Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry Letters
Bioorganic & Medicinal Chemistry Letters 17 (2007) 1117–1121

β-Secretase (BACE-1) inhibitors: Accounting for 10s loop flexibility using rigid active sites

Georgia B. McGaughey,[a,*] Dennis Colussi,[b] Samuel L. Graham,[c] Ming-Tain Lai,[b] Sanjeev K. Munshi,[d] Philippe G. Nantermet,[c] Beth Pietrak,[b] Hemaka A. Rajapakse,[c] Harold G. Selnick,[c] Shaun R. Stauffer[c] and M. Katharine Holloway[a]

[a]*Department of Molecular Systems, Merck Research Laboratories, PO Box 4, West Point, PA 19486, USA*
[b]*Department of Biological Chemistry, Merck Research Laboratories, PO Box 4, West Point, PA 19486, USA*
[c]*Department of Medicinal Chemistry, Merck Research Laboratories, PO Box 4, West Point, PA 19486, USA*
[d]*Department of Structural Biology, Merck Research Laboratories, PO Box 4, West Point, PA 19486, USA*

Received 30 September 2006; revised 1 November 2006; accepted 1 November 2006
Available online 6 November 2006

Abstract—BACE-1 is a flexible enzyme with experimentally determined motion in the flap region, the catalytic aspartates, and the 10s loop. Four in-house crystallographically determined complexes of tertiary carbinamine inhibitors revealed 10s loop motion in the S_3 pocket. These X-ray structures were used to correlate K_i values, which span over five orders of magnitude, with the calculated interaction energy, using the Merck Molecular Force Field for a series of 19 tertiary carbinamine inhibitors.

β-Secretase (BACE-1), a transmembrane, aspartyl protease, is a promising therapeutic target for the treatment of Alzheimer's Disease (AD) since it is recognized as the rate-limiting step in the production of Aβ40–42.[1] The Aβ monomers and oligomers produced through BACE-1 cleavage of APP have been implicated as a causative factor in neuronal degeneration. One of the major challenges for the development of a viable therapeutic agent is the need for a low molecular weight compound which can cross the blood–brain barrier. This is a formidable challenge in the BACE-1 inhibitor program due to the large, open, hydrophilic active site (>1000 Å^3) and the requirement of CNS inhibition.

Publicly available crystal structures of inhibitors binding to BACE-1 have demonstrated that this enzyme is inherently flexible[2,3] and computational chemists have tried to capture this movement through the application of molecular dynamics and essential dynamic analysis.[4] Additionally, there are numerous reports where computational chemists have attempted to predict the biological activity of various BACE-1 inhibitors.[5–10] Accounting for varied protein flexibility in a scoring function is a challenge due to the difficulty in reliably predicting entropic costs, water motion, and concerted protein–ligand movement, particularly when one is rank-ordering many compounds in a drug discovery program to prioritize compounds for synthesis. The goal of the present study is to account for this flexibility while more accurately predicting binding of inhibitors in a high throughput manner.

We recently described the discovery of tertiary carbinamines as novel aspartyl protease inhibitors, and were interested in modeling and scoring variants with different S_3 groups (pdb code: 2IRZ).[11]

Examination of the X-ray co-crystal structure of **1** in the BACE-1 active site revealed a unique geometric shape of the primary tertiary carbinamine transition state isostere with the catalytic aspartates.[11] The catalytic aspartic acids, D32 and D228, adopt an orthogonal orientation relative to one another which creates an optimal trigonal hydrogen bonding network between the primary amine and G230, D32, and D228.[11] The benzyl occupies the S_1 pocket and the methanesulfonyl makes hydrogen bonding interactions with T232, T233, and R235 in S_2.[11,12] The amide carbonyl oxygen is within hydrogen

Keywords: BACE; Scoring functions; Molecular mechanics; Tertiary carbinamine inhibitors; Molecular modeling.
* Corresponding author. Tel.: +1 215 652 7183; fax: +1 215 652 4625; e-mail: georgia_mcgaughey@merck.com

