Crystal Structure of Memapsin 2 (β-Secretase) in Complex with an Inhibitor OM00-3†

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

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

Received June 4, 2002; Revised Manuscript Received July 8, 2002

ABSTRACT: The structure of the catalytic domain of human memapsin 2 bound to an inhibitor OM00-3 (Glu-Leu-Asp-Leu*Ala-Val-Glu-Phe, $K_i = 0.3$ nM, the asterisk denotes the hydroxyethylene transition-state isostere) has been determined at 2.1 Å resolution. Uniquely defined in the structure are the locations of S_3' and S_4' subsites, which were not identified in the previous structure of memapsin 2 in complex with the inhibitor OM99-2 (Glu-Val-Asn-Leu*Ala-Ala-Glu-Phe, $K_i = 1$ nM). Different binding modes for the P_2 and P_4 side chains are also observed. These new structural elements are useful for the design of new inhibitors. The structural and kinetic data indicate that the replacement of the P_2' alanine in OM99-2 with a valine in OM00-3 stabilizes the binding of P_3' and P_4'.

The hallmark of Alzheimer's disease (AD) is a progressive degeneration of the brain caused by the accumulation of amyloid β-peptide (Aβ) (*1*). The first step in the production of Aβ is the cleavage of a membrane protein called the amyloid precursor protein (APP) by a protease known as β-secretase, which has been identified as a membrane-anchored aspartic protease (*2*−*6*) termed memapsin 2 (or BACE or ASP-2). The feasibility of an inhibitor drug targeting this protease for treatment of AD is substantiated by success in the design of inhibitor drugs against HIV protease, also an aspartic protease, and the apparent tolerance of memapsin 2 gene deletion in mice (*7*−*9*). On the basis of preliminary substrate specificity information (*6*), we designed a first-generation inhibitor OM99-2 (*10*), an eight-residue transition-state analogue, Glu-Val-Asn-Leu*Ala-Ala-Glu-Phe, with a K_i near 1 nM (*11*). A 1.9 Å crystal structure of the catalytic unit of memapsin 2 bound to OM99-2 (*12*) revealed that the conformation of the protease and the main features of its active site are those of the aspartic proteases of the pepsin family. All eight residues of OM99-2 were accommodated within the substrate-binding cleft of memapsin 2. The locations and structures of six memapsin 2 subsites for the binding of residues P_4−P_2' of OM99-2 were clearly defined in the structure (*12*). This part of the inhibitor assumed an essentially extended conformation with the active site aspartyls positioned near the transition-state isostere between P_1 and P_1'. Unexpectedly, the backbone of the inhibitor turned at P_2' Ala, departing from the extended conformation, to produce a kink. With less defined electron density, the side chains of P_3' Glu and P_4' Phe appeared to be located on the molecular surface and to have little interaction with the protease. These observations led to the idea that the S_3' and S_4' subsites in memapsin 2 were not well formed and perhaps contributed little to the interaction with substrates and inhibitors (*12*). Subsequently, we determined the detailed subsite preferences of memapsin 2 (*13*), and by using preferred binding residues selected from a combinatorial inhibitor library, a second-generation inhibitor OM00-3, Glu-Leu-Asp-Leu*Ala-Val-Glu-Phe, was designed and found to have a K_i of 0.3 nM. Here we report the structure of the catalytic unit of memapsin 2 in complex with OM00-3. The new structure clearly defines the locations and structures of subsites S_3' and S_4' and provides new insight into their functions. Novel inhibitor−enzyme interactions were also observed in other subsites.

MATERIALS AND METHODS

Crystallization. Human memapsin 2 was expressed in *Escherichia coli* as inclusion bodies, refolded, and purified as previously described (*6*, *11*). Crystallization of memapsin 2 in complex with OM00-3 was carried out using the procedures described previously (*12*) with minor modifications. Crystals were grown at 20 °C in 22.5% PEG 8000 and 0.2 M $(NH_4)_2SO_4$ buffered with 0.1 M sodium cacodylate (pH 6.2) using the hanging drop vapor diffusion method with a 1:1 volume ratio of well to sample solution. Orthorhombic crystals were obtained under these conditions.

