JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Potency Prediction of β-Secretase (BACE-1) Inhibitors Using Density Functional Methods

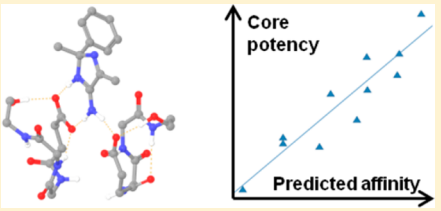
Katarina Roos,*[,†] Jenny Viklund,[‡,⊥] Johan Meuller,[§] Karin Kaspersson,[§] and Mats Svensson[‡,⊥]

[†]Department of Medicinal Chemistry and [§]Discovery Sciences, AstraZeneca R&D Mölndal, SE-431 83 Mölndal, Sweden

[‡]Department of Medicinal Chemistry, AstraZeneca R&D Södertälje, SE-151 85 Södertälje, Sweden

Supporting Information

ABSTRACT: Scoring potency is a main challenge for structure based drug design. Inductive effects of subtle variations in the ligand are not possible to accurately predict by classical computational chemistry methods. In this study, the problem of predicting potency of ligands with electronic variations participating in key interactions with the protein was addressed. The potency was predicted for a large set of cyclic amidine and guanidine cores extracted from β-secretase (BACE-1) inhibitors. All cores were of similar size and had equal interaction motifs but were diverse with respect to electronic substitutions. A density functional theory approach, in combination with a representation of the active site of a protein using only key residues, was shown to be predictive. This computational approach was used to guide and support drug design, within the time frame of a normal drug discovery design cycle.



INTRODUCTION

Predicting Binding Affinity in Drug Design. Target potency is an important optimization parameter in drug discovery. High potency is advantageous for obtaining efficacy with a low dose drug, which helps to avoid nonspecific side effects. Predicting protein−ligand interaction energies on an absolute level for diverse ligands requires control of complicated enthalpic and entropic contributions for very complex systems. The treatment of solvation for ligands and proteins makes the problem even more challenging. However, in drug design an absolute potency prediction is often not necessary—instead an equally useful tool would be the ability to rank the potency of similar ligands and, thus, prioritize the most promising for synthesis. In a drug discovery optimization phase, this could, e.g., be small variations in a specific part of the lead series.

For an accurate electronic description of the protein−ligand interaction, and the effect on the affinity by subtle electronic variations in the ligand, quantum chemical methods are required. Quantum chemical methods have rarely been employed in drug design because of the relatively high computational cost compared to other methods. However, with improved computer power the field has evolved over the last years.[1,2] Some examples of approaches that have been developed to study protein−ligand interactions in drug design are semiempirical methods, for example in a study of kinases,[3] dispersion-corrected density functional theory (DFT)[4] and methods such as PMISP,[5] that combine quantum mechanics with force field based methods. For recent reviews of the field see refs 1 and 2.

In this paper we are interested in inhibitors that are diverse with respect to electronic substitutions, and for which the relative potency is driven by specific interactions rather than lipophilicity or size. The focus is to go beyond force field methods and develop a quantum chemical approach applicable to drug discovery projects. For systems with specific binding sites with conserved key interactions, and where the ligands occupy a similar pose, the affinity prediction problem is less complex. In a simplified model of the protein, including only a smaller set of residues participating in key interactions, the relative protein−ligand binding energies can be calculated using normal quantum chemical approaches, such as density functional theory (DFT) methods. A similar DFT cluster approach has previously been successfully employed to describe enzymatic reaction mechanisms.[6]

In the present study, a DFT approach was employed to predictively rank the potency of an electronically diverse set of drug molecule cores. The core is defined as the central part of the ligand, participating in key interactions with the protein. The definition of a chemical series is often connected to the core definition.[7,8] The approach described here was used to guide the prioritization between different types of chemical cores in a structure based drug discovery project, where the aim was to synthesize potent compounds with novel cores.[9]

Alzheimer Target BACE-1. The drug target in this study was β-site APP cleaving enzyme (BACE-1), which is a major drug target for finding a treatment for Alzheimer's disease (AD).[10] BACE-1 is an aspartyl protease, with the catalytic residues Asp32 and Asp228 located in the center of the substrate binding cleft.[11] By hydrolyzing the amyloid precursor protein (APP), BACE-1 is responsible for the first, rate-limiting step in the formation of toxic, amyloid plaques found in the diseased AD brain.[12,13] So far, there are no disease-modifying

Received: June 28, 2013
Published: January 23, 2014

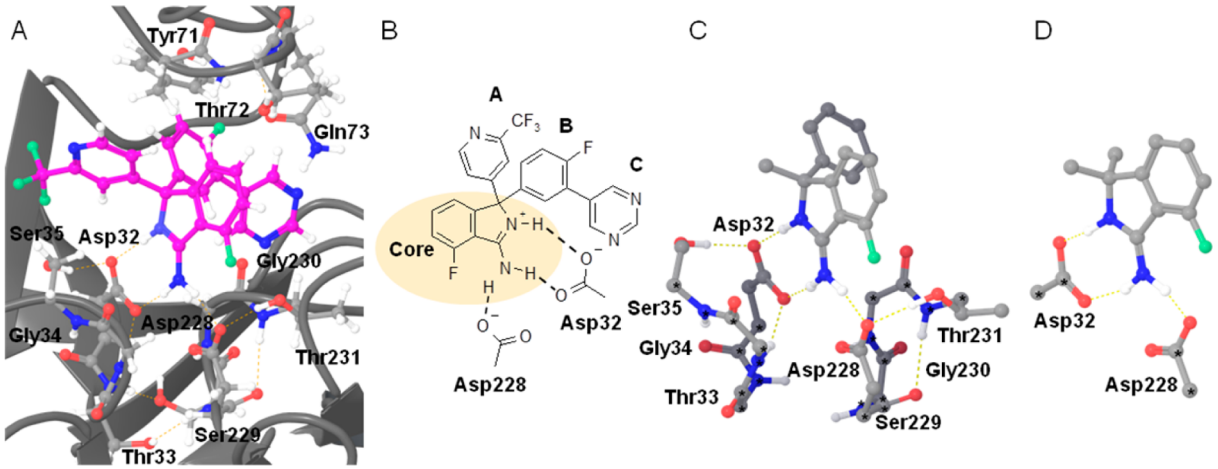


Figure 1. (A) Example of a crystal structure of BACE-1, in complex with a cyclic amidine inhibitor (purple), represented by core number 20, pdb code 4azy (Figure 2).[18] (B) 2D representation of the same inhibitor. The B and C aryls bind in the S1 and S2 pockets, respectively. The A-aryl binds on the prim-side of the catalytic aspartates 32 and 228. (C) Medium sized model of the protein−ligand complex including the key residues of the binding pocket was constructed based on the crystal structure. The side chains of Thr33 and Ser229 were truncated and excluded in the model. Atoms marked with stars were kept fixed in the optimizations. The ligands were truncated with methyl groups in A, phenyl groups in B, and nothing in C. (D) Smaller cluster model including only the catalytic aspartates constructed for a fast track method. The dihedral planes of the aspartates and atoms marked with stars were kept fixed in the optimizations. The ligands were truncated with methyl groups in A and B.

