Structural Locations and Functional Roles of New Subsites S_5, S_6, and S_7 in Memapsin 2 (β-Secretase)[†,‡]

Robert T. Turner, III,[§,||,⊥] Lin Hong,[§,||,⊥] Gerald Koelsch,[§,||] Arun K. Ghosh,[@] and Jordan Tang*[,§]

Protein Studies Program, Oklahoma Medical Research Foundation, Department of Biochemistry and Molecular Biology, University of Oklahoma Health Science Center, Oklahoma City, Oklahoma 73104, Zapaq Inc., Oklahoma City, Oklahoma 73104, and Department of Chemistry, University of Illinois, Chicago, Illinois 60607

Received September 2, 2004; Revised Manuscript Received October 22, 2004

ABSTRACT: Memapsin 2 (β-secretase) is the membrane-anchored aspartic protease that initiates the cleavage of β-amyloid precursor protein (APP), leading to the production of amyloid-β (Aβ), a major factor in the pathogenesis of Alzheimer's disease. The active site of memapsin 2 has been shown, with kinetic data and crystal structures, to bind to eight substrate residues (P_4-P_4'). We describe here that the addition of three substrate residues from P_7 to P_5 strongly influences the hydrolytic activity by memapsin 2 and these subsites prefer hydrophobic residues, especially tryptophan. A crystal structure of memapsin 2 complexed with a statine-based inhibitor spanning $P_{10}-P_4'$ revealed the binding positions of P_5-P_7 residues. Kinetic studies revealed that the addition of these substrate residues contributes to the decrease in K_m and increase in k_{cat} values, suggesting that these residues contribute to both substrate recognition and transition-state binding. The crystal structure of a new inhibitor, OM03-4 ($K_i = 0.03$ nM), bound to memapsin 2 revealed the interaction of a tryptophan with the S_6 subsite of the protease.

Memapsin 2 (*1*), also known as β-secretase, BACE (*2*), and ASP-2 (*3*), is the aspartic protease that initiates the cleavage of β-amyloid precursor protein (APP).[1] Following a second cut by γ-secretase, a peptide called amyloid-β (Aβ) is generated. The excess level of Aβ in the brain is regarded as an important factor in the pathogenesis of Alzheimer's disease (AD) (*4*). Therefore, memapsin 2 is an important target for the development of inhibitor drugs against AD. For this reason, the structure−function properties of this protease are of great interest at present.

Memapsin 2 is a single-chain class I transmembrane protein that contains an N-terminal ectodomain, a single-strand transmembrane domain, and a cytosolic domain (*1*−*3*, *5*, *6*). The cytosolic domain has been shown to mediate the trafficking of the protease in the endocytic and recycling pathways (*7*−*9*). The catalytic activity resides completely in the ectodomain, which is homologous to other aspartic proteases in overall folding and catalytic apparatus (*10*). The crystal structures of memapsin 2 (*10*−*12*) show that the catalytic unit is folded into the N-terminal and C-terminal lobes with the substrate binding site in a long cleft between the lobes, and the open position of the "flap" may contribute to the substrate specificity. In two crystal structures of enzyme−inhibitor complexes (*10*, *11*), the positions of eight side chains of the inhibitors also define the locations of eight subsites (from S_4 to S_4') in the protease. Although the overall arrangement of these subsites is similar to those present in all other mammalian aspartic proteases of the pepsin family, the subsite specificity is quite different. The complete residue specificity of each of these subsites has been determined (*13*). Together, these results provide insight to the structural basis of memapsin 2 activity in familial AD of the APP Swedish mutation and also provide a template for the rational design of memapsin 2 inhibitors (*14*, *15*).

Several lines of evidence suggest that memapsin 2 may contain additional subsites on the N-terminal side of the substrate beyond subsite S_4. When compared to the structure of other mammalian aspartic proteases, the crystal structure of memapsin 2 (*10*, *11*) contains several extra loops, comprising an area that is structurally unique among aspartic proteases and can house extra subsites extending from S_4. Moreover, additional binding beyond subsite P_4 is corroborated since the inclusion of substrate residues extending beyond P_4 increased the hydrolytic activity of memapsin 2 (*16*) and the potency of statine-based inhibitors (*17*). Here we describe the kinetic and structural evidence for the presence of subsites S_5, S_6, and S_7 in memapsin 2 and the roles played by these subsites in the catalytic activity of memapsin 2.

EXPERIMENTAL PROCEDURES

Design of Substrate Templates and Extended Peptide Substrates. The peptide sequence EVNL/AAEF, successfully utilized in the previous residue preference analysis for memapsin 2 (*13*), was used as the base template peptide for the analysis of the N-terminal extension effect. For the initial

[†] This study was supported in part by the National Institutes of Health (Grant AG-18933) and Alzheimer's Association Pioneer Award to J.T.

[‡] Atomic coordinates for the crystal structures described here have been deposited in the Protein Data Bank as entries 1XN2 and 1XN3.

* To whom correspondence should be addressed. E-mail: Jordan-Tang@omrf.ouhsc.edu. Phone: (405) 271-7291. Fax: (405) 271-7249.

[§] Oklahoma Medical Research Foundation and University of Oklahoma Health Science Center.

[||] Zapaq Inc.

[⊥] Co-first authors.

[@] University of Illinois.

[1] Abbreviations: APP, β-amyloid precursor protein; Aβ, amyloid-β; AD, Alzheimer's disease; rms, root-mean-square; PEG, polyethylene glycol.

10.1021/bi048106k CCC: $30.25
Published on Web 12/13/2004

series of analyses, three peptides were created using solid-phase peptide synthesis (Research Genetics, Invitrogen, Huntsville, AL). These peptides [EVNL/AAEF*WHDR* (designated WHDR), *RWHH*EVNL/AAEF (designated RWHH), and *REIS*EVNL/AAEF (designated REIS), in which the slant denotes the cleavage site in each peptide] incorporated four-residue sequence extension WHDR from the C-terminus, the sequence RWHH from the N-terminus, and the native APP sequence extension REIS from the N-terminus, respectively. Additionally, three peptide mixtures were synthesized on the basis of the extended native APP sequence (RTEEIXEVNL/AAEF for P_5, RTEEXSEVNL/AAEF for P_6, and RTEXISEVNL/AAEF for P_7, where X denotes a mixture of nine amino acid residues at that position) to examine the residue preference of the three upstream binding sites. To facilitate MALDI-TOF detection, an arginine was added to the N-terminus of the peptides. For each of the three subsites, these peptides were created through solid-phase peptide synthesis with equimolar amounts of a mixture of nine amino acids (see Figure 4 for a list of these amino acids) added at the appropriate cycle of the synthesis (*13*). The resulting mixture of nine peptides differed by only one amino acid at a single subsite. The amino acid corresponding to the native APP sequence substrate was included in each mixture to serve as an internal standard.

