

Discovery of a Novel Warhead against β -Secretase through Fragment-Based Lead Generation[§]

Stefan Geschwindner,^{†,||} Lise-Lotte Olsson,^{†,||} Jeffrey S. Albert,[#] Johanna Deinum,[⊥] Philip D. Edwards,[#] Tonny de Beer,^{†,‡} and Rutger H. A. Folmer^{*,†}

Global Structural Chemistry, AstraZeneca R&D, S-431 83, Mölndal, Sweden, CVGI Molecular Pharmacology, AstraZeneca R&D, S-431 83, Mölndal, Sweden, and CNS Lead Generation, AstraZeneca R&D, 1800 Concord Pike, Wilmington, Delaware 19850

Received July 9, 2007

Fragment-based lead generation was applied to find novel small-molecule inhibitors of β -secretase (BACE-1), a key target for the treatment of Alzheimer's disease. Fragment hits coming from a 1D NMR screen were characterized by BIAcore, and the most promising compounds were soaked into protein crystals to help the rational design of more potent hit analogues. Problems arising due to our inability to grow BACE-1 crystals at the biologically relevant pH at which the screen was run were overcome by using endothiapepsin as a surrogate aspartyl protease. Among others, we identified 6-substituted isocytosines as a novel warhead against BACE-1, and the accompanying paper in this journal describes how these were optimized to a lead series of nanomolar inhibitors.¹

Introduction

Although screening large collections of relatively drug-like compounds in high-throughput mode has been the paradigm of drug discovery ever since HTS^a was introduced in the late 1980s, there remain fundamental issues with HTS that limit its scope. First, even in an ideal world where reliable assays can be developed against all target proteins, HTS will only pull out compounds that already exist in corporate collections. This not only puts a strain on novelty and downstream intellectual property of the hits themselves but also limits medicinal chemistry to existing themes and knowledge; exploration of compound chemical space in truly new directions is prohibited. An additional concern is that corporate libraries are continuously being filled with compounds that have been synthesized in late-phase discovery projects and hence have drug-like rather than lead-like properties, e.g., high molecular weights and lipophilicity.^{2–4}

Second, and not completely unrelated, even the largest compound libraries in the industry will sample compound space only incredibly sparsely. Whatever reported estimate of the number of possible chemical structures with lead-like sizes one would like to accept, it will be larger than the number of atoms on Earth. For many target proteins, suitable lead molecules will simply be absent from our collections.

Fragment-based lead generation (FBLG, recently reviewed^{5–7}) addresses most of the issues above and is increasingly being accepted as a valuable complement to high-throughput screening. The sampling problem is addressed through initially

screening collections of quite small and simple compounds, usually consisting of no more than one or two rings with a few substituents (the term fragment is used solely to denote that smallness). Doing so, compound space is sampled effectively, and it is often possible to find small ligands that bind very efficiently to the target. Intuitively one probably feels this must be true, but it also follows from an analysis by Hann and co-workers.⁸ These fragment hits (FRITs) can be considered building blocks that can be combined (merged, linked) to form larger and potentially much more potent and druglike lead compounds.^{9,10} Alternatively, FRITs can be considered seeds or anchor points, which can be synthetically expanded into lead compounds, picking up increasingly more interactions with the target protein. These fragment linking or expansion stages usually have a significant design component, hence allowing for a high amount of novelty in the final lead compound.

Herein, we present our fragment-based lead generation approach against β -secretase (BACE-1), a target for which no nonpeptidic inhibitors were known at the time we initiated our work, and where our HTS and peptide-mimetic approaches had been unsuccessful.¹ BACE-1, an aspartyl protease, is a well-established target for the treatment of Alzheimer's disease or possibly even for preventing the disease from occurring. Being responsible for cleaving amyloid precursor protein, BACE-1 plays a key and rate-limiting role in the generation of β -amyloid fibrils, which in turn aggregate to form the neurofibrillary plaques causing dementia.¹¹ While current Alzheimer's treatments only treat the disease symptoms, inhibition of BACE-1 is widely believed to be a promising way to address the underlying neuropathology.

In this paper, we will show the discovery of a novel warhead against BACE-1. In the accompanying paper¹ it will be demonstrated how this warhead was progressed to a lead series of nanomolar inhibitors, using a combination of structure-based drug design and traditional medicinal chemistry.

Results and Discussion

The Fragment Screen. Even though fragment hits can be very efficient binders, they can have only a limited number of interaction points with the protein, which will limit their absolute affinity. Consequently, fragment screening hits are sometimes referred to as “weak binders”, which is unfortunate, as their

[§] Coordinates for the EP complex structure with compound **2** have been deposited in the Protein Data Bank, with the accession code 2v00

^{*} To whom correspondence should be addressed. E-mail: rutger.folmer@astrazeneca.com, tel: +46-31-776 2043.

[†] Global Structural Chemistry.

[⊥] CVGI Molecular Pharmacology.

[#] CNS Lead Generation.

^{||} These authors contributed equally to this work.

[‡] Present address: Proctor & Gamble Pharmaceuticals, Health Care Research Center, Mason, OH 45040-8006.

^a Abbreviations: FBLG, fragment-based lead generation; FRIT, fragment hit; BACE, β -site amyloid-precursor protein cleaving enzyme, β -secretase; K_D , dissociation constant; FRET, fluorescence resonance energy transfer; RU, resonance units; SPR, surface plasmon resonance; TDC, target definition compound; HTS, high-throughput screening; ISA, inhibition in solution assay; EP, endothiapepsin; SAR, structure–activity relationship; LE, ligand (binding) efficiency; HAC, heavy atom count.