0960-894X/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2006.11.003

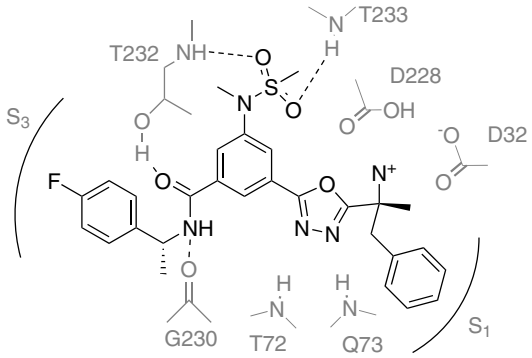


Figure 1. Representative tertiary carbinamine derived inhibitor, compound **1**. Key sidechain residues are colored in gray.

bonding distance to the T232 hydroxyl sidechain and the amide NH hydrogen bonds to G230. Finally, the α-methyl group gears the fluorophenyl moiety to occupy the S_3 pocket (Fig. 1).

In addition to the unique conformation of the catalytic aspartates, multiple X-ray structures of bound tertiary carbinamines revealed the 10s loop could reside in an 'up' conformation or in a 'down' conformation.[13] This motion is also seen in the published literature.[14]

Figure 2 depicts the 10s loop in both an up and down conformation. In the 10s loop up conformation, the heavy atom distance between the T232 hydroxyl side-chain and S10 backbone carbonyl oxygen is 9 Å. In the 10s loop down conformation this distance is 2.7 Å.

The observation that the 10s loop could reside in at least two distinct conformations led us to hypothesize that the movement was modulated by the ligand and could be described as an induced effect. Scoring ligands with varying substituents accessing S_3 in one fixed active site could be wrought with errors. How would we choose which 10s loop conformation to use for scoring and to ultimately prioritize compounds for synthesis?

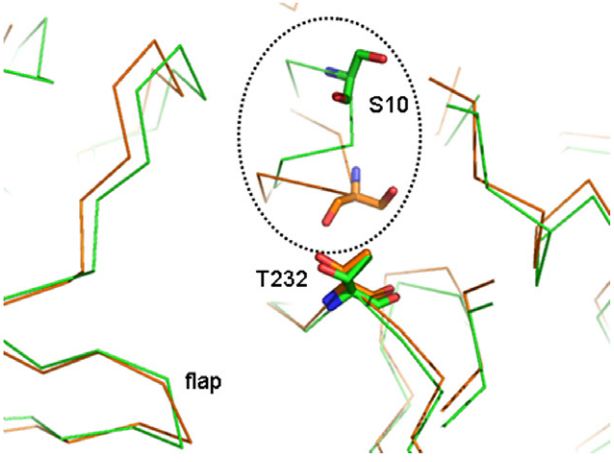


Figure 2. Two poses of the BACE-1 structure are superposed. The 10s up conformation is colored in green and the 10s down conformation is colored in orange. Residues S10 and T232 are depicted and colored by atom type. The 10s loop is circled and the flap is shown for reference. BACE-1 complexed with compound **1** reveals the 10s up conformation, whereas the 10s down structure was obtained from the 1fkn structure.[2] The figure was created in PyMol.[15]

A plot of the MMFFs E_{inter} versus the inverse log of the biological K_i (pK_i) is depicted in Figure 3.[16,17] The lower (or more negative) E_{inter} value correlates with a higher (or more potent) pK_i value. For instance, an E_{inter} value of −70 kcal/mol is indicative of a molecule with a binding constant in the nanomolar range, whereas a value of −60 kcal/mol suggests data in the micromolar range.