Kinetic Analysis Using MALDI-TOF Mass Spectroscopy. Substrate mixtures described above were prepared so that a reaction mixture with peptide in the micromolar range at pH 4.0 could be obtained. The reactions were allowed to proceed for 60 min with aliquots removed periodically. Aliquots were mixed with an equal volume of MALDI matrix (α-hydroxycinnamic acid in acetone), and immediately spotted on a 96 dual-well Teflon-coated analysis plate. The MALDI data collection and analysis were performed on a PE Biosystems Voyager DE instrument as previously described (*13*). Data were analyzed using the Voyager Data Explorer module to obtain ion intensity values. The relative product produced per unit time was obtained from nonlinear regression analysis of the data representing the initial 15% of product formation, and these data were used to determine the relative k_{cat}/K_m values (*13*).

Crystallization. Cocrystals of the catalytic domain of memapsin 2 (*10*) bound to inhibitor $P_{10}-P_4'$StatVal (*6*) [structure, Lys-Thr-Glu-Glu-Ile-Ser-Glu-Val-Asn-(statine)-Val-Ala-Glu-Phe; $K_i = 40$ nM] purchased from ICN Pharmaceuticals, Inc., were grown at 17 °C by the hanging-drop method, in which 5 μL of the 25 mg/mL protein solution was mixed with 5 μL of a reservoir solution containing 18% PEG 8000 (pH 6.5) with sodium cacodylate buffer. Because of the poor inhibitor solubility, we could only achieve an inhibitor:enzyme molar ratio of 0.8:1 in the crystal trials. Good-quality crystals appeared in 2−3 weeks.

Inhibitor OM03-4 (structure, Arg-Glu-Trp-Trp-Ser-Glu-Val-Asn-Leu*Ala-Ala-Glu-Phe, where the asterisk represents the hydroxylethylene isostere) was synthesized with the protected Leu*Ala isostere using a previously described procedure (*18*). Cocrystals of OM03-4 bound to memapsin 2 were obtained under conditions similar to those described above, but with a protein concentration of 20 mg/mL, at pH 7.0, and 16% PEG 8000. The inhibitor:enzyme molar ratio was 2:1 during crystal growth.

Diffraction Data Measurement and Structure Determination. For cryoprotection, crystals were soaked for a few minutes in crystallization solutions containing 20% (v/v) glycerol. The crystals were then mounted on a fiber loop and flash-frozen in a cryo-nitrogen gas stream. For the $P_{10}-P_4'$StatVal−menapsin 2 structure, the diffraction data were collected to a resolution of 2.0 Å under cryogenic conditions (100 K) on a Mar 345 image plate mounted on an Msc-Rigaku RU-300 X-ray generator with Osmic focusing mirrors. Diffraction data for the OM03-4−memapsin 2 complex were collected to 1.9 Å resolution at beamline A1 of Cornell High Energy Synchrotron Source. Both data sets were processed with DENZO and SCALEPACK (*19*). Initial phases were determined for both structures by molecular replacement using AmoRe (*20*) with a resolution range of 15−3.5 Å. The atomic coordinates of a single memapsin 2 molecule bound to a peptidic inhibitor OM00-3 (*11*) were used as the search model (excluding the inhibitor and the solvent atoms). Model phases were then calculated and used for initial $2F_o - F_c$ and $F_o - F_c$ electron density maps. The inhibitor was clearly visible, and its structure was built into the electron density map. Crystallographic refinement was carried out with CNS (*21*) using simulated annealing procedures. Eight percent of the diffraction data were excluded from the refinement at the beginning of the process to monitor the R_{free} values. Iterative model building was carried out with O (*22*). The final model of the structure includes all residues for the four protein molecules in the asymmetric unit. Among those, only one memapsin 2 molecule was found to bind to a $P_{10}-P_4'$StatVal inhibitor, whereas all four protein molecules were found to bind to an OM03-4 inhibitor. Data and refinement statistics are given in Table 1.

RESULTS

Effects of Substrate Extension on the Hydrolytic Activity of Memapsin 2. In the investigation of memapsin 2 subsite specificity (*13*), we employed a substrate sequence of ELDL/AVEF. The sequences of the substrates were extended either at the N-terminus (RWHH) or at the C-terminus (WHDR) to facilitate the identification of hydrolytic products in MALDI-TOF. While the C-terminal extension had little effect on the relative hydrolytic rate from the native parental peptide (1.34×10^{-3} min^{-1}), the substrate with an N-terminal extension displayed a 50−60-fold decrease in the level of hydrolysis (0.019×10^{-3} min^{-1}). This inhibitory effect was alleviated through replacement of the N-terminal extension with the native APP sequence (REIS), yielding a relative cleavage rate of 1.04×10^{-3} min^{-1}. These results suggested that additional subsites may be present at the N-terminal side of the substrate but not at the C-terminal side. An examination of the crystal structure of memapsin 2 (*10*, *11*) suggested that the active site cleft can indeed accommodate at least three additional residues at the N-terminus, i.e., P_5-P_7 (results not shown). To define these additional binding sites in greater detail and aid the inhibitor drug design for memapsin 2, we initiated structural and kinetic studies of these new subsites.