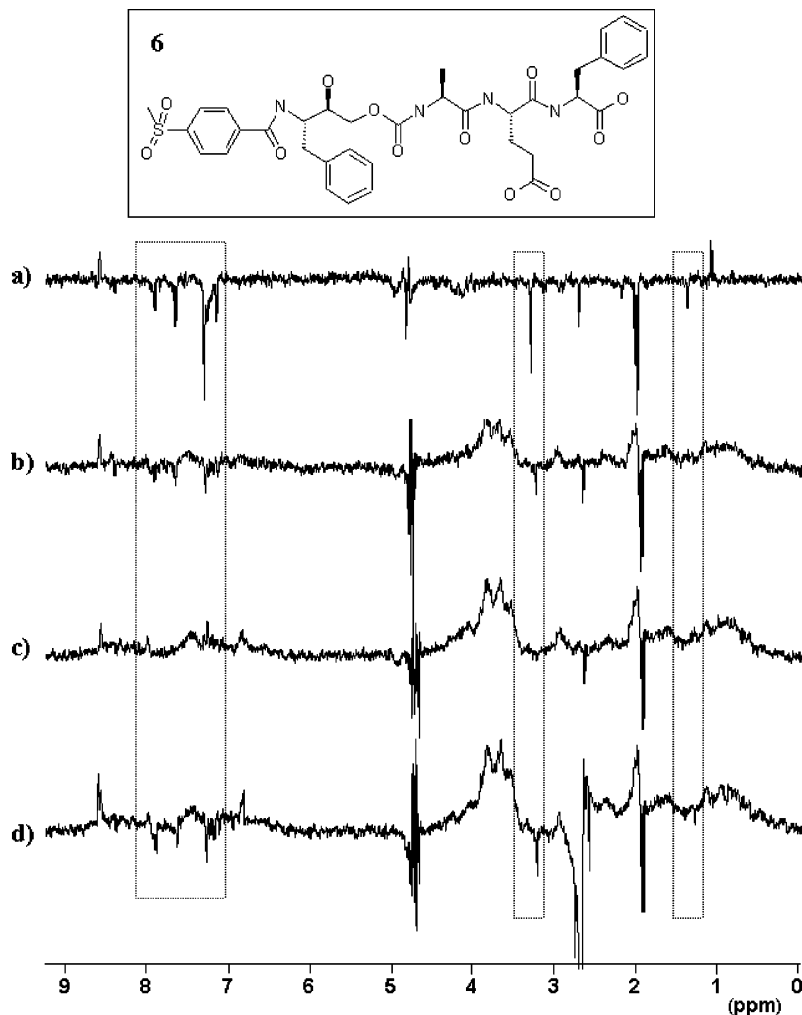


Figure 1. Series of NMR waterLOGSY experiments, performed to establish the NMR screening conditions. Boxes mark the signals from compound **6**, whose structure is shown in the insert. Compound **6** binds in fast exchange, with low affinity (IC_{50} 30 μ M), to the active site of BACE-1. (a) Spectrum of **6** in absence of protein, showing strong negative peaks for the unbound compound. (b) Spectrum after addition of 1.8 μ M BACE-1. Signals from **6** are significantly less negative, indicating fast-exchange (i.e., μ M affinity) binding to the protein. (c) Spectrum at 3.2 μ M protein. Signals from **6** become even more positive, indicating increased interaction with the protein. (d) Spectrum after addition of the high affinity binder OM99-2, displacing compound **6** (NMR signals become more negative again, a clear indication of **6** competing for the same binding site as OM99-2).

binding is often highly efficient if one takes into account the compound size.¹² Usually, dissociation constants (K_D) for FRITs are in the 20–2000 μ M range. Such weak affinities require robust screening methods for the reliable detection of binding.

We use NMR (1D or 2D) for fragment screening, mainly for two reasons: we believe it to be the best balance between information richness and throughput, and we and others have experienced NMR to be highly trustworthy and robust in identifying weak binding.^{13,14} An often-overlooked advantage of 1D NMR for screening purposes is that it can reliably demonstrate binding using compound concentrations much below the K_D of the interaction. For example, t1rho-based methods¹⁵ can reproducibly detect binding with millimolar affinity, while using only tens of micromolar compound concentrations in the experiment. Most other biochemical or biophysical methods require compound concentrations around or above the IC_{50} or K_D , in order to detect binding. This can be challenging if one is aiming for affinities of around 1 mM and is a serious concern, particularly in biochemical screening at high concentration, potentially causing either false positives (nonspecific binding) or false negatives (limited compound solubility).

Prior to performing the primary screen, we established the experimental conditions required for detection of ligand binding.

Moreover, we tested competition with a highly potent, well-characterized inhibitor. Here, we used OM99-2 (K_i of 1.6 nM),¹⁶ a noncleavable substrate analogue (Glu-Val-Asn-Leu*Ala-Ala-Glu-Phe, where the asterisk designates the hydroxyethylene transition-state isostere replacing the substrate's amide bond). This involved testing a series of known binders, with affinities (K_D) in the range where we expected the fragments to bind, typically 20–2000 μ M. Figure 1 shows how the experimental setup was validated using one such known binder. Comparison between spectra a and b shows the binding event, and comparison between spectra c and d confirms that the binding is competitive with OM99-2.

When performing the actual fragment screen, we always recorded two experiments per compound mixture; those corresponding to spectrum c and d in Figure 1. We would indeed stress the importance of running competition experiments in ligand-detected screening, that is, identifying hits through comparison of the experiments run in the absence and presence of a known, potent binder.¹⁷ In our hands, almost any compound at 100s of μ M concentration will display some (nonspecific) interaction with any protein. Without a proper competition experiment, a fragment screen is likely to yield a large proportion of false positives, i.e., nonspecific binders. The NMR

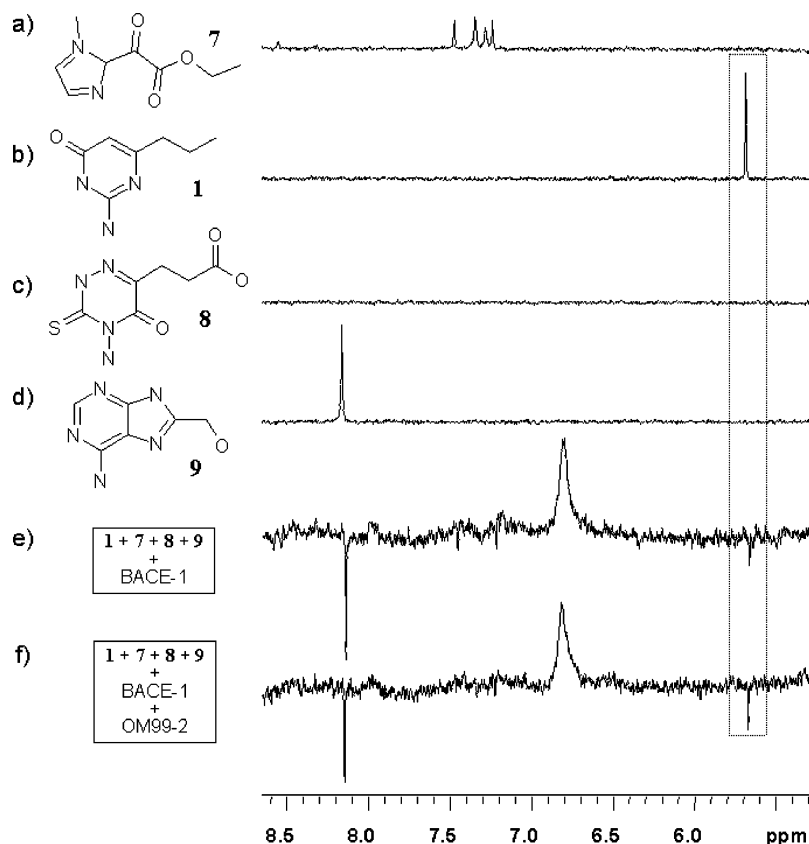


Figure 2. Screenshot from our spectrum analysis setup. (a–d) Straight 1D spectra of the corresponding compounds to the left. They were acquired in the absence of protein, and served as reference spectra during the analysis of the actual screening spectra (e, f). (e) waterLOGSY experiment run on a mixture of the four compounds with protein. (f) Same as e but with saturating amounts of OM99-2. For details, see Experimental Section. The dotted box denotes the signal from the isocytosine aromatic H⁵ proton. Its waterLOGSY signal becomes more negative in the presence of OM99-2, indicating that the isocytosine and OM99-2 compete for the same binding site on BACE-1. This is not true for the signals from compounds 7 and 9, indicating that only 1 is a fragment hit in this fragment mixture (compound 8 does not have aromatic protons, hence the empty spectrum in this region, but the aliphatic portion of the spectrum showed it is not a hit). The overall somewhat lower intensity of f compared to e is caused by sample dilution during addition of OM99-2.