Data Collection and Processing. For data collection at 100 K, a crystal was first cryoprotected by transferring to well solution containing 20% (v/v) glycerol and then quickly frozen with liquid nitrogen. Diffraction data were collected on a Mar 345 image plate mounted on a Msc-Rigaku RU-

† This work was supported in part by NIH Grant AG-18933 and a Pioneer Award from the Alzheimer's Association. R.T.T. is a recipient of a Glenn/American Foundation of Aging Research Scholarship. G.K. is a Scientist Development Grant Awardee of the American Heart Association. J.J.N.T. is holder of the J. G. Puterbaugh Chair in Biomedical Research at the Oklahoma Medical Research Foundation.

* To whom correspondence should be addressed. E-mail: Jordan-Tang@omrf.ouhsc.edu.

‡ Oklahoma Medical Research Foundation.

§ Zapaq, Inc.

|| University of Oklahoma Medical Center.

⊥ University of Illinois at Chicago.

10.1021/bi026232n CCC: $22.00 © 2002 American Chemical Society
Published on Web 08/15/2002

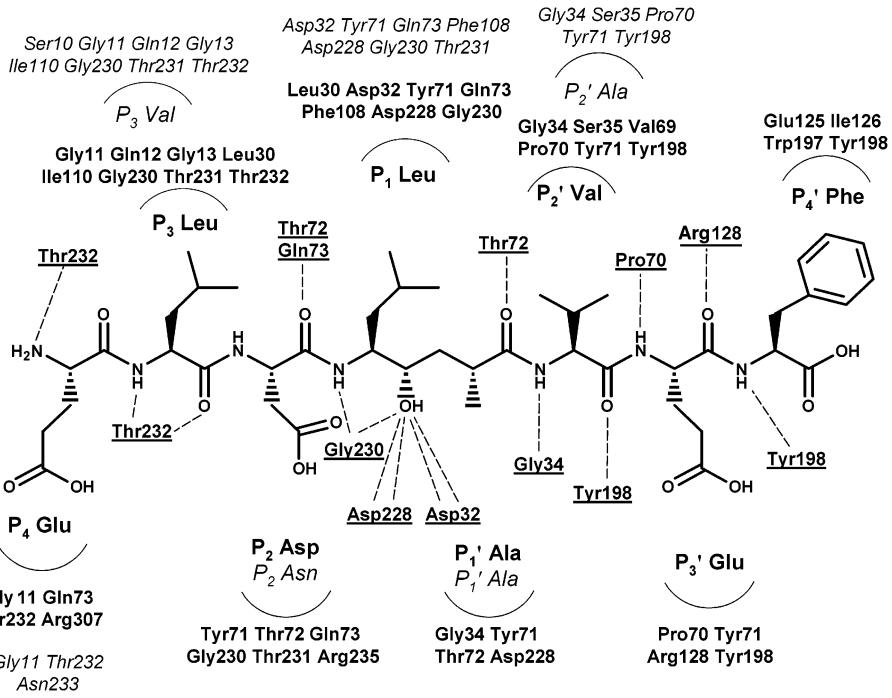

FIGURE 1: Schematic representation of OM00-3. The memapsin 2 residues which contact OM00-3 (distance of <4 Å) are shown in bold. Dashed lines represent hydrogen bonds. The corresponding residues of the OM99-2 structure are in italics for both the inhibitor and the enzyme. No enzyme residues that contact OM99-2 at P_3' and P_4' were identified.

Table 1: Data Collection and Refinement Statistics

	OM003−memapsin 2 complex
space group	$P2_12_12_1$
unit cell dimensions *a*, *b*, and *c* (Å)	86.5, 88.8, 131.0
resolution (Å)	25.0−2.1
no. of observed reflections	190727
no. of unique reflections	58864
R_{merge}[a]	0.119
data completeness (%)	98.8 (25.0−2.1 Å) 97.1 (2.18−2.10 Å)
$I/\sigma(I)$	7.3 (25.0−2.1 Å) 2.2 (2.18−2.10 Å)
R_{work}[b]	0.216
R_{free}[b]	0.271
rms deviation from ideal values	
bond lengths (Å)	0.009
bond angles (deg)	1.6
no. of water molecules	450
average *B* factor (Å^2)	
protein	22.8
solvent	26.1

[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.

