Department of Pharmacology, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California at San Diego, La Jolla, California 92093-0657, United States

ABSTRACT: The use of computer-aided structure-based drug design prior to synthesis has proven to be generally valuable in suggesting improved binding analogues of existing ligands.[1] Here we describe the application of the program AutoDock[2] to the design of a focused library that was used in the "click chemistry *in-situ*" generation of the most potent noncovalent inhibitor of the native enzyme acetylcholinesterase (AChE) yet developed (K_d = ~100 fM).[3] AutoDock version 3.0.5 has been widely distributed and successfully used to predict bound conformations of flexible ligands. Here, we also used a version of AutoDock which permits additional conformational flexibility in selected amino acid side chains of the target protein.



INTRODUCTION

The "click chemistry" *in situ* approach[3] depends on the simultaneous binding of molecular building blocks to adjacent subsites in a target protein, thereby selecting moieties with reactive functionalities that are within the close distances and suitable geometries. This proximity promotes a reaction between the reactive groups, which otherwise remain "silent": they neither degrade nor react with protein functionalities under physiological conditions. In principle, the enzyme could either repeatedly catalyze the formation of a privileged product with multiple turnovers or generate a product that inhibits the enzyme. In the latter case, the higher affinity products would then serve as lead compounds for drug discovery.[4] This approach involves the target enzyme directly in the selection and synthesis of its own inhibitors from a given array of molecular building blocks.

Computational methods can assist in the selection and design of appropriate candidate building blocks and can potentially save significant time in eliminating the synthesis of nonproductive reagents. For this purpose, we have applied the automated docking software package AutoDock version 3.0[2] and 4.0, which allows side chains in the target protein to change conformation while docking a flexible ligand. The AutoDock results correctly selected appropriate building blocks prior to synthesis, by showing which ones could bind productively.

Our particular approach involves the 1,3-dipolar cycloaddition reaction of azides and acetylenes to give 1,2,3-triazoles (Figure 1 and Figure 2A).[5] This reaction employs functional groups — azides and acetylenes — that are generally compatible with enzymes under physiological conditions and are readily incorporated into a wide range of organic molecules.[6]

We have explored the concept of *in situ* inhibitor assembly[7] using the enzyme acetylcholinesterase, or AChE, which plays a key role in signaling events in the central and peripheral nervous systems. The enzyme displays a binding pocket that is approximately 20 Å deep and consists of a catalytic active site at its base, and a "peripheral" site at its outer rim, linked by a narrow gorge lined with aromatic residues.[8] The active site is comprised of the acylation cleft and the choline binding site. Small molecule ligands for each of these two subsites (catalytic active site and peripheral site) are known,[9−11] and, in addition, inhibitors spanning the catalytic center gorge have also been shown to exhibit tighter binding than their individual components.[9]

Acetylcholinesterase from the electric organ of the eel *Electrophorus electricus* (*Ee*AChE)[12] was used to select and

Received: November 14, 2012
Published: March 1, 2013

   dx.doi.org/10.1021/ci300545a | *J. Chem. Inf. Model.* 2013, 53, 898−906

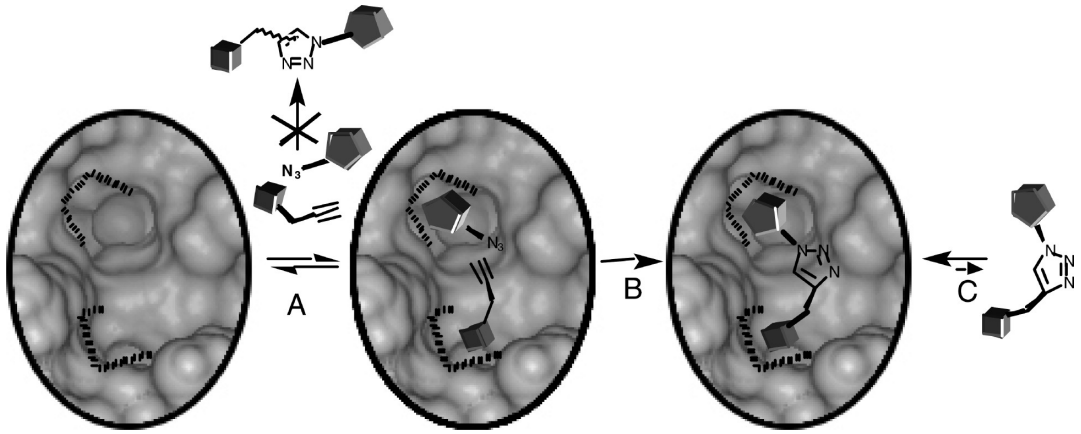


Figure 1. Target-directed assembly of ditopic inhibitors *via* azide−alkyne 1,3-dipolar cycloaddition: (A) molecular building blocks react very slowly in the absence of protein but bind to adjacent subsites; (B) the proximity of azide and alkyne groups facilitates their cycloaddition; (C) the resulting triazole shows very high product inhibition.