screen was performed on a 2000-compound general fragment library and identified a variety of fragments binding competitively with OM99-2. The affinities of these FRITs were modest; mostly in the low millimolar region. The overall hit rate was about 0.5%, which in our experience is comparatively low in fragment screening. This agreed well with the fact that no high-affinity small molecule ligand previously was known for BACE-1, and the fact that our HTS approach did not deliver suitable lead compounds.¹ This relationship between low hit rates in fragment screens and lack of prior existence of a potent ligand has recently been demonstrated by Hajduk and co-workers.¹⁸

For the purpose of the work described here, we focused on an isocytosine fragment hit, which was the first FRIT chosen for chemistry optimization, and eventually progressed to a successful lead series.¹ Figure 2 shows an example of the data collected in the fragment screen and highlights the isocytosine primary fragment hit, compound 1. The difference in the waterLOGSY signal between the bound (Figure 2e) and displaced state (Figure 2f) is small, which reflects the low affinity of this fragment. Similarly small effects were seen for all FRITs in the primary screen, and they were usually smaller than the differences observed between waterLOGSY read-out in absence and presence of protein (cf. Figure 1). This underpins the need for performing the screen as a series of displacement experiments; contributions from nonspecific binding are effectively absent in the difference between the NMR readout in the bound and displaced states.

Fragment Hit Characterization Using BIAcore. An earlier reported FBLG study performed at Astex in collaboration with AstraZeneca used X-ray crystallography as the screening method and an enzymatic fluorescence resonance energy transfer (FRET) assay to follow up the output from the screen.^{19,20} The assay was needed to address the functionality of the hits, i.e., to ensure inhibition of the enzyme and to quantify their affinity. However, FRET assays may suffer from autofluorescence of compounds, and, in general, artifacts arising from nonspecific inhibition are likely to appear in biochemical assays that are run at the high concentrations ($\gg 100 \mu\text{M}$) needed to detect binding of the low-affinity fragment hits. We decided therefore not to use a biochemical assay but instead surface plasmon resonance (SPR) for the characterization of our screening output.

SPR is a widely used label-free detection technique that is usually associated with BIAcore technology and aims at studying the interaction of proteins with other molecules.^{21–25} BIAcore has also been reported as a technique for screening fragment libraries.^{26,27} It is important to point out that BIAcore, like NMR, is able to monitor and evaluate the relatively weak interactions that FRITs normally display with the target protein.^{26,27} However, we feel that NMR has its key strength in detecting very weak interactions (mM range) and can be optimally employed in screening of larger libraries (1000s of compounds) in an automated fashion, whereas BIAcore can deliver precise affinity values and kinetic parameters, at the cost of lower throughput. This is of particular importance for the ranking and

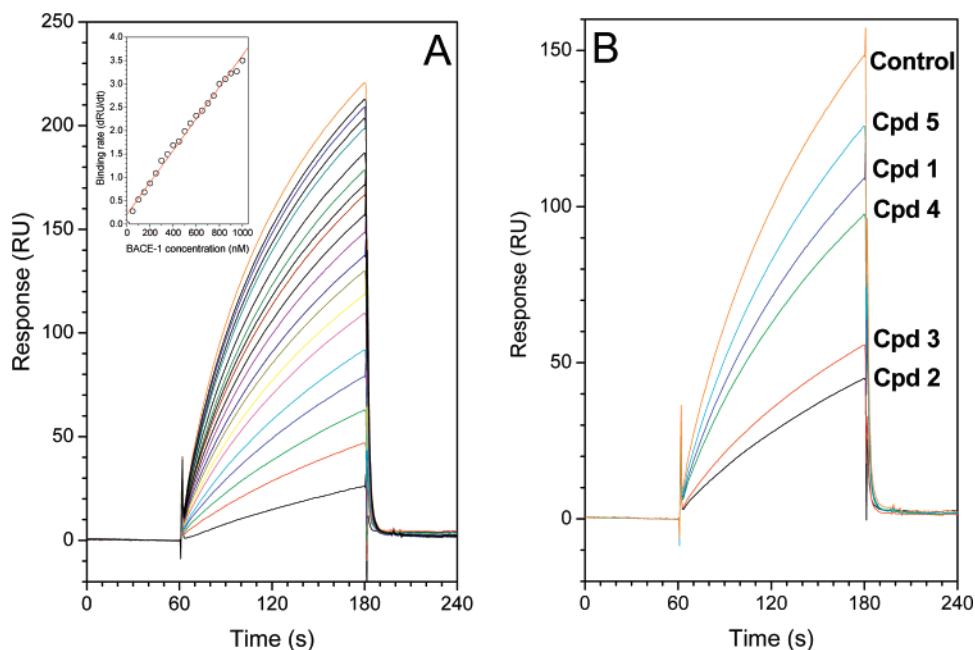


Figure 3. (A) Calibration of BIAcore ISA. Concentration series (0.05 to 1 μM) of BACE-1 were run at 20 $\mu\text{L}/\text{min}$ at 25 $^{\circ}\text{C}$ over a CM5 surface with 360 RU amine-coupled substrate analogue (P₁ (S)-statin-substituted) as TDC. The running buffer was composed of 25 mM sodium acetate, 200 mM NaCl, 0.005% Tween 20, pH 4.5. After each cycle at 180 s, the surface was regenerated with a pulse of 50 mM Tris/HCl at pH 8.5 and 0.5% SDS. The inset depicts the concentration dependency of the initial binding rate (dRU/dt) with the protein concentration. The red line represents the linear regression of the data ($R = 0.988$). (B) Build-up of SAR using BIAcore ISA. BACE-1 protein at a concentration of 0.5 μM was injected over the modified CM5 surface (see conditions as described above), after preincubation for 10 min with 1 mM of the respective compounds (identified with “cpd” and number, cf Table 1). The control contained only BACE-1 protein in the absence of any compound and served as a reference for the calculation of the free BACE-1 concentration.

selection of fragment hits and the FRIT analogues we study subsequently.

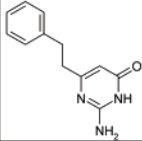
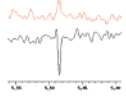
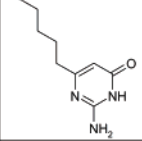
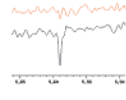
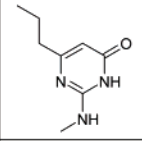
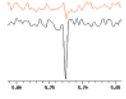
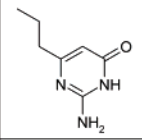
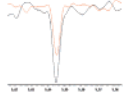
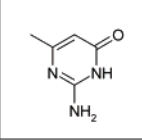
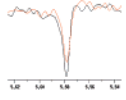
The traditional, best-known BIAcore assay format is the direct binding assay, where the protein target is immobilized on the BIAcore sensor chip and compounds are passed over the surface. As the work with fragment-sized compounds brings in limitations and challenges because of their low affinity and small analyte masses,²⁸ alternative assays such as the surface competition assay or the inhibition in solution assay (ISA) are important alternatives to the direct binding assay.²¹ The ISA setup consists of a target definition compound (TDC) that is immobilized on the sensor surface, which serves as a probe for a defined binding site. The protein target is kept at a constant concentration, and FRITs are mixed with the protein thus lowering the concentration of the free protein in a concentration-dependent manner. The concentration of the free protein is subsequently measured by passing the mixture over the sensor surface, which should result in a signal decrease if the FRIT is competing for the defined binding site with the TDC. The application of ISA has several advantages that are particularly important for the rapid and accurate analysis of FRIT binding. Using just a single compound concentration, this approach enables a fast and qualitative affinity ranking of FRITs without the need to determine the kinetic parameters or the absolute affinity for this interaction. Acquiring binding data at varying concentrations of FRITs allows determination of their K_D values. A critical factor is the choice of the target definition compound. It needs to be covalently linked to the chip, while maintaining high affinity binding to the site of interest on the target. Furthermore, proper TDCs should have high on-rates to be able to perform the binding studies under conditions of mass transport limitation and low off-rates so that the dissociation in the initial association phase is negligible.

