J|A|C|S
ARTICLES
Published on Web 03/30/2006

In Silico Discovery of β-Secretase Inhibitors

Danzhi Huang,[†] Urs Lüthi,[‡] Peter Kolb,[†] Marco Cecchini,[†] Alcide Barberis,[‡] and Amedeo Caflisch*,[†]

Contribution from the Department of Biochemistry, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland, and ESBATech AG, Wagistrasse 21, CH-8952 Zürich-Schlieren, Switzerland

Received October 26, 2005; E-mail: caflisch@bioc.unizh.ch

Abstract: Alzheimer's disease, the most common amyloid-associated disorder, accounts for the majority of the dementia diagnosed after the age of 60. The cleavage of the β-amyloid precursor protein is initiated by β-secretase (BACE-1), a membrane-bound aspartic protease, which has emerged as an important but difficult protein target. Here, an in silico screening approach consisting of fragment-based docking, ligand conformational search by a genetic algorithm, and evaluation of free energy of binding was used to identify low-molecular-weight inhibitors of BACE-1. More than 300 000 small molecules were docked and about 15 000 prioritized according to a linear interaction energy model with evaluation of solvation by continuum electrostatics. Eighty-eight compounds were tested in vitro, and 10 of them showed an IC_{50} value lower than 100 μM in a BACE-1 enzymatic assay. Interestingly, the 10 active compounds shared a triazine scaffold. Moreover, four of them were active in an assay with mammalian cells ($EC_{50} < 20$ μM), indicating that they are cell-permeable. Therefore, these triazine derivatives are very promising lead candidates for BACE-1 inhibition. The discoveries of this series and two other series of nonpeptidic BACE-1 inhibitors demonstrate the usefulness of our in silico high-throughput screening approach.

Introduction

Insoluble, extracellular amyloid plaques, a histopathological hallmark in the post-mortem brain of Alzheimer's disease patients,[1] consist mainly of fibrillar aggregates of the amyloid-β (Aβ) peptide, which is a proteolytic cleavage product of the β-amyloid precursor protein (APP). Two enzymes, γ- and β-secretase (β-site APP cleaving enzyme, or BACE-1), are responsible for the sequential processing of APP.[2] Genetic deletion of BACE-1 in mice has been shown to abolish β-amyloid formation with an otherwise normal, i.e., healthy, phenotype.[3] Although there is no definitive evidence whether the plaques or oligomeric prefibrillar species are responsible for neuronal loss and dementia,[4] the pepsin-like aspartic protease BACE-1 is considered an important target for the development of small-molecule inhibitors to fight Alzheimer's disease.[5,6] The relatively small number of known nonpeptide inhibitors indicates that BACE-1 is not an easy target to block.[5−8] In fact, not a single BACE-1 inhibitor was found in a library containing more than 1800 renin inhibitors,[9] despite the fact that both BACE-1 and renin are pepsin-like enzymes. Furthermore, only a single molecule (1,3,5-trisubstituted benzene) emerged as BACE-1 inhibitor from a multimillion compound library submitted to a high-throughput in vitro screening campaign.[10] It is also important to note that the recently reported peptidomimetics with low nanomolar affinity in BACE-1 enzymatic assays are not active in cell-based assays because of limited penetration across cell membranes.[11] Here, we report the successful application of our in silico high-throughput docking approach in the screening of more than 300 000 existing compounds, which has resulted in the discovery of a series of nonpeptide BACE-1 inhibitors with a common (1,3,5-triazin-2-yl)hydrazone scaffold. The fragment-based docking procedure, which takes into account electrostatic solvation, shows a high hit rate and generates few false positives. Most notably, the combination of in silico screening with validation by enzymatic and cell-based assays has led to the identification of several molecules with excellent potential as lead compounds against BACE-1.

Methods

The essential elements of our in silico screening are a fragment-based docking procedure and an efficient evaluation of binding free energy with electrostatic solvation. The latter is presented first because of its importance for ranking compounds, which is the most challenging part of the in silico approach.

[†] University of Zürich.
[‡] ESBATech AG.

(1) Selkoe, D. J. *Nature* **1999**, *399*, A23−A31.
(2) Lin, X.; Koelsch, G.; Wu, S.; Downs, D.; Dashti, A.; Tang, J. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 1456−1460.
(3) Luo, Y.; et al. *Nat. Neurosci.* **2001**, *4*, 231−232.
(4) Petkova, A. T.; Leapman, R. D.; Guo, Z.; Yau, W. M.; Mattson, M. P.; Tycko, R. *Science* **2005**, *307*, 262−265.
(5) Citron, M. *Trends Pharmacol. Sci.* **2004**, *25*, 92−97.
(6) Cumming, J. N.; Iserloh, U.; Kennedy, M. E. *Curr. Opin. Drug. Discovery Dev.* **2004**, *7*, 536−556.
(7) Roggo, S. *Curr. Top. Med. Chem.* **2002**, *2*, 359−370.
(8) Middendorp, O.; Lüthi, U.; Hausch, F.; Barberis, A. *Biol. Chem.* **2004**, *385*, 481−485.
(9) Grüninger-Leitch, F.; Schlatter, D.; Küng, E.; Nelböck, P.; Döbeli, H. *J. Biol. Chem.* **2002**, *277*, 4687−4693.
(10) Coburn, C. A.; Stachel, S. J.; Li, Y. M.; Rush, D. M.; Steele, T. G.; Chen-Dodson, E.; Holloway, M. K.; Munshi, S.; Simon, A. J.; Kuo, L.; Vacca, J. P. *J. Med. Chem.* **2004**, *47*, 6117−6119.
(11) Hanessian, S.; et al. *J. Med. Chem.* **2005**, *48*, 5175−5190.