Structure of Subsites S_5-S_7. To identify the locations of the new subsites on memapsin 2, we determined the crystal structure of memapsin 2 catalytic domain complexed to a long inhibitor called P10−P4′StatVal (*6*) ($K_i = 40$ nM). The

Table 1: Data Collection and Refinement Statistics

	P10−P4′StatVal−memapsin 2	OM03-4−memapsin 2
space group	$P2_1$	$P2_1$
unit cell		
a, *b*, *c* (Å)	86.2, 130.9, 88.6	86.2, 130.8, 88.7
α, *β*, *γ* (deg)	90, 97.3, 90	90, 97.7, 90
resolution (Å)	50.0−2.0	30−1.9
no. of observed reflections	500659	560667
no. of unique reflections	130092	148428
R_{merge}[a]	0.079	0.079
	0.384 (2.07−2.0 Å)	0.433 (1.97−1.90 Å)
data completeness (%)	99.7 (50.0−2.0 Å)	99.9 (30.0−2.0 Å)
	99.2 (2.07−2.0 Å)	99.4 (1.97−1.9 Å)
$I/\sigma(I)$	14.5 (50.0−2.0 Å)	17.5 (30.0−2.0 Å)
	3.3 (2.07−2.0 Å)	3.1 (1.97−1.90 Å)
$R_{working}$[b]	0.202	0.194
R_{free}[b]	0.237	0.226
rms deviation from ideal values		
bond lengths (Å)	0.01	0.01
bond angles (deg)	1.6	1.6
no. of water molecules	744	723
average *B*-factor (Å^2)		
protein	25.5	26.4
solvent	27.5	29.3

[a] $R_{merge} = \sum_{hkl}\sum_i|I_{hkl,i} - \langle I_{hkl}\rangle|/\sum_{hkl}\langle I_{hkl}\rangle$, where $I_{hkl,i}$ is the intensity of the *i*th measurement and $\langle I_{hkl}\rangle$ is the weighted mean of all measurements of I_{hkl}. [b] R_{work} and $R_{free} = \sum||F_o| - |F_c||/\sum|F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively. Reflections with an $F_o/\sigma(F_o)$ of ≥0.0 are included in the refinement and *R*-factor calculation.

structure of this inhibitor is Lys-Thr-Glu-Glu-Ile-Ser-Glu-Val-Asn-(statine)-Val-Ala-Glu-Phe. Since statine is a mimic of the transition state of aspartic protease catalysis (*23*), it usually occupies both S_1 and S_1' subsites. Thus, the six N-terminal residues of the inhibitor represent subsites P_{10}−P_5, respectively. The crystal of the memapsin 2−inhibitor complex is monoclinic with the $P2_1$ space group. The structure was determined at 2.0 Å resolution (Table 1) using the molecular replacement method. There are four molecules in the asymmetric unit, but only one memapsin 2 molecule was observed to have a bound inhibitor. Although the inhibitor has a K_i value of 40 nM, it could not reach an equal molar concentration with the protease (see Experimental Procedures) because of poor solubility. Hence, there is ∼55−65% unbound enzyme present in the cryatallization solution. Upon examining the crystal packing and modeling the inhibitor into the unbound enzyme molecule in the asymmetric unit, we found that the amino acid residues at the N-terminus of the inhibitor (from P_8 to P_{10}) would experience steric interference in the crystal packing of another memapsin 2 molecule or its bound inhibitor within the same asymmetric unit. These are likely to be the reasons that a bound inhibitor was only observed for one protein molecule among the four molecules in the asymmetric unit.

In the crystal structure of the P10−P4′StatVal−memapsin 2 complex (Figure 1), the protein part is very similar to the structure of memapsin 2 bound to inhibitors OM99-2 (*10*) or OM00-3 (*11*) with a main chain rms deviation of 0.6 or 0.7 Å, respectively. The calculated electron density clearly defined the inhibitor structure from P_7 to P_4' [Glu-Ile-Ser-Glu-Val-Asn-(statine)-Val-Ala-Glu] (Figure 2). The first three residues (P_{10}−P_8, Lys-Thr-Glu) from the N-terminus and the P_5' Phe side chain at the C-terminus of the inhibitor have no definable electron density. Current and previous structures (*10*, *11*) clearly show that both P_7 and P_4' are situated at the ends of the binding cleft. Therefore, the residues beyond these two subsites are likely without significant interactions with the enzyme, and thus mobile. The visible part of the inhibitor is in an extended conformation. The inhibitor backbone from P4 to P4′ has essentially the same configuration as that of inhibitors OM99-2 and OM00-3 bound to memapsin 2 (*10*). As predicted, statine occupies both P_1 and P_1' and its binding position is similar to that previously reported for the inhibitor complex of other aspartic proteases (Protein Data Bank entry 1LYB). In the absence of a side chain at P_1', the S_1' pocket is filled with several stabilized water molecules that interact with one another and form hydrogen bonds to the enzyme residues as reported previously (*10*, *11*). The conformation and the interaction of previously observed side chains of Glu at P_4, Val at P_3, Asn at P_2, and Val at P_2' are essentially the same

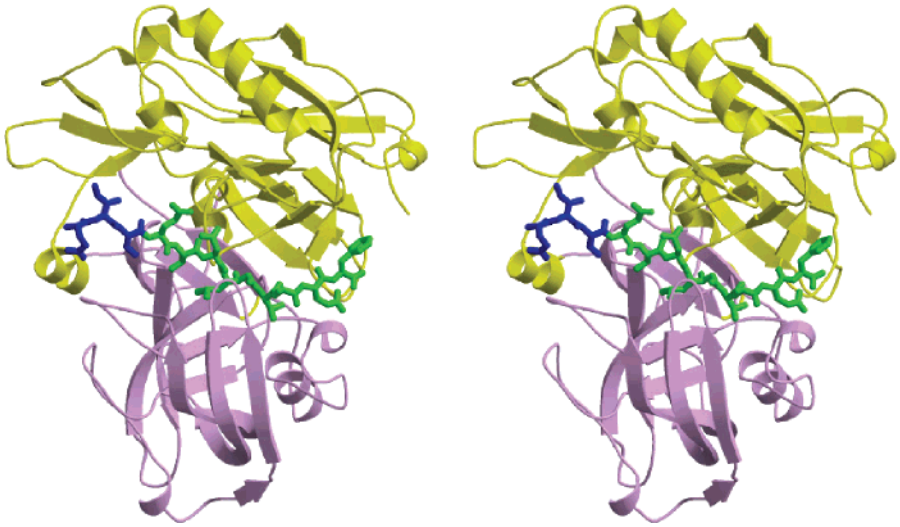

FIGURE 1: Stereopair presentation of inhibitor P10−P4′StatVal bound to the catalytic unit of memapsin 2. The N-terminal domain of the enzyme is colored pink, and the C-terminal domain is colored yellow as a ribbon diagram. The inhibitor, represented by the stick model, is colored green from P_4 to P_4' and blue from P_7 to P_5.