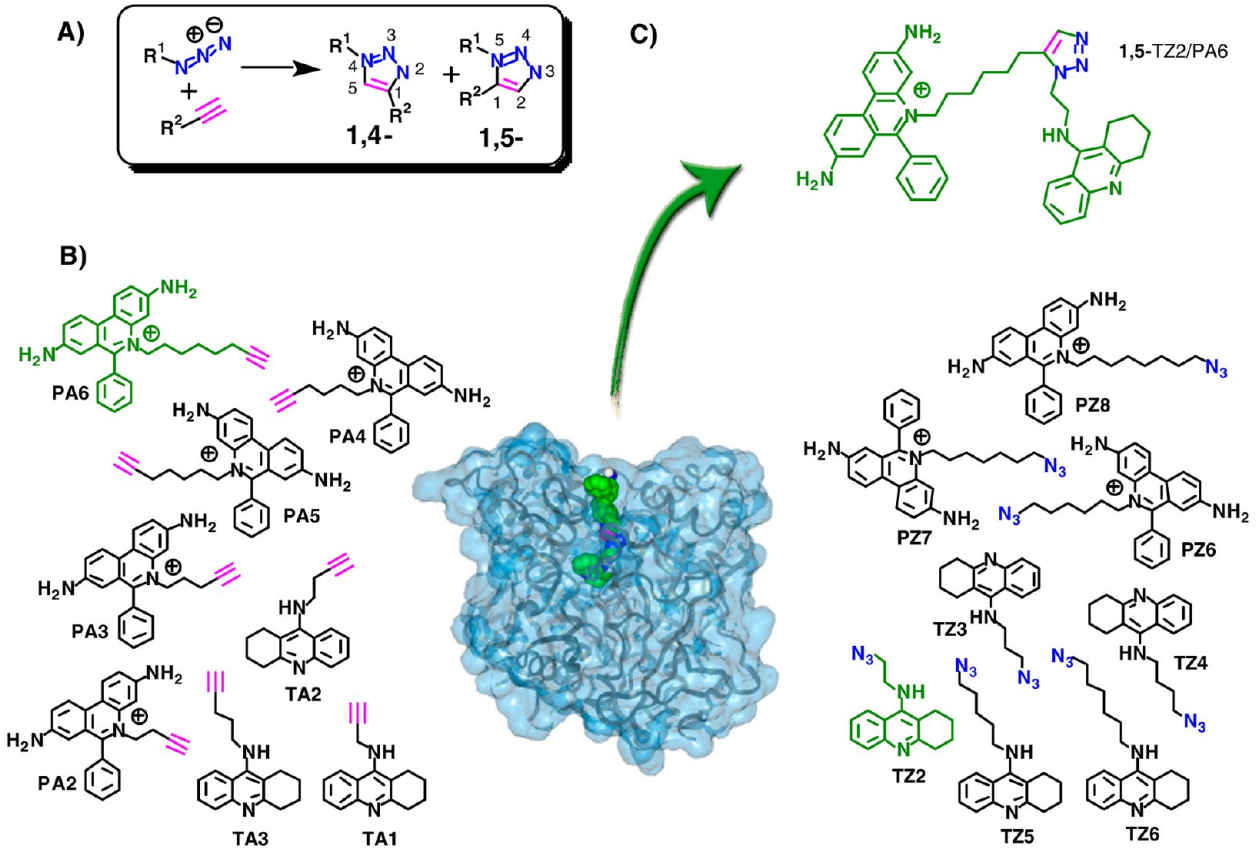
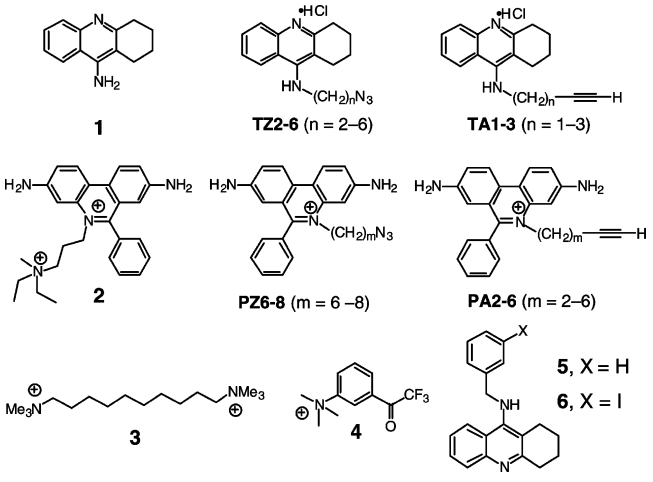


Figure 2. A) The nontemplated 1,3-dipolar cycloaddition reaction of an azide and an acetylene affords an *ca.* 1:1 mixture of triazole regioisomers. B) The array of building blocks used *in vitro*, and *Tc*AChE as a partially transparent blue surface with a ribbon representation of the protein backbone. The acetylene PA6 (green carbons, magenta alkyne group) and the azide TZ2 (green carbons, blue nitrogens) are also highlighted. C) Only one product was detected −1,5-PA6/TZ2 − and is the most tightly binding known noncovalent inhibitor of both the template protein *Ee*AChE and *Tc*AChE.

synthesize a triazole linked ditopic inhibitor using site-specific ligands as building blocks. A selection of subsite binders, inspired by known inhibitors, decorated with alkyl-azides and alkyl-acetylenes of varying chain lengths was envisioned. These building blocks were based on tacrine, aiming at the acylation/choline site (**1**),[9,10] and propidium (**2**),[10,11] directed toward the peripheral site. The key feature for the success of the "click chemistry *in situ*" approach is the simultaneous presence of the building blocks in the binding cleft of the enzyme, while presenting the reactive termini in such a geometry that promotes their ligation.

Two crystal structures of tetrameric *Ee*AChE were solved at 4.5 Å (PDB code 1C2B) and 4.2 Å (PDB code 1C2O).[13] Resolution at these levels is not sufficient to provide substantial detailed information on the topography of the binding site. Given that the sequence homology between *Torpedo californica* AChE (*Tc*AChE) and *Ee*AChE is about 60% and the active center gorge of these two isozymes differs only in a Phe 330/

dx.doi.org/10.1021/ci300545a | *J. Chem. Inf. Model.* 2013, 53, 898−906



Tyr 337 mutation (*Tc*AChE/*Ee*AChE), it would be reasonable to assume that the inhibitor binding modes to *Tc*AChE and *Ee*AChE are similar. Therefore, docking studies of inhibitors using *Tc*AChE at 2.8 Å resolution[14] instead of *Ee*AChE at 4.2 Å resolution may shed light on the binding modes in both systems.

Inspection of the crystal structure of *Tc*AChE complexed with (**1**) (PDB code 1ACJ)[14] suggests that an *N*-alkyl derivative (TZ2-6 or TA1-3) could, in principle, direct the reacting azide or acetylene group upward through the gorge. In addition, the crystal structure of *Tc*AChE complexed with decamethonium (**3**) (PDB code 1ACL)[14] suggests that an alkyl chain attached to the phenanthridinium (PZ6-8 or PA2-6) building block could in principle span from the peripheral site downward through the gorge. The corresponding alkyl chains decorated with reactive azide and acetylene functionalities need to be long enough and, at the same time, be able to adapt to the geometry of the binding site to permit the functional groups to react, thus yielding the triazole linker. Before investing a significant synthetic effort in the chemical synthesis of a library of building blocks, it was important to determine the appropriate length(s) of the alkyl chains of the putative molecular building blocks and whether the active site gorge would accommodate the proposed ligation.

■ COMPUTATIONAL METHODS

The program AutoDock[2] was used to identify putative binding modes of the tacrine and phenanthridinium building blocks of various chain lengths, to select which building blocks could potentially form a reactive ternary complex, and to predict the possible binding modes of the resulting triazole inhibitors. Both AutoDock version 3.0 and the newer version 4.0 have a variety of search methods, including Simulated Annealing (SA), Genetic Algorithm (GA), and Larmarckian Genetic Algorithm (LGA). In our experience, the most capable search method is the hybrid search method, LGA. This couples a typical Darwinian genetic algorithm for global searching with the Solis and Wets algorithm for local searching. LGA allows improvements in the phenotype (i.e., conformations with lower energies) to be passed on *via* the genotype (the system's translation, orientation, and torsion angles) to subsequent generations: this improves the efficiency of the calculation compared with traditional GAs. Version 4.0 of AutoDock introduces side chain flexibility into the target macromolecule. Selected side chains are allowed to change their conformations at the same time as the ligand that is being docked. However, the rest of the macromolecule remains rigid, so certain kinds of larger-scale conformational changes in the tertiary structure cannot be predicted with this method. It *is* possible, however, to model hinged-domain flexibility, but this is beyond the scope of this paper. The additional degrees of freedom added by these macromolecule side chains must also be taken into consideration when docking ligands with many rotatable bonds: the method is limited not only to a finite number of 32 torsions but also by the search capabilities of the particular search method used. As reported earlier, the LGA is much more efficient than the SA method, being able to handle search problems with more degrees of freedom. The standard error in both the AutoDock 3 and 4 scoring functions is ca. 2.5 kcal/mol, and this value was used to determine the significance of binding energy differences.