10.1021/ja0573108 CCC: $33.50

Evaluation of Binding Free Energy with LIECE. The linear interaction energy with continuum electrostatics (LIECE) approach was recently reported elsewhere.[12] Here, only a brief overview of the method is presented, while the in-depth validation for BACE-1 is given in the Results and Discussion. The essential aspect of the linear interaction energy (LIE) method is that the free energy of binding can be calculated by considering only the end points of the thermodynamic cycle of ligand binding, i.e., bound and free states. For this purpose, Åqvist and coworkers proposed to calculate average values of interaction energies from molecular dynamics (MD) simulations of the isolated ligand and the ligand/protein complex.[13,14] They approximated the free energy of binding by

$$\Delta G_{\text{bind}} = {}^1/_2 \left(\langle E^{\text{elec}} \rangle_{\text{bound}} - \langle E^{\text{elec}} \rangle_{\text{free}} \right) + \alpha \left(\langle E^{\text{vdW}} \rangle_{\text{bound}} - \langle E^{\text{vdW}} \rangle_{\text{free}} \right) \quad (1)$$

where E^{elec} and E^{vdW} are the electrostatic and van der Waals interaction energies between the ligand and its surroundings. The surroundings are either the solvent (free) or the solvated ligand/protein complex (bound). The $\langle\ \rangle$ denotes an ensemble average sampled over a molecular dynamics (MD)[13] or Monte Carlo[15] trajectory, and the parameter α is determined empirically.[13] The LIE method is faster than rigorous free energy perturbation techniques and has been successfully applied in the design of a series of inhibitors of the malarial aspartic proteases Plm I and II.[16] Yet, LIE cannot be used for high-throughput docking because of its computational requirements (about 1 day for each compound). Therefore, we have replaced the MD sampling with a simple energy minimization and combined the LIE method with a rigorous treatment of solvation within the continuum electrostatics (CE) approximation,[12] i.e., the numerical solution of the Poisson equation by the finite-difference technique.[17] The LIECE approach is about 2 orders of magnitude faster than previous LIE methods and shows a similar precision on the targets tested. In fact, a predictive accuracy of about 1.0 kcal/mol was observed for 13 and 29 peptidic inhibitors of BACE-1 and HIV-1 protease, respectively.[12]

Preparation of the BACE-1 Structure. The X-ray structure of BACE-1 from its complex with a nanomolar peptidic inhibitor (PDB code 1M4H[18]) was used for the in silico screening because BACE-1/nonpeptide inhibitor structures were not available when this work was initiated. The side chain of Asp228 was protonated[12] (see also below), while all other Asp and Glu side chains were considered negatively charged and the Lys and Arg positively charged. Further details on the protein preparation for docking can be found in the Supporting Information.

Preparation of the Compound Libraries. Two unrelated libraries were screened in silico. The first contains about 10 000 molecules with an average molecular weight of 497.3 ± 42.8 g/mol (Chemdiv Inc., 2002). The second library is a subset of about 300 000 molecules (424.9 ± 71.4 g/mol) selected from a collection of chemical libraries of about six million compounds (Chemnavigator Inc., 2004). For this selection, the size and physicochemical character of the substrate binding site were taken into account by filtering out compounds with molecular weight smaller than 200 g/mol or larger than 700 g/mol, and molecules without at least one hydrogen bond donor and acceptor. The 2D-to-3D conversion was performed using CORINA.[19] This step was followed by the determination of the protonation state and hydrogen coordinates generation with BABEL,[20] the assignment of CHARMm atom types[21] and partial charges,[22,23] and energy minimization with a distance-dependent dielectric function.

High-Throughput Fragment-Based Docking. The library-docking approach consists of four consecutive steps: (1) decomposition of each molecule of the library into mainly rigid fragments, (2) fragment docking with evaluation of electrostatic solvation, (3) flexible docking of each molecule of the library using the position and orientation of its fragments as anchors, and (4) LIECE evaluation of the binding free energy for the best poses. The first three steps are performed by in-house-developed computer programs, while CHARMM[24] is used for the energy minimization and finite-difference Poisson calculations in the fourth step. The main aspects of the docking approach are illustrated in the four following subsections, while the details are given in the Supporting Information.

(1) Decomposition of Library Compounds into Fragments. The decomposition of a molecule into mainly rigid substructures and the selection of the three anchor fragments for docking are performed by the program DAIM (Decomposition And Identification of Molecules, P. Kolb and A. Caflisch, unpublished results). The major rules are listed in the Supporting Information. The decomposition generates mainly rigid fragments which can be docked very efficiently (see below).

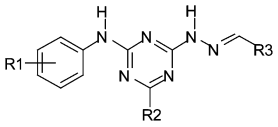
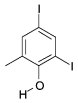
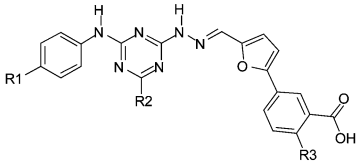
(2) Fragment Docking with Evaluation of Electrostatic Solvation. The docking approach implemented in the program SEED determines optimal positions and orientations of small to medium-size molecular fragments in the binding site of a protein.[25,26] Apolar fragments are docked into hydrophobic regions 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 values of receptor and fragment desolvation.[27]

(3) Flexible Docking of Library Compounds. The flexible-ligand docking approach FFLD uses a genetic algorithm and a very efficient but approximate scoring function.[28,29] FFLD requires three not necessarily different fragments to place a flexible ligand unambiguously in the binding site, e.g., the fluorobenzene, piperidine, and phenol of compound **5** (Table 1). Solvation effects are implicitly accounted for as the binding modes of the fragments are determined with electrostatic solvation in SEED. Each molecule was docked by three independent FFLD runs using a population of 100 members for each run and different initial values for the random number generator.

