Supplementary Information

Discovery of cell-permeable nonpeptide inhibitors of β-secretase by high-throughput docking and continuum electrostatics calculations

Danzhi Huang[a†], Urs Lüthi[b†], Peter Kolb[a], Karin Edler[a,b], Marco Cecchini[a], Stephan Audetat[a,b], Alcide Barberis[b], and Amedeo Caflisch[a*]

Contents

1 Computational approach

The *in silico* screening approach consists of four steps which are presented in the following four subsections. In subsection **1.1** the first step (DAIM) is explained in detail whereas for the remaining three steps the main differences with respect to the original methods are highlighted (subsections **1.2-1.4**).

1.1 DAIM (Decomposition and Identification of Molecules)

The decomposition of a ligand into fragments and the choice of the anchor fragments for FFLD (see **1.3**) have been automatized recently (P. Kolb and A. Caflisch, manuscript in preparation). The major rules are listed here. The decomposition is guided by the fact that SEED (see **1.2**) treats all molecules as rigid. Hence, preference is given to aromatic rings but also other small rings and molecules that contain several amidic, double or triple bonds. The fact that non-aromatic ring systems might have several distinct conformations can be accounted for by the ability of SEED to dock multiple (predefined) conformations at the same time. If one of these conformations can be docked with a lower binding energy than the others, it will automatically be chosen in the subsequent steps, since it will receive higher ranks. FFLD requires three not-necessarily different fragments to place a flexible ligand unambiguously in the binding site. The fragment determination and selection of the three most suitable fragments for flexible ligand docking by SEED-FFLD follows a few simple rules:

1. All atoms in a fragment must be connected by rigid or terminal bonds.

2. Large fragments are preferred since there are more steric constraints for large entities, as a consequence these should be positioned first.

3. Cyclic fragments are preferred because they usually are more rigid than acyclic moieties.

4. Since the fragments should be involved in the most significant interactions, those that contain hydrogen bond donors and acceptors are selected. Charged

groups usually do not make such good anchors, since they tend to be positioned at the borders of the binding site, which are more exposed to the solvent. (However, there are exceptions as in the case of thrombin, where a very favorable electrostatic interaction is provided by a charged aspartic acid in the specificity pocket).

5. Fragments that are close to the center of the molecule are omitted, especially if they have a high number of substituent groups. Such "central" or "scaffold" fragments will hardly ever form significant interactions.

6. Finally, fragments should not overlap (i.e. one atom should not be part of two fragments), since this would mean that there are no rotatable bonds in between, so their relative position can not be changed.

The DAIM rules for fragment identification and selection of the three most suitable fragments for flexible docking by SEED-FFLD can be exemplified with the molecule XK263 (Dupont Merck, Fig. 1). In principle, there are three fragment types that

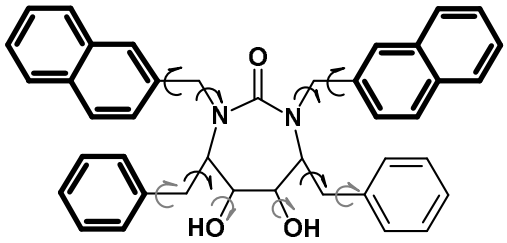

Figure 1: *XK263 (Dupont Merck) is a nanomolar inhibitor of HIV-1 aspartic protease (PDB accession code of the complex: 1HVR). Fragments selected by DAIM for SEED-FFLD docking are bold. Curly arrows denote rotatable bonds.*

could be chosen: naphthalene, benzene and the cyclic urea in the center. The largest fragment would be the cyclic urea. According to rule 5, this is not a good choice however, as it is the core fragment and has 4 substituents. Furthermore, it is the most flexible of the three types, which is another point against its choice. The remaining two types are aromatic and thus a recommended choice (rule 1). Finally, DAIM selects two naphthalenes and one benzene and not vice versa (rule 2).