The macromolecule side chains chosen to be flexible are separated from the rigid portion of the macromolecule and actually become part of the input ligand (PDBQT) file. These side chains are distinguished from the ligand by special records new to AutoDock 4.0, that indicate that these residues may only change conformation and, unlike the ligand, may neither change translation nor orientation. What remains of the rigid portion of the macromolecule is treated as before, by storing it in a PDBQT-formatted file with the corresponding Kollman united atom partial atomic charges and AutoDock atom types. Lennard-Jones, Goodford-directional hydrogen bonding, and electrostatic potential grid maps are precomputed using AutoGrid. The moving portion includes the ligand, which may translate, change orientation and/or conformation, and the flexible side chains in the macromolecule, which, as just mentioned, may only change conformation. During the docking, the "docked energy" of the moving portion is computed as before.[2] AutoDock uses trilinear interpolation of the precomputed AutoGrid maps between the moving and rigid portions of the system; *within* the moving portion, it uses the 'intramolecular' energy calculation, summing the nonbonded and electrostatic terms between atom-pairs whose separations vary depending on the selected rotatable bonds. For each pair of atoms at the boundary between the rigid and flexible portions of the macromolecule, the vector that defines the rotatable bond emanating from the rigid region is used for rotating the side chain's atoms. These same two atoms, however, are omitted from the trilinear intermolecular nonbonded energy calculation, even though they are technically in the moving portion of the system.[15]

AutoDock requires precomputed grid maps, and since the free energy scoring function incorporates solvation free energies, Stouten atomic solvation parameters and fractional volumes were assigned to the acteylcholinesterase atoms using AddSol, an AutoDock utility. Atomic affinity and electrostatic potential grid maps were calculated using AutoGrid 3.0. A grid map with 81 × 81 × 81 points and a grid point spacing of 0.375 Å generously included the whole binding site of the acetylcholinesterase (see Figure 4). All the dockings performed in this work were carried out using the Lamarckian Genetic Algorithm (LGA). The specific AutoDock parameters for the LGA were set as follows: the population size was 50 individuals, each of whose initial translations, orientations, and torsions were set to random values. The limit of the number of energy evaluations was set to 1×10^6 for each docking. The maximum number of generations was set to 27,000. The mutation and crossover rates were set to 0.02 and 0.80 respectively. The maximum number of iterations per local search was set to 300,

dx.doi.org/10.1021/ci300545a | J. Chem. Inf. Model. 2013, 53, 898–906

and the probability of performing such a local search on an individual was 0.06. The maximum number of consecutive successes or failures before doubling or halving the local search step size was 4. Elitism was applied, with the top-scoring individual in the current generation automatically surviving into the next generation. A set of 100 docking calculations was performed for each ligand, and the resulting docked conformations were clustered into families of like-conformations, using a root mean squared positional deviation cluster tolerance of 1.0 Å. These same docked conformations were reclustered at a more forgiving value of 3.0 Å.

Prior to docking to *Tc*AChE, all three-dimensional models of the candidate click chemistry building blocks were minimized using the CVFF force field using the Discover module in the InsightII package (Accelrys, San Diego, California). Conjugate gradients energy minimization was performed until a maximum derivative of 0.01 was accomplished: a Morse potential was used for bond energies, and cross terms were included in the energy calculations. AutoDock reports the interaction energy between the docked ligand and the protein. The minimizations were performed using a relative dielectric of 1. Finally, Gasteiger partial charges were calculated for the ligand and Kollman charges were assigned to the protein. The geometry for the azide group was extracted from the crystal structure of the complex of *E. coli* thymidylate kinase with the bisubstrate inhibitor AZTP5A (PDB code 5TMP),[16] and the charges were adapted to each system from an *ab initio* calculation on methyl azide.[17]

Comparison of the active center gorge of *Tc*AChE complexed with tacrine (PDB code 1ACJ) and that of *Tc*AChE with decamethonium (PDB code 1ACL) reveals that the side chain conformations are highly conserved between these two crystal structures in this region. Nevertheless, two aromatic residues display significant conformational differences. In the **1**-*Tc*AChE complex, Phe 330 swings out from the bottom of the gorge allowing effective aromatic interactions with the tacrine inhibitor. In the **3**-*Tc*AChE complex, however, Phe 330 is buried deeper in the catalytic cleft; while Trp 279, located close to the outer rim of the binding pocket, changed its conformation to engage the ammonium ion in a cation-π interaction. A hybrid structure combining the conformations of Phe 330 in 1ACJ and Trp 279 in 1ACL was created prior to the docking of the building block inhibitors and product inhibitors.

Models of both regioisomers of the TZ2/PA6-product were built using the Biopolymer module of InsightII (Accelrys, San Diego). The geometries were energy minimized using the MOPAC module of InsightII and the PM3 Hamiltonian for 300 SCF (self-consistent field) iterations or until an energy convergence of 0.0001 was reached. The full-length triazole TZ2/PA6 inhibitors could not be minimized due to limitations of the semiempirical calculation method. For that reason, the TZ2/PA6 inhibitors were divided into three overlapping partial structures: tacrine, phenanthridinium, and triazole, which were minimized separately. The final inhibitors were then assembled in keeping with the corresponding calculated geometries and partial charges.