300 X-ray generator with Osmic focusing mirrors. The data were processed using the HKL program package (*14*). Data processing and refinement statistics are shown in Table 1.

Structure Determination and Refinement. Molecular replacement solutions were obtained with the program AmoRe (*15*) using the previously determined memapsin 2 structure (PDB entry 1FKN) as the search model. A translation search confirmed that the crystal form belongs to space group $P2_12_12_1$ (Table 1) with two memapsin 2−inhibitor complexes per crystallographic asymmetric unit. The refinement was completed with iterative cycles of manual model fitting using graphics program O (*16*) and model refinement using CNS (*17*). Water molecules were added at the later stages of refinement as identified in $|F_o| - |F_c|$ maps contoured at the 3σ level. Ten percent of the diffraction data were excluded from the refinement at the beginning of the process to monitor the R_{free} values. The two memapsin 2−inhibitor complexes in the crystallographic asymmetric unit were found to be essentially identical.

Coordinates. The coordinates for the structure reported in this paper have been deposited in the Protein Data Bank (entry 1M4H).

Kinetic Measurements. The measurements of relative k_{cat}/K_m values for the determination of residue preference at P_3' and P_4' were carried out as previously described (*13*). Two template substrate sequences, *WHDR*EVNLA**A**EF and *WHDR*EVNLA**V**EF, were used. The former had a P_3' Ala and the latter a P_3' Val. Four N-terminal residues, *WHDR*, were added to the substrate to facilitate the analysis using mass spectrometry. For each template, three groups of each of the peptides containing a total of 11 representative residues (in single-letter code, A, D, E, F, L, M, R, T, V, W, and Y) for both P_3' and P_4' were designed and synthesized. The initial velocities for memapsin 2 hydrolysis of each peptide in the mixtures were determined with a MALDI-TOF mass spectrometer as previously described (*13*). The internal standards and the calculation of relative k_{cat}/K_m values were also as reported previously (*13*).

RESULTS AND DISCUSSION

Structure and Inhibitor Binding. The structure of the OM00-3−memapsin 2 complex in space group $P2_12_12_1$ was determined at 2.1 Å using the molecular replacement method (Table 1). The structure of the enzyme, the interactions of the P_1/P_1' (Leu*Ala) region of OM00-3 with the substrate binding cleft of memapsin 2, and the backbone conformation