(4) Clustering and LIECE Binding Energy Evaluation. For each compound, the best 150 FFLD poses (50 poses from each FFLD run) were clustered by using a leader algorithm with a similarity cutoff of 0.7.[25,30] The representative of each cluster was selected for further

(12) Huang, D.; Caflisch, A. *J. Med. Chem.* **2004**, *47*, 5791−5797.
(13) Åqvist, J.; Medina, C.; Samuelsson, J.-E. *Protein Eng.* **1994**, *7*, 385−391.
(14) Hansson, T.; Åqvist, J. *Protein Eng.* **1995**, *8*, 1137−1144.
(15) Jones-Hertzog, D. K.; Jorgensen, W. H. *J. Med. Chem.* **1996**, *40*, 1539−1549.
(16) Ersmark, K.; Nervall, M.; Hamelink, E.; Janka, L. K.; Clemente, J. C.; Dunn, B. M.; Blackman, M. J.; Samuelsson, B.; Åqvist, J.; Hallberg, A. *J. Med. Chem.* **2005**, *48*, 6090−6106.
(17) Warwicker, J.; Watson, H. C. *J. Mol. Biol.* **1982**, *157*, 671−679.
(18) Hong, L.; Turner, R. T.; Koelsch, G.; Shin, D.; Ghosh, A. K.; Tang, J. *Biochemistry* **2002**, *41*, 10963−10967.
(19) Sadowski, J.; Gasteiger, J. *Chem. Rev.* **1993**, *93*, 2567−2581.
(20) Stahl, M. T. *Drug Discovery Today* **2005**, *10*, 219−222.
(21) Momany, F.; Rone, R. *J. Comput. Chem.* **1992**, *13*, 888−900.
(22) No, K.; Grant, J.; Scheraga, H. *J. Phys. Chem.* **1990**, *94*, 4732−4739.
(23) No, K.; Grant, J.; Jhon, M.; Scheraga, H. *J. Phys. Chem.* **1990**, *94*, 4740−4746.
(24) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(25) Majeux, N.; Scarsi, M.; Apostolakis, J.; Ehrhardt, C.; Caflisch, A. *Proteins: Struct., Funct., Bioinformatics* **1999**, *37*, 88−105.
(26) Majeux, N.; Scarsi, M.; Caflisch, A. *Proteins: Struct., Funct., Bioinformatics* **2001**, *42*, 256−268.
(27) Scarsi, M.; Apostolakis, J.; Caflisch, A. *J. Phys. Chem. A* **1997**, *101*, 8098−8106.
(28) Budin, N.; Majeux, N.; Caflisch, A. *Biol. Chem.* **2001**, *382*, 1365−1372.
(29) Cecchini, M.; Kolb, P.; Majeux, N.; Caflisch, A. *J. Comput. Chem.* **2004**, *25*, 412−422.
(30) Kearsley, S. K.; Smith, G. M. *Tetrahedron Comput. Methodol.* **1990**, *3*, 615−633.

Table 1. BACE-1 Inhibitors Identified by High-Throughput Fragment-Based Docking

COMPOUND	STRUCTURE R1	STRUCTURE R2	STRUCTURE R3	MW (g mol^{-1})	BACE-1[a] IC_{50} (μM)	Abeta(sw)[b] EC_{50} (μM)	Cytotoxic[c] CC_{50} (μM)	LIECE[d] K_i (μM)
								
1	m-H, p-H			466	11.2 ± 0.2	>10	19.1	20.2
2	m-H, p-H			511	11.9 ± 4.9	>10	16.7	74.2
3	m-H, p-H			649	25.5 ± 8.4	9.4	31.2	38.8
4	m-Cl, p-CH_3			598	20.6 ± 1.8	>10	>50	49.4
								
5	F		OH	518	27.9 ± 4.2	16.9 ± 1.8	>50	2.9
6	H		Cl	518	151.8 ± 14.5	18.0 ± 7.6	>50	6.6
7	O-CH3		H	522	66.6 ± 11.0	10.9 ± 5.4	28.0	12.6
8	F		Cl	538	7.1 ± 1.2	>25	>50	12.8

[a] The BACE-1 fluorescence resonance energy transfer assay kit was purchased from PanVera (Madison, WI; catalog no. P2985). BACE-1 activity assays were carried out according to the manufacturer's instructions.Values of average and standard deviation are from three independent experiments. [b] Cell-based assay.[45] [c] Cytotoxic concentration in HEK293 cells (not transgenic).[47] [d] See ref 12.

CHARMM minimization with distance-dependent dielectric function. During minimization, the protein was kept rigid. In the larger of the two screening experiments (306 022 compounds, see Results and Discussion), the minimized poses were re-ranked by the LIECE model using first a spacing of 1.0 Å for the finite-difference Poisson calculation as a filter and finally a grid spacing of 0.3 Å for the top 8000 poses, i.e., 2880 compounds.

Computational Requirements. The LIECE approach requires 26 min (mainly for the finite-difference Poisson calculations with grid spacing of 0.3 Å) of a CPU of a single Opteron 244 (1.8 GHz) for each pose in BACE-1. The total CPU time can be further reduced by first using a coarse grid spacing of 1.0 Å in the finite-difference Poisson calculation, which takes about half a minute. The in silico screening of the 306 022 compound library, i.e., docking and LIECE energy evaluation, took about 10 days on a Beowulf cluster of 100 Opteron 1.8 GHz CPUs.

BACE-1 Enzymatic Assay. The BACE-1 fluorescence resonance energy transfer (FRET) assay was performed as described by the manufacturer (PanVera, P2985) with an incubation time of 30 min. Additional measurements were performed in the presence of detergent or with an incubation time of only 3 min to check for nonspecific effects (e.g., compound aggregation[31,32]). Briefly, fluorescence progress curves of 30 μL reaction volumes were measured on a Tecan GENios reader (Männedorf, Switzerland) upon excitation at 535 nm and emission at 580 nm in 384-well microtiter plates (Corning, 3654). Linear regression analysis was calculated with the Magellan 5.0 software (Tecan Austria GmbH, Salzburg, Austria).

