Structural Comparison Suggests That Thermolysin and Related Neutral Proteases Undergo Hinge-Bending Motion during Catalysis†

D. R. Holland,‡ D. E. Tronrud,‡ H. W. Pley,§ K. M. Flaherty,§ W. Stark,⊥ J. N. Jansonius,⊥ D. B. McKay,§ and B. W. Matthews*,‡

Institute of Molecular Biology, Howard Hughes Medical Institute and Department of Physics, University of Oregon, Eugene, Oregon 97403, Department of Cell Biology, Beckman Laboratories for Structural Biology, Stanford University School of Medicine, Stanford, California 94305, and Biozentrum der Universitat Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

Received June 16, 1992; Revised Manuscript Received August 28, 1992

ABSTRACT: Crystal structures are known for three members of the bacterial neutral protease family: thermolysin from *Bacillus thermoproteolyticus* (TLN), the neutral protease from *Bacillus cereus* (NEU), and the elastase of *Pseudomonas aeruginosa* (PAE), both in free and ligand-bound forms. Each enzyme consists of an N-terminal and C-terminal domain with the active site formed at the junction of the two domains. Comparison of the different molecules reveals that the structure within each domain is well conserved, but there are substantial hinge-bending displacements (up to 16°) of one domain relative to the other. These domain motions can be correlated with the presence or absence of bound inhibitor, as was previously observed in the specific example of PAE [Thayer, M. M., Flaherty, K. M., & McKay, D. B. (1991) *J. Biol. Chem. 266*, 2864–2871]. The binding of inhibitor appears to be associated with a reduction of the domain hinge-bending angle by 6–14° and a closure of the "jaws" of the active site cleft by about 2 Å. Crystallographic refinement of the structure of thermolysin suggests that electron density seen in the active site of the enzyme in the original structure determination probably corresponds to a bound dipeptide. Thus, the crystal structure appears to correspond to an enzyme–inhibitor or enzyme–product complex, rather than the free enzyme, as has previously been assumed.

There are three members of the bacterial neutral protease family whose three-dimensional structures are known. These are thermolysin from *Bacillus thermoproteolyticus* (TLN) (Matthews et al., 1972), the neutral protease from *Bacillus cereus* (NEU) (Pauptit et al., 1988; Stark et al., 1992), and the elastase of *Pseudomonas aeruginosa* (PAE) (Thayer et al., 1991; McKay et al., 1992). The three enzymes have homologous amino acid sequences, have a catalytically essential zinc, and have their active sites located in a well-defined cleft between the N- and C-terminal domains.

A comparison of the native structure of the bacterial elastase (PAE) with two inhibitor-bound forms of the enzyme showed that the active site cleft of the free enzyme was more "open" than that of the inhibitor-bound form, suggesting that the enzyme might undergo a conformational change during catalysis (Thayer et al., 1991; McKay et al., 1992). Determination of the structure of the *Bacillus cereus* enzyme also showed that its active site cleft is more open than that of thermolysin (Pauptit et al., 1988; Stark et al., 1992).

These apparent "hinge-bending" displacements have prompted a comparison of the respective structures and are the subject of the present communication. One question to be addressed is whether the observed "hinge bending" is relevant for function, whether it simply reflects a difference in the respective structures, or whether it might be a consequence of different crystal packing interactions in the respective crystal structures. The present work suggests that the neutral metalloproteases may, in fact, undergo hinge bending during catalysis.

MATERIALS AND METHODS

The four crystallographic structures used for the comparison are summarized in Table I. All three proteases are quite similar and can be considered as made up of two domains. The N-terminal domain consists predominantly of β-sheet and is defined by residues 1–77, 91–134, and 195–199. [Numbering for thermolysin (Figure 1) is used throughout the text.] The C-terminal domain is mostly helical and is defined by residues 78–90, 135–194, and 200–316. Domain definitions are based on those suggested by Thayer et al. (1991) and McKay et al. (1992) and on optimal superposition of the four structures.

Corresponding domains in the respective structures were compared using the method of Rossmann and Argos (1975). The comparison is based on superposition of α-carbon atoms. Parts of the backbone that are present in one molecule but have no counterpart in the structure being compared are ignored. In the case of TLN, the coordinates used were those of a complex with Cbz-Phe-(PO_2)-Leu-Ala since this is believed to be an excellent mimic of the presumed transition state (Holden et al., 1987).

An overlay of the four structures, based on superposition of the C-terminal domains, is shown in Figure 2. In each case, the PAE structures and the NEU structure have been superimposed onto TLN. Overall, the corresponding domains of TLN and NEU are structurally more similar than those of TLN and PAE. This is consistent with the greater sequence identity between TLN and NEU (73%) relative to TLN and PAE (28%). The four C-terminal domains superimpose somewhat better than is the case for the four N-terminal

† This work was supported in part by grants to B.W.M. from the NSF (PCM8312151) and the Lucille P. Markey Charitable Trust, to D.B.M. from the NIH (AI-19762 and Research Career Development Award AI-00631), and to J.N.J. from the Swiss National Science Foundation (3.098-085 and 31-25712.88).

* Author to whom correspondence should be addressed.
‡ University of Oregon.
§ Stanford University.
⊥ Biozentrum der Universitat Basel.