Based on the strong correlation between the computed E_{inter} values and experimentally determined K_i numbers, ($R^2 = 0.89$), we evaluated the inhibitor structural elements present in the most preferred inhibitor-protein poses to establish a set of criteria to develop a guide for predicting the configuration of the 10s loop when bound to a ligand. Apo BACE-1 depicts the 10s loop in an up configuration where the distance from the S10 carbonyl oxygen to the T232 hydroxyl oxygen is 6.8 Å.[3] If the S_3 substituent is large, the 10s loop will remain up to accommodate the steric bulk (compound **1**, $K_i = 42.4$ nM, Table 1). If the S_3 substituent is small, but the linker atoms between the central scaffold and the S_3 substituent interact with the T232 hydroxyl, then the 10s loop will remain up and activity is compromised (compound **2**, $K_i = 16.5$ μM, Table 1). The carbonyl oxygen in this linker region is within hydrogen bonding distance to the T232 hydroxyl hydrogen. Hence, we believe that the carbonyl oxygen accepts a hydrogen bond from the T232 hydroxyl hydrogen. In such a case, the T232 hydroxyl may not hydrogen bond to the S10 backbone carbonyl oxygen. If such a hydrogen bonding interaction could occur, the carbonyl oxygen from S10 would displace a crystallographically determined water molecule which resides in the S_3 pocket. If one reversed the amide, or introduced an aniline nitrogen, the hydrogen on the aniline nitrogen could donate a hydrogen bond to the T232 hydroxyl oxygen allowing the hydrogen of the hydroxyl group to engage in hydrogen bonding interactions with S10 thereby keeping the 10s loop 'down.'

An alternative linker is exemplified in compound **3**, Table 1. Here, the *cis* double bond contains no polar atom thereby eliminating the ability to establish a hydrogen

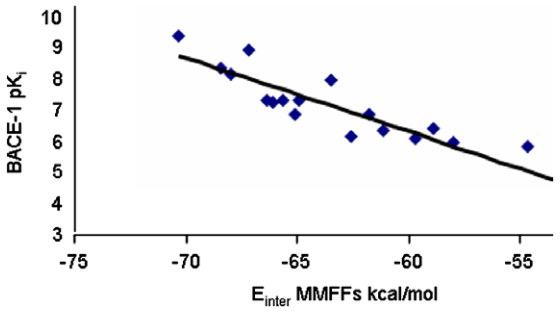
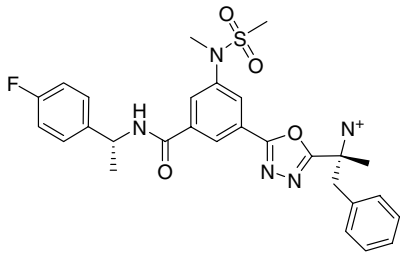
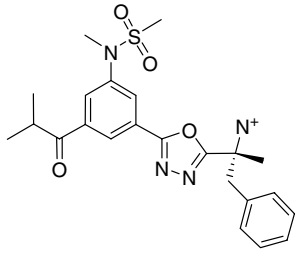
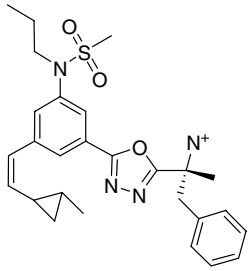


Figure 3. Graph showing the correlation between the biological pK_i value and the computed E_{inter} using MMFFs.

Table 1. Structure, biological activity, and calculated interaction energy of three tertiary carbinamines differing in the S_3 region

Compound	Structure	K_i (nM)	pK_i	Interaction energy (kcal/mol)
1		42.4	7.4	−66.4
2		16,500	4.8	−52.1
3		51.9	7.3	−66.1

bonding interaction with the T232 sidechain. In addition, the S_3 cyclopropyl group is small. This affords a unique opportunity for the 10s loop to come down.

A 2D representation of compound **3** in the BACE-1 active site is depicted in Figure 4. Here, one can see that the inhibitor has no hydrogen bonding capabilities with T232. Further, the small S_3 group affords the T232 hydroxyl to hydrogen bond with the S10 carbonyl allowing the 10s loop to be in a down configuration.