Abeta(sw) (Amyloid β40 ELISA) Cell-Based Assay. Swedish APP695 transgenic human embryonic kidney 293 cells (HEK 293) 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.[5] Briefly, a 400× compound stock solution (dissolved in DMSO) was resuspended in 140 μL of medium lacking G418 and distributed in poly-L-lysine-precoated 96-well cell culture plates (final DMSO concentration 0.25%). Immediately thereafter, 50 000 transgenic HEK 293 cells, resuspended in 20 μL of medium lacking G418, were added to each well. After 2 days of incubation at 37° C and 5% CO_2,

(31) Ryan, A. J.; Gray, N. M.; Lowe, P. N.; Chung, C. W. *J. Med. Chem.* **2003**, *46*, 3448−3451.
(32) McGovern, S. L.; Helfand, B. T.; Feng, B.; Shoichet, B. K. *J. Med. Chem.* **2003**, *46*, 4265−4272.

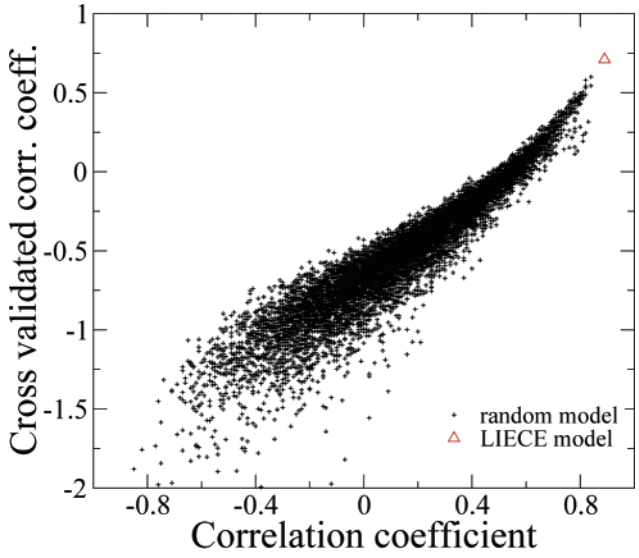


Figure 1. Statistical test to assess the predictive power of the LIECE two-parameter model for BACE-1 compared to 10 000 random models (see main text for details). The fact that the LIECE model data point (red triangle) is on the right-top indicates that LIECE not only better fits the data than the random models (black crosses) but has also a better predictive ability.

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.

Results and Discussion

Validation of the LIECE Model on BACE-1. As in our previous works,[12,33] a two-parameter LIECE model is used here: $\Delta G_{\text{bind}} = 0.2737\ \Delta E^{\text{vdW}} + 0.1795\ \Delta G^{\text{elec}}$, where ΔE^{vdW} is the ligand/protein van der Waals interaction energy and ΔG^{elec} is the sum of the ligand/protein Coulombic energy in vacuo and the change in solvation energy of ligand and protein upon binding. Note that the values of the two parameters ($\alpha = 0.2737$ and $\beta = 0.1795$) were obtained by using a training set of 13 peptidic inhibitors[34] in our previous work[12] and have not been modified since. To further evaluate the predictive power of the LIECE model for BACE-1, three additional tests were performed.

First, a statistical test based on the randomization of the data points was used to analyze an eventual chance correlation.[35,36] The binding free energies of 13 peptidic inhibitors[34] were randomized within the same range as the experimental values, i.e., from -14 to -6 kcal/mol, and the two multiplicative parameters for ΔE^{vdW} and ΔG^{elec} were determined by fitting to random "data points". The randomization and fitting were repeated 10 000 times, and Figure 1 shows the cross-validated correlation coefficient (obtained by the leave-one-out procedure) plotted versus the correlation coefficient. The LIECE model with the two parameters fitted to the real data points is located in the top right corner and is significantly separated from the models generated by the randomization of the binding free energies. This separation provides further evidence that the LIECE two-parameter model not only fits the experimental data but also has very good predictive ability, i.e., chance correlation is not present.

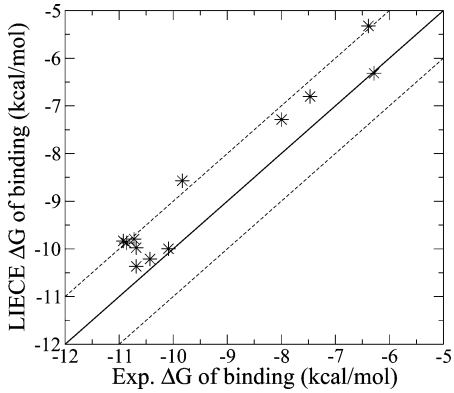


Figure 2. Cross-validation of the LIECE two-parameter model on 12 BACE-1 inhibitors consisting of a 1,3,5-trisubstituted benzene scaffold.[10,38] These 12 BACE-1 inhibitors were not used to derive the LIECE two-parameter model. The dashed lines emphasize the region corresponding to a 1 kcal/mol accuracy.

Second, the recent publication of two X-ray structures of BACE-1 in the complex with nonpeptide inhibitors (PDB codes 1W51[37] and 1TQF[10]) allowed us to perform additional tests of the two-parameter model and its robustness with respect to different protein structures. The LIECE-predicted binding affinity of the 1W51 nonpeptide inhibitor was calculated using two BACE-1 structures, 1W51 and 1M4H.[18] Both calculations gave a LIECE K_i of 0.49 μM, which is close to the experimental IC_{50} of 0.2 μM. Furthermore, we tested a series of 12 inhibitors of BACE-1 based on a 1,3,5-trisubstituted benzene scaffold,[10,38] which adopt a nontraditional binding mode with a displacement of the 10s loop with respect to the 1M4H conformation. These compounds were manually docked into the binding site of 1TQF using the 1TQF inhibitor as a template.[10] The LIECE binding free energy values are plotted versus the corresponding experimental values in Figure 2 (see also Table 1 in the Supporting Information). Remarkably, the root-mean-square of the error and maximal error are 0.78 and 1.3 kcal/mol, respectively, and the correlation coefficient is 0.89. In addition, the LIECE model successfully reproduces the binding energy change between two compounds which differ only in the stereochemistry at the α-methyl group pointing toward the P_3 pocket (compounds **3** and **4** of ref 10). Therefore, the LIECE two-parameter model derived from a single structure (1M4H) shows good predictive ability on a different class of inhibitors, even when the calculations are based on slightly different BACE-1 conformations. This result agrees with the previous work on HIV-1 protease inhibitors binding free energy calculation, where the parameters derived from a single structure were used to accurately predict the activity of a different series of inhibitors.[12] Furthermore, a recent application on 48, 62, and 41 inhibitors of Lck, CDK2, and p38 kinases, respectively, indicates that the LIECE model is transferable among enzymes which share a similar ATP binding site (P. Kolb, D. Huang, F. Dey, and A. Caflisch, manuscript in preparation). Transferability of LIECE parameters between slightly different structures of a given