RESULTS AND DISCUSSION

As a check of the parameters implemented in AutoDock, initial calculations were carried out to reproduce the binding mode of **1** in *Tc*AChE as seen in the crystal structure 1ACJ. In this crystal structure, the inhibitor participates in hydrophobic contacts and in aromatic interactions, mainly with Trp 84 and Phe 330. In the protonated **1**, a hydrogen-bond from the quinoline NH^+ to the backbone carbonyl oxygen of His 440 seems probable (the N−O distance is 3.2 Å). This interaction would also contribute to the overall binding of **1**. In addition, hydrogen-bonding of the amino group of **1** to water molecules bound to *Tc*AChE has also been shown to be of importance.[14] Interestingly, the structure of a trifluoromethyl ketone (**4**) (an acetylcholine transition state-mimic inhibitor which covalently links to the catalytic Ser 200, PDB code 1AMN) showed that the quaternary ammonium group interacts with the aromatic Trp 84 (the N-to-ring distance is ∼4.2 Å), rather than with the His 440 backbone oxygen (where the N−O distance is 5.52 Å). Moreover, two tacrine derivatives (**5** and **6**) cocrystallized with *Drosophila melanogaster* acetylcholinesterase (*Dm*AChE, PDB codes 1DXA and 1QON, respectively)[18] show the tetrahydroacridine moiety is in a similar position to the one it occupies in 1ACJ, but the lack of a correlating hydrogen-bond gives no clear indication of their protonation states (with N−O distances 3.68 Å and 3.70 Å, respectively). The dockings of protonated and neutral **1** to *Tc*AChE, depleted of all water molecules, performed with AutoDock 3.0.5 are practically superimposable with the crystal structure, suggesting that the hydrophobic contacts and the aromatic interactions may be the main driving force in the mode of binding of this inhibitor. Based on these observations, and due to the hydrophobic nature of the binding site, this would favor an unprotonated **1**. It is possible that a positively charged tacrine[19] protonated at the endocyclic nitrogen at physiological pH accounts for the fast k_{on}, but there is no experimental evidence that would directly indicate whether or not this molecule remains charged in its final bound conformation at the bottom of the active center gorge. Docking studies were performed using neutral tacrine derivatives.[20] When the independent docking results of **1** (using 500,000 evaluations) to *Tc*AChE were clustered with an RMSD tolerance of 1.0 Å, all 20 dockings fell within the same cluster; the average RMSD from the crystal structure (PDB entry 1ACJ) was 0.581 Å, with a standard deviation of 0.006 Å.

Various tacrine derivatives with including 1−5 methylene units (TZ1-TZ5 and TA1-TA5) were tested computationally and assessed for their ability to place the reactive azide or acetylene side chain in a reactively competent location, pointing up through the gorge. From the docking of the tacrine derivatives, it was predicted that a 3-carbon chain for the acetylene building block (i.e., TA3) and a 2-carbon chain for the azide building block (i.e., TZ2) would result in suitable compounds with the best fit. It is interesting to note that in both cases, the chain length is roughly the same: 5 heavy atoms. The docking results for TZ2 can be divided conformationally into three major groups. The two clusters with lower energy put the inhibitor in the catalytic site, while the third group placed the tacrine ring at the rim of the gorge. Of the two clusters occupying the base of the binding site, only one group presents the alkyl chain in a conformation that would be productive in the "click" context, i.e. pointing upward through the gorge and toward the peripheral binding site. From the docking of the tacrine derivatives, it was predicted that a 3-carbon chain for the acetylene building block (i.e., TA3) and a 2-carbon chain for the azide building block (i.e., TZ2) would result in the suitable compounds with the best fit. It is interesting to note that in both cases, the chain length is roughly the same: 5 heavy atoms.

dx.doi.org/10.1021/ci300545a | J. Chem. Inf. Model. 2013, 53, 898−906

In addition, various phenylphenanthridinium derivatives with linear alkyl chains having 5−7 carbons (PZ5-PZ7 and PA1-PA7) were also tested computationally with AutoDock 3.0.5, by docking to *Tc*AChE alone. The results indicated that all the compounds would prefer to bind to the peripheral site, but the lack of clustering here suggests that many different conformations are possible, unlike the tacrine-based building blocks. The phenylphenanthridinium docking results could be roughly grouped into two clusters. In one group, the alkyl side chain would be solvent exposed, while in the other group (which includes the lowest energy conformer for each chain length) the alkyl chain was buried toward the active site. So according to the docking results for all the tested chain lengths, all could generate "click"-productive conformations: however, it was much harder to discriminate between the various chain lengths than was the case for the tacrine building blocks. Six carbons for the acetylene-based derivative (i.e., PA6) and 5 carbons for the azide-based building block (i.e., PZ5) would represent slightly better choices than the other chain lengths. Anything shorter than these would have been too short to reach the previously selected tacrine building block and could never react.

Thus, the AutoDock results emphasized that the chain-length pairs TZ2-PA6 and TA3-PZ5 were central to the design of the synthetic library (Figure 2B). Hence, subsite inhibitors based on tacrine and phenanthridinium motifs decorated with alkyl-azides and alkyl-acetylenes were prepared synthetically by variations of known methods.[10,21]

These synthetic building blocks allow for the preparation of 98 potential ditopic inhibitors to *Tc*AChE: 34 regioisomeric pairs of tacrine/phenanthridinium triazoles, namely TZ2-6/PA2-6 and TA1-3/PZ6-8, and 15 regioisomeric pairs of tacrine/tacrine triazoles, i.e. TA1-3/TZ2-6. Phenanthridinium/phenanthridinium triazole pairs were not considered, as they would not span the catalytic center gorge. Each one of the binary mixtures was incubated in the presence of *Ee*AChE at room temperature. The individual reaction mixtures were analyzed using Desorption Ionization On porous Silicon (DIOS) mass spectrometry.[22] Only *one* triazole derivative was assembled by the enzyme *Ee*AChE. Its identity was established by comparison with authentic samples of possible triazole products, which were prepared and characterized independently; these results indicated that this compound was 1,5-TZ2/PA6 (Figure 2C). Detailed kinetic analysis of the binding and inhibitory properties of 1,5- and 1,4-TZ2/PA6 were carried out against *Ee*AChE, *Tc*AChE, and mouse AChE, or mAChE (Table 1).[3] The dissociation constant, K_d, of the compound 1,5-TZ2/PA6 was determined to be 77 fM with *Tc*AChE and 410 fM with mAChE. This made 1,5-TZ2/PA6 the most potent noncovalent AChE inhibitor by approximately 2 orders of magnitude.[23] The 1,4-TZ2/PA6 isomer exhibited K_d values of between 720 fM with *Tc*AChE and 14 pM with *Ee*AChE. Thus, the 1,5-triazole regioisomer, which binds more tightly than the 1,4-triazole isomer, is the same structure that is preferentially assembled by the enzyme *Ee*AChE.

Table 1. Kinetic Parameters Derived for the Binding of TZ2/PA6 to AChE from Various Species and Reported Data for mAChE

inhibitor	k_{on} (10^{10} M^{-1} min^{-1})	k_{off} (min^{-1})	K_d	AChE species
1,5-TZ2/PA6[3]	1.5	0.0015	99 fM	*E. electricus*
	1.3	0.0011	77 fM	*T. californica*
	1.3	0.0079	410 fM	mouse
1,4-TZ2/PA6[3]	1.8	0.25	14,000 fM	*E. electricus*
	3.2	0.026	720 fM	*T. californica*
	2.4	0.30	8,900 fM	mouse
tacrine[9]	0.78	138	18 nM	mouse
propidium[9]	1.4	15,000	1,100 nM	mouse

The dissociation constants for both 1,5- and 1,4-TZ2/PA6 are substantially lower, i.e. have substantially higher affinity, than their parent components (estimated to be ∼10−100 nM for **1** and low μM for **2**). In this case, the ditopic inhibitors exhibit significantly tighter binding than their individual components, suggesting that the triazole group may interact favorably with the enzyme (e.g., as a weak H-bond acceptor, involved in aromatic interactions or through its relatively large dipole moment). Interestingly, the inhibitor TZ2/PA6, which was assembled by *Ee*AChE, has a similar affinity to the *Tc*AChE enzyme.