For the characterization and analysis of FRIT binding to BACE-1, see Figure 3, we designed an inhibition in solution assay with a target definition compound derived from a P₁ (S)-statin substituted substrate reported to be a potent inhibitor of BACE-1.²⁹ In order to test whether identified FRITs might act as broad aspartyl protease warheads we also configured a BIAcore ISA using the renin-inhibitor H-142^{30–32} as immobilized TDC and probed the FRIT binding to the aspartyl protease endothiapepsin (EP) in a similar fashion. H-142 has also previously been shown to bind and inhibit EP.³² Both the BACE-1 and EP TDCs form stable complexes with their target proteins and fulfill the aforementioned criteria for the dissociation and association rates. Experimental details of the BIAcore assays are included in the Experimental Section.

We ensured that the BIAcore measurements were done under conditions of mass transport limitations, by applying high immobilization densities of the TDCs (>300 resonance units, RU) and a low flow rate (20 $\mu\text{L}/\text{min}$). The response in the SPR assay, expressed as the initial binding rate, was found to be linear with the free protein concentration in solution, in the range from 0.1 to 1000 nM, both with BACE-1 and EP, see Figure 3. Moreover, in the presence of excess inhibitor none of the proteins bound to the immobilized compound. This indicated that the ISA setup was suitable to determine the concentration of the fraction of noninhibited protein in solution as a measure of inhibition by the FRIT.

Fragment Hit Analogues. After the primary fragment screen, most FBLG approaches enter what is often referred to as the ‘analoguing phase’, during which corporate or commercial libraries are harnessed for nearest neighbors to the fragment hits. Table 1 lists four such neighbors that were present in our corporate compound library. All compounds 1–5 were then tested both in the NMR waterLOGSY experiment and in

Table 1. Structure–Activity Relation of Five Isocytosine Analogues

Compound	Structure	Response with BACE-1 in BIAcore ISA ^a	Response in waterLOGSY ^b	BACE-1 K_D (mM) ^c	EP K_D (mM) ^c
2		31 %		0.66	0.22
3		38 %		n.d. ^d	n.d.
4		65 %		n.d.	n.d.
1 (NMR Hit)		72 %		4.45	n.d.
5		83 %		n.d.	n.d.

^a % free BACE-1 in presence of 1 mM compound, i.e., smaller numbers indicate higher affinity, determined from the BIAcore assay with amine-coupled substrate analogue (P₁ (*S*)-statin-substituted) as TDC. ^b The NMR spectra show the H⁵ proton of the isocytosine ring. Red spectra are waterLOGSY spectra of the ligand plus protein, showing the bound state of the ligand. Black spectra are of ligand plus protein plus OM99-2, showing the displaced ligand (cf. spectra c and d in Figure 1). ^c K_D determined from BIAcore ISA. ^d n.d. = not determined.

BIAcore. The results of both techniques correlate very well and indicate among others that larger substituents at the C⁶ position afford increasingly more affinity. In particular, the additional phenyl group on compound **2** increased the affinity by 7-fold over the NMR hit **1**. Although at this stage absolute affinities of the FRITs were still rather low, it was very comforting to see this crude structure activity–relationship (SAR) emerging. In the absence of crystallographic information, this SAR was a crucial observation, encouraging us to further pursue the isocytosine FRIT.

In order to appreciate the absolute affinities of our best compounds, we calculated ligand binding efficiencies (LE).¹² In this report, we define LE as the free energy of binding divided by the heavy atom count (HAC): $LE = -\Delta G/HAC = -RT \ln(K_D)/HAC$. If the K_D is not known, we use K_i instead. For compounds **1** and **2** we find $LE = 0.29$ and 0.27 respectively. In comparison, OM99-2 and **6** ($K_i = 30 \mu M$, in-house data) have $LE = 0.19$ and 0.12 , respectively. This clearly shows that even though the absolute affinities of our best compounds are rather modest (Table 1), they do bind efficiently to BACE-1 and should be suitable starting points (“anchors”) for chemical optimization.

Endothiapepsin as Structural Surrogate. After characterizing the available nearest neighbors to compound **1**, we identified compound **2** as our most promising ligand (Table 1). Unfortunately, we did not succeed in getting crystallographic information of that compound complexed with BACE-1. At that point in time, our in-house crystallization conditions only allowed cocrystallization with compounds that displayed high affinities (μM to nM) toward BACE-1. An additional concern was that both of our crystallization systems, and those reported

by others,³³ have buffer conditions close to neutral pH, while substrate catalysis of BACE-1 is optimal around pH 4.5.³⁴ It is well-established that both substrate catalysis and inhibitor binding are pH-dependent processes in aspartyl proteases, strongly linked to the ionization state of the two aspartic acid residues. We had therefore chosen to perform the NMR fragment screen at pH 4.5, the same pH also later used in the BIAcore assay. The higher pH in the crystals is likely to affect the ionization states of the Asp residues, which could be disadvantageous for the (already low) affinities of the FRITs, thus hindering successful complex formation in the crystal.

In order to assess this apparent incompatibility of the neutral pH in the crystals, with the lower, and biologically more relevant pH at which the screen was run, we measured the pH dependency of binding of compound **2** to BACE-1. This pH dependency could not be assessed using the ISA format because of the pH-dependent binding of BACE-1 to the TDC, which resulted in a loss of the SPR signal at pH > 6.0. We chose, therefore, to monitor the thermal denaturation of the BACE-1–ligand complex at varying pH-values in order to probe the binding of compound **2**. While we observed increased melting temperature for BACE-1 in the presence of compound **2** at low pH, there was no detectable stabilization at pH > 5.8. This suggests there is no appreciable interaction between **2** and the protein at pH > 5.8, which indeed explained why we failed in obtaining crystals of the complex of BACE-1 and the isocytosines at near neutral pH.