(33) Huang, D.; Lüthi, U.; Kolb, P.; Edler, K.; Cecchini, M.; Audetat, S.; Barberis, A.; Caflisch, A. *J. Med. Chem.* **2005**, *48*, 5108−5111.
(34) Ghosh, A. K.; Bilcer, G.; Harwood, C.; Kawahama, R.; Shin, D.; Hussain, K. A.; Hong, L.; Loy, J. A.; Nguyen, C.; Koelsch, G.; Ermolieff, J.; Tang, J. *J. Med. Chem.* **2001**, *44*, 2865−2868.
(35) Zoete, V.; Michielin, O.; Karplus, M. *J. Comput.-Aided Mol. Des.* **2003**, *17*, 861−880.
(36) So, S.; Karplus, M. *J. Med. Chem.* **1999**, *39*, 5246−5256.
(37) Patel, S.; Vuillard, L.; Cleasby, A.; Murray, C. W.; Yon, J. *J. Mol. Biol.* **2004**, *343*, 407−416.
(38) Stachel, S. J.; et al. *J. Med. Chem.* **2004**, *47*, 6447−6450.

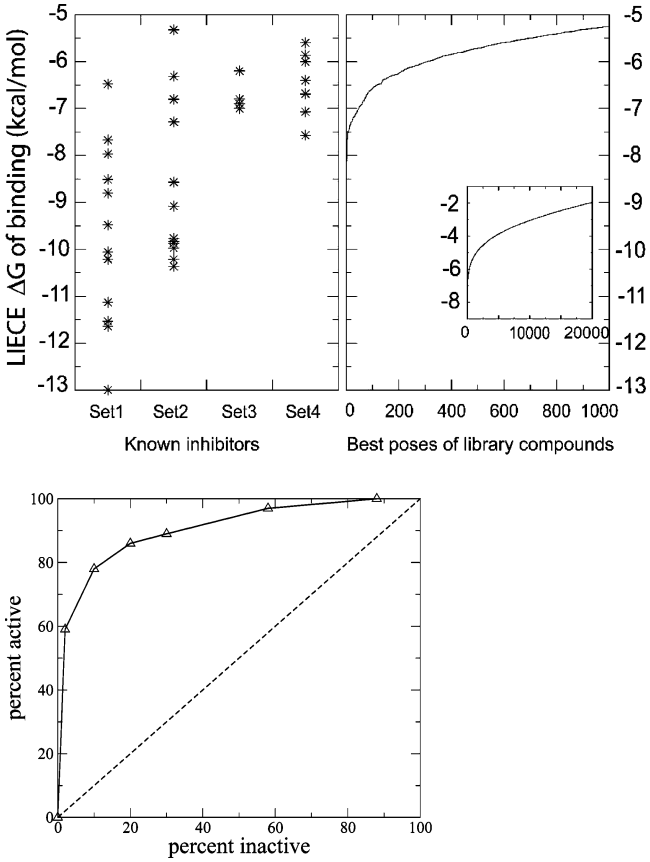


Figure 3. (Top) The LIECE two-parameter model does not generate too many false positives. Comparison between a composite set of 37 known inhibitors of BACE-1 (left) and the 1000 poses with the most favorable LIECE energy from a library of 200 000 small molecules (right), the vast majority of which are not expected to bind to BACE-1. Set1, 13 peptidic inhibitors developed in Tang's group;[34] Set2, 12 derivatives of a 1,3,5-trisubstituted benzene scaffold[10,38] (see also Supporting Information); Set3, four phenylurea derivatives;[33] Set4, eight compounds listed in Table 1. The inset on the right plot zooms out on the first 20 000 poses of the library compounds. (Bottom) Thirty-seven known inhibitors compared to the top 1000 poses as a ROC curve (solid line with triangles). The dashed line of slope 1 shows the behavior of a random model as a basis of comparison. The area under the ROC curve is close to the ideal value, which indicates that the LIECE model generates few false positives.

protein is a useful property, which could be used to take into account binding site flexibility during in silico screening or hit explosion. In this context, we have generated a set of low-energy conformations of BACE-1 using molecular dynamics with explicit water.[39] Because of the transferability of the LIECE parameters, a virtual screening based on these structures may find inhibitors with new binding modes.

Third, it is useful to estimate the amount of false positives, i.e., compounds with good predicted affinity which in reality do not bind. For this purpose, LIECE binding energies of a composite set of 37 BACE-1 inhibitors were compared with those of a library of about 200 000 small molecules (average value of molecular weight of 407.2 ± 72.2 g/mol; this library is unrelated to the libraries used for the in silico screening described here) under the reasonable assumption that very few of the 200 000 compounds inhibit BACE-1. The compounds were docked by FFLD, minimized by CHARMM in the rigid 1M4H structure, and the resulting poses were filtered according to two criteria: the van der Waals intermolecular energy (more favorable than −40 kcal/mol) and the van der Waals intermolecular energy divided by molecular weight (quotient more favorable than −0.1 kcal/g). Figure 3 (top) shows a comparison between the 37 inhibitors (left) and the library of 200 000 compounds (right). Remarkably, 78% and 100% of the known inhibitors have a LIECE energy in the range of values of the 111 and 1000 library compound poses with the most favorable LIECE energy, respectively. The 111 and 1000 poses originate from 100 and 651 different compounds, respectively. In other words, the large majority of the 200 000 compounds are predicted to be worse than most of the known inhibitors. Furthermore, a receiver operating characteristic (ROC) plot[40] for the known 37 compounds over the top 1000 poses (Figure 3 bottom) confirms that the LIECE model of BACE-1 does not generate many false positives.