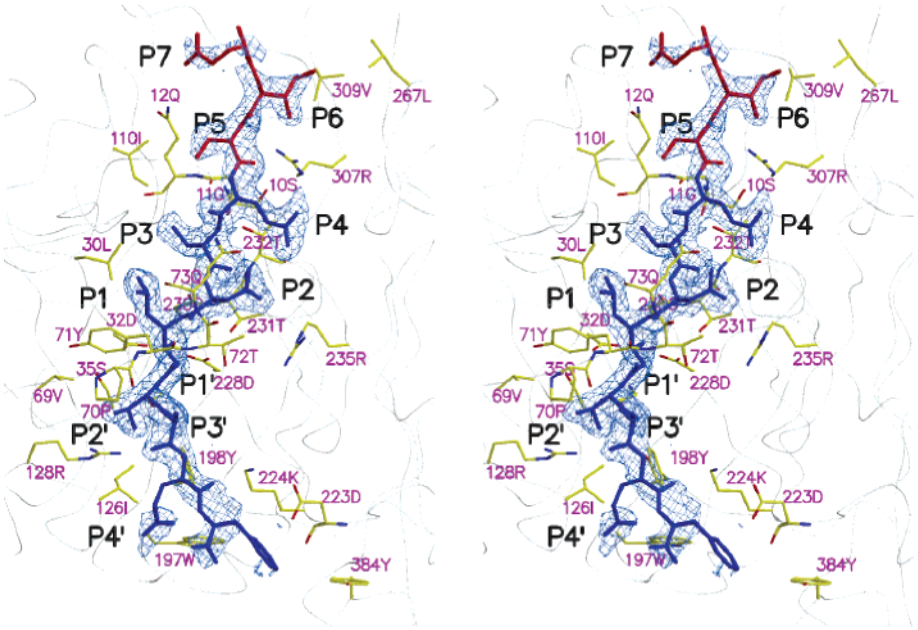

FIGURE 2: Active site interactions of inhibitor P10−P4′StatVal with memapsin 2. The inhibitor is represented by a stick model where the region from P_4 to P_4' is colored blue and from P_7 to P_5 red. Contacting residues of memapsin 2 are colored yellow. The electron density is contoured at the 1σ level using the $2F_o - F_c$ map.

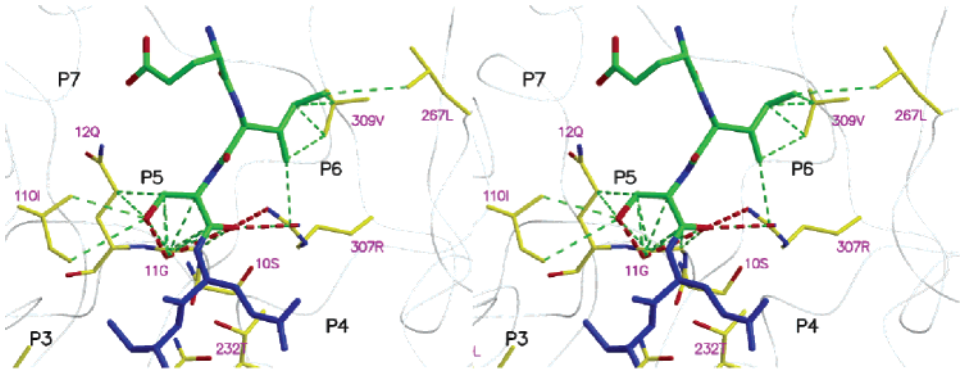

FIGURE 3: Close-up view of new subsites P_7−P_5 of P10−P4′StatVal. Subsites P_7−P_5 are colored green. The contacting residues of memapsin 2 (<4 Å) are depicted as a yellow stick model. van del Waals contacts between the inhibitor upstream structure and memapsin 2 are represented by green dashed lines and hydrogen bonds by red dashed lines.

as described previously (*10*, *11*). The current P_3' and P_4' residues, Ala and Glu, respectively, take the positions corresponding to Glu and Phe in inhibitor OM00-3 reported previously (*11*).

The conformation of P_5−P_7 of the inhibitor is essentially a direct extension of the extended structure of the rest of the inhibitor (Figures 1 and 2). The Cα atom of the side chain of Ser at P_5 makes van der Waals contacts with Ile 110, Gln 12, and Gly 11 of the enzyme (Figure 3). The carbonyl oxygen of Ser at P_5 contacts the backbone atoms of Gly 11 and hydrogen bonds to both granido nitrogen atoms of Arg 307. Side chain atoms of Ile at P_6 contact memapsin 2 residues Leu 267, Arg 307, and Val 309. Glu at P_7 does not interact with the enzyme atoms nearby. It produces a higher *B*-factor and loosely defined electron density as compared to the rest of the inhibitor residues. It is possible that the absence of significant interaction at P_7 is due to the fact that the Glu side chain is not favored in the S_7 binding pocket (*13*). Despite this, the structural evidence from subsites P_5−P_7 suggests that they are part of the substrate recognition for memapsin 2.

Residue Preference on Subsites P_5−P_7. To further define the activity of new subsites P_5−P_7, we determined the residue preference of representative amino acids in these sites. We employed a previously devised method (*13*) in which relative rates of peptides in a mixture were determined using MALDI-TOF mass spectrometry. Peptide mixtures having the basic sequence RTEEISEVNL/AAEF (in which the slant denotes the cleavage site) were synthesized. The substrate sequence from P_4 to P_4' (EVNLAAEF) contains memapsin 2-preferred residues (*13*). Residues at P_5−P_9 (TEEIS) were taken from the corresponding residues at the β-site of APP, while the arginine at P_{10} was chosen to facilitate analysis using a mass spectrometer. In each substrate group, one of the positions in the sequence of EIS, representing subsites P_7−P_5, contained a mixture of representative residues, including A, E, G, H, S, T, R, W, and Y. The relative rates of hydrolysis of the peptides in each of the three groups were determined using a MALDI-TOF mass spectrometry method as previously described (*13*). The resulting substrate side chain preferences, reported as the preference index (essentially relative k_{cat}/K_m), for these three positions are