This is reflected in the K_i value of compound **3**, K_i = 51.9 nM and the E_{inter} value is −66.1 kcal/mol. If the 10s loop were up, the biological activity is predicted to be much higher (i.e., μM range) due to the size of the S_3 group and its inability to make optimal van der Waals interactions. This conjecture is quantified when we scored compound **3** in the 10s up active site. The E_{inter} of −58.5 kcal/mol is indicative of a micromolar K_i.

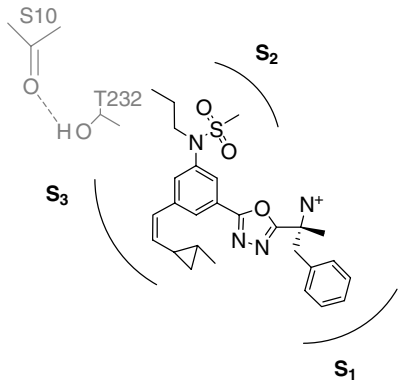


Figure 4. Proposed binding pose of compound **3** depicting the resultant impact on T232-S10 hydrogen bonding capabilities.

This 10s loop motion has been observed in other BACE-1 transition state isosteres such as the hydroxyethylamine[18] and aminoethylene series[19] leading us to believe the 10s loop movement is not dictated based on ligand interactions with the catalytic aspartic acids. Rather, the orientation of the linker region on the tri-substituted central aryl ring and its concomitant interaction with T232 modulates the 10s loop via an induced fit.

Nineteen ligands which differed primarily in the S_3 region were selected whose activity spanned many orders of magnitude. Because we knew that the 10s loop was flexible, four protein crystal structures were chosen exhibiting 10s up and down conformations. The 19 tertiary carbinamine ligands were scored in these four, fixed BACE-1 active sites. In four cases, the cognate was known eliminating the necessity for conformational searching within those active sites. In the other 15 cases where the ligand:protein complex (cognate) was not known, multiple conformations of the ligand were generated and each was scored in the four poses of the protein. All ligands occupied the non-prime side of the BACE-1 active site, spanning the S_1, S_2, and S_3 pockets. Thus, we held the tertiary carbinamine, oxadiazole, and

S_1 benzyl fixed, and generated conformations only of the S_2 and S_3 groups. We generated 10–30 conformations using a distance geometry method.[20,21] The exact number of conformations computed depended on the number of chiral atoms and rotatable bonds. The entire protein structure was kept, crystallographic waters were removed, and all amino acid charges reflected neutral pH. The charge on D32 was anionic and D228 was neutral.[22]

The resultant interaction energy (E_{inter}), which is a measure of the ligand:protein interaction and represents the enthalpy of binding, was then used as a 'score' based on minimizing each pose in the entire BACE-1 active site using a distance-dependent dielectric constant of two.[9,23,24] No solvation was utilized in the calculations. For each ligand, there were four scores resulting from scoring studies in four protein poses. We first visually inspected all minimized poses to ensure known, key inhibitor:protein binding contacts were still evident. Then, we took the lowest, most negative E_{inter} as the score only if the intramolecular energy, E_{intra}, a measure of the conformational strain, was also the lowest or within a few kcals/mol of the lowest energy conformer based on this analysis. If this was not the case, we chose the next lowest E_{inter} pose. Both the E_{intra} and E_{inter} energetic values were determined using the MMFFs forcefield.[16]

Merely scoring each ligand in one protein pose would not have yielded an optimal correlation. The R^2 correlation for the compounds scored in a representative 10s down active site was 0.01 and the R^2 correlation for the compounds scored in a 10s up active site is 0.32 (Table 2). This underscores the importance of including multiple active sites. The result makes intuitive sense as large S_3 pieces on the ligand would not be accommodated in a 10s down protein pose. On the other hand, a small S_3 ligand group could be physically accommodated in a 10s up protein pose; however, the van der Waals contact with the protein would be insufficient for tight binding to occur.