Dockings of 1,4- and 1,5-TZ2/PA6 in *Tc*AChE using AutoDock 3.0.5 show that the tacrine portion of the inhibitor can be accommodated at the bottom of the active center gorge, in a position that is virtually identical to that of **1** observed in its crystal structure (1ACJ). The tacrine is "sandwiched" by Trp 84 and Phe 330: in 1,5-TZ2/PA6-*Tc*AChE the quinoline presents $\pi-\pi$ interactions to Trp 84, while in 1,4-TZ2/PA6-*Tc*AChE the quinoline presents $\pi-\pi$ interactions to Phe 330 as seen in the TZ2 and **1** complexes. The phenanthridinium moiety, on the other hand, is located in the peripheral site at the rim of the gorge and participates in a positive charge to aromatic interaction, with Trp 279. Interestingly, in both cases the triazole moiety is predicted to lie below (i.e., deeper than) the narrowest point of the gorge defined by residues Phe 330, Tyr 334, Phe 331, Phe 288, Trp 233, Phe 290, and Tyr 121. In the docked conformations, the triazole rings of 1,4- and 1,5-TZ2/PA6 could interact *via* weak hydrogen-bonds with the OH groups of Tyr 121 and Ser 122, respectively (Figure 3A).

Notably, the conformations of 1,4- and 1,5-TZ2/PA6 triazole products docked to TcAChE appear reasonably similar to the X-ray crystal structures solved for both TZ2/PA6 complexes with mouse AChE (Figure 3B), keeping in mind cross-species differences in the AChE sequences between the predicted and 3D X-ray structures. Similarity is highest at the base of the AChE gorge where the tricyclic tacrine system forms a nearly identical sandwich between Trp86 and Tyr337 of wt mouse AChE (Trp84 and Phe330 in TcAChE). The triazole ring remains in the vicinity of Tyr337 (Phe330 in TcAChE) and Tyr 124 (Tyr121 in TcAChE). Phenanthridinium rings of TZ2/PA6 remain stabilized by peripheral site aromatic clusters (Tyr72, Ty124, Trp286 in mouse AChE and Tyr70, Ty121, Trp279 in TcAChE), but their conformations show largest divergence in the two enzymes. This is not surprising since the distribution of charged residues in peripheral sites of the two enzymes vary causing a clear difference in electrostatic interactions with the cationic phenantridinium moiety. As a matter of fact, both propidium[9] and TZ2/PA6[3] bind tighter to TcAChE than to mouse AChE. It is worth mentioning that in place of anionic Glu73, Glu247, Glu350, and Asp 351 in TcAChE uncharged Thr75, Leu250, Ala357, and Gln354 are found. Also, two potentially cationic histidines found in the peripheral site of mouse AChE (His284 and His287) are replaced by uncharged Thr277 and Asn280 in TcAChE. Furthermore, a degree of flexibility is inherent to residues of the AChE active center gorge including the peripheral site, as

dx.doi.org/10.1021/ci300545a | *J. Chem. Inf. Model.* 2013, 53, 898−906

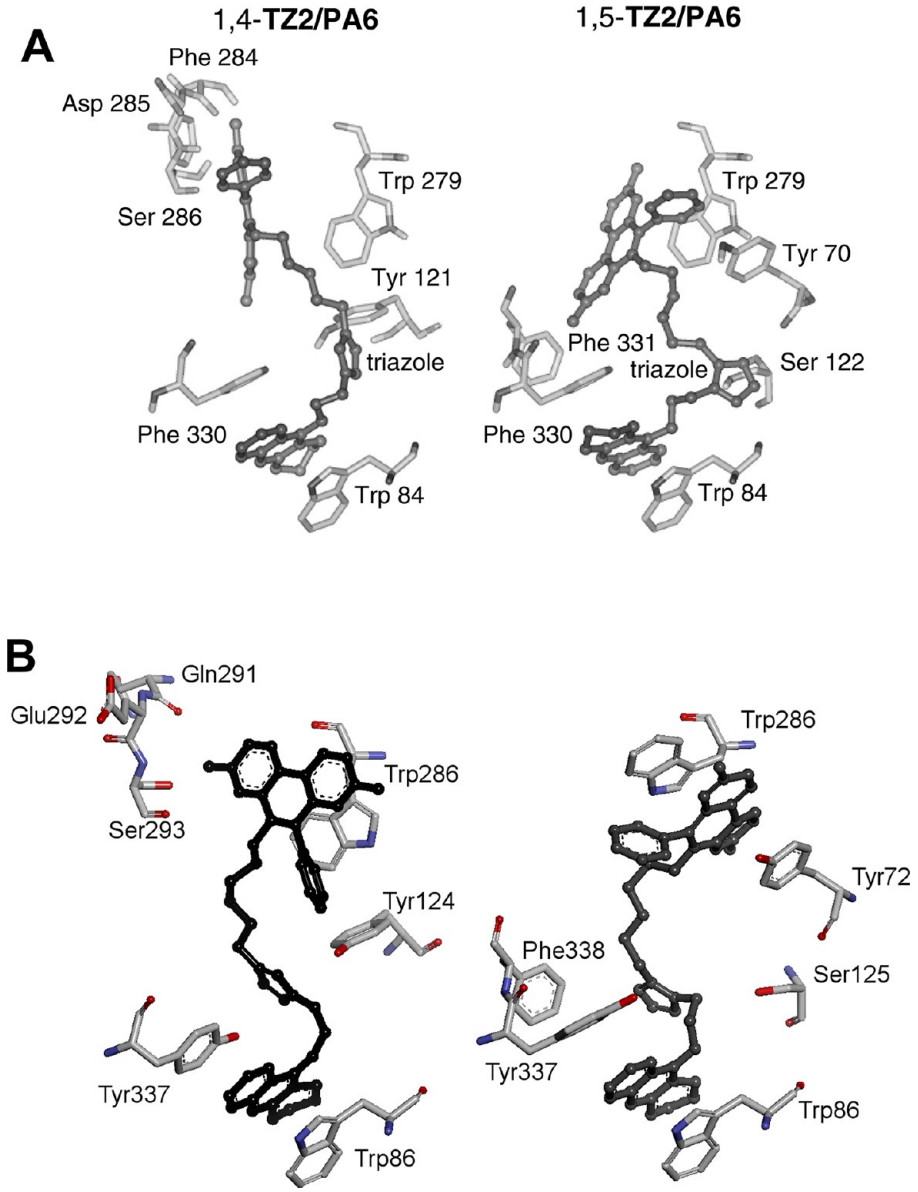


Figure 3. The 1,4- and 1,5-triazole products TZ2/PA6 (left and right, respectively) bound within the active center gorge of AChE. **A)** Lowest energy dockings found by AutoDock for TcAChE. **B)** Crystal structure determined for mouse AChE (PDB id: 1Q84 and 1Q83).[24] Inhibitors are depicted as black ball-and-stick models. AChE residues that are within 5 Å of the AutoDock docked inhibitors which are depicted in gray. Oxygen and nitrogen atoms are highlighted. Most of the hydrogen atoms were omitted for clarity.