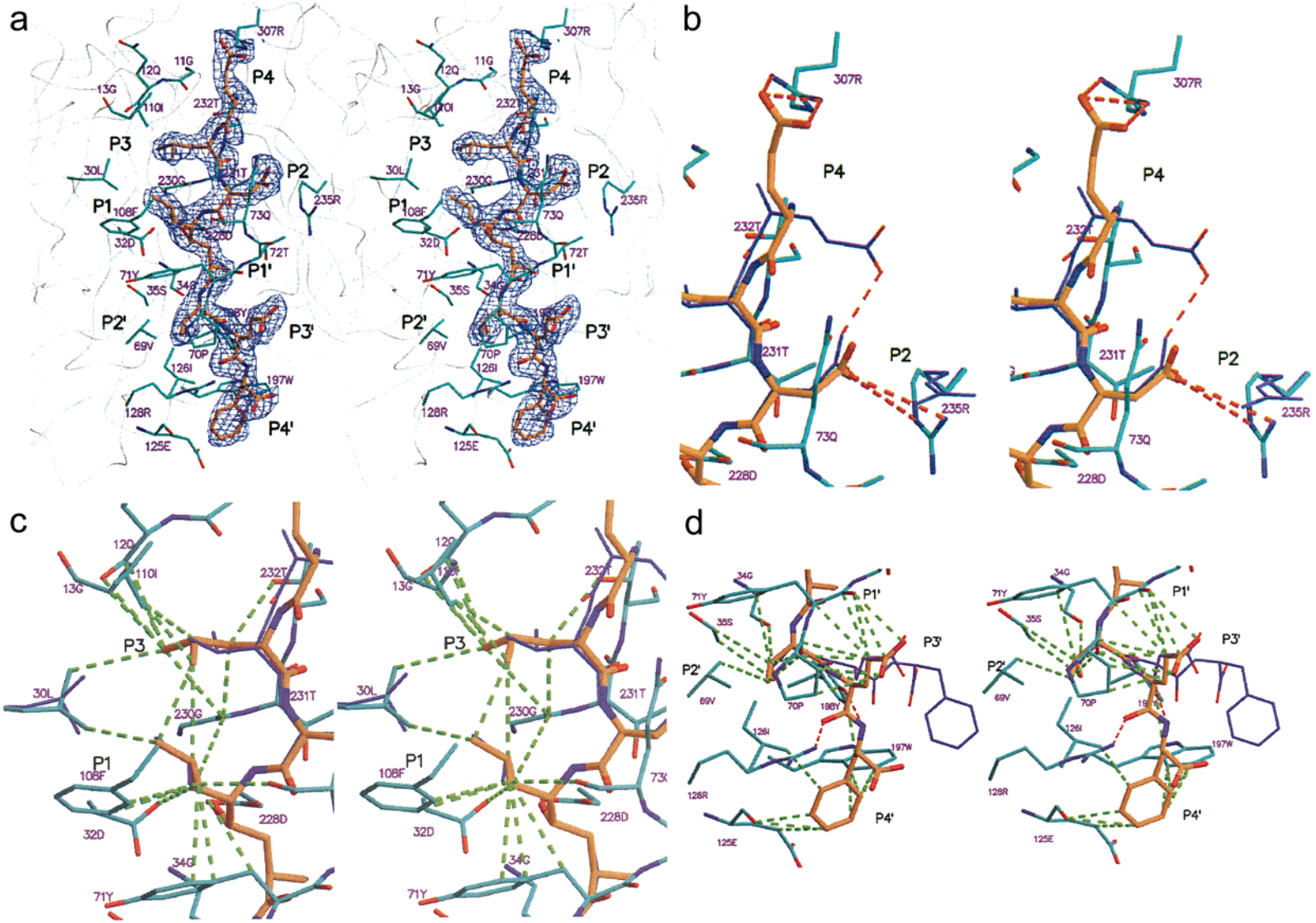

FIGURE 2: (a) Stereoview of the electron density and subsite residues for OM00-3 bound to memapsin 2. The inhibitor is shown in brown. Memapsin 2 residue side chains in direct contact with the inhibitor (<4 Å) are shown in blue, and main chain atoms are included if they have contact with the inhibitor. The $2F_o - F_c$ electron density map for the inhibitor was contoured at the 1σ level. (b) Stereorepresentation of the OM00-3 P_4 and P_2 subsites (brown) bound to memapsin 2. Red dashed lines represent the salt bridges from the P_4 and P_2 side chains to the enzyme. OM99-2 from the previous crystal structure is superimposed onto OM00-3 in blue. The hydrogen bond between the P_4 and P_2 side chains of OM99-2 is also shown as a red dashed line. (c) van der Waals contacts (<4 Å) from the side chains of P_3 and P_1 of OM00-3 to memapsin 2 are shown as green dashed lines. The contacts between the side chains of P_3 and P_1 are also included. OM99-2 is superimposed onto OM00-3 in blue. (d) van der Waals contacts from the side chains of P_2', P_3', and P_4' of OM00-3 to memapsin 2 are shown as green dashed lines. Hydrogen bonds from the P3′ and P4′ backbone to the enzyme are shown as red dashed lines. OM99-2 is superimposed in blue.