This would explain why the R^2 correlation was better for scoring all 19 compounds in one 10s up active site rather than scoring all compounds in the 10s down active site. Regardless, the use of four 10s loop poses was required for an optimal correlation. However, we found that a minimum of two protein poses representing one 10s up (pdb code: 2IRZ) and one 10s down (pdb code: 2NTR) protein structure is required. In such a case, utilizing two divergent poses yielded a R^2 correlation of 0.84. Finally, both van der Waals and electrostatics are contributing to the overall interaction. Utilizing merely the van der Waals components, a R^2 correlation of 0.71 is obtained; however, using only the electrostatics component yields a R^2 correlation of 0.59. This support the hypothesis that the dominant component for activity is modulated with an appropriately 'sized' group in S_3.

Table 2. R^2 values correlating pK_i with E_{inter} scores within different BACE-1 active sites

	Multiple sites	10s up	10s down
R^2	0.89	0.32	0.01

The use of in-house X-ray crystallography was central to the success of correlating a score with the biologically determined K_i values in the tertiary carbinamine series of BACE-1 inhibitors. We have found that through the incorporation of multiple protein structures which differ in 10s loop conformations, one can predict the biological activity of a series of tertiary carbinamine BACE-1 inhibitors using the E_{inter} as determined with the MMFFs force field. The more negative a score, the more likely the molecule will be potent in the primary in vitro BACE-1 assay. We believe 10s loop motion is governed by the choice of ligand atoms: non-interacting hydrogen bond acceptors in the linker region combined with small S_3 groups will likely result in a 10s loop down conformation. More sophisticated computational tools have been used by us and others to predict biological activity; however, we have found this simple scoring protocol to be an efficient means of rank-ordering compounds for synthesis, particularly when one is computationally scanning a pocket for the incorporation of reagents to occupy a specific portion of the BACE-1 active site.

Acknowledgments

The authors acknowledge the contributions of additional biological data led by A. Simon. In addition, M.K.H. and G.B.M. are grateful to J. C. Culberson for insightful suggestions during the course of this work. Finally, we thank J. Barrow for a critical reading of this manuscript.

References and notes

1. Cummings, J. L. *N. Eng. J. Med.* **2004**, *351*, 56.
2. Hong, L.; Koelsch, G.; Lin, X.; Wu, S.; Terzyan, S.; Ghosh, A. H.; Zhang, X. C.; Tang, J. *Science* **2000**, *290*, 150.
3. Patel, S.; Vuillard, L.; Cleasby, A.; Murray, C. W.; Yon, J. *J. Mol. Biol.* **2004**, *343*, 407.
4. Xiong, B.; Huang, X.; Shen, L.; Shen, J.; Luo, X.; Shen, X.; Jiang, H.; Chen, X. *Acta Pharmacol. Sin.* **2004**, *25*, 705.
5. Huang, D.; Lüthi, U.; Kolb, P.; Cecchini, M.; Barberis, A.; Caflisch, A. *J. Am. Chem. Soc.* **2006**, *128*, 5436.
6. Gorfe, A.; Caflisch, A. *Structure* **2005**, *13*, 1487.
7. Rajamani, R.; Reynolds, C. H. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 4843.
8. Tounge, B. A.; Reynolds, C. H. *J. Med. Chem.* **2003**, *46*, 2074.
9. Holloway, M. K.; McGaughey, G. B.; Coburn, C. A.; Stachel, S. J.; Jones, K. G.; Stanton, E. L.; Gregro, A. R.; Lai, M.-T.; Ming-Chih Crouthamel, M.-C.; Pietrak, B. L.; Munshi, S. *Bioorg. Med. Chem. Lett.* (in press).
10. Moitessier, N.; Therrien, E.; Hanessian, S. *J. Med. Chem.* **2006**, *49*, 5885.
11. Rajapakse, H. A.; Nantermet, P. G.; Selnick, H. G.; Munshi, S.; McGaughey, G. B.; Lindsley, S. R.; Young, M. B.; Lai, M.-T.; Shi, X.-P.; Colussi, D.; Pietrak, B.;