Effect of Different Protonation States. The protonation state of the catalytic dyad has been investigated by different groups recently.[41,42] The main observation is that only one of the two aspartate side chains should be protonated in the presence of an inhibitor. However, it is still under debate which of the two side chains should be protonated. All calculations in the present study and previous works[12,33] have been performed with Asp228 protonated and Asp32 negatively charged. To test the robustness of this choice, docking of the 306 022 compounds of the second library was repeated using the BACE-1 structure with Asp32 protonated and Asp228 negatively charged. The range of LIECE energies of the top 500 poses (332 compounds) was −8.40 to −5.14 kcal/mol, which is comparable to the previous screening (−8.52 to −5.75 kcal/mol). Importantly, there were 194 compounds in common between the two lists (58%). The four active compounds (**5**−**8**) were ranked in the top 500 list upon docking with Asp32 protonated, and their LIECE affinities were 18.9, 116.8, 105.5, and 54.7 μM, respectively. These values are about an order of magnitude less favorable than those obtained with Asp228 protonated (see below and Table 1).

In Silico Screening and Enzymatic Assay. The DAIM decomposition of the 10 067 and 306 022 compound libraries yielded 469 and 4917 unique fragments, respectively. In the first in silico screening (Figure 4, left), 10 067 compounds were docked, 1000 poses were further evaluated by LIECE energy (1000 unique molecules), 64 compounds (19 of which with a (1,3,5-triazin-2-yl)hydrazone scaffold) were tested in an enzymatic assay, and seven (11%) showed an IC_{50} for BACE-1 smaller than 100 μM. The LIECE ranking of the seven active compounds was among the first 24 of 1000 molecules. Strikingly, the high hit rate was achieved using solely the LIECE energy ranking without manual intervention or visual inspection.

In the second in silico screening (Figure 4, right), 306 022 compounds were docked, 58 000 poses were further evaluated by LIECE (14 085 unique molecules), and 24 compounds (six of which with a (1,3,5-triazin-2-yl)hydrazone scaffold and a carboxy group which was negatively charged for docking and LIECE) were tested in an enzymatic assay. Three of them (12%) showed an IC_{50} smaller than 100 μM, and a fourth compound showed an IC_{50} of 152 μM. Remarkably, these compounds have LIECE ranks of first, fourth, seventh, and eighth.

To obtain information on the mechanism of inhibition and provide evidence against nonspecific effects (e.g., aggrega-

(39) Gorfe, A. A.; Caflisch, A. *Structure* **2005**, *13*, 1487−1498.
(40) Zweig, M. H.; Campbell, G. *Clin. Chem.* **1993**, *39*, 561−577.
(41) Park, H.; Lee, S. *J. Am. Chem. Soc.* **2003**, *125*, 16416−16422.
(42) Rajamani, R.; Reynolds, C. H. *J. Med. Chem.* **2004**, *47*, 5159−5166.

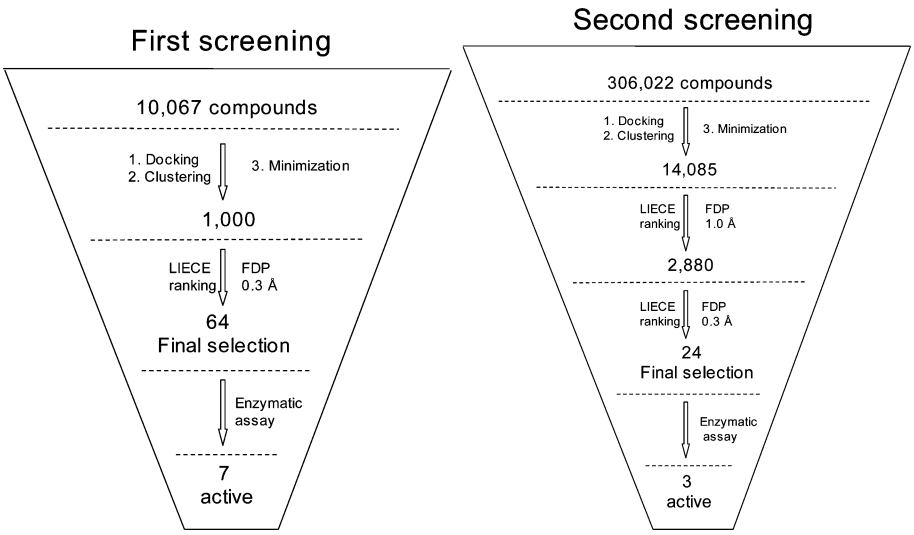


Figure 4. Schematic picture of the two applications of the in silico screening approach. FDP stays for finite-difference Poisson calculations.[17]