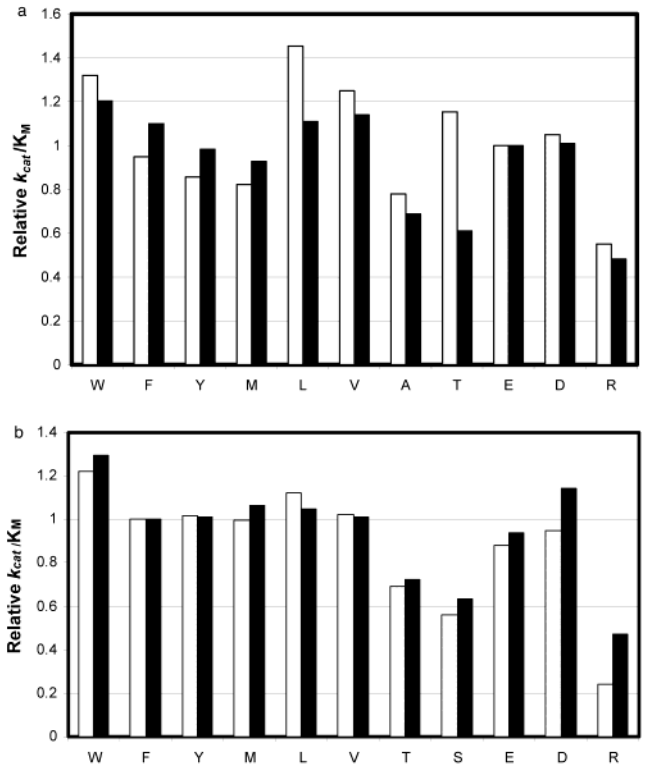

FIGURE 3: Residue preference at positions P_3' (a) and P_4' (b) with the P_2' residue being alanine (white bars) or valine (black bars). The template sequence was EVNLAAEF which encompassed P_4-P_4'. The bars show relative k_{cat}/K_m values with the native template residues glutamic acid (E) as 1.0 at P_3' and phenylalanine (F) as 1.0 at P_4'. In both panels a and b, the two sets of k_{cat}/K_m values were normalized separately for the substrates with P2′ alanine and P2′ valine.

of the inhibitor from P_3 to P_2' are essentially the same as in the structure of the OM99-2−memapsin 2 complex. However, the current structure shows different side chain configurations within the S_4, S_3, and S_2 subsites (Figures 1 and 2) when compared to those of the OM99-2 structure (*12*). In addition, the locations and nature of S_3' and S_4' binding pockets are clearly defined for the first time.

S_4, S_3, and S_2 Subsites. The new S_4 pocket in the current structure involves memapsin 2 residues Glu^{11}, Gln^{73}, Thr^{232}, and Arg^{307}. The latter forms several ionic bonds to the carboxylate oxygen atoms of inhibitor P_4 Glu (Figure 2b). In the previous OM99-2−memapsin 2 structure, the main chain torsion angle, ψ, of the P_4 Glu is different from the current one by 152°. Thus, in OM99-2 structure, there is a hydrogen bond between the side chains of P_4 Glu and P_2 Asn, and the P_4 Glu side chain in that structure has little interaction with the protease (Figure 2b). In the OM00-3 structure, however, P_2 is an Asp, and thus, its interaction with P_4 Glu is unfavorable. Although the average *B* factor of P_4 Glu is somewhat higher (42 Å^2) than those of the interior residues of P_3-P_2' (17−20 Å^2), its multiple interactions with the protease residues suggest that the newly observed S_4 pocket contributes significantly to inhibitor binding.

In the OM00-3 structure, Leu^{30} in S_3 of the protease has contacts with the leucines of the inhibitor at P_3 and P_1. These two side chains also contact each other, contributing to the further stabilization of the inhibitor conformation (Figure 2c). These productive interactions are not present in the OM99-2 structure where P_3 is Val rather than Leu and the conformation of Leu^{30} is different as a result of a 60° rotation around χ_2 (Figure 2c).