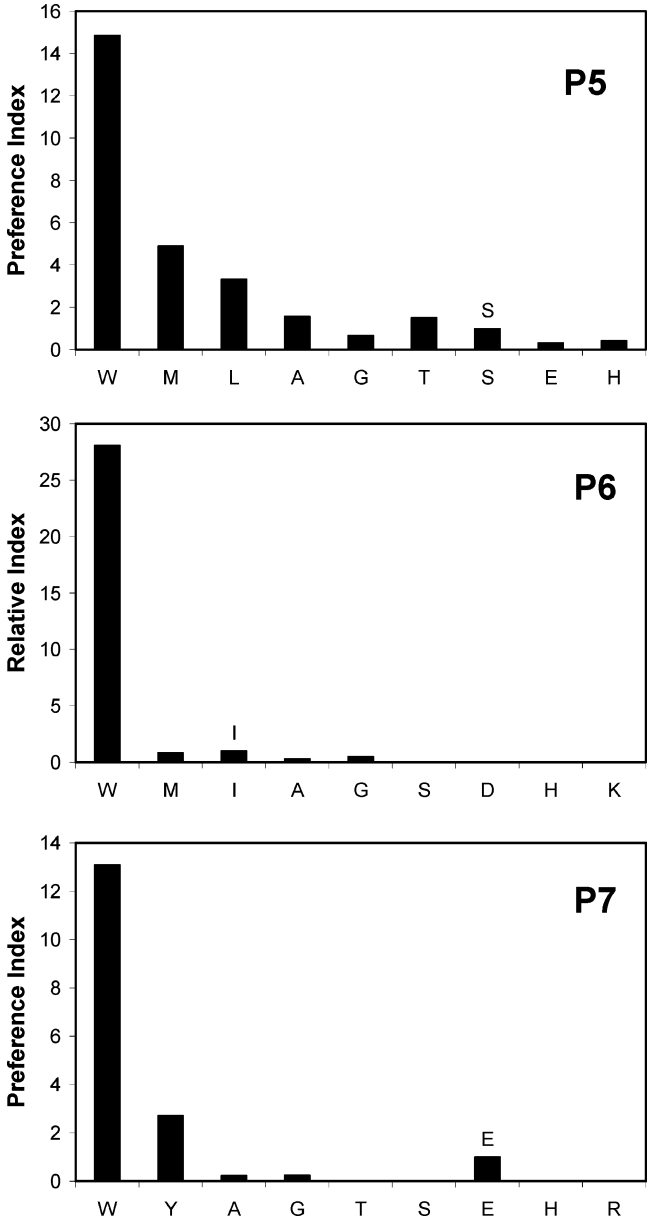

FIGURE 4: Amino acid residue preference for three new subsites (S_5-S_7) of memapsin 2. The preference index was calculated from the relative initial hydrolytic rates of mixed substrates and is proportional to the relative k_{cat}/K_m. The amino acids listed appear in the position designated in each panel. The residues labeled above the bar indicate the residues found in the native APP sequence.

presented in Figure 4. Interestingly, Trp appears to be the most preferred residue in all three sites, with Tyr and Met also demonstrating improved catalytic efficiencies. The position with the greatest observed effect is clearly the P_6 subsite, with the Trp substitution producing a 50-fold increase in catalytic efficiency over the native APP residue, Ile. When His was present at either P_5 or P_6, no detectable hydrolytic product was found. This observation suggests that histidine is strongly unfavorable at these two subsites and gave an explanation of why the addition of an N-terminal sequence (RWHH) decreased the hydrolytic rate (see above). These results suggest that subsites P_5-P_7 exhibit residue preference and they prefer in general the hydrophobic residues, especially Trp.

Influence of Subsite P_5 and P_6 Residues on the Hydrolytic Activity of Memapsin 2. To examine the influence of these new subsites on the catalysis of memapsin 2, we compared kinetic data of substrates with different P_6 and P_7 residues. Substrate Aswe was taken from the β-site sequence of APPswedish, RTEEISEVNL/AAEF (in which the slant denotes the cleavage site). Arg at P_8 is included to assist the determination of the hydrolytic rate by MALDI-TOF. Another substrate, Aswe-WW (RTEWWSEVNL/AAEF), has both Glu at P_6 and Ile at P_7 changed to Trp, the most favorable residue for these positions. The kinetic parameters of these two substrates are shown in Table 2. The replacement of P_6 and P_7 residues with tryptophans caused a decrease in the K_m of ~3-fold and an increase in k_{cat} of ~4-fold and an increase in k_{cat}/K_m of ~14-fold.

Table 2: Kinetic Parameters for Memapsin 2 Substrates and Inhibitors

Kinetic Constants of the Extended Trp-Containing Substrate

peptide	sequence[a]	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M ($min^{-1}\,\mu M^{-1}$)
Aswe-WW	RTEWWSEVNL/AVEF	15.55	24.248	0.87
Aswe	RTEEISEVNL/AAEF	3.75	82.6	0.045

Inhibition Constants of the Extended Memapsin 2 Inhibitors

inhibitor	sequence[b]	K_i (nM)
OM99-2	EVNL*AAEF	1.6
OM00-3	ELDL*AVEF	0.31
OM03-4	REWWSEVNL*AVEF	0.03

[a] The cleavage site is denoted with a slant. [b] The isostere position is denoted with an asterisk.

Transition-State Inhibitor OM03-4 with Trp in P_6 and P_7. The kinetic data mentioned above suggest that the inclusion of Trp in subsites P_6 and P_7 would enhance the binding of a transition-state inhibitor of memapsin 2. We therefore designed and synthesized a 14-residue inhibitor, OM03-4, with the primary structure Arg-Glu-Trp-Trp-Ser-Glu-Val-Asn-L*A-Ala-Glu-Phe (in which L*A represents Leu and Ala linked by a hydroxylethylene isostere) with a tryptophan residue at P_6 and P_7. Inhibitor OM03-4 has a K_i value of 0.03 nM (Table 2, bottom section) as compared to K_i value of 0.3 nM for the specificity-optimized eight-residue inhibitor OM00-3 (*13*). This comparison confirmed that subsites P_5-P_7 contributed to the binding of a transition-state analogue.