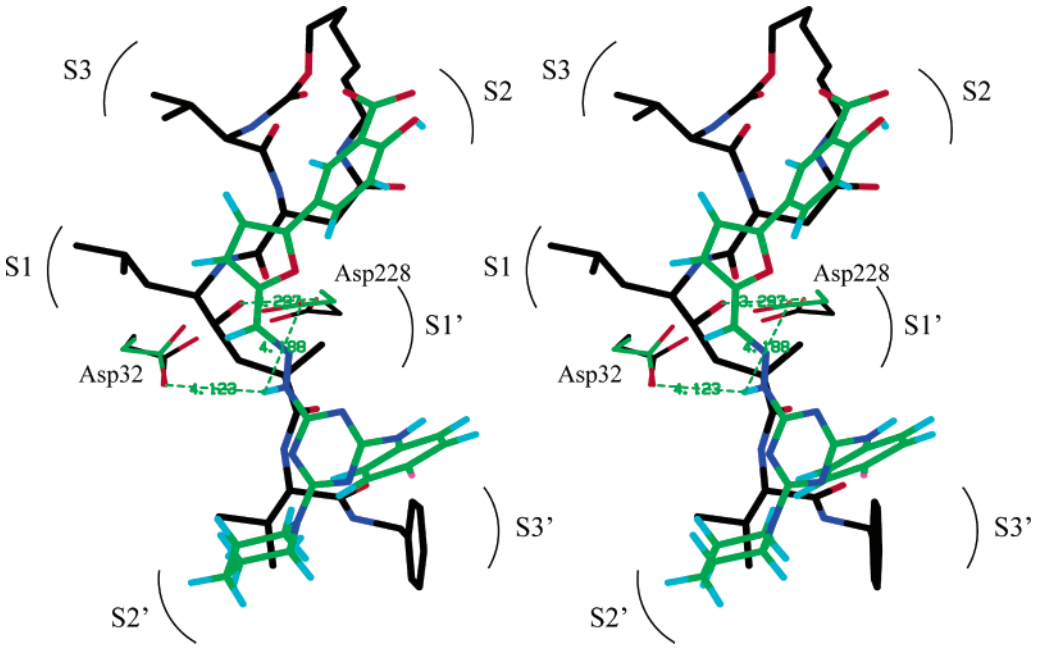


Figure 5. Stereoview of the superposition of a known nanomolar inhibitor of BACE-1[43] (carbon atoms in black) and compound **5** (carbon atoms in green). The side chains of the catalytic residues Asp32 and Asp228 are shown together with the distances between their oxygen atoms and the hydrazone NH of compound **5** or the hydroxyl group of the known nanomolar inhibitor (dashed lines). The structural alignment was generated taking into account only the C_α atoms of the two BACE-1 structures 1XS7[43] and 1M4H,[18] which overlap with a deviation of only 0.3 Å. The 1M4H structure was used for the high-throughput docking.

tion[31,32] or covalent modification), two additional experiments were performed. First, detergent (0.05% (v/v) Triton X-100) was added in the enzymatic assay. No significant reduction of activity was observed. Second, the effect of two different incubation times, 3 vs 30 min, was investigated for compounds **1** and **5**. After the shorter incubation time, the percentage inhibition at 20 μM concentration of compound **1** is 57%, and that at 50 μM concentration of compound **5** is 66%. These values are consistent with the IC_{50} values measured with 30 min incubation time (Table 1). Furthermore, two compounds were tested using another commercially available FRET assay kit (SIGMA CS0010, which includes Triton X-100 0.08% at final concentration). With the SIGMA kit, IC_{50} values of 7 and 32 μM were measured for compounds **7** and **8**, respectively. Compound **7** shows 1 order of magnitude difference in the IC_{50} value measured with two different kits. This discrepancy is likely to be a consequence of differences in substrate, protein, and assay buffer.

Interestingly, the 11 active compounds (10 with $IC_{50} < 100$ μM and one with $IC_{50} = 152$ μM) from the two in silico screenings have a common (1,3,5-triazin-2-yl)hydrazone scaffold. Table 1 shows structure as well as experimental and predicted affinity of eight compounds, four from each screening. Compounds **1**−**3** differ only in the substituents of the ring at R3, and compounds **5** and **6** are also very similar.

Binding Mode. The predicted binding mode of compound **5** is shown in Figure 5, overlapped with the cycloamide-urethane-derived peptidic inhibitor **2c** of ref 43. The hydrogen atom of the hydrazone NH group is at a distance of about 4 Å from two oxygen atoms in the catalytic aspartates. Such distance suggests

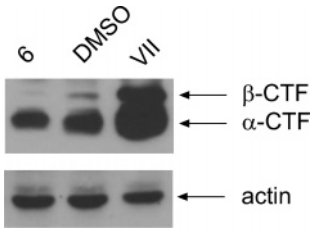


Figure 6. Western blot analysis of compound **6**, confirming reduction of the β-site cleavage of APP. (Top) The β-carboxy-terminal fragment (CTF) and α-CTF bands indicate that cleavage of β-CTF is decreased compared to DMSO negative control without affecting α-CTF. On the other hand, the γ-secretase inhibitor VII[46] causes an accumulation of both β-CTF and α-CTF. (Bottom) The actin bands indicate equal loading of the samples in each lane.

the presence of a water-bridged hydrogen bond as observed in the X-ray structure of BACE-1 in the complex with an oxyacetamide compound ($IC_{50} = 1.4\ \mu M$, PDB code 1TQF[10]) and in the structure of plasmepsin II complexed with an inhibitor featuring a tertiary amino group close to the two catalytic aspartates.[44] The S2, S2′, and S3′ pockets of BACE-1 are occupied by the 2-hydroxybenzoic acid, piperidine, and fluorobenzene substituents of compound **5**, respectively. The 2-hydroxybenzoic acid and piperidine of **5** overlap with part of the macrocycle and P2′-propyl side chain of **2c**, respectively. On the other hand, the fluorobenzene of **5** has a slightly different orientation compared to the benzyl group of **2c**. The furan and (1,3,5-triazin-2-yl)hydrazone mimic part of the peptidic backbone of **2c**.

The predicted binding mode of compounds **6**−**8** is essentially identical to the one of **5**, while the R3 substituent of compounds **1**−**4**, which lack the furan linker, points toward the S1 pocket. Given the small range of measured IC_{50} values and uncertainties in the details of the predicted binding mode (determined by FFLD docking and energy minimization in the rigid BACE-1 structure), it is not possible to obtain a detailed structural explanation of the measured relative affinities. The rather small differences in measured affinities is consistent with the fact that compounds **1**−**8** are similar among each other and show similar predicted binding modes.