illustrated by the difference in conformations of TZ2/PA6 obtained in crystals of the complex with Tyr337Ala mouse AChE mutant formed over various intervals of crystallization time.[25]

It is possible that $\pi-\pi$ interactions between the triazole ring and some of the aromatic residues, in the narrowest section of the gorge, contribute to the slow k_{off}. The 1,5-TZ2/PA6 lowest docked energy is −9.06 kcal/mol (the mean docked energy of the conformationally similar cluster was −7.86 kcal/mol, with 8 out of the 100 dockings finishing in this cluster). The lowest docked energy of the regioisomer 1,4-TZ2/PA6 is −8.48 kcal/mol (mean docked energy = −7.58 kcal/mol, 14 conformers in this cluster). The difference in the estimated K_i values computed from these "*in silico*" lowest docked energies is about 1 order of magnitude (1,5-TZ2/PA6: 2.30×10^{-7} M and 1,4-TZ2/PA6: 6.11×10^{-7} M, at T = 298.15 K). This calculated difference agrees very well with the *in vitro* experimental results (Table 1).

Since there is a high degree of similarity between *Tc*AChE and *Ee*AChE in both sequence and binding of the TZ2/PA6 inhibitors, it is intriguing whether *Tc*AChE would assemble the 1,5-TZ2/PA6, in the same way *Ee*AChE did. AutoDock 4.0, which allows flexibility in both the ligand and the enzyme, was used to predict the outcome of such a reaction. Assuming there are no major rearrangements in the enzyme, the lowest energy docked conformation of the ternary complex, formed by the enzyme and the two "un-reacted" building blocks, should indicate whether the key cycloaddition reaction would afford the 1,5- or the 1,4-triazole regioisomer. For the 1,3-dipolar cycloaddition to take place, the tacrine building block should first bind to the bottom of the AChE binding pocket and then quietly "wait" for the complementary phenanthridinium building block to subsequently bind to this binary complex.

dx.doi.org/10.1021/ci300545a | *J. Chem. Inf. Model.* 2013, 53, 898−906

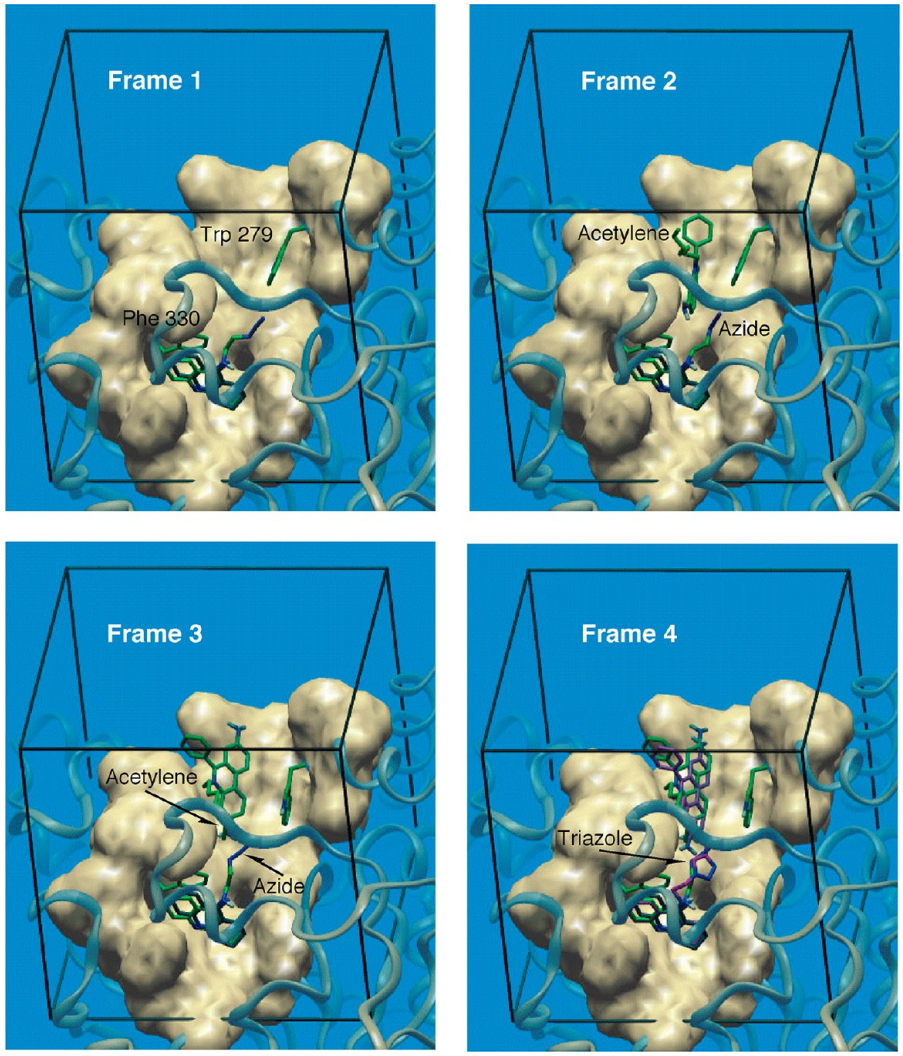


Figure 4. These images show the predicted sequence of events leading to the generation of 1,4-TZ2/PA6 by *Tc*AChE. The crystallographic structure of **1** is in black tubes. Azide, acetylene building blocks, and selected AChE flexible residues are in green tubes. 1,4-TZ2/PA6 is in purple tubes. Nitrogen atoms are highlighted in blue. Only polar hydrogens are shown, in white. Frame 1: Lowest energy conformer of TZ2 (docked with AutoDock 3.0.5) superimposed with crystallographic **1**. Frame 2: Representative high-energy conformer of the ternary complex (docked with AutoDock 4.0). Note the solvent exposed acetylene group. Frame 3: Lowest energy conformer of the ternary complex is shown. Frame 4: Lowest energy conformer of 1,4-TZ2/PA6 superimposed with frame 3 and **1**.