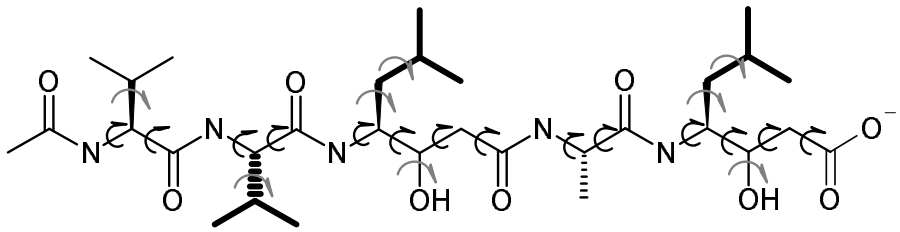

Figure 2: *Acetyl-pepstatin is a micromolar inhibitor of HIV-1 aspartic proteinase (PDB accession code of the complex: 5HVP). Fragments selected by DAIM for SEED-FFLD docking are bold. Curly arrows denote rotatable bonds.*

A more difficult choice is presented by acetyl-pepstatin (Fig. 2), since it has no rings and almost no rigid bonds. All the fragments obtained by the application of rule 1 are therefore very small. All the larger fragments with a rigid bond (the amide groups) are located in the backbone and will not make good anchors (rule 5). One of the few choices remaining is to select three i-butanes (the "side chains") which are preferable with respect to the terminal carboxylic group, which is charged (rule 4).

1.2 SEED (Solvation Energy for Exhaustive Docking)

The docking approach implemented in the program SEED [1] determines optimal positions and orientations of small to medium-size molecular fragments in the binding site of a protein. Apolar fragments are docked into hydrophobic regions of the receptor while polar fragments are positioned such that at least one intermolecular hydrogen bond is formed. Each fragment is placed at several thousand different positions with multiple orientations (for a total of in the order of 10^6 conformations) and the binding energy is estimated whenever severe clashes are not present (usually about 10^5 conformations). The binding energy is the sum of the van der Waals interaction and the electrostatic energy. The latter consists of screened receptor-fragment interaction, as well as receptor and fragment desolvations.

As an improvement with respect to previous versions of SEED [1, 2, 3] the screened electrostatic interaction as well as fragment desolvation energy were evaluated using an empirical correction of the Coulomb field approximation, i.e., equation 8 of

Ref. [4]. The SEED input parameters used for this application to BACE-1 are identical to those in Table I of the original SEED article [1], except for the following three: (i) The interior dielectric constant is set to 2.0 to partially account for the electronic polarizability and dipolar reorientation effects of the solute. (ii) The number of apolar points on the receptor is increased from 100 to 300 because of the very large substrate-binding site of BACE-1. (iii) To discard polar and apolar receptor vectors that point outside of the binding site, a selection using an angle criterion is performed. Initially, the minimal and maximal distances between the end points of the vectors and a set of points in the binding site (e.g., the positions of the heavy atoms of the ligand OM-003 are evaluated. A vector is discarded if the angle it spans with the closest point is larger than a cutoff. This selection uses a permissive cutoff of 100 degrees for vectors close to the binding site points and a stricter one (70 degrees) for distant vectors. SEED version 3.0 of March 2003 was used in this study.

1.3 FFLD (Fragment-based Flexible Ligand Docking)

The flexible-ligand docking approach FFLD uses a genetic algorithm and a very efficient scoring function [5]. The genetic algorithm perturbations affect only the conformation of the ligand; its placement in the binding site is determined by the SEED anchors and a least square fitting method [6]. In this way the position and orientation of the ligand in the binding site are determined by the best binding modes of its fragments previously docked using an accurate energy function with electrostatic solvation [7]. The scoring function used in FFLD is based on van der Waals and hydrogen bond terms and does not explicitly include solvation for efficiency reasons. Solvation effects are implicitly accounted for as the binding mode of the fragments are determined with electrostatic solvation.

The FFLD version 2.1 of July 2003, which contains the improvements presented in Ref. [3] as well as the ligand dihedral energy was used in the present application to BACE-1.

The FFLD poses are postprocessed by minimization with CHARMM [8].

1.4 LIECE (Linear Interaction Energy with Continuum Electrostatics)

The LIECE method was used exactly as described recently [9].

2 Experimental tests

2.1 BACE-1 enzymatic assay

The BACE-1 fluorescence resonance energy transfer (FRET) assay was performed as described by the manufacturer (PanVera, P2985). Briefly, fluorescence progress curves of 30 μl reaction volumes were measured on a Tecan *GENios* reader (Maennedorf, Switzerland) upon excitation at 535 nm and emission at 580 nm in 384-well microtiter plates (Corning, 3654). Linear regression analysis was calculated with Magellan 5.0 software (Tecan Austria GmbH, Salzburg).