Cellular Assay. To assess the potential for further development, e.g., hit explosion, it is important to verify that the compounds which are active in the enzymatic assay are also cell-permeable and able to inhibit BACE-1 in mammalian cells. For this purpose, 7 and 24 compounds from the first and second screening, respectively, were submitted to a cell-based test in which reduction of Aβ peptide secretion was measured as reported previously by others.[45] Table 1 shows that compounds **3** and **5**−**7** from the first and second screenings are active, with $EC_{50} < 10\ \mu M$ and $EC_{50} < 20\ \mu M$, respectively.

It is interesting to note that compound **8**, with the highest potency in the enzymatic assay for BACE-1 ($IC_{50} = 7.1\ \mu M$ in the Panvera kit and 32 μM in the SIGMA kit), is not active in the cell-based assay at a concentration of 25 μM. These data have to be compared with the corresponding ones for compound **6**, which has a very poor activity in the enzymatic assay but a cell-based EC_{50} value of 18.0 μM. These discrepancies might be due to several reasons, including differences in cell permeability, cleavage efficiency of full-length BACE-1 (cell-based assay) versus the lumenal domain only (enzymatic assay), different substrates, and assay conditions (e.g., pH 4.5 in the enzymatic assay).

To provide further evidence that compound **6** inhibits BACE-1 activity in cells, Western blot analysis was used to detect differentially cleaved carboxy-terminal fragments (CTFs) of APP (Figure 6). Compound **6** lowered β-CTF without affecting α-CTF compared to the dimethyl sulfoxide (DMSO) negative control. On the other hand, the γ-secretase inhibitor VII[46] caused an accumulation of both β-CTF and α-CTF compared to DMSO. This result indicates that reduced secretion of Aβ from cells was caused by β-site cleavage inhibition.

Finally, it is important to note that the peptidic inhibitors OM99-2 (molecular weight 897.2 g/mol),[34] its cycloamide-urethane derivative **2c** (731.1 g/mol),[43] and the peptidomimetic **57** (699.4 g/mol) of ref 11 have low-nanomolar affinity for BACE-1 in enzymatic tests but show only micromolar activity in cellular assays because of limited ability to cross cell membranes.[11,43] Despite their more than 3 orders of magnitude worse inhibitory activity in the enzymatic assay, the four triazine derivatives **3** and **5**−**7** have cellular activity similar to that of the three known peptidic inhibitors mentioned above. Given their smaller size, the triazine derivatives are likely to be more suitable for further development than the peptidic inhibitors.

Conclusions

High-throughput, fragment-based docking into the BACE-1 active site and LIECE binding free energy evaluation were used to select 88 compounds for experimental validation from an initial set of more than 300 000 molecules. Ten of the 88 compounds inhibit BACE-1 activity in an enzymatic assay ($IC_{50} < 100\ \mu M$), and four of them are active in a mammalian cell-based assay ($EC_{50} < 20\ \mu M$). Taken together, the discoveries of three novel series of BACE-1 inhibitors, i.e., phenylurea derivatives,[33] triazine derivatives (this work), and a set of five cell-permeable, nonpeptide, low-micromolar inhibitors of BACE-1 with a different scaffold (D. Huang and A. Caflisch, unpublished results), are a proof-of-principle of our in silico high-throughput screening approach. Furthermore, the present study represents a successful combination of computational predictions and experimental validation of inhibitors of a pharmaceutically relevant enzyme for which few nonpeptidic inhibitors have been already discovered, despite the availability of the X-ray structure of BACE-1 for more than 5 years. We are currently applying high-throughput docking and the LIECE approach to identify kinase inhibitors from very large collections of low-molecular-weight compounds. For protein targets of known three-dimensional structure, the efficient in silico approach presented in this paper is a cost-effective alternative to high-throughput in vitro screening campaigns.

Availability of the Software. The software suite of programs for high-throughput docking (DAIM, SEED, FFLD), including

(43) Ghosh, A.; Devasamudram, T.; Hong, L.; DeZutter, C.; Xu, X.; Weerasena, V.; Koelsch, G.; Bilcer, G.; Tang, J. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 15−20.
(44) Prade, L.; Jones, A. F.; Boss, C.; Richard-Bildstein, S.; Meyer, S.; Binkert, C.; Bur, D. *J. Biol. Chem.* **2005**, *280*, 23837−23843.
(45) Dovey, H. R.; Suomensaari-Chrysler, S.; Lieberburg, L.; Sinha, S.; Keim, P. S. *NeuroReport* **1993**, *4*, 1039−1042.
(46) Durkin, J. T.; Murthy, S.; Husten, E. J.; Trusko, S. P.; Savage, M. J.; Rotella, D. P.; Greenberg, B. D.; Siman, R. *J. Biol. Chem.* **1999**, *274*, 20499−20504.
(47) Scudiero, D. A.; Shoemaker, R. H.; Paull, K. D.; Monks, A.; Tierney, S.; Nofziger, T. H.; Currens, M. J.; Seni, D.; Boyd, M. R. *Cancer Res.* **1988**, *48*, 4827−4833.

input files, test cases, and documentation, are available from the corresponding author (at no expense for not-for-profit institutions).

Acknowledgment. We are grateful to Stephan Audetat, Fabian Dey, and Nicolas Majeux for interesting discussions and Karin Edler for help with the in vitro experiments. The calculations were performed on Matterhorn, a Beowulf Linux cluster at the Informatikdienste of the University of Zurich, and we thank C. Bolliger, T. Steenbock, and A. Godknecht for installing and maintaining the Linux cluster. This work was supported by grants from KTI (Kommission Technologie und Innovation) and the National Competence Center for Research (NCCR) on Neural Plasticity and Repair.

Supporting Information Available: Experimental procedures, information on 88 tested compounds, the LIECE energy values of 37 known inhibitors and the top 100 library compounds of Figure 3, and complete refs 3, 11, and 38. This material is available free of charge via the Internet at http://pubs.acs.org.

JA0573108