In the S_2 pocket, the P_2 Asp of OM00-3 forms two ionic bridges to the Arg^{235} side chain (Figure 2b). The conformation of Arg^{235} is different from that in the OM99-2 structure where the P_2 residue is Asn (Figure 2b). Flexibility within the S_2 pocket allows interaction with either Asp or Asn at P_2 and is consistent with the observation that these two residues are the most preferred substrate and inhibitor residues for this subsite (*13*).

S_3' and S_4' Subsites. In contrast to the OM99-2−memapsin 2 structure, the conformation of the P_3' and P_4' subsites is well defined by electron density in the OM00-3−memapsin 2 structure (Figure 2a). The backbone at P_3' and P_4' of OM00-3 assumes an extended conformation which is stabilized by a hydrogen bond from the P_3' backbone carbonyl to Arg^{128} (Figure 2d). A very weak hydrogen bond from the P_4' backbone nitrogen to Tyr^{198} may make small contributions to the binding. The S_3' subsite, defined by several direct van der Waals interactions (<4 Å), comprises residues Pro^{70}, Tyr^{71}, Arg^{128}, and Tyr^{198}. The S_4' pocket consists of Glu^{125}, Ile^{126}, Trp^{197}, and Try^{198} (Figures 1 and 2d). By virtue of their location at the C-terminus of the inhibitor, both P_3' and P_4' residues have average *B* factor values somewhat higher (28 and 37 Å^2, respectively) than those of the residues in the region from P_3 to P_2'. In the presence of easily interpretable electron density, these higher temperature factors do not compromise the validity of the structural information and the analysis of the interactions for subsites S_3' and S_4'.

Contribution of P_2' to the Binding of P_3' and P_4'. Inhibitors OM99-2 and OM00-3 have identical P_3' and P_4' residues. It was therefore unexpected that the P_3' and P_4' residues are better defined for the latter structure. Kinetics studies have shown that, compared to the other subsites, the P_3' and P_4' subsites have a considerably broader range of amino acid preference (*13*). Because the P_2' Val in OM00-3 has several more contacts with the enzyme than the Ala in OM99-2 (*12*), we reasoned that a better binding of P_2' Val may contribute to the stability of P_3' and P_4' residues in OM00-3. A possibility that was considered was that P_2' Val shifted residue preference at P_3' and P_4' toward Glu and Phe, respectively. To test this hypothesis, we measured the relative residue preference at the P_3' and P_4' subsites for two sets of substrates, EVNL/A**A**EF and EVNL/A**V**EF (*13*), which differed only in the Ala or Val at P_2'. Ten representative residues were chosen for each of the P_3' and P_4' subsites. The relative k_{cat}/K_m values of these 10 substrates were determined by their relative initial hydrolytic rate using a mass spectrometric method (*13*). Our results show that the differences in residue preferences at the P_3' (Figure 3a) and P_4' (Figure 3b) subsites for two sets of substrates with P_2' Ala and P_2' Val are small. More significantly, there is not a shift of preference at the P_3' and P_4' subsites toward Glu and Phe, respectively, when P_2' is Val. Yet peptide substrates with Val at P_2' have on average ~30% higher k_{cat}/K_m values than their counterparts with Ala at P_2' (*13*). To determine which kinetic parameter contributes to this difference, we measured individual k_{cat} and K_m values for two substrates that differed at only P_2' by Val or Ala. Substrate EVNL/A**V**EFWHDR (*13*) produced a K_m of 83 ± 8.9 μM and a

k_{cat} of 1007 ± 106 s^{-1} (n = 3), while substrate EVNL/A**A**EFWHDR had a K_m of 125 ± 11 μM and a k_{cat} of 274 ± 23 s^{-1} (n = 2). The differences in kinetic parameters between P_2' Val and P_2' Ala substrates are much greater in k_{cat} (~4-fold) than in K_m (~1.5-fold). These observations suggest that, compared with P_2' Ala, P_2' Val primarily improves the transition-state binding of P_3' and P_4' residues, but does not alter their specificity.