drugs for AD on the market, despite a huge medical need and extensive, ongoing research.[14] A large number of diverse, highly potent, small-molecule BACE-1 inhibitors have been published.[10,15,16] There are also several crystal structures showing that these ligands make close interactions with the catalytic aspartates, usually via a basic amine, amidine, or guanidine binding motif (Figure 1).[7,13,17−22]

Experimental Core Potency. The data set in this study was a large set (661 compounds) of monocyclic and bicyclic amidine and guanidine BACE-1 inhibitors. These compounds were measured for BACE-1 activity using a FRET based assay. All ligands with a pIC50 lower than 4.5 were also measured with SPR, to confirm specific binding and to increase the activity range in the low potency area (Supporting Information). Together, these compounds represent 35 electronically diverse amidine and guanidine cores, all decorated with matched paired variations in the A, B, and C aryls (Figures 1 and 2). According to a Free-Wilson analysis, the structural elements of the core, A, B, and C aryls were separated, and an additive SAR was shown for target activity, with a strong correlation factor R^2 of 0.95 and an RMSE of 0.2. Thus, the relative contributions of the cores to the target potency could be extracted, with activities representing a potency range of 4.6 kcal/mol (Table 1). This analysis is further described in the Supporting Information. To evaluate the performance of the quantum chemical approach, the predicted binding affinities were compared to the extracted core potencies.

Computational Details. The quantum chemical calculations were performed using density functional theory with the B3LYP, B3LYP-D3, and M06-2X functionals using Jaguar.[34−36] All structures were optimized using the 6-31G** basis set. The solution phase energies were derived from single point calculations with the 6-31+G** basis set on the previously optimized structures, using the Poisson−Boltzmann finite element solvation model (PBF)[37,38] and the SM8 solvation model.[39,40] For comparison, a number of alternative computational chemistry methods, for example docking and MM-GBSA,[41] were employed on the same data set, following standard procedures as described in the Supporting Information.

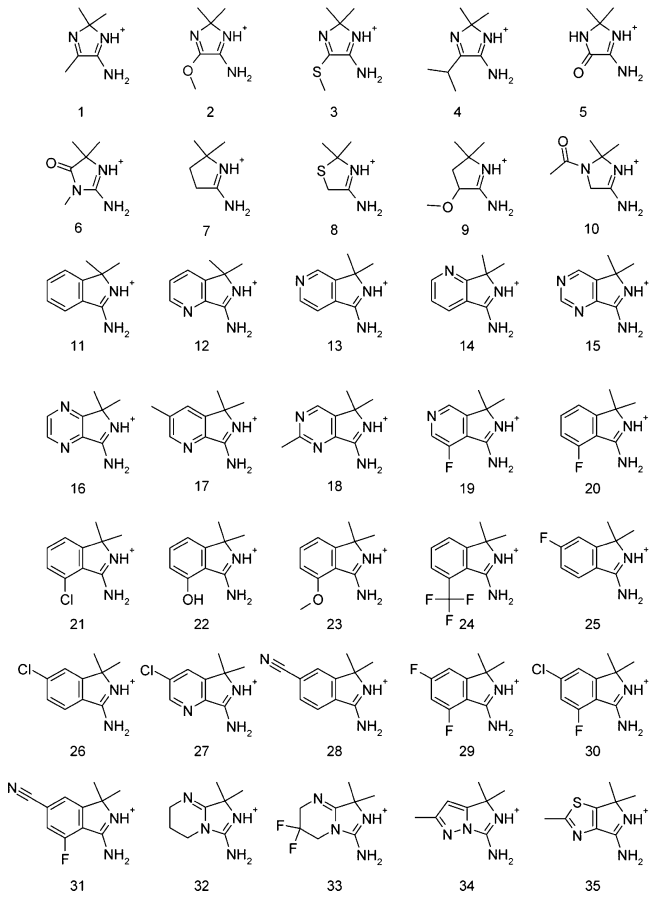


Figure 2. Data set of cyclic amidine and guanidine cores included in the study. Potency data on the cores and references to the full sized compounds representing the cores are found in Table 1.

dx.doi.org/10.1021/ci400374z | J. Chem. Inf. Model. 2014, 54, 818−825

Table 1. Experimental Core Potency and Calculated Protein−Ligand Binding Affinities

ID	type[a]	core potency[b]	med m. PBF pred affinity[c]	med m. SM8 pred affinity[d]	small m. PBF pred affinity[e]	ref
1	mono	0.4	3.7	−1.1	3.1	23
2	mono	−0.4	0.8	−2.9	1.7	*f*
3	mono	0.3	2.3	0.9	3.0	*f*
4	mono	−1.0	1.9	1.2	3.9	23
5	mono	−1.5	−1.2	−3.9	−0.5	*f*
6	mono	1.2	4.6	2.8	4.8	24, 25
7	mono	−1.5	0.4	−3.0	−0.1	9
8	mono	0.0	3.6	−0.2	3.4	9
9R	mono	−1.3	−1.1	−3.1	−1.0	9
9S	mono	−2.4	−2.8	−5.7	−1.1	9
10	mono	−0.3	2.4	−0.2	2.5	9
11	bi	1.1	0.7	2.5	0.8	26, 27
12	bi	0.4	−3.2	−1.5	−2.9	28−30
13	bi	1.1	1.6	0.3	2.3	31
14	bi	0.9	2.2	1.5	2.2	31
15	bi	0.6	−1.3	−1.2	−1.3	*f*
16	bi	0.2	−0.1	−1.2	−0.5	29
17	bi	−0.7	−2.3	−0.6	−3.7	28, 30
18	bi	−0.5	−2.2	−0.8	−1.4	*f*
19	bi	1.0	0.9	−0.1	0.4	*f*
20	bi	0.8	−0.9	0.9	−2.0	18, 32
21	bi	−1.6	−2.7	−0.5	−3.8	*f*
22	bi	−1.4	−3.6	−1.0	−6.4	*f*
23	bi	−1.5	−4.7	−1.8	−5.8	*f*
24	bi	−3.3	−7.0	−5.7	−5.1	*f*
25	bi	1.3	1.7	1.8	1.8	26, 27
26	bi	1.3	1.6	4.2	2.2	26, 27
27	bi	0.6	−1.1	1.1	−1.3	28, 30
28	bi	1.1	1.6	2.5	1.4	26, 27
29	bi	0.7	0.3	0.8	−0.5	18
30	bi	1.1	0.9	3.1	−0.4	32
31	bi	1.3	0.1	1.1	−0.9	*f*
32	bi	1.0	2.6	2.0	2.3	17
33	bi	1.0	4.7	4.6	4.5	17
34	bi	0.2	−0.5	4.0	0.2	9
35	bi	0.0	−3.6	−0.9	−1.9	9, 33

[a]Monocyclic and bicyclic amidine and guanidine cores. [b]Experimental core contribution to potency in kilocalories per mole extracted from FRET and SPR data of full-sized compounds by a Free-Wilson analysis. [c]Calculated binding affinities in kilocalories per mole for the medium sized model using the quantum chemical approach with B3LYP and the PBF solvation model, also plotted in Figure 5A. [d]Calculated binding affinities in kilocalories per mole for the medium sized model using the quantum chemical approach with B3LYP and the SM8 solvation model, also plotted in Figure 5B. [e]Calculated binding affinities in kilocalories per mole for the small fast track model using the quantum chemical approach with B3LYP and the PBF solvation model. [f]In-house data.