The lowest energy of the ternary complex would represent a snapshot of the building blocks prior to the fusion reaction and therefore should give an indication of the reaction path that gives the favored triazole regioisomer. Thus, TZ2 was first docked in *Tc*AChE using AutoDock 3.0.5 (Figure 4, frame 1). The lowest energy conformer was selected and integrated to the enzyme as an extra residue. Then PA6 was docked to the TZ2-*Tc*AChE complex, while conformational flexibility was introduced into the enzyme as the tacrine alkyl chain as well as the side chains of residues Trp 279 and Phe 330 (*vide supra*). The partial charges of PA6 were computed using the PM3 semiempirical method of MOPAC, giving acetylene-carbon atom charges that were approximately neutral (−0.08e and +0.02e). The partial charges of TZ2 were computed using *ab initio* methods,[17] and this produced values that were slightly negative (−0.114e and −0.167e). In order to emphasize the affinity of these reacting groups, we modified the partial charges in order to "drive" the reacting functional groups toward one another and thus improve the efficiency of the simulation. The carbon atoms in the terminal acetylene were empirically modified by adding 0.5e; similarly, the first and third nitrogens of the azide group were modified by subtracting 0.5e from their partial charges. The docking results show that in the highest energy conformations the acetylene group in PA6 resides in the peripheral site, far away from the azide group in TZ2 (Figure 4, frame 2). This observation suggests that the added charges are a reasonable choice that resulted in an efficient calculation but did not introduce a strong bias to the simulations. The lowest energy conformation of this ternary complex (Figure 4, frame 3) was compared with the docked structures of triazole inhibitors 1,5- and 1,4-TZ2/PA6 with *Tc*AChE. A surprisingly high resemblance was found between the predicted conformation of the two reactant-building blocks and the lowest energy docked conformation of the cycloaddition product 1,4-TZ2/PA6 (Figure 4, frame 4). If the predicted ternary complex is considered as the step before the formation of the final triazole inhibitor, these results would suggest that *Tc*AChE might preferentially catalyze the formation of the 1,4-TZ2/PA6 rather than the 1,5-regioisomer; the latter 1,5-adduct is a better binder by about 1 order of magnitude, however. The

dx.doi.org/10.1021/ci300545a | *J. Chem. Inf. Model.* 2013, 53, 898−906

conformational freedom observed for the tacrine side chain and that of the phenanthridinium side chain within the binding site gorge may account for the time that the protein requires to produce a *circa* equimolar amount of inhibitor (five days incubation at room temperature).

In summary, we have shown that AutoDock was successful in predicting the length of the alkyl chains of the azide-containing and actylene-containing building blocks for a productive click chemistry reaction. Furthermore, this approach helped to generate the best-known noncovalent inhibitor of AChE. The predicted chain lengths, while being not completely obvious, might otherwise have been overlooked in the design of a small library of building blocks. The AutoDock-calculated energies also correctly ranked the relative binding affinities of the triazole inhibitors. The prediction of the 1,4-TZ2/PA6 regioisomer as the favored templated product from *Tc*AChE may indicate that the host enzyme, in the conformation used in these dockings, would not necessarily generate the "best" inhibitor. The "click chemistry *in situ*" triazole product has a very high affinity to the target enzyme and represents an exquisite lead compound in the quest for the generation of novel pharmaceuticals.

AutoDock has been used extensively (e.g., the AutoDock 3 paper, Morris et al. 1998, has been cited more than 4,800 (as of March 2013) times according Google Scholar), and it has been shown to be effective in a large number of diverse cases, including systems that range from predominantly hydrophobic to predominantly polar in nature, and in ligand-protein, ligand-DNA, and even protein−protein docking. Based on this body of literature, and our experience, we believe that the methodology described here should be generally applicable. It should be noted that AutoDock 4 tends to be able to predict binding modes accurately for ligands with up to about 12 rotatable bonds, while AutoDock Vina, for example, can predict molecules with more torsions.

AUTHOR INFORMATION

Corresponding Author

*E-mail: flaviog@ariel.ac.il.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This publication was supported by National Institutes of Health Grant P01-GM48870 (G.M.M. and A.J.O.) RO1 GM18360-41 (P.T.) and UO1 NS 05846 (P.T., K.B.S.), and the California Breast Cancer Research Program for an IDEA Award (F.G.).

REFERENCES

(1) (a) Matter, H.; Baringhaus, K. H.; Naumann, T.; Klabunde, T.; Pirard, B. Computational approaches towards the rational design of drug-like compound libraries. *Comb. Chem. High Throughput Screening* **2001**, *4*, 453−457. (b) Mezey, P. G. Computer aided drug design: some fundamental aspects. *J. Mol. Model.* **2000**, *6*, 150−157. (c) Drews, J. Drug discovery: a historical perspective. *Science* **2000**, *287*, 1960−1963. (d) Wess, G.; Urman, M.; Sickenberger, B. Medicinal chemistry: challenges and opportunities. *Angew. Chem., Int. Ed.* **2001**, *40*, 3341−3350.

(2) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using a Lamarckian genetic algorithm and and empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639−1662.

(3) (a) Lewis, W. G.; Green, L. G.; Grynszpan, F.; Radić, 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**, *114*, 1095−1099. (b) Mamidyala, S. K.; Finn, M. G. In situ click chemistry: probing the binding landscapes of biological molecules. *Chem. Soc. Rev.* **2010**, *39*, 1252−1261.

(4) Kirby, A. J. Effective molarities for intramolecular reactions. *Adv. Phys. Org. Chem.* **1980**, *17*, 183−278. (b) Jencks, W. P. On the attribution and additivity of binding energies. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 4046−4050. (c) Mammen, M.; Choi, S.-K.; Whitesides, G. M. Polyvalent interactions in biological systems: implications for design and use of multivalent ligands and inhibitors. *Angew. Chem., Int. Ed.* **1998**, *37*, 2755−2794.

(5) (a) Huisgen, R. *Profiles, Pathways, and Dreams*; Seeman, J. I., Ed.; American Chemical Society, Washington, DC, 1994. (b) Mock, W. L. Cucurbituril. *Top. Curr. Chem.* **1995**, *175*, 1−24. (c) Chen, J.; Rebek, J., Jr. Selectivity in an encapsulated cycloaddition reaction. *Org. Lett.* **2002**, *4*, 327−329.

(6) Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R. Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 19−24.

(7) Ramström, O.; Lehn, J.-M. Drug discovery by dynamic combinatorial libraries. *Nat. Rev. Drug Discovery* **2002**, *1*, 26−36.

(8) Sussman, J. L.; Harel, M.; Frolow, F.; Oefner, C.; Goldman, A.; Toker, L.; Silman, I. Atomic structure of acetylcholinesterase from torpedo californica: a prototypic acetylcholine-binding protein. *Science* **1991**, *253*, 872−879.

(9) Radić, Z.; Taylor, P. Interaction kinetics of reversible inhibitors and substrates with acetylcholinesterase and its fasciculin 2 complex. *J. Biol. Chem.* **2001**, *7*, 4622−4633.

(10) Carlier, P. R.; Han, Y. F.; Chow, E. S.-H.; Li, C. P.-L.; Wang, H.-S.; Lieu, T. X.; Wong, H. S.; Pang, Y.-P. Evaluation of short-tether bis-THA AChE inhibitors. A further test of the dual binding site hypothesis. *Bioorg. Med. Chem.* **1999**, *7*, 351−357.