Crystal Structure of OM03-4. To define the binding of Trp residues in the active site of memapsin 2, the crystal structure of the OM03-4−memapsin 2 complex was determined at 1.9 Å resolution (Table 1). This crystal form was also monoclinic with the $P2_1$ space group. The structure of OM03-4 bound to memapsin2 was clearly defined from P_6 to P_4' (Trp-Ser-Glu-Val-Asn-L*A-Ala-Glu-Phe) by electron density (Figures 5 and 6). Unlike the case of inhibitor P10−P4′StatVal, a bound OM03-4 was observed for each of the four memapsin 2 molecules in the asymmetric unit. The conformations and the interactions of side chains from Glu at P_4 to Phe at P_4' are essentially the same as those of OM99-2 and OM00-3 described previously (*10*, *11*). As in the structure of P10−P4′StatVal, the first two residues (P_9 and P_8), Arg and Glu, respectively, in OM03-4 are not visible while the electron density for P_7 is weak (Figure 6). Also similar are the interactions of Ser at P_5 with the enzyme (see above). Trp at P_6 has extensive interactions with memapsin 2 residues Gln 265, Glu 266, Leu 267, Arg 307, and Lys

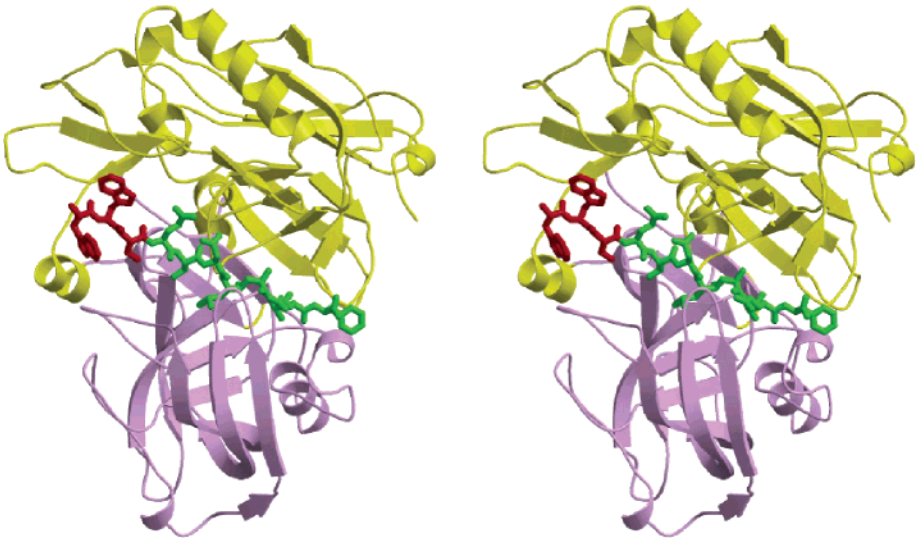

FIGURE 5: Stereopair presentation of inhibitor OM03-2 bound to memapsin 2. The N-terminal domain of the enzyme is colored pink and the C-terminal domain yellow, depicted as a ribbon diagram. The inhibitor, represented by the stick model, is colored green from P_4 to P_4' and red from P_7 to P_5.

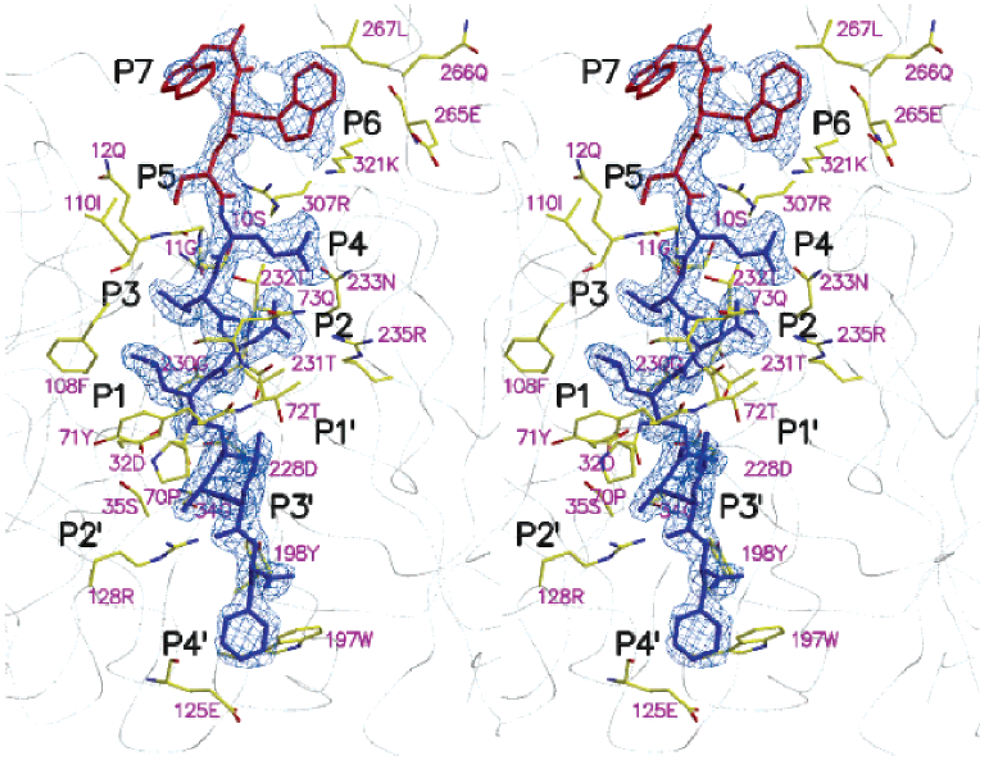

FIGURE 6: Stereoview of the detailed interactions between inhibitor OM03-2 and the active site cleft of memapsin 2. The inhibitor is represented by a stick model with P_4-P_4' colored red and P_7-P_5 blue. Contacting residues of memapsin 2 are colored yellow. The electron density is contoured at the 1σ level using the $2F_o - F_c$ map.

321 (Figure 7). These contacts lead to significant backbone positional changes at residues 265 and 266 and cause the side chains of Leu 267 and Lys 321 to adopt different rotomer positions (Figure 7). These structural changes in the enzyme are necessary to accommodate the large and hydrophobic Trp side chain at P_6. Two hydrogen bonds are formed to the new position of the amino nitrogen of Lys 321, one from the indole nitrogen of Trp at P_6 and another from one of the carboxyl atoms of Glu at P_4 with bond distances of 2.5 and 2.6 Å, respectively (Figure 7). A positional shift of 2.5 Å at the P_4 Glu side chain toward that of Lys 321 takes place prior to fromation of the hydrogen bond (Figure 7). These interactions correlate with the greatly improved inhibitor affinity for a Trp at P_6. The Trp at P_7 has higher *B*-factors and less well defined electron density. It appears to make several van der Waals contacts with Asn 111.