RESULTS AND DISCUSSION

Computational Methods for Predicting Affinity. There are many computational approaches used in the drug discovery phase to predict protein ligand interactions, ranging from descriptor based statistical methods to force field based methods and different types of docking protocols.[42,43] Two properties often shown to correlate to ligand binding affinities are the size and lipophilicity of the molecules.[44] However, for the present BACE-1 data set, neither the size (solvent accessible surface area), nor the lipophilicity (clogP) of the cores correlated to the experimental core potency (Figure 3). The calculated partial charges on the protons participating in hydrogen bonding did not vary either. In addition, the pK_a could be expected to correlate to hydrogen bond strength. However, the experimental pK_a values of the amidine and guanidine cores did not correlate to the core potency (Figure 3).[7] This indicates that the variation in potency in the data set is not described by these common molecular descriptors. Poses from docking using Glide,[41] were correctly reproduced compared to X-ray crystal structures for all compounds. However, neither of the GlideScore or XP GlideScore could be used to rank the potency of the cores in the data set, with an R^2 worse than 0.1 for all scores to core potency (Figure 3). The relative free energy of binding calculated by MM-GBSA did not correlate to the core potency either.[41] The R^2 for all cores was 0.17 (Figure 3).

One fundamental problem of scoring functions and force-field based methods in general is their limited electronic description of the ligand and the protein. This problem will be pronounced when we attempt to predict the relative binding affinity for ligands with structural variations affecting the electronics present in the core of the molecules, as for this data set. All cores make the exact same hydrogen-bonding interactions to the catalytic aspartates of the target, but the electronic variations in the cores affect the strengths of these interactions.

Quantum Chemical Approach. If electronic effects are hypothesized to be important factors for the relative potency in a data set (rather than lipophilicity or size), but yet a fast prediction method is needed, as often in drug discovery, a quantum chemical cluster approach is instead a good choice. In the present study a DFT cluster approach to predict core potency for BACE-1 was employed. The relative protein−ligand binding affinities (the negative of the binding energies) were predicted in a simplified model of the protein, including only a smaller set of residues participating in the key interactions. The binding affinity was calculated as the difference in energy between a model of the protein−ligand complex and the sum of the energy of the ligand in solution and the energy of a model of the apoprotein. The energy of the apoprotein is a constant term for all ligands and was omitted. The predicted affinity is reported as offset from the mean value for the N = 36 cores (Figure 2).

$$\Delta E_{\text{predicted affinity},i} = (E_{\text{ligand}} - E_{\text{protein-ligand complex}})_i - \sum_{N}^{i=1} (E_{\text{ligand}} - E_{\text{protein-ligand complex}})_i/N$$

$$E_{\text{ligand}} = E_{\text{gas phase,6-31+}G^{**}} + E_{\text{solv},\varepsilon=80.0,\text{6-31+}G^{**}} (+E_{\text{dispersion D3}})$$

$$E_{\text{protein-ligand complex}} = E_{\text{gas phase,6-31+}G^{**}} + E_{\text{solv},\varepsilon=4.0,\text{6-31+}G^{**}} (+E_{\text{dispersion D3}})$$

Constructing Protein Models. A medium sized model of the protein, including only a smaller set of residues participating in the key interactions, was constructed based on the X-ray crystal structure of BACE-1 in complex with a ligand representing core number 20 (Figure 1).[18] The catalytic