We realized that without information on mode of binding it would be difficult to progress the isocytosine fragment hit to more potent lead compounds. But we needed potent compounds to be successful in crystallizing them in BACE-1. To circumvent

Table 2. Data Collection and Refinement Statistics for the Crystal Structure of EP with Compound 2

data collection	
resolution (Å)	1.55
no. of unique reflections	47864
completeness (%)	99.1 (98.8) ^a
data redundancy	3.9 (3.7)
<i>R</i> _{merge} (%)	4.7 (19.3)
<i>I</i> /σ(<i>I</i>)	9.4 (3.7)
model and refinement	
<i>R</i> _{work} (%)	16.7
<i>R</i> _{free} (%)	21.0
rmsd ^b bond lengths (Å)	0.011
rmsd bond angles (deg)	1.3
average <i>B</i> -factor (Å ²)	15.0
Protein Data Bank ID	2v00

^a Values in parentheses are for the highest resolution shell, 1.55–1.59 Å. ^b Root-mean-square deviation.

this situation, we decided to use a surrogate aspartyl protease to replace BACE-1 in crystallographic studies. The intention was that the structure information obtained with the other protease would initially give us sufficient knowledge on the binding mode of the isocytosines, to guide the design of more potent analogues that eventually could be crystallized with our real target protein, BACE-1. We chose endothiapepsin (EP) as the surrogate, based on the in-house availability of large amounts of protein from a previous drug-discovery project, as well as that the crystallization of EP has been reported to be successful under acidic conditions (pH 4.5).³⁵ The core of the active sites of BACE-1 and EP are similar, both having two catalytic aspartates and a nearby, flexible β hairpin loop, referred to as ‘the flap’. However, away from these aspartic acids, the amino acid sequences of the two proteins are no longer homologous, and EP would be less suitable as a BACE-1 surrogate to study compounds binding to those regions. Using SPR, we measured efficient binding of compound 2 to EP (Table 1). Subsequent soaking of an EP crystal with compound 2 at pH 4.6 readily resulted in a crystal structure of the complex.

Structure of Compound 2 Bound to EP. High-quality apo EP crystals were produced, which could be used for soaking. Data from an EP crystal soaked with compound 2 was collected to 1.55 Å resolution, and the structure was solved by molecular replacement (see Experimental Section). The compound is well defined in the electron density, and data reflecting the quality of the structure are compiled in Table 2. Figure 4 shows that the exocyclic primary amine and the N¹ nitrogen atom in the core of the ligand interact with the active site aspartic acids via hydrogen bonds. These interactions explain the reduced affinity of the isocytosines at higher pH, as the charged aspartic acid side chains lose their capability to act as hydrogen-bond donors. The ligand binds to a closed flap conformation, and there is also a hydrogen bond between the ligand carbonyl and the main-chain nitrogen of Asp81 in the flap (corresponds to Gln73 in BACE-1). An additional hydrogen bond is found between the N³ nitrogen atom of compound 2 and a glycerol molecule. Glycerol is present at 15% in the cryo solution, and its presence in the active site says therefore little about its potency as a BACE-1 ligand. It likely binds nonspecifically to the EP active site. Still, this hydrogen bond does show the potential for expanding compound 2 into this direction to gain potency.

Figure 5 compares the EP complex to the structure of the BACE-1 complex with OM99-2 (PDB code: 1FKN).³³ The primary amine of compound 2 is in the same position as the secondary alcohol of the hydroxy ethylene transition-state isostere of OM99-2, also interacting with the aspartic acids.

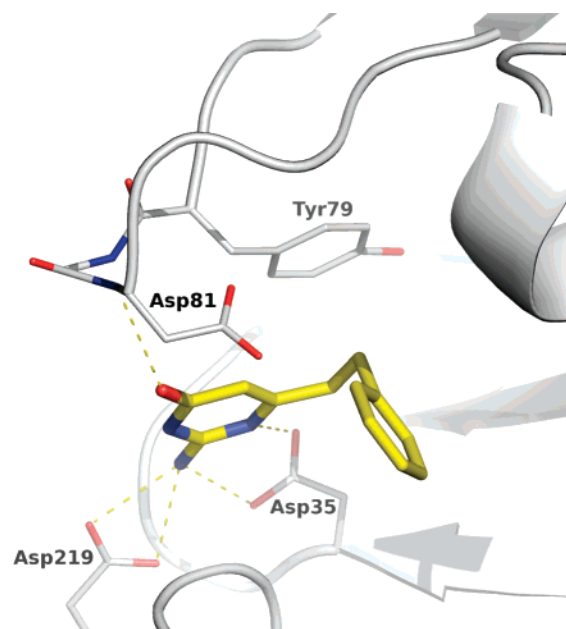


Figure 4. Compound 2 bound to endothiapepsin. There are four hydrogen bonds (shown as dashed lines) between compound 2 and the active site aspartic acids. A fifth, weaker hydrogen bond is present between the ligand and Asp81 in the flap. Compound 2 carbon atoms are shown in yellow and protein carbon atoms in gray, oxygen atoms red, and nitrogen atoms blue. Residue numbering according to EP amino acid sequence.

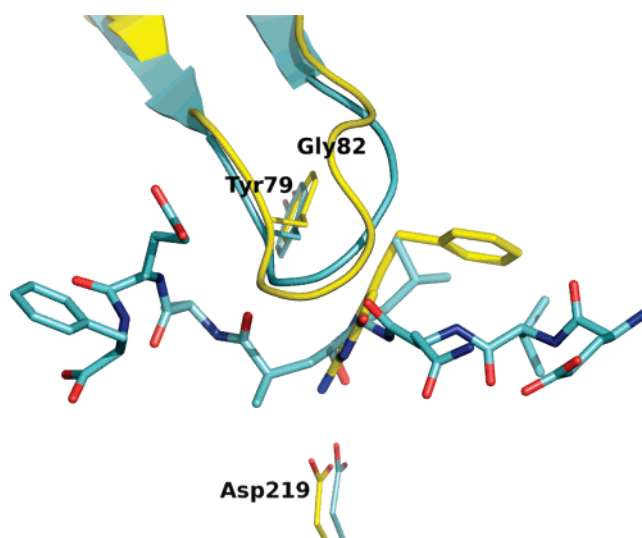


Figure 5. Endothiapepsin with compound 2 (yellow) superposed on the published OM99-2/BACE-1 1FKN structure (cyan). The flap is closed in both structures and the flap tyrosines take up similar positions. The extra flap residue in EP is labeled (Gly82). The nonprime subsites are shown to the right with the phenyl group on compound 2 positioned in the S₁/S₃ pocket. Residue numbering according to EP sequence. The two proteins were super-imposed using the ‘Secondary Structure Matching’ algorithm⁴⁸ in Coot.⁴⁶

Despite EP having an extra residue in the flap (Gly82), both structures show a similar, closed flap conformation, and a hydrogen bond to a flap amide is present in both complexes. The role of the tyrosine residue on the tip of the flap is often discussed in the stabilization of the flap conformation in aspartic proteases.³⁶ Tyr79 in EP and the corresponding Tyr71 in BACE-1 have the same orientation in both structures (Figure 5), and this orientation is characteristic for the closed-flap conformation.³⁶ The phenyl group on compound 2 occupies the S₁/S₃ pocket of the enzyme. The structure reveals that there is

room for expansion of the fragment both deeper into the S₃ pocket and on the prime side of the active site. This structure was of great help in the subsequent optimization of the isocytosines, which will be further described in the accompanying paper.¹

In conclusion, we succeeded in using EP structural information in the design of optimized inhibitors, and when the compounds became more potent, we switched to the higher-pH BACE-1 crystals for soaking. The binding mode for compound **2** found in EP, including a closed flap, was later found to be very similar in BACE-1 crystal structures of the isocytosine compounds. At this point, our compounds had grown into areas of the protein (beyond the catalytic Asp residues) where EP and BACE-1 are too dissimilar to rely on EP structures for design purposes.