DISCUSSION

Current kinetic data and crystal structures suggest that the active site of memapsin 2 interacts with 11 substrate residues, from P_7 to P_4'. Eight of these subsites, P_4-P_4', had been well defined previously for specificity (*13*) and structural interac-

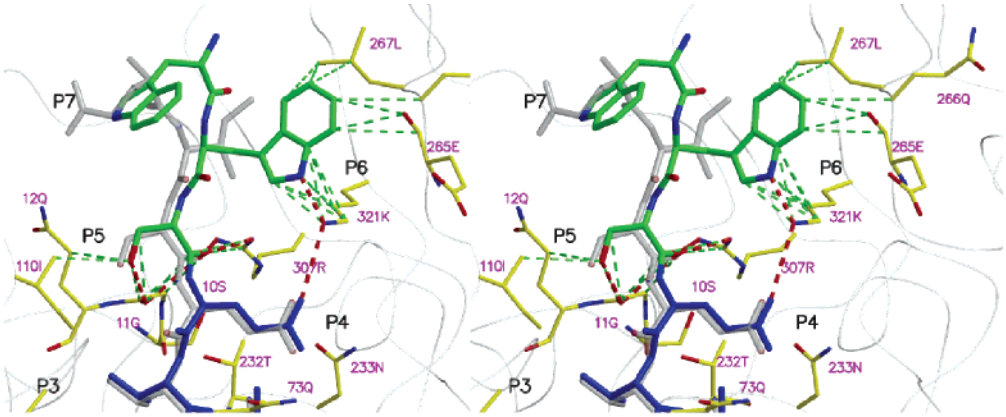

FIGURE 7: Close-up view of the binding mode of new subsites P_7-P_5 (green) of inhibitor OM03-4. The contacting residues of memapsin 2 (<4 Å) are represented as a yellow stick model. van del Waals contacts between the inhibitor upstream structure and memapsin 2 are represented by green dashed lines and hydrogen bonds by red dashed lines. The structure of P10−P4′StatVal as the gray stick model is superimposed onto that of OM03-2.

tion with memapsin 2 (*10*, *11*). Observations described above support the conclusion that the active site of memapsin 2 contains three additional subsites, P_7-P_5. These new subsites are unique among mammalian aspartic proteases, with the possible exception of memapsin 1 (BACE 2) which is highly similar to memapsin 2 in both amino acid sequence (*1*) and specificity (*13*).

Crystal structures revealed that the locations of the new subsites (S_7-S_5) are in the part of the active site cleft which extends directly from the eight previously known subsites. This is a region (residues 158−167) in the vicinity of an insertion helix (*10*), a unique region for memapsin 2 not seen in the crystal structures of other aspartic proteases. Interestingly, for the P10−P4′StatVal inhibitor, side chains P_7-P_5 interact exclusively with protein side chains and are devoid of any interaction with the peptide backbone. Since the structure of this area of the protease comprises several surface loops, there may be considerable flexibility in the observed interactions stemming from the relative ease of the protease to reorient its interacting side chains or even the backbone. This flexibility may contribute to the residue preference observed for subsites P_7-P_5.

Clearly, residues in these three subsites have a strong influence on the overall catalytic efficiency. Kinetic data indicate that the hydrophobic residues, especially tryptophan, are preferred in all three positions. The structural basis of this preference at P_6 is in clear agreement with the extensive interaction of the Trp side chain with residues in S_6, as seen in the crystal structure of the memapsin 2−OM03-4 complex (Figure 7). The structural basis for the preference of Trp at P_7 or P_5 is, however, less obvious. Kinetic studies suggest favorable binding of tryptophan residues in S_7 and S_5 pockets. To answer this question, we built a model based on the crystal structure of the OM03-4−memapsin 2 complex in which the Ser at P_5 was changed to a Trp (not shown). Our model demonstrates that a Trp at P_5 would be able to form hydrophobic interactions with Ile 110 in the S_5 pocket. The indole nitrogen of the P_5 residue has the potential to from a hydrogen bond to the amide oxygen of the Gln 12 side chain. The side chain of Trp at P_7 in OM03-4 has only limited interaction with the enzyme. Our model indicates that a large Trp side chain in P_7 can make extensive interactions with a hydrophobic side chain at P_5. The additional flexibility due to the lack of binding with the enzyme at P_7 should facilitate the optimization of this hydrophobic interaction. This type of binding would explain why both P_7 and P_5 subsites prefer Trp and hydrophobic residues (Figure 4).

It is also clear that some charged residues negatively impact the catalytic efficiency. This is exemplified by the diminished hydrolytic activity when histidine is in one of these three positions (Figures 4 and 7). The His side chain at P_6 would come very close to Lys 321 and Arg 307. At the assay pH of memapsin 2, the His side chain would be at least partially protonated and, therefore, unfavorably positioned. The same unfavorable situation can be extended to all basic residues, and together, their presence constitutes a virtual "veto power" for the hydrolytic activity (Figure 4). The decrease in K_m and the increase in k_{cat}, when the residues are favorable in these subsites (Table 2), indicate that the new subsites are involved in substrate recognition as well as the efficiency of transition-state binding. Since these three subsites are farther from the flap, which has been shown to produce a restricted access to the active site (*12*), it seems possible that these subsites are involved in substrate recognition at an early stage of the catalytic cycle.

Given that the residues in subsites P_7-P_5 contribute significantly to the catalysis by memapsin 2, these subsites likely contribute to the cellular activity of this protease in general. The β-secretase cleavage site of APP does not contain the most preferred residues in these three subsites. It would be interesting to determine if some other physiological substrates of memapsin 2 incorporate more favorable residues in these subsites. Another intriguing possibility is that genetic mutation in these residues of APP may accelerate the production of Aβ and thus manifest the familial type of AD. However, mutations of these subsites from Ser, Ile, and Glu to Trp would require two or three codon changes at each position and, therefore, would be very rare.