dx.doi.org/10.1021/ci400374z | J. Chem. Inf. Model. 2014, 54, 818−825

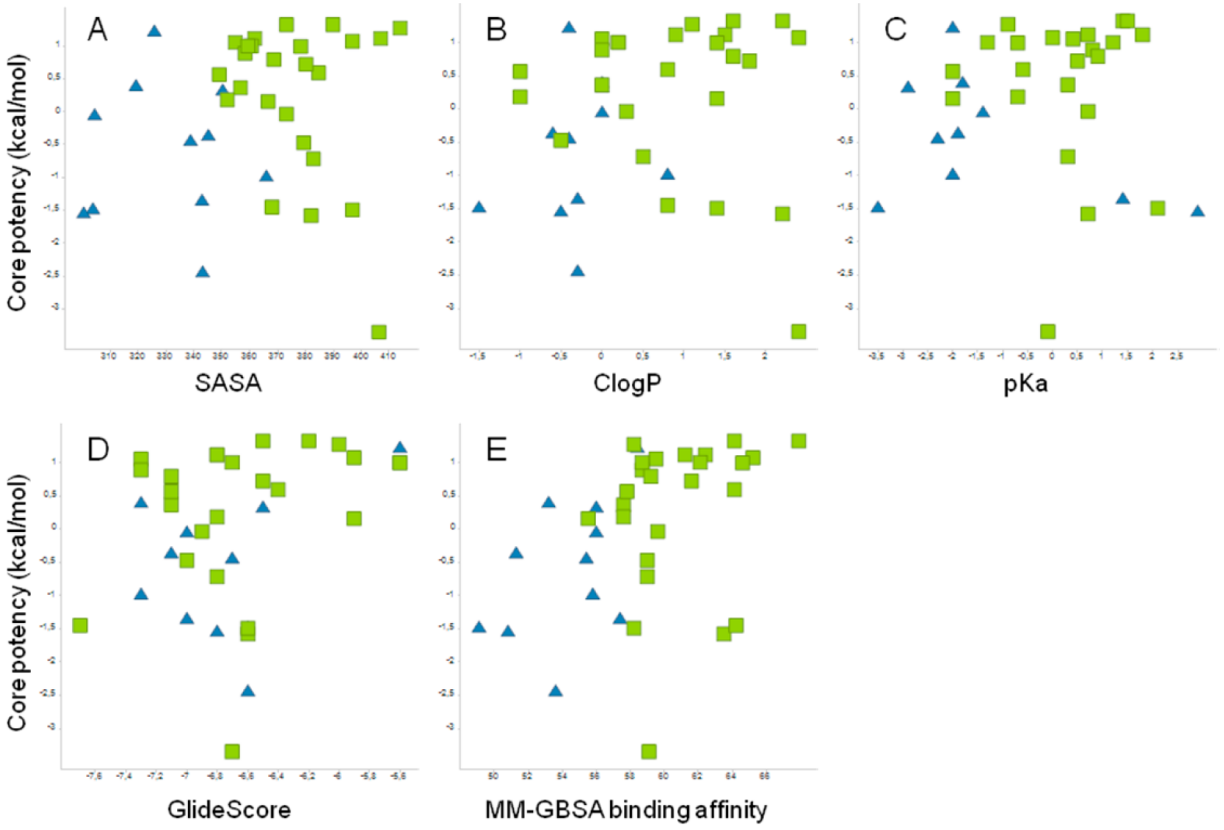


Figure 3. Experimental core potency versus: (top left) size (calculated solvent accessible surface in squared angstroms),[34] (top middle) lipohilicity (calculated clogP),[45] (top right) experimental pK_a,[7] (down left) GlideScore,[41] and (down right) MM-GBSA binding affinity.[41] Green squares are bicyclic and blue triangles are monocyclic cores.

aspartates forming the key interaction with the ligand, Asp32 and Asp228, were included in the model. Both aspartates were assumed to be deprotonated.[17] These two aspartates accept hydrogen bonds from a serine, Ser35, and a threonine, Thr231, respectively. By including these amino acids, together with the backbones of the Ser-Gly-Thr triads for each aspartate, the correct side chain conformations of the aspartates were stabilized, and thus, artificial fixations could be avoided (Figure 1). To account for the strain from the surrounding protein, the backbone atoms were fixed in the optimizations. The ligands were truncated with a methyl in A, a phenyl in B, and nothing in C (Figure 1). All cores were positively charged on the amidine nitrogen.[7,13] If nothing else is stated reported values are with this medium sized model of the protein and ligands.

A larger model was constructed, by extending to a sphere of 4 Å from the core fragment, including Tyr71, Thr72 and Gln73 (Figure 1). These residues are situated in a protein flap, a flexible loop situated on top of the binding pocket. The backbone position of the included flap residues were held fixed in the optimizations.

For a fast track method, a smaller cluster model was constructed, where only the two aspartates were included and modeled as acetate groups (Figure 1). In order to keep the aspartates from rotating, the dihedral plane was held fixed. The ligands were truncated with methyl groups in A and B.

Workflow for Predicting Affinity. A conformational search was performed for all core structures in water, using extended, mixed torsional/low-mode sampling in Macromodel, and minimization using the OPLS2005 force field (Figure 4).[41] The geometry of all conformations within a 5 kcal/mol energy window were further optimized in gas phase using DFT and the 6-31G** basis set. If nothing else is stated, reported values were calculated with the B3LYP functional. In a model of the protein−ligand complex, the generated core conformations were aligned with the crystal structure position of a cocrystallized reference ligand, and the phenyl group in B was oriented as in the crystal structure, before the DFT gas phase optimization. The optimized geometries were solvated using the PBF and the SM8 continuum solvation models and the 6-31+G** basis set. For PBF, the free ligand was solvated in water using a dielectric constant of 80. For the more hydrophobic protein environment, a dielectric constant of 4 was used, in line with previous modeling.[6] For SM8, the ligand was solvated in water and the protein−ligand complex was solvated in chloroform as a model for a protein environment. For both the free ligand and the protein−ligand complex, the conformation with the lowest calculated energy including solvation was selected. The predicted affinity was calculated as the difference in energy between the solvated free ligand and the solvated protein−ligand complex and reported as offset from mean. The entropy loss upon binding was assumed to be similar for all the fairly rigid cores and hence neglected.

Predicted Affinity Correlates to Core Potency. The positions of the ligands after DFT optimization in the medium sized model of the protein−ligand complex were consistent with the position of core 20 in the crystal structure, with an average RMSD of 0.6 Å and a maximum RMSD of 1.3 Å. The calculated binding affinities using the cluster model approach and the PBF solvation model for the two types of amidine and guanidine cores correlated to the measured core potency with

dx.doi.org/10.1021/ci400374z | J. Chem. Inf. Model. 2014, 54, 818−825

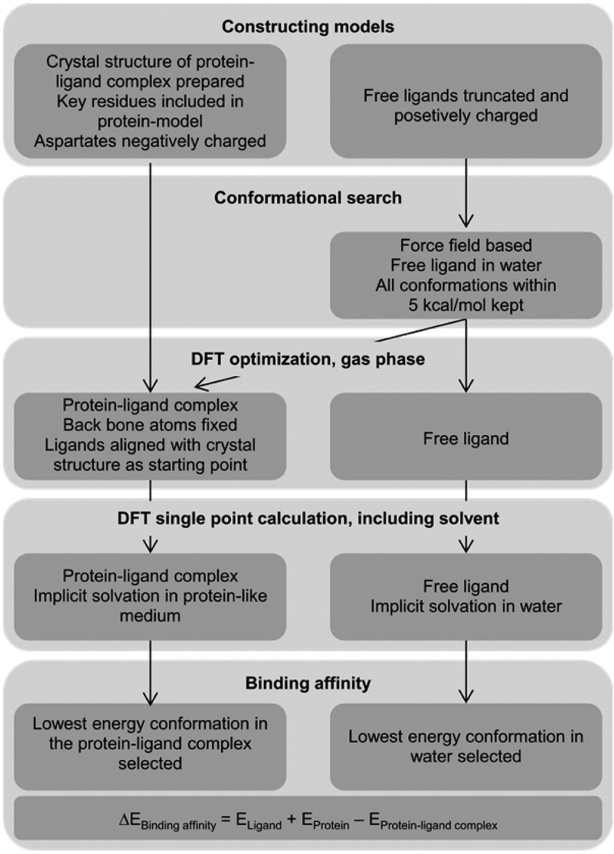


Figure 4. Workflow for predicting the affinity by the quantum chemical approach.

an R^2 value of 0.86 and 0.73 for the monocyclic and bicyclic types, respectively (Figure 5A). The mean absolute deviations were 0.4 and 0.4 kcal/mol, and the maximum deviations were 0.6 and 1.3 kcal/mol for the monocyclic and bicyclic types, respectively.

Solvation is Important. An important step in the protein−ligand complex formation is the movement of the ligand from water to a protein environment. The calculated gas phase affinity gave a significantly lower correlation to the experimental core affinity, with R^2 values of 0.58 and 0.55, mean absolute deviations of 0.5 and 0.5 kcal/mol, and maximum deviations of 1.4 and 2.1 kcal/mol for the monocyclic and bicyclic cores, respectively; hence, including solvation was important for accurate results. If both the free ligand and the protein−ligand complex were solvated in water using the PBF model, the correlation was better than in the gas phase, with R^2 values of 0.69 and 0.72, mean absolute deviations of 0.5 and 0.4 kcal/mol, and maximum deviations of 1.3 and 1.2 kcal/mol for the monocyclic and bicyclic cores, respectively. However, the best result was achieved when the dielectric continuum was different for the free ligand and the protein−ligand complex.