(11) Lappi, S.; Taylor, P. Interaction of fluorescence probes with acetylcholinesterase. Site and specificity of propidium binding. *Biochemistry* **1975**, *14*, 1989−1997.

(12) Purchased from Sigma.

(13) Bourne, Y.; Grassi, J.; Bourgis, P. E.; Marchot, P. Conformational flexibility of the acetylcholinesterase tetramer suggested by X-ray crystallography. *J. Biol. Chem.* **1999**, *274*, 30370−30376.

(14) Sussman, J. L.; Silman, I.; Axelsen, P. H.; Hirth, C.; Goeldner, M.; Bouet, F.; Ehret-Sabatier, L.; Schalk, I.; Harel, M. Quaternary ligand binding to aromatic residues in the active-site gorge of acetylcholinesterase. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 9031−9035.

(15) When AutoDock 3.0's new force field was created, we introduced constant energy penalties to the oxygen and polar hydrogen maps designed to penalize any hydrogen bonds not formed upon binding. This was based on the assumption that a small molecule when it is unbound is fully hydrated and all possible hydrogen bonds are satisfied. However, when a very large molecule such as a protein is used as the 'ligand', as in protein-protein docking (Ollman-Saffire, E.; Parren, P. W. H. I.; Pantophlet, R.; Zwick, M. B.; Morris, G. M.; Rudd, P. M.; Dwek, R. A.; Stanfield, R. L.; Burton, D. R.; Wilson, I. A. *Science* **2001**, *293*, 1155−1159. Legge, G. B.; Morris, G. M.; Sanner, M. F.; Takada, Y.; Olson, A. J.; Grynszpan, F. *Proteins: Structure, Func., Genetics* **2002**, *48*, 151−160), it is necessary to set these AutoGrid map constants to zero. This is because the assumption breaks down within the protein: no internal hydrogen bonds are lost upon binding within the rigid protein core. If this is not done, the significant number of oxygens and polar hydrogens within the protein generate a large energy penalty, which although constant during the docking, precludes final dockings with negative energies.

(16) Lavie, A.; Ostermann, N.; Brundiers, R.; Goody, R. S.; Reinstein, J.; Konrad, M.; Schlichting, I. Structural basis for efficient phosphorylation of 3′-azidothymidine monophosphate by *Escherichia coli* thymidylate kinase. *Proc. Natl. Acad. Sci. USA.* **1998**, *95*, 14045−14050.

(17) The partial charges for the azide group were calculated by Dr. Fahmi Himo and Prof. Lou Noodleman, The Scripps Research Institute, La Jolla, California. Personal communication.
(18) Harel, M.; Kryger, G.; Rosenberry, T. L.; Mallender, W. D.; Lewis, T.; Fletcher, R. J.; Guss, J. M.; Silman, I.; Sussman, J. L. Three-dimensional structures of Drosophila melanogaster acetylcholinesterase and of its complexes with two potent inhibitors. *J. Prot. Sci.* **2000**, *9*, 1063−1072.
(19) Desai, M. C.; Thadeio, P. F.; Lipinski, C. A.; Liston, D. R.; Spencer, R. W.; Williams, I. H. Physical parameters for brian uptake: optimizing log P, log D and pKa of THA. *Bioorg. Med. Chem. Lett.* **1991**, *8*, 411−414.
(20) Remarkably, protonated forms of predicted TZ2 building block and the triazole inhibitors TZ2/PA6 adopted virtually the same conformations of the corresponding neutral docked compounds. In all cases, clustering was consistently better using neutral molecules.
(21) Ross, S. A.; Pitié, M.; Meunier, B. A straightforward preparation of primary alkyl triflates and their utility in the synthesis of derivatives of ethidium. *J. Chem. Soc., Perkin Trans. 1* **2000**, 571−574.
(22) (a) Wei, J.; Buriak, J.; Siuzdak, G. Desorption-ionization mass spectrometry on porous silicon. *Nature* **1999**, *401*, 243−246. (b) Thomas, J. J.; Shen, Z.; Crowell, J. E.; Finn, M. G.; Siuzdak, G. Desorption/ionization on silicon (DIOS): A diverse mass spectrometry platform for protein characterization. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4932−4937. (c) Shen, Z.; Thomas, J. J.; Averbuj, C.; Broo, K. M.; Engelhard, M.; Crowell, J. E.; Finn, M. G.; Siuzdak, G. Porous silicon as a versatile platform for laser desorption/ionization mass spectrometry. *Anal. Chem.* **2001**, *73*, 612−619.
(23) (a) Nair, H. K.; Lee, K.; Quinn, D. M. m-(N,N,N-Trimethylammonio)trifluoroacetophenone: a femtomolar inhibitor of acetylcholinesterase. *J. Am. Chem. Soc.* **1993**, *115*, 9939−9941. (b) Harel, M.; Quinn, D. M.; Nair, H. K. The X-ray structure of a transition state analog complex reveals the molecular origins of the catalytic power and substrate specificity of acetylcholinesterase. *J. Am. Chem. Soc.* **1996**, *118*, 2340−2346. (c) Marchot, P.; Khelif, A.; Ji, Y. H.; Mansuelle, P.; Bougis, P. E. Binding of 125I-fasciculin to rat brain acetylcholinesterase. The complex still binds diisopropyl fluorophosphate. *J. Biol. Chem.* **1993**, *268*, 12458−12467. (d) Duran, R.; Cerveñansky, C.; Dajas, F.; Tipton, K. F. Fasciculin inhibition of acetylcholinesterase is prevented by chemical modification of the enzyme at a peripheral site. *Biochim. Biophys. Acta* **1994**, *1201*, 381−388. (f) Radić, Z.; Duran, R.; Vellom, D. C.; Li, Y.; Cerveñansky, C.; Taylor, P. Site of fasciculin interaction with acetylcholinesterase. *J. Biol. Chem.* **1994**, *269*, 11233−11239. (g) Eastman, J.; Wilson, E. J.; Cerveñansky, C.; Rosenberry, T. L. Fasciculin 2 binds to the peripheral site on acetylcholinesterase and inhibits substrate hydrolysis by slowing a step involving proton transfer during enzyme acylation. *J. Biol. Chem.* **1995**, *270*, 19694−19701. (h) Puu, G.; Koch, M. Comparison of kinetic parameters for acetylthiocholine, soman, ketamine and fasciculin towards acetylcholinesterase in liposomes and in solution. *Biochem. Pharmacol.* **1990**, *40*, 2209−2214.
(24) Bourne, Y.; Kolb, H. C.; Radić, 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.
(25) Bourne, Y.; Radić, Z.; Taylor, P.; Marchot, P. Conformational remodeling of femtomolar inhibitor−acetylcholinesterase complexes in the crystalline state. *J. Am. Chem. Soc.* **2010**, *132*, 18292−18300.