The enhancing of inhibitor binding potency by the inclusion of subsites P_7-P_5 residues raises the interesting question of whether the newly observed binding area on memapsin 2 can be incorporated in the design of new inhibitor drugs. The new sites are, however, distant from the transition-state isostere. The inclusion of structures at P_6 and P_5, for example, would likely create inhibitors that are too bulky for penetration of the blood−brain barrier. On the other hand, the cavity around subsites P_7-P_5 appears to be flexible and receptive to ligand interactions. Thus, it seems possible that non-

transition-state inhibitors can be designed for this new region to compete with the substrate APP.

ACKNOWLEDGMENT

We thank Drs. J. A. Hartsuck for critical reading of this paper and J. Yohannan and V. Weerasena for memapsin 2 preparation.

REFERENCES

1. Lin, X., Koelsch, G., Wu, S., Downs, D., Dashti, A., and Tang, J. (2000) Human aspartic protease memapsin 2 cleaves the β-secretase site of β-amyloid precursor protein, *Proc. Natl. Acad. Sci. U.S.A. 97*, 1456−1460.
2. 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. C., Collins, F., Treanor, J., Rogers, G., and Citron, M. (1999) β-Secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE, *Science 286*, 735−741.
3. Hussain, I., Powell, D., Howlett, D. R., Tew, D. G., Meek, T. D., Chapman, C., Gloger, I. S., Murphy, K. E., Southan, C. D., Ryan, D. M., Smith, T. S., Simmons, D. L., Walsh, F. S., Dingwall, C., and Christie, G. (1999) Identification of a novel aspartic protease (Asp 2) as β-secretase, *Mol. Cell Neurosci. 14*, 419−427.
4. Selkoe, D. J. (2001) Alzheimer's disease: Genes, proteins, and therapy, *Physiol. Rev. 81*, 741−766.
5. Yan, R., Bienkowski, M. J., Shuck, M. E., Miao, H., Tory, M. C., Pauley, A. M., Brashier, J. R., Stratman, N. C., Mathews, W. R., Buhl, A. E., Carter, D. B., Tomasselli, A. G., Parodi, L. A., Heinrikson, R. L., and Gurney, M. E. (1999) Membrane-anchored aspartyl protease with Alzheimer's disease β-secretase activity, *Nature 402*, 533−537.
6. 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., Suomensaari, S. M., Wang, S., Walker, D., John, V., et al. (1999) Purification and cloning of amyloid precursor protein β-secretase from human brain, *Nature 402*, 537−540.
7. Huse, J. T., Pijak, D. S., Leslie, G. J., Lee, V. M., and Doms, R. W. (2000) Maturation and endosomal targeting of β-site amyloid precursor protein-cleaving enzyme. The Alzheimer's disease β-secretase, *J. Biol. Chem. 275*, 33729−33737.
8. He, X., Chang, W. P., Koelsch, G., and Tang, J. (2002) Memapsin 2 (β-secretase) cytosolic domain binds to the VHS domains of GGA1 and GGA2: Implications on the endocytosis mechanism of memapsin 2, *FEBS Lett. 524*, 183−187.
9. He, X., Zhu, G., Koelsch, G., Rodgers, K. K., Zhang, X. C., and Tang, J. (2003) Biochemical and structural characterization of the interaction of memapsin 2 (β-secretase) cytosolic domain with the VHS domain of GGA proteins, *Biochemistry 42*, 12174−12180.
10. Hong, L., Koelsch, G., Lin, X., Wu, S., Terzyan, S., Ghosh, A. K., Zhang, X. C., and Tang, J. (2000) Structure of the protease domain of memapsin 2 (β-secretase) complexed with inhibitor, *Science 290*, 150−153.
11. Hong, L., Turner, R. T., III, Koelsch, G., Shin, D., Ghosh, A. K., and Tang, J. (2002) Crystal structure of memapsin 2 (β-secretase) in complex with an inhibitor OM00-3, *Biochemistry 41*, 10963−10967.
12. Hong, L., and Tang, J. (2004) Flap position of free memapsin 2 (β-secretase), a model for flap opening in aspartic protease catalysis, *Biochemistry 43*, 4689−4695.
13. Turner, R. T., III, Koelsch, G., Hong, L., Castanheira, P., Ermolieff, J., Ghosh, A. K., Tang, J., Castenheira, P., and Ghosh, A. (2001) Subsite specificity of memapsin 2 (β-secretase): Implications for inhibitor design, *Biochemistry 40*, 10001−10006.
14. Ghosh, A. K., Bilcer, G., Harwood, C., Kawahama, R., Shin, D., Hussain, K. A., Hong, L., Loy, J. A., Nguyen, C., Koelsch, G., Ermolieff, J., and Tang, J. (2001) Structure-based design: Potent inhibitors of human brain memapsin 2 (β-secretase), *J. Med. Chem. 44*, 2865−2868.
15. Park, H., and Lee, S. (2003) Determination of the active site protonation state of β-secretase from molecular dynamics simulation and docking experiment: Implications for structure-based inhibitor design, *J. Am. Chem. Soc. 125*, 16416−16422.
16. Tomasselli, A. G., Qahwash, I., Emmons, T. L., Lu, Y., Leone, J. W., Lull, J. M., Fok, K. F., Bannow, C. A., Smith, C. W., Bienkowski, M. J., Heinrikson, R. L., and Yan, R. (2003) Employing a superior BACE1 cleavage sequence to probe cellular APP processing, *J. Neurochem. 84*, 1006−1017.
17. Tung, J. S., Davis, D. L., Anderson, J. P., Walker, D. E., Mamo, S., Jewett, N., Hom, R. K., Sinha, S., Thorsett, E. D., and John, V. (2002) Design of substrate-based inhibitors of human β-secretase, *J. Med. Chem. 45*, 259−262.
18. Ghosh, A. K., Shin, D., Downs, D., Koelsch, G., and Tang, J. (2000) Design of potent inhibitors for human brain memapsin 2 (β-secretase), *J. Am. Chem. Soc. 122*, 3522−3523.
19. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode, *Methods Enzymol. 276*, 307−326.
20. Navaza, J. (2001) Implementation of molecular replacement in AMoRe, *Acta Crystallogr. D57*, 1367−1372.
21. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. D54*, 905−921.
22. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for binding protein models in electron density maps and the location of errors in these models, *Acta Crystallogr. A47*, 110−119.
23. Marciniszyn, J., Jr., Hartsuck, J. A., and Tang, J. (1976) Mode of inhibition of acid proteases by pepstatin, *J. Biol. Chem. 251*, 7088−7094.

BI048106K