Predicted Affinity Is Offset between Core Types. A constant shift of about 4 kcal/mol between the predicted-to-experimental affinity curves of the two core types was observed (Figure 5A). The smaller monocyclic cores were predicted to have a higher affinity relative to the bicyclic cores. The extra ring in the bicyclic type is in closer contact with a protein flap, a flexible loop situated on top of the binding pocket (Figure 1). These interactions were not included in the model described here, and this may be a reason for the constant shift between

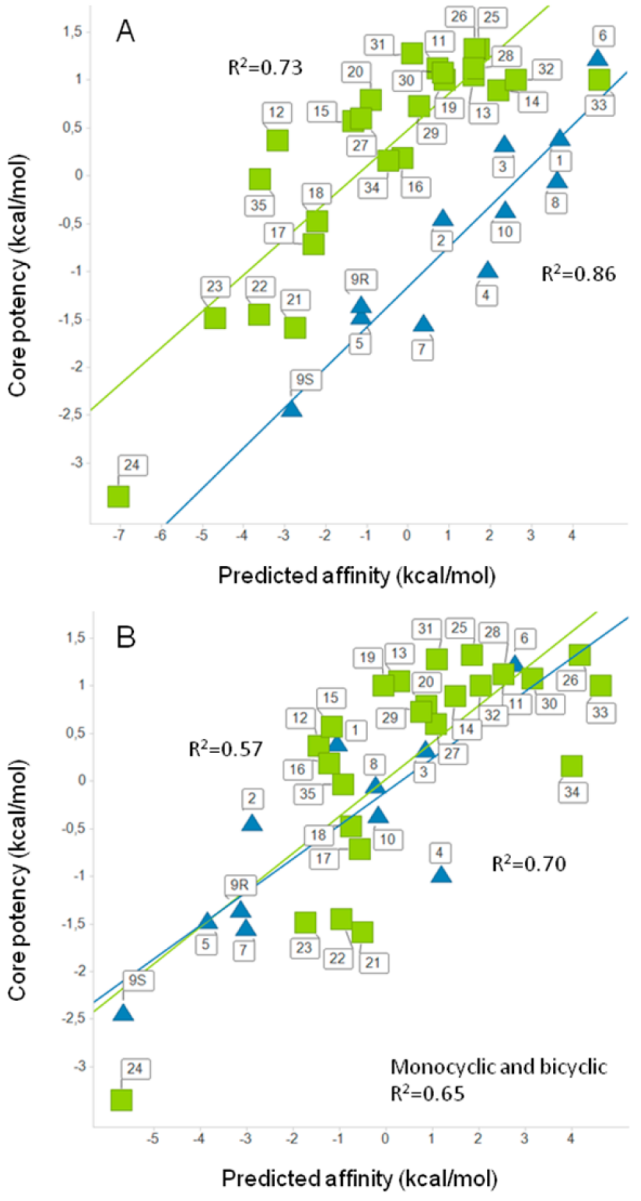


Figure 5. Calculated binding affinity versus core potency for the monocyclic cores (blue triangles) and bicyclic cores (green squares) with the B3LYP functional and (A) the PBF solvation model or (B) the SM8 solvation model.

the predicted affinities for the two different core types. In an attempt to include the flap region, the model was extended to include Tyr71, Thr72 and Gln73 (Figure 1). These amino acids in the flap region are however disordered and could therefore not easily be included in the cluster model without introducing uncertainties. Cores 17 and 25−31 with substitutions in the 5 position artificially clashed with Gln73, and therefore, their predicted binding affinities were underestimated (Figure 2). Cores 13, 15, 18 and 19 with nitrogen in the 5 position formed artificial hydrogen bonds with Gln73 and core 10 formed an artificial hydrogen bond with Tyr71 in the models. Because the flexibility of the flap region was not included in the cluster model, conclusions on the effect on the shift could not be drawn.

It is interesting to note that if the SM8 solvation model was instead employed, the shift decreased (Figure 5B). The correlation within the two types of cores decreased to R^2

dx.doi.org/10.1021/ci400374z | J. Chem. Inf. Model. 2014, 54, 818−825

values of 0.70 and 0.57, with mean absolute deviations of 0.3 and 0.5 kcal/mol and maximum deviations of 1.3 and 1.4 kcal/mol for the monocyclic and bicyclic cores respectively; whereas, the overall correlation with no separation into core types was an R^2 value of 0.65, with mean absolute deviation of 0.4 kcal/mol and maximum deviations of 1.5 kcal/mol. The term mainly responsible for the reduced shift was the solvation of the protein−ligand complex and not the solvation of the free ligand. The cavity-dispersion-solvent structure free energy contributed to a smaller degree to the decrease in the shift, and the electron-nuclear-polarization free energy corresponded to the major effect.

Predicted Relative Binding Affinity Is Exaggerated. The slope of the core potency to predicted affinity for the two core types was the same; however, flatter than expected, with a gradient of about 0.4 for both the PBF and SM8 solvation models. A similar effect has previously been observed for quantum mechanical predicted pK_a and been suggested to be due to difficulties to accurately treat solvation of ionic species.[46] Including an explicit water molecule in the solvation of the free ligand did not affect the slope in our case. In general, MM-GBSA also produces exaggerated relative binding affinities.[47] A possible reason may be due to enthalpy−entropy compensation, by the partial capture of the reorganization free energy of the solvent and the insufficient screening at short distances, leading to overestimation of charge−charge interactions. These effects are treated accurately by free energy perturbation methods, which generally show a slope near one.

Dispersion Has No Relative Effect. A limitation of DFT with the B3LYP functional is the inability to describe weak interactions such as dispersion. Using the B3LYP-D3 functional, including semiempirical dispersion corrections by Grimme,[35] did not improve the correlation in our set between calculated values and measured core potencies. The gas phase binding affinities calculated with the B3LYP-D3 and the B3LYP functionals correlated with an R^2 value of 0.99 and a slope of 0.99; hence, the effect of including the dispersion was small. In general, dispersion is important for describing binding affinities. However, in our data set, the interactions between the catalytic aspartates and the amidines are dominated by electrostatic interactions. Dispersion did not seem to be of importance for describing the relative affinity in this case. The same conclusions could be drawn from the correlation between the gas phase binding affinity calculated with the M06-2X and the B3LYP functionals with an R^2 value of 0.96 and a slope of 0.98.