Conclusions

We presented a fragment-based lead generation approach to find novel small-molecule inhibitors of BACE-1. Our work stream consisted of 1D NMR for the primary fragment screening, and both NMR and BIAcore in the subsequent analoguing phase to characterize our FRIT series and map some initial structure–activity relationships. We saved crystallography efforts for the established binders and believe that to be the optimal balance and order of using these biophysical methods, considering their different throughputs *versus* the richness of the delivered information. Problems arising due to our inability to grow crystals at the biologically relevant pH at which the screen was run were overcome by using endothiapepsin as a surrogate aspartyl protease. We identified a series of 6-substituted isocytosines as a novel scaffold for BACE-1 inhibitors, and the accompanying paper describes how these were optimized to a lead series of nanomolar inhibitors.¹

With this result in hand, we wondered why our HTS program had not delivered similar compounds, and we searched our corporate collection for compounds bearing the isocytosine substructure. We did find several such compounds, but they were either inappropriately substituted such that they could not bind to BACE-1 or were not large enough to afford sufficient potency to lit up in our HTS screen. This underpins the claim of FBLG that working with fragment-sized molecules allows for a more efficient sampling of compound space.

Experimental Section

Materials. All chemicals were purchased from Sigma-Aldrich or BIAcore AB (Uppsala, Sweden), unless stated otherwise. All the peptidic transition state isosteres used in the BIAcore assay and the NMR displacement experiments have been either produced in-house or obtained from California Peptide Research.

Protein Expression and Purification. Active BACE-1 protein was produced in HEK-293 cells as a C-terminal fusion protein with the Fc part of human IgG1 as described previously.³⁷ In brief, the BACE-1 Fc construct was produced in either batch or perfusion cultures of stably transfected HEK-293 cells to levels of about 3 mg/L and purified on rProtein A beads (GE Healthcare) utilizing expanded bed capturing. The final step included an ion-exchange chromatography step on Resource Q (GE Healthcare). Endothiapepsin was purified to homogeneity in two steps from Suparen liquid (Pfizer) according to a modified method³⁸ by affinity chromatography on pepstatin-agarose, followed by FPLC ion-exchange chromatography on Mono Q (GE Healthcare).

BIAcore Assay. All BIAcore studies have been performed at 25 °C using a BIAcore 3000 instrument (BIAcore AB). Coupling of a P₁ (S)-statin substituted substrate analogue (KTEEISEVN-statin-DAEF) to the CM5 surface was performed using the amine-coupling kit from BIAcore according to the instructions of the

manufacturer. In brief, the substrate analogue was dissolved to 0.2 mg/mL in a 20 mM Na-acetate buffer at pH 4.5 and clarified by centrifugation to remove particulates. The functional groups of the CM5 chip surface were first activated by injecting at 10 μ L/min for 10 min a freshly prepared 1:1 mixture of 0.5 M *N*-ethyl- *N'*-(3-dimethylaminopropyl)carbodiimide and 0.5 M *N*-hydroxysuccinimide, immediately followed by injecting the substrate analogue for 7 min at a flow rate of 10 μ L/min to 360 RU. Remaining activated carboxyl groups on the surface were blocked with 0.5 M ethanolamine. For internal referencing and to correct for unspecific binding during the binding experiments, a scrambled version of the substrate analogue (KFES-statin-ETIAEVENV) was coupled to a reference flow channel in a similar fashion. The renin-inhibitor H-142 (PHFHL^RVIHK, where R depicts the reduced isostere of the scissile peptide bond between residue Leu10 and Val11 in human angiotensinogen) was coupled to a CM5 surface to 410 RU using a similar immobilization procedure. The running buffer used in all subsequent binding and inhibition experiments was composed of 25 mM sodium acetate, 200 mM NaCl, 0.005% Tween 20, pH 4.5 with varying concentrations (1–5%) of DMSO, dependent on the concentration and solubility of the tested compound. For the analysis of FRIT binding to BACE-1 and EP, the flow-rate was 20 μ L/min and protein samples in running buffer in the presence of the FRITs were injected for 1–2 min over both the reference and the TDC flow channel. To remove bound protein and unspecific bound compounds from the biosensor surfaces the flow channels were regenerated by a 15 s pulse of 50 mM Tris/HCl at pH 8.5 and 0.5% SDS. We used the initial binding rate (dRU/dt) determined within the first 15 s after the injection start of the subtracted sensorgrams as a measure of the free protein concentration. As this approach is insensitive toward deviations of DMSO concentrations between the sample and the running buffer, as well as that the binding of a large protein to the sensor surface resulted in large responses (see Figure 3), a solvent correction as usually done in the course of a direct binding assay was not necessary. The calibration procedure included the injection of varying concentrations of BACE-1 or EP enzyme (0.01–1 μ M) in the absence of FRITs over the reference and TDC surfaces and plotting the initial binding rate, dRU/dt, against the protein concentration. For the primary characterization of FRIT binding, selected candidates at a concentration of 1 mM were preincubated for 10 min with either 0.5 μ M BACE-1 or 20 nM EP, and the free protein concentration was determined as described above. The same procedure was adopted for the determination of the K_D, but using varying concentrations of the selected FRITs (0.1–5 mM), dependent on the initial response at 1 mM and the solubility of the FRIT.

NMR Screening. NMR samples used for screening contained 3–5 μ M protein, in a buffer consisting of 50 mM sodium acetate pH 4.6, plus 20 μ M dTMSP for signal referencing. WaterLOGSY experiments^{17,39} were performed at 20 °C on a Bruker DRX600 instrument, equipped with a conventional ('warm') probehead. Typical interscan delay was 2 s, mixing time 1.4 s, and 640 scans were accumulated giving an acquisition time of about 40 min per experiment. In order to speed up the screening process, the compounds were screened in mixtures of four or six, at an individual concentration of 300 μ M. Compounds were stored as 100 mM DMSO-*d*₆ stock solutions on 384-well plates. A SampleRail system (Bruker) connected to a Genesis robot (Tecan) was used for automated, just-in-time sample preparation.⁴⁰ Each experiment was run in two steps. The first waterLOGSY experiment was run on a sample containing the protein and ligands. Then a strong inhibitor, usually OM99–2 (K_i of 1.6 nM),¹⁶ was added to the sample and a second experiment was recorded. This strong binder will displace any competitive ligands binding during the first experiment, and the difference between the first and second spectrum will therefore correctly identify competitive hit compounds. Spectra were processed and analyzed with VnmrX (Varian Inc.) supplemented with a suite of in-house written macros (Per-Olof Eriksson, unpublished). As ligand-observed NMR was used, there was no need for deconvolution of mixtures containing a hit. Nevertheless, we

decided to rerun hits as individual compounds, to eliminate the risk of compound–compound interactions interfering with the original observation.