2.2 Abeta(sw) (Amyloid β40 ELISA) assay

Swedish APP695 transgenic HEK 293 cells were maintained in Dulbecco's modified Eagle's medium (Sigma) supplemented with 10% fetal calf serum (Gibco) and 200 μg/ml G418 (Gibco) for continued selection of the stably integrated transgene, as described elsewhere [10]. Briefly, a 400x compound stock solution (dissolved in DMSO) was resuspended in 140 μl medium lacking G418 and distributed in poly-lysine-precoated 96-well cell culture plates (final DMSO concentration 0.25%). Immediately thereafter, 50,000 transgenic HEK 293 cells resuspended in 20 μl medium lacking G418 were added to each well. After 2 days of incubation at 37 °C and 5% CO2, an ELISA assay to measure Aβ40 in the supernatant was performed according to the protocol of the manufacturer of the assay kit (The Genetics Company, Switzerland). In parallel, an XTT assay of the cells was performed to measure cell viability, thus verifying that a reduction in the Aβ40 signal is not due to compound toxicity.

2.3 SEAP (secreted alkaline phosphatase) assay

HEK 293 cells were cultivated in Dulbecco's modified Eagle's medium (Sigma) supplemented with 10% fetal calf serum (Gibco). 50,000 cells resuspended in 100 μl medium were seeded per S2 well in poly-lysine precoated 96-well cell culture plates. After 6 hrs, 0.5 μl of a 400x compound stock solution (dissolved in DMSO) was resuspended in 49.5 μl medium and added to the wells (final DMSO concentration 0.25%). Thereafter, 0.3 μg of pBUDCE4.1/lacZ/CAT (Invitrogen) and 0.3 μg of SEAP-APP(sw) were resuspended in 45 μl medium lacking FCS, supplemented with the PolyFect transfection reagent and co-transfected according to the manufacturer (Qiagen). After 2.5 days incubation at 37 °C and 5% CO2, 50 μl of the supernatant was heated for 30 min at 65 °C , and SEAP activity was assayed according to the protocol of the manufacturer (phospha light SEAP reporter gene assay system, Applied Biosystems, T1017). The remaining cell monolayers in the 96-well plates were lysed in 0.2% Triton X-100 buffer, and a galactosidase assay was performed. Measured luminescence values were then normalized to galactosidase activity in order to account for transfection efficiency.

References

[1] Majeux, N.; Scarsi, M.; Apostolakis, J.; Ehrhardt, C.; Caflisch, A. Exhaustive docking of molecular fragments on protein binding sites with electrostatic solvation. *Proteins: Structure, Function and Genetics* **1999**. *37*, 88–105.

[2] Majeux, N.; Scarsi, M.; Caflisch, A. Efficient electrostatic solvation model for protein-fragment docking. *Proteins: Structure, Function and Genetics* **2001**. *42*, 256–268.

[3] Cecchini, M.; Kolb, P.; Majeux, N.; Caflisch, A. Automated docking of higly flexible ligands by genetic algorithms: A critical assessment. *J. Comput. Chem.* **2004**. *25*, 412–422.

[4] Lee, M. S.; Salsbury, F. R.; Brooks III, C. L. Novel generalized born methods. *J. Chem. Phys.* **2002**. *116*, 10606–10614.

[5] Budin, N.; Majeux, N.; Caflisch, A. Fragment-based flexible ligand docking by evolutionary opimization. *Biol. Chem.* **2001**. *382*, 1365–1372.

[6] Kabsch, W. A solution for the best rotation to relate two sets of vectors. *Acta Cryst.* **1976**. *A32*, 922–923.

[7] Scarsi, M.; Apostolakis, J.; Caflisch, A. Continuum electrostatic energies of macromolecules in aqueous solutions. *J. Phys. Chem. A* **1997**. *101*, 8098–8106.

[8] Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; et al. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983**. *4*, 187–217.

[9] Huang, D.; Caflisch, A. Efficient evaluation of binding free energy using continuum electrostatic solvation. *J. Med. Chem.* **2004**. *47*, 5791–5797.

[10] Citron, M. β-Secretase inhibition for the treatment of Alzheimer's disease: promise and challenge. *Trends in Pharmacological Sciences* **2004**. *25*, 92–97.