Small Fast Track Model: Applicable in a Drug Design Project. For a fast track method, the smaller cluster model was used, including only the two aspartates modeled as acetate groups (Figure 1). The small model was not as good at reproducing the positions of the ligands in the crystal structures.[17−22] The average RMSD for the B3LYP optimized cores compared to the crystal structure position of core 20 was 1.3 Å and the maximum RMSD was 2.2 Å. When the ligands were free to move without constraints, they bound in a position in the plane of the aspartates. In the crystal structure, the ligands were slightly out of plane by about 15°. The small model gave somewhat worse correlation compared to the larger model, with R^2 values of 0.74 and 0.66, mean absolute deviations of 0.3 and 0.4 kcal/mol, and maximum deviations of 1.3 and 2.2 kcal/mol for the monocyclic and bicyclic types, respectively (Table 1). However, the method could be used to rank within and between series to a lower computational cost (of approximately 30 min on a 4 CPU machine per protein−ligand complex calculation). Due to the high-throughput capacity, the small model was used in a drug design project to rank de novo designed cores and, thus, to guide medicinal chemists in prioritizing the most promising for synthesis. A number of novel compounds were made and tested as highly active in BACE-1 in vitro assays.[7,9] The fact that a large part of the experimental variation in core affinity was already reproduced in the small model also showed that the interaction to the catalytic aspartates was relevant to affinity in this case.

Core number 9 has an additional chiral center, and the separated stereoisomers 9*R* and *S* show a difference of 0.8 log-units in core enzymatic potency, with 9*R* being the most potent isomer (Table 1).[9] The two stereoisomers have the same electronic effect on the key interaction, and the potency could not be reproduced in the small model, in which they had the same predicted affinity value. With the larger model, the calculations reproduced the observed higher activity of 9*R* compared to 9*S* (Figure 5).

Electronic Variations in the Ligands. On the basis of the quantum chemical calculations, the inductive effects of the electronic variations in the ligands on the binding affinity could be quantified. Comparing for example core numbers 5 and 6, number 6 had a larger calculated polarizability, which may contribute to the ability to form strong hydrogen bonds to the aspartates (Figure 2). Number 5 on the other hand had larger calculated solvation energy in water, which gives a negative contribution to the binding affinity. Together these properties may give a rationalization to the higher potency of core number 6 compared to core number 5. However, on a more general level, for the monocyclic cores, core number 4 had the largest calculated polarizability, and core number 10 had together with 5 the largest calculated solvation energy in water, and both cores 4 and 10 had intermediate potency. Hence, even though the difference in potency between individual pairs of cores may be rationalized by a number of different properties, the general trend was more complicated. By performing the binding affinity calculations using the cluster model approach, a predictive model for the complete data set was achieved (Figure 5).

DFT and CCSD(T) Predicted Affinities Correlate. In order to check the reliability of the DFT B3LYP approach, CCSD(T) single point gas phase binding affinities were calculated for the monocyclic cores using the small protein model.[48] The R^2 value of the gas phase binding energies between the different approaches was 0.96 with a slope of 1.06, showing that DFT was accurate enough for describing the relative binding energies of the compounds in the data set.

SUMMARY AND CONCLUSIONS

In the present work, the specific interactions between a set of positively charged amidine and guanidine cores and the negatively charged catalytic aspartates in BACE-1 were studied using the B3LYP, B3LYP-D3, and M06-2X functionals, the PBF and SM8 solvation models, and three protein models of different size. A large part of the experimental variation was present in the small model; whereas, the medium sized model gave a higher correlation. Extending the model to include the flap region gave less accuracy due to the flexibility of this part of the protein, an effect that could not easily be included in the modeling approach. For the system studied in the present work, where the interactions between the catalytic aspartates and the amidines were dominated by electrostatic interactions, dispersion did not seem to be of importance for describing the variation in relative binding affinity and the choice of

dx.doi.org/10.1021/ci400374z | J. Chem. Inf. Model. 2014, 54, 818−825

functional did not affect the results. In another system where dispersive forces are expected to be relevant, the choice of a dispersion-corrected functional should be more important. The PBF solvation model gave a higher signal within the congeneric series whereas, if diverse classes of cores were studied, bicyclic and monocyclic in this case, the SM8 model resulted in a higher overall correlation.

The affinity models described here were initially developed in conjunction with a drug discovery project, where the aim was to synthesize BACE-1 inhibitors with novel cores.[7,9] The predictive models were primarily applied in order to avoid wasting synthetic effort on making examples with very low core potencies, such as core numbers 21−24 (Figure 2, Table 1). Core numbers 7−10, 34, and 35 were prioritized for synthesis via a multiparameter in silico workflow, where one of the key parameters was the affinity predictions presented in this report.[7] These six cores were synthesized as full-sized lead molecules and tested active in both biochemical and cell-based BACE assays.[9] All of these new cores were quite diverse, and thus, the results could be used to both confirm the stability of the models and increase the confidence in the correlation factors. In summary, a quantum chemical cluster approach has been shown to be predictive for ranking potency of different cores with diverse electronic substitutions, where other structure based techniques fail. This is a promising approach to guide and support medicinal chemistry design that could be used to predict relative binding affinity for ligands with complicated electronic variations participating in key interactions with the protein. Further studies will focus on exploring the applicability of this approach for other targets than BACE-1.

■ ASSOCIATED CONTENT

Supporting Information

FRET and SPR validation of benchmark set, Free-Wilson analysis, additional data, comparisons to other prediction methods, predicted affinities with the small model, comparisons with CCSD(T), effect of different solvation treatment, and effect of including dispersion, coordinates. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Tel.: +46(0)317761296. E-mail: katarina.roos@astrazeneca.com.

Notes

The authors declare no competing financial interest.

[⊥]M.S. and J.V. were employees at AstraZeneca when the work was performed.

■ REFERENCES

(1) Mucs, D.; Bryce, R. A. The application of quantum mechanics in structure-based drug design. *Expert Opin Drug Discov.* **2013**, *8*, 263−276.

(2) Raha, K.; Peters, M. B.; Wang, B.; Yu, N.; Wollacott, A. M.; Westerhoff, L. M.; Merz, K. M., Jr. The role of quantum mechanics in structure-based drug design. *Drug Discovery Today* **2007**, *12*, 725−731.

(3) Brahmkshatriya, P. S.; Dobes, P.; Fanfrlik, J.; Rezac, J.; Paruch, K.; Bronowska, A.; Lepsik, M.; Hobza, P. Quantum Mechanical Scoring: Structural and Energetic Insights into Cyclin-Dependent Kinase 2 Inhibition by Pyrazolo[1,5-a]pyrimidines. *Curr. Comput.-Aided Drug Des.* **2013**, *9*, 118−129.

(4) Antony, J.; Grimme, S.; Liakos, D. G.; Neese, F. Protein-Ligand Interaction Energies with Dispersion Corrected Density Functional Theory and High-Level Wave Function Based Methods. *J. Phys. Chem. A.* **2011**, *115*, 11210−11220.