Implications of New Subsites in Inhibitor Design. The first structure of the memapsin 2 catalytic domain complexed to inhibitor OM99-2 (*12*) has been shown to be useful in the structure-based design of smaller and potent memapsin 2 inhibitors (*18*). The new structure described here provides improved versatility for inhibitor design. Memapsin 2 inhibitors with clinical potentials should be potent, selective, and small enough to penetrate the blood−brain barrier. It is known that the HIV protease inhibitor drug indinavir, 614 Da, can cross the blood−brain barrier (*19*). A memapsin 2 inhibitor of similar size would bind to approximately five subsites consecutively. We have demonstrated that inhibitors with K_i in the low nanomolar range can be designed without evoking binding at the P_3' and P_4' subsites (*18*). The new binding modes at P_4 and P_2 could be utilized for the design of inhibitors of this type. The new subsite structures of S_3' and S_4' described above might be incorporated in the design of inhibitors with P_3' and P_4' but without P_4 and P_3 residues. Current findings also suggest that such a design should have a strong binding side chain, such as Val, at P_2'.

ACKNOWLEDGMENT

We thank Dr. Simon Terzyan for assistance in data collection and Drs. Xuejun C. Zhang and Jean Hartsuck for critical reading of the manuscript.

REFERENCES

1. Selkoe, D. J. (2001) *Physiol. Rev. 81*, 741−766.
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) *Science 286*, 735−741.
3. 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) *Nature 402*, 537−540.
4. 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) *Nature 402*, 533−537.
5. 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) *Mol. Cell. Neurosci. 14*, 419−427.
6. Lin, X., Koelsch, G., Wu, S., Downs, D., Dashti, A., and Tang, J. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 1456−1460.
7. Cai, H., Wang, Y., McCarthy, D., Wen, H., Borchelt, D. R., Price, D. L., and Wong, P. C. (2001) *Nat. Neurosci. 4*, 233−234.
8. Luo, Y., Bolon, B., Kahn, S., Bennett, B. D., Babu-Khan, S., Denis, P., Fan, W., Kha, H., Zhang, J., Gong, Y., Martin, L., Louis, J. C., Yan, Q., Richards, W. G., Citron, M., and Vassar, R. (2001) *Nat. Neurosci. 4*, 231−232.
9. Roberds, S. L., Anderson, J., Basi, G., Bienkowski, M. J., Branstetter, D. G., Chen, K. S., Freedman, S. B., Frigon, N. L., Games, D., Hu, K., Johnson-Wood, K., Kappenman, K. E., Kawabe, T. T., Kola, I., Kuehn, R., Lee, M., Liu, W., Motter, R., Nichols, N. F., Power, M., Robertson, D. W., Schenk, D., Schoor, M., Shopp, G. M., Shuck, M. E., Sinha, S., Svensson, K. A., Tatsuno, G., Tintrup, H., Wijsman, J., Wright, S., and McConlogue, L. (2001) *Hum. Mol. Genet. 10*, 1317−1324.
10. Ghosh, A. K., Shin, D., Downs, D., Koelsch, G., and Tang, J. (2000) *J. Am. Chem. Soc. 122*, 3522−3523.
11. Ermolieff, J., Loy, J. A., Koelsch, G., and Tang, J. (2000) *Biochemistry 39*, 12450−12456.
12. Hong, L., Koelsch, G., Lin, X., Wu, S., Terzyan, S., Ghosh, A. K., Zhang, X. C., and Tang, J. (2000) *Science 290*, 150−153.
13. Turner, R. T., III, Koelsch, G., Hong, L., Castanheira, P., Ermolieff, J., Ghosh, A. K., and Tang, J. (2001) *Biochemistry 40*, 10001−10006.
14. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
15. Navaza, J. (2001) *Acta Crystallogr. D57*, 1367−1372.
16. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard (1991) *Acta Crystallogr. A47*, 110−119.
17. 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) *Acta Crystallogr. D54*, 905−921.
18. 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) *J. Med. Chem. 44*, 2865−2868.
19. Martin, C., Sonnerborg, A., Svensson, J. O., and Stahle, L. (1999) *AIDS 13*, 1227−1232.

BI026232N