Crouthamel, M.-C.; Graham, S.; Vacca, J., *J. Med. Chem.* (in press).
12. Stachel, S. J.; Coburn, C. A.; Steele, T. G.; Jones, K. G.; Loutzenhiser, E. F.; Gregro, A. R.; Rajapakse, H. A.; Lai, M.-T.; Crouthamel, M.-C.; Xu, M.; Tugusheva, K.; Lineberger, J. E.; Pietrak, B. L.; Espeseth, A. S.; Shi, X.-P.; Chen-Dodson, E.; Holloway, M. K.; Munshi, S.; Simon, A. J.; Kuo, L.; Vacca, J. P. *J. Med. Chem.* **2004**, *47*, 6447.
13. The terminology '10s loop' was first used by Patel et al. (Ref. 3) and refers to amino acids 9–14 at the base of the S_3 pocket.
14. The Protein Data Bank code, 1fkn[2] (i.e., peptide complex), reveals the 10s loop in a down conformation, whereas the 1w50[3] apo-crystal structure reveals the 10s loop in an up conformation.
15. PyMol, v.99 distributed by DeLano Scientific LLC was utilized to generate this image.
16. (a) Halgren, T. A. J. *Comput. Chem.* **1996**, *17*, 490, 520, 553, and 616; (b) Halgren, T. A.; Nachbar, R. B. *J. Comput. Chem.* **1996**, *17*, 587.
17. Biological values were determined as described in: Pietrak, B. L.; Crouthamel, M. C.; Tugusheva, K.; Lineberger, J. E.; Xu, M.; DiMuzio, J. M.; Steele, T.; Espeseth, A. S.; Stachel, S. J.; Coburn, C. A.; Graham, S. L.; Vacca, J. P.; Shi, X.-P.; Simon, A. J.; Hazuda, D. J.; Lai, M.-T. *Anal. Biochem.* **2005**, *342*, 144.
18. Stachel, S. J.; Coburn, C. A.; Steele, T. G.; Crouthamel, M.-C.; Pietrak, B. L.; Lai, M.-T.; Holloway, M. K.; Munshi, S. K.; Graham, S. L.; Vacca, J. P. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 641.
19. Stauffer, S. R.; Stanton, M. G.; Gregro, A. R.; Steinbeiser, M. A.; Shaffer, J. R.; Nantermet, P. G.; Barrow, J. C.; Rittle, K. E.; Colussi, D.; Espeseth, A. S.; Lai, M.-T.; Peitrak, B. L.; Holloway, M. K.; McGaughey, G. B.; Munshi, S. K.; Hochman, J. H.; Simon, A. J.; Selnick, H. G.; Graham, S. L.; Vacca, J. P. *Bioorg. Med. Chem. Lett.* (submitted for publication).
20. Crippen, C. M.; Havel, T. F. D. Bawden, Ed.; Research Studies Press, Wiley, New York, 1988.
21. Kuszewski, J.; Nilges, M.; Brunger, A. T. *J. Biomol. NMR* **1992**, *2*, 33.
22. Rajamani, R.; Reynolds, C. H. *J. Med. Chem.* **2004**, *47*, 5159.
23. Holloway, M. K.; Wai, J. M.; Halgren, T. A.; Fitzgerald, P. M. D.; Vacca, J. P.; Dorsey, B. D.; Levin, R. B.; Thompson, W. J.; Chen, L. J., et al. *J. Med. Chem.* **1995**, *38*, 305.
24. McGaughey, G. B.; Culberson, J. C.; Feuston, B. P.; Kreatsoulas, C.; Maiorov, V.; Shpungin, J. *Mol. Divers.* **2006**, *10*, 341.