(5) Söderhjelm, P.; Kongsted, J.; Ryde, U. Ligand Affinities Estimated by Quantum Chemical Calculations. *J. Chem. Theory Comput.* **2010**, *6*, 1726−1737.

(6) Siegbahn, P. E.; Himo, F. Recent Developments of the Quantum Chemical Cluster Approach for Modeling Enzyme Reactions. *J. Biol. Inorg. Chem.* **2009**, *14*, 643−651.

(7) Viklund, K.; Kolmodin, K.; Nordvall, G.; Swahn, B. M.; Svensson, M.; Gravenfors, Y.; Rahm, F. Creation of Novel Cores for β-Secretase (BACE-1) Inhibitors: A Multi-Parameter Lead Generation Strategy. *ACS Med. Chem. Lett.*, DOI: 10.1021/ml5000433.

(8) Zhao, H. Scaffold selection and scaffold hopping in lead generation: a medicinal chemistry perspective. *Drug Discov. Today* **2007**, *12*, 149−155.

(9) Ginman, T.; Viklund, J.; Malmström, J.; Blid, J.; Emond, R.; Forsblom, R.; Johansson, A.; Kers, A.; Lake, F.; Sehgelmeble, F.; Sterky, K. J.; Bergh, M.; Lindgren, A.; Johansson, P.; Jeppsson, F.; Fälting, J.; Gravenfors, Y.; Rahm, F. Core Refinement towards Permeable β-Secretase (BACE-1) Inhibitors with Low hERG Activity. *J. Med. Chem.* **2013**, *56*, 4181−4205.

(10) Ghosh, A. K.; Brindisi, M.; Tang, J. Developing beta-secretase inhibitors for treatment of Alzheimer's disease. *J. Neurochem.* **2012**, *120*, 71−83.

(11) Hong, L.; Koelsch, G.; Lin, X.; Wu, S.; Terzyan, S.; Ghosh, A K; Zhang, X C; Tang, J. Structure of the Protease Domain of Memapsin 2 (β-Secretase) Complexed with Inhibitor. *Science* **2000**, *290*, 150−153.

(12) Hardy, J.; Selkoe, D. J. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* **2002**, *297*, 353−356.

(13) Vassar, R.; Bennett, B. D.; Babu-Khan, S.; Kahn, S.; Mendiaz, E. A.; Denis, P.; Teplow, D. B.; Ross, S.; Amarante, P.; Loeloff, R.; Luo, Y.; Fisher, S.; Fuller, J.; Edenson, S.; Lile, J.; Jarosinski, M. A.; Biere, A. L.; Curran, E.; Burgess, T.; Louis, J.; Collins, F.; Treanor, J.; Rogers, G.; Citron, M. beta-Secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science* **1999**, *286*, 735−741.

(14) Alzheimer's Association. Alzheimer's disease facts and figures. *Alzheimer's Dement.* **2012**, *8*, 131−168.

(15) Evin, G.; Lessene, G.; Wilkins, S. BACE inhibitors as potential drugs for the treatment of Alzheimer's disease: focus on bioactivity. *Recent. Pat. CNS Drug Discov.* **2011**, *6*, 91−106.

(16) Probst, G.; Xu, Y. Small-molecule BACE1 inhibitors: a patent literature review (2006−2011). *Expert Opin. Ther. Pat.* **2012**, *22*, 511−540.

(17) Swahn, B. M.; Holenz, J.; Kihlstrom, J.; Kolmodin, K.; Lindstrom, J.; Plobeck, N.; Rotticci, D.; Sehgelmeble, F.; Sundstrom, M.; Berg, S.; Falting, J.; Georgievska, B.; Gustavsson, S.; Neelissen, J.; Ek, M.; Olsson, L. L.; Berg, S. Aminoimidazoles as BACE-1 inhibitors: the challenge to achieve in vivo brain efficacy. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 1854−1859.

(18) Swahn, B. M.; Kolmodin, K.; Karlström, S.; von Berg, S.; Söderman, P.; Holenz, J.; Berg, S.; Lindström, J.; Sundström, M.; Turek, D.; Kihlström, J.; Slivo, C.; Andersson, L.; Pyring, D.; Rotticci, D.; Öhberg, L.; Kers, A.; Bogar, K.; Bergh, M.; Olsson, L.; Janson, J.; Eketjäll, S.; Georgievska, B.; Jeppsson, F.; Fälting, J. Design and Synthesis of β-Secretase (BACE1) Inhibitors with In Vivo Brain Reduction of β-Amyloid Peptides. *J. Med. Chem.* **2012**, *55*, 9346−9361.

(19) Malamas, M. S.; Erdei, J.; Gunawan, I.; Turner, J.; Hu, Y.; Wagner, E.; Fan, K.; Chopra, R.; Olland, A.; Bard, J.; Jacobsen, S.; Magolda, R. L.; Pangalos, M.; Robichaud, A. J. Design and Synthesis of 5,5′-Disubstituted Aminohydantoins as Potent and Selective Human beta-Secretase (BACE1) Inhibitors. *J. Med. Chem.* **2010**, *53*, 1146−1158.

(20) May, P. C.; Dean, R. A.; Lowe, S. L.; Martenyi, F.; Sheehan, S. M.; Boggs, L. N.; Monk, S. A.; Mathes, B. M.; Mergott, D. J.; Watson, B. M.; Stout, S. L.; Timm, D. E.; Smith Labell, E.; Gonzales, C. R.; Nakano, M.; Jhee, S. S.; Yen, M.; Ereshefsky, L.; Lindstrom, T. D.; Calligaro, D. O.; Cocke, P. J.; Greg Hall, D.; Friedrich, S.; Citron, M.;