For the work described in this paper, we screened about 2000 compounds from our general fragment library, with molecular mass in the 150–250 Da range. Stability and toxicity of functional groups have been taken into account when building the library, as well as the suitability of the compounds for downstream medicinal chemistry. As we need hundreds of micromolar compound solubility for some of our NMR screens, we used an upper limit of 2.5 for the calculated log *D* when selecting compounds to be incorporated into the library.⁷

Crystallization and Structure Determination. Apo crystals of endothiapepsin were grown at 20 °C using the hanging drop/vapor diffusion method, the drops containing 2 μL of protein solution (5 mg/mL) and 2 μL of mother liquor. The 500-μL reservoir solution consisted of 32.5% PEG 4000, 200 mM NH₄Ac, and 100 mM sodium acetate pH 4.5. Crystals were ready to use for soaking after 1–2 days, and they were soaked for approximately 18 h and cryoprotected at the same time in 9.5 μL cryosolution (27.5% PEG 4000, 200 mM NH₄Ac, 100 mM sodium acetate pH 4.6, 15% glycerol) + 0.5 μL compound solution (200 mM in DMSO), giving a final compound concentration of 10 mM in the soaking drop. Apo crystals could be stored for several years without loss of quality. Data from soaked crystals were typically collected to 1.55 Å resolution in-house at 100 K on a RU300 rotating copper anode (Rigaku/MSC) using a MAR345 image plate. The data was processed with MOSFLM⁴¹ and SCALA.⁴² The crystals belonged to space group *P*₂₁ with the cell dimensions for the compound **2** complex being *a* = 45.5 Å, *b* = 73.6 Å, *c* = 53.4 Å, and β = 110.3°. The structure was solved by molecular replacement using the program MOLREP⁴³ and the endothiapepsin coordinates with the PDB code 1GVU. The first rounds of refinement and model rebuilding were performed using CNX⁴⁴ and O.⁴⁵ Final refinement and model rebuilding were done using CCP4⁴² and Coot.⁴⁶ All figures illustrating structures were generated using PyMOL.⁴⁷ A summary of data collection and refinement statistics is presented in Table 2.

Acknowledgment. We thank Per-Olof Eriksson for development and supplying the spectrum analysis software.

References

- Edwards, P. D.; Albert, J. S.; Sylvester, M.; Aharony, D.; Andisik, D.; Campbell, J.; Chessari, G.; Congreve, M.; Folmer, R. H. A.; Geschwindner, S.; Koether, G.; Kolmodin, K.; Krumrine, J.; Mauger, R. C.; Olsson, L.-L.; Patel, S.; Spear, N.; Tian, G. Application of fragment-based lead generation to the discovery of novel, cyclic amidine β-secretase inhibitors with nanomolar potency, cellular activity and high ligand efficiency. *J. Med. Chem.* **2007**, *50*, 5912–5925.
- Hann, M. M.; Oprea, T. I. Pursuing the leadlikeness concept in pharmaceutical research. *Curr. Opin. Chem. Biol.* **2004**, *8*, 255–263.
- Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **1997**, *23*, 3–25.
- Teague, S. J.; Davis, A. M.; Leeson, P. D.; Oprea, T. The design of leadlike combinatorial libraries. *Angew. Chem., Int. Ed.* **1999**, *38*, 3743–3748.
- Leach, A. R.; Hann, M. M.; Burrows, J. N.; Griffen, E. J. Fragment screening: an introduction. *Mol. Biosyst.* **2006**, *2*, 430–446.
- Rees, D. C.; Congreve, M.; Murray, C. W.; Carr, R. Fragment-based lead discovery. *Nat. Rev. Drug Discovery* **2004**, *3*, 660–672.
- Albert, J. S.; Blomberg, N.; Breeze, A. L.; Brown, A. J. H.; Burrows, J. N.; Edwards, P. D.; Folmer, R. H. A.; Geschwindner, S.; Griffen, E. J.; Kenny, P. W.; Nowak, T.; Olsson, L.-L.; Sanganev, H.; Shapiro, A. B. An Integrated Approach to Fragment Based Lead Generation: Philosophy, Strategy and Case Studies from AstraZeneca's Drug Discovery Programmes. *Curr. Topics Med. Chem.* **2007**, *7*, 1600–1629.
- Hann, M. M.; Leach, A. R.; Harper, G. Molecular complexity and its impact on the probability of finding leads for drug discovery. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 856–864.
- Murray, C. W.; Verdonk, M. L. The consequences of translational and rotational entropy lost by small molecules on binding to proteins. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 741–753.
- Jencks, W. P. On the attribution and additivity of binding energies. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 4046–4050.
- 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.
- Hopkins, A. L.; Groom, C. R.; Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discovery Today* **2004**, *9*, 430–431.
- Moore, J. M. NMR screening in drug discovery. *Curr. Opin. Biotechnol.* **1999**, *10*, 54–58.
- Coles, M.; Heller, M.; Kessler, H. NMR-based screening technologies. *Drug Discovery Today* **2003**, *8*, 803–810.
- Hajduk, P. J.; Olejniczak, E. T.; Fesik, S. W. One-dimensional relaxation- and diffusion-edited NMR methods for screening compounds that bind to macromolecules. *J. Am. Chem. Soc.* **1997**, *119*, 12257–12261.
- Ermolieff, J.; Loy, J. A.; Koelsch, G.; Tang, J. Proteolytic activation of recombinant pro-memapsin 2 (pro-β-secretase) studied with new fluorogenic substrates. *Biochemistry* **2000**, *39*, 12450–12456.
- Dalvit, C.; Fogliatto, G.; Stewart, A.; Veronesi, M.; Stockman, B. WaterLOGSY as a method for primary NMR screening: practical aspects and range of applicability. *J. Biomol. NMR* **2001**, *21*, 349–359.
- Hajduk, P. J.; Huth, J. R.; Fesik, S. W. Druggability indices for protein targets derived from NMR-based screening data. *J. Med. Chem.* **2005**, *48*, 2518–2525.
- Congreve, M.; Aharony, D.; Albert, J.; Callaghan, O.; Campbell, J.; Carr, R. A. E.; Chessari, G.; Cowan, S.; Edwards, P. D.; Frederickson, M.; McMenamin, R.; Murray, C. W.; Patel, S.; Wallis, N. Application of fragment screening by X-ray crystallography to the discovery of aminopyridines as inhibitors of beta-secretase. *J. Med. Chem.* **2007**, *50*, 1124–1132.
- Murray, C. W.; Callaghan, O.; Chessari, G.; Cleasby, A.; Congreve, M.; Frederickson, M.; Hartshorn, M. J.; McMenamin, R.; Patel, S.; Wallis, N. Application of fragment screening by X-ray crystallography to beta-secretase. *J. Med. Chem.* **2007**, *50*, 1116–1123.
- Karlsson, R.; Kullman-Magnusson, M.; Hamalainen, M. D.; Remaues, A.; Andersson, K.; Borg, P.; Gyzander, E.; Deinum, J. Biosensor analysis of drug-target interactions: Direct and competitive binding assays for investigation of interactions between thrombin and thrombin inhibitors. *Anal. Biochem.* **2000**, *278*, 1–13.
- Huber, W. A new strategy for improved secondary screening and lead optimization using high-resolution SPR characterization of compound-target interactions. *J. Mol. Recogn.* **2005**, *18*, 273–281.
- Huber, W.; Mueller, F. Biomolecular interaction analysis in drug discovery using surface plasmon resonance technology. *Curr. Pharm. Design* **2006**, *12*, 3999–4021.
- McDonnell, J. M. Surface plasmon resonance: towards an understanding of the mechanisms of biological molecular recognition. *Curr. Opin. Chem. Biol.* **2001**, *5*, 572–577.
- Deinum, J.; Gustavsson, L.; Gyzander, E.; Kullman-Magnusson, M.; Edstrom, A.; Karlsson, R. A thermodynamic characterization of the binding of thrombin inhibitors to human thrombin, combining biosensor technology, stopped-flow spectrophotometry, and microcalorimetry. *Anal. Biochem.* **2002**, *300*, 152–162.
- Metz, G.; Otteleben, H.; Vetter, D. Small molecule screening on chemical microarrays. *Meth. Principles Med. Chem.* **2003**, *19*, 213–236.
- Neumann, T.; Junker, H.-D.; Keil, O.; Burkert, K.; Otteleben, H.; Gamer, J.; Sekul, R.; Deppe, H.; Feurer, A.; Tomandl, D.; Metz, G. Discovery of thrombin inhibitor fragments from chemical microarray screening. *Lett. Drug Des. Discovery* **2005**, *2*, 590–594.
- Papalia, G. A.; Leavitt, S.; Bynum, M. A.; Katsamba, P. S.; Wilton, R.; Qiu, H. W.; Steukers, M.; Wang, S. M.; Bindu, L.; Phogat, S.; Giannetti, A. M.; Ryan, T. E.; Pudlak, V. A.; Matusiewicz, K.; Michelson, K. M.; Nowakowski, A.; Pham-Baginski, A.; Brooks, J.; Tieman, B. C.; Bruce, B. D.; Vaughn, M.; Baksh, M.; Cho, Y. H.; De Wit, M.; Smets, A.; Vandersmissen, J.; Michiels, L.; Myszk, D. G. Comparative analysis of 10 small molecules binding to carbonic anhydrase II by different investigators using Biacore technology. *Anal. Biochem.* **2006**, *359*, 94–105.
- Sinha, S.; Anderson, J. P.; Barbour, R.; Basi, G. S.; Caccavello, R.; Davis, D.; Doan, M.; Dovey, H. F.; Frigon, N.; Hong, J.; Jacobson-Croak, K.; Jewett, N.; Keim, P.; Knops, J.; Lieberburg, I.; Power, M.; Tan, H.; Tatsuno, G.; Tung, J.; Schenk, D.; Seubert, P.; Suomensari, S. M.; Wang, S.; Walker, D.; Zhao, J.; McConlogue, L.; John, V. Purification and cloning of amyloid precursor protein [beta]-secretase from human brain. *Nature* **1999**, *402*, 540.