Audia, J. E. Robust central reduction of amyloid-beta in humans with an orally available, non-peptidic beta-secretase inhibitor. *J. Neurosci.* **2011**, *31*, 16507−16516.
(21) Tresadern, G.; Delgado, F.; Delgado, O.; Gijsen, H.; Macdonald, G. J.; Moechars, D.; Rombouts, F.; Alexander, R.; Spurlino, J.; Gool, M. V.; Vega, J. A.; Trabanco, A. A. Rational design and synthesis of aminopiperazinones as beta-secretase (BACE) inhibitors. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 7255−7260.
(22) Cumming, J. N.; Smith, E. M.; Wang, L.; Misiaszek, J.; Durkin, J.; Pan, J.; Iserloh, U.; Wu, Y.; Zhu, Z.; Strickland, C.; Voigt, J.; Chen, X.; Kennedy, M. E.; Kuvelkar, R.; Hyde, L. A.; Cox, K.; Favreau, L.; Czarniecki, M. F.; Greenlee, W. J.; McKittrick, B. A.; Parker, E. M.; Stamford, A. W. Structure based design of iminohydantoin BACE1 inhibitors: Identification of an orally available, centrally active BACE1 inhibitor. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 2444−2449.
(23) Gravenfors, Y.; Viklund, J.; Blid, J.; Ginman, T.; Karlström, S.; Kihlström, J.; Kolmodin, K.; Lindström, J.; von Berg, S.; von Kieseritzky, F.; Slivo, C.; Swahn, B. M.; Olsson, L.-L.; Johansson, P.; Eketjäll, S.; Fälting, J.; Jeppsson, F.; Strömberg, K.; Janson, J.; Rahm, F. New Aminoimidazoles as β-Secretase Inhibitors Showing Aβ-Lowering in Brain. *J. Med. Chem.* **2012**, *55*, 9297−9311.
(24) Albert, J.; Arnold, J.; Chessari, G.; Congreve, M. S.; Edwards, P.; Murray, C.; Patel, S. Novel 2-amino-imidazole-4-one compounds and their use in the manufacture of a medicament to be used in the treatment of cognitive impairment, alzheimer's disease, neurodegeneration and dementia. WO2007058601A1, 2007.
(25) Berg, S.; Burrows, J.; Chessari, G.; Congreve, M. S.; Hedstroem, J.; Hellberg, S.; Hoegdin, K.; Kihlstroem, J.; Kolmodin, K.; Lindstroem, J.; Murray, C.; Patel, S. Novel 2-amino-imidazole-4-one compounds and their use in the manufacture of a medicament to be used in the treatment of cognitive impairment, alzheimer's disease, neurodegeneration and dementia. WO2007058602A2, 2007.
(26) Holenz, J.; Kers, A.; Kolmodin, K.; Rotticci, D.; Oehberg, L.; Hellberg, S. Aryl and heteroaryl substituted isoindole derivatives as bace inhibitors. WO2009005470A1, 2009.
(27) Holenz, J.; Kers, A.; Kolmodin, K.; Rotticci, D.; Rakos, L.; Hellberg, S. Aryl and heteroaryl substituted isoindole derivatives as bace inhibitors. WO2009005471A1, 2009.
(28) Holenz, J.; Karlstroem, S.; Kihlstroem, J.; Kolmodin, K.; Rakos, L.; Soederman, P.; Swahn, B.; Von Berg, S., Von Kieseritzky, F. New compounds 575. WO2010056195A1, 2010.
(29) Holenz, J.; Karlstroem, S.; Kihlstroem, J.; Kolmodin, K.; Lindstroem, J.; Rakos, L.; Rotticci, D.; Swahn, B.; Von Berg, S. 5h-Pyrrolo[3,4-L]pyrazin-7-amine derivatives inhibitors of beta-secretase. WO2011002409A1, 2011.
(30) Holenz, J.; Karlstroem, S.; Kihlstroem, J.; Kolmodin, K.; Lindstroem, J.; Rakos, L.; Rotticci, D.; Soederman, P.; Sundstroem, M.; Swahn, B.; Von Berg, S. 5h-Pyrrolo[3,4-b]pyridin derivatives and their use. WO2010056194A1, 2010.
(31) Chessari, G.; Congreve, M. S.; Holenz, J.; Murray, C.; Patel, S.; Rakos, L.; Rotticci, D. New Compounds 834. US20080287462A1, 2008.
(32) Holenz, J.; Karlstroem, S.; Kolmodin, K.; Lindstroem, J.; Rakos, L.; Rotticci, D.; Soederman, P.; Swahn, B.; Von, B., Stefan. New compounds 578. WO2010056196A1, 2010.
(33) Dillard, L. W.; Yuan, J.; Leftheris, K.; Venkatraman, S.; Wu, G.; Jia, L.; Xu, Z.; Cacatian, S.; Morales-Ramos, A.; Singh, S.; Zheng, Y. Inhibitors of beta-secretase.WO2011106414A1, 2011.
(34) *Jaguar*, version 7.9; Schrodinger, LLC: New York, NY, 2012.
(35) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
(36) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functional. *Theor. Chem. Acc.* **2008**, *120*, 215−241.
(37) Tannor, D. J.; Marten, B.; Murphy, R.; Friesner, R. A.; Sitkoff, D.; Nicholls, A.; Ringnalda, M.; Goddard, W. A., III; Honig, B. Accurate First Principles Calculation of Molecular Charge Distributions and Solvation Energies from Ab Initio Quantum Mechanics and Continuum Dielectric Theory. *J. Am. Chem. Soc.* **1994**, *116*, 11875−11882.
(38) Marten, B.; Kim, K.; Cortis, C.; Friesner, R. A.; Murphy, R. B.; Ringnalda, M. N.; Sitkoff, D.; Honig, B. New Model for Calculation of Solvation Free Energies: Correction of Self-Consistent Reaction Field Continuum Dielectric Theory for Short-Range Hydrogen-Bonding Effects. *J. Phys. Chem.* **1996**, *100*, 11775−11788.
(39) Marenich, A. V.; Olson, R. M.; Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Self-Consistent Reaction Field Model for Aqueous and Nonaqueous Solutions Based on Accurate Polarized Partial Charges. *J. Chem. Theor. Comp.* **2007**, *3*, 2011−2033.
(40) Olson, R. M.; Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Charge Model 4 and Intramolecular Charge Polarization. *J. Chem. Theor. Comp.* **2007**, *3*, 2046−2054.
(41) Schrödinger Suite 2012: *Epik*, version 2.3; *Glide*, version 5.8; *Impact*, version 5.8; *MacroModel*, version 9.9; *Prime*, version 3.1; *Protein Preparation Wizard*; Schrödinger, LLC: New York, NY, 2012.
(42) Reddy, A. S.; Pati, S. P.; Kumar, P. P.; Pradeep, H. N.; Sastry, G. N. Virtual screening in drug discovery - a computational perspective. *Curr. Protein Pept. Sci.* **2007**, *8*, 329−351.
(43) Gohlke, H.; Klebe, G. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angew. Chem., Int. Ed.* **2002**, *41*, 2644−2676.
(44) Hann, M. M. Molecular obesity, potency and other addictions in drug discovery. *Med. Chem. Comm.* **2011**, *2*, 349−355.
(45) *ClogP*, version 4.3; Daylight/Biobyte, BioByte Corp.: Claremont, CA, 2008.
(46) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Adding Explicit Solvent Molecules to Continuum Solvent Calculations for the Calculation of Aqueous Acid Dissociation Constants. *J. Phys. Chem. A* **2006**, *110*, 2493−2499.
(47) Ravindranathan, K.; Tirado-Rives, J.; Jorgensen, W. L.; Guimarães, C. R. W. Improving MM-GB/SA Scoring through the Application of the Variable Dielectric Model. *J. Chem. Theory Comput.* **2011**, *7*, 3859−3865.
(48) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A. Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision B.01; Gaussian, Inc.: Wallingford CT, 2010.