- (30) Webb, D. J.; Cumming, A. M.; Leckie, B. J.; Lever, A. F.; Morton, J. J.; Robertson, J. I.; Szelke, M.; Donovan, B. Reduction of blood pressure in man with H-142, a potent new renin inhibitor. *Lancet* **1983**, *2*, 1486–1487.
- (31) Hemmings, A. M.; Foundling, S. I.; Sibanda, B. L.; Wood, S. P.; Pearl, L. H.; Blundell, T. Energy calculations on aspartic proteinases: human renin, endothiapepsin and its complex with an angiotensinogen fragment analogue, H-142. *Biochem. Soc. Trans.* **1985**, *13*, 1036–1041.
- (32) Foundling, S. I.; Cooper, J.; Watson, F. E.; Cleasby, A.; Pearl, L. H.; Sibanda, B. L.; Hemmings, A.; Wood, S. P.; Blundell, T. L.; Valler, M. J.; et al. High resolution X-ray analyses of renin inhibitor-aspartic proteinase complexes. *Nature* **1987**, *327*, 349–352.
- (33) Hong, L.; Koelsch, G.; Lin, X. L.; Wu, S. L.; Terzyan, S.; Ghosh, A. K.; Zhang, X. C.; Tang, J. Structure of the protease domain of memapsin 2 (beta-secretase) complexed with inhibitor. *Science* **2000**, *290*, 150–153.
- (34) Lin, X. L.; Koelsch, C.; Wu, S. L.; Downs, D.; Dashti, A.; Tang, J. Human aspartic protease memapsin 2 cleaves the beta-secretase site of beta-amyloid precursor protein. *Proc. Natl Acad. Sci. U.S.A.* **2000**, *97*, 1456–1460.
- (35) Aqvist, J.; Medina, C.; Samuelsson, J. E. A new method for predicting binding affinity in computer-aided drug design. *Protein Eng.* **1994**, *7*, 385–391.
- (36) Hong, L.; Tang, J. Flap position of free memapsin 2 (beta-secretase), a model for flap opening in aspartic protease catalysis. *Biochemistry* **2004**, *43*, 4689–4695.
- (37) Lullau, E.; Kanttinen, A.; Hassel, J.; Berg, M.; Haag-Alvarsson, A.; Cederbrant, K.; Greenberg, B.; Fenge, C.; Schweikart, F. Comparison of batch and perfusion culture in combination with pilot-scale expanded bed purification for the production of soluble recombinant beta-secretase. *Biotechnol. Prog.* **2003**, *19*, 37–44.
- (38) Kobayashi, H.; Kusakabe, I.; Murakami, K. Rapid isolation of microbial milk-clotting enzymes by N-acetyl (or N-isobutyryl)-pepstatin-aminohexylagarose. *Anal. Biochem.* **1982**, *122*, 308–312.
- (39) Dalvit, C.; Pevarello, P.; Tato, M.; Veronesi, M.; Vulpetti, A.; Sundström, M. Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water. *J. Biomol. NMR* **2000**, *18*, 65–68.
- (40) Folmer, R. H. A.; Fetzer, J. Automated NMR screening in drug development. *BIOSpektrum* **2004**, *10*, 687–688.
- (41) Leslie, A. G. W. Integration of macromolecular diffraction data. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1999**, *55*, 1696–1702.
- (42) Collaborative Computational Project Number 4 The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1994**, *50*, 760–763.
- (43) Vagin, A.; Teplyakov, A. Molrep - an Automated Program for Molecular Replacement. *J. Appl. Crystallogr.* **1997**, *30*, 1022–1025.
- (44) Brunger, A. T.; Adams, P. D.; Clore, G. M.; Delano, W. L.; Gros, P.; Grossekunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. Crystallography and NMR System - a New Software Suite for Macromolecular Structure Determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1998**, *54*, 905–921.
- (45) Jones, T. A.; Zou, J. Y.; Cowan, S. W.; Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **1991**, *47*, 110–119.
- (46) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126–2132.
- (47) PyMol Molecular Graphics System: Delano, W. L. The PyMOL Molecular Graphics System, Delano Scientific, Palo Alto, CA.
- (48) Krissinel, E.; Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2256–2268.

JM070825K