0006-2960/92/0431-11310$03.00/0

Table I: Crystallographic Structures Compared[a]

enzyme	ligand	resolution (Å)	R-factor	space group	cell dimensions
TLN	Cbz-Phe-(PO_2)-Leu-Ala	1.7	0.163	$P6_122$	a = 94.1 Å, b = 131.4 Å, γ = 120.0°
NEU	none	2.0	0.175	$P6_522$	$a = b$ = 76.5 Å, c = 201.0 Å, γ = 120.0°
PAE	none	1.5	0.190	$P2_1$	a = 41.9 Å, b = 90.9 Å, c = 40.8 Å, β = 113.8°
PAE-I	S-HomoPhe[N$\alpha\alpha$]Phe-IsoAsn	2.1	0.150	$P6_1$	$a = b$ = 116.4 Å, c = 44.5 Å, γ = 120.0°

[a] TLN is thermolysin from *B. thermoproteolyticus* (Holden et al., 1987); NEU is neutral protease from *B. cereus* (Stark et al., 1992). PAE is the native form of the elastase of *P. aeruginosa* (previously identified as PAE1) and PAE-I is the enzyme–inhibitor complex previously identified as PAE2 (Thayer et al., 1991; unpublished results of H. W. Pley, K. M. Flaherty, and D. B. McKay).

```
            βββββββ  β      ββββ    βββββ  ββ  β ββ          βββββ         βββββ
TLN    iTGTstVGvG r-GVLGDqKn iNTTyS-tyY YLQDNTRGdg IFTYDAKyRt TLPGsLWADA  58
NEU    vTGTnkVGtG K-GVLGDtKs lNTTlSgssY YLQDNTRGat IFTYDAKnRs TLPGtLWADA  59
PAE    aeaggpgGnq KiGkytygsd ygplivnaac emdDgnvitv dmnsstddsk TtPfrfacpt  60
       ...................................................................

          ββ      αααααααα   ααααααααα  α               βββββ         ββββ
TLN    D--NQFfAsY DAPAVDAHYY AGVTYDYYKn vhNRlSydgn nAairSsVHY sqgYNNAFWN 116
NEU    D--NvFNAAY DAaAVDAHYY AGkTYDYYKa tFNRnSInda gApLkStVHY GsnYNNAFWN 117
PAE    ntykQvNgAY -sPlnDAHff gGVvfklYrd wFgtsplthk ---LymkVHY GasveNAyWd 117
       ........................             ............................

         ββββ                ααααααααα  ααααα          ααααααα   ααααααααα
TLN    GSeMVYGDGD GqTFiPLSGG IDVvAHELTH AVTdytaGLI YQNESGAiNE AISDIFGTLV 176
NEU    GSqMVYGDGD GvTFtsLSGG IDVigHELTH AVTEnsSnLI YQNESGAlNE AISDIFGTLV 177
PAE    GtaMlfGDGa tm-FyPLvs- lDVaAHEvsH gfTEqnSGLI YrgqSGgmNE AfSDmaGeaa 174
       ...................

       αααα         ββ ααα        βββ                          αα  ααα    ααα
TLN    EFYaNkNPDW EIGEDvYTPG isGDsLRSMS DPaKYGDPDH YSKRYTGtqD NGGVHiNSGI 236
NEU    EFYdNrNPDW EIGEDiYTPG KaGDALRSMS DPtKYGDPDH YSKRYTGssD NGGVHtNSGI 237
PAE    EFYmrganDf lIG---Ydik KgSgALRyMd qPsrdGrsid nasqYyngiD ---VHhsSGv 228
                          .. ...

       ααααααααα    βββ   βββ      αααααα  ααααααα          ααααα  ααααααααα
TLN    INKAAYLIsq GGTHYGVsVv GIGrDKLGkI FYRAlTQYlT pTSNFSQlRA aAVQSAtDLY 296
NEU    INKqAYLLAN GGTHYGVtVt GIGkDKLGaI yYRANTQYfT qsttFSQaRA GAVQaAaDLY 297
PAE    yNrAfYLLAN s--------p Gwdtraafev FvdANryYwT aTSNynsgac GvirSAqn-- 278

           ααααα   αααααα
TLN    GstSqEVAsV KQAFdAVGVk -- 316
NEU    GanSaEVAaV KQsFSAVGVn -- 317
PAE    --rnysaAdV trAFStVGVt cp 298
```

FIGURE 1: Sequence alignment of TLN (Titani et al., 1972), NEU (Sidler et al., 1986), and PAE (Bever & Iglewski, 1988). The sequences are aligned according to the structural correspondence. Residues that are identical between sequences are denoted by capital letters; those that are not identical are shown in lower case. Residues underscored by a dot (●) comprise the N-terminal domain, and the remainder are within the C-terminal domain. Secondary structure for TLN is indicated by α for α-helix and β for β-sheet.

domains (Table II). Although superposition of the individual domains is quite good, it is apparent that the relative positioning of the N- and C-terminal domains differs in each of the four structures, with TLN having the most "closed" and native PAE having the most "open" interdomain conformation.

RESULTS

One method to analyze the overall differences in conformation of the different molecules is to use a difference C^{α} distance plot (Nishikawa et al., 1972). Such a plot shows the changes in intramolecular distances between α-carbon atoms. Figure 3 compares NEU with TLN and PAE with PAE-I. Featureless regions indicate those parts of the proteins being compared that have essentially identical backbone structures. These are consistent with the N- and C-terminal domains remaining unchanged, as defined above. The fact that features in the lower right of Figure 3 are similar to those in the upper left indicates that the interdomain movement of TLN relative to NEU is similar to that of PAE relative to PAE-I, although of slightly different magnitude. The difference distance plots

Table II: Analysis of Domain Hinge-Bending[a]

comparison	change in hinge-bending angle (deg)	root-mean-square difference between C^{α} atoms in domains compared: amino-terminal domain (Å)	carboxyl-terminal domain (Å)
TLN vs PAE	16	2.3 (101)	1.3 (167)
TLN vs PAE-I	4	2.3 (101)	1.3 (168)
TLN vs NEU	6	0.7 (120)	0.5 (192)
PAE vs PAE-I	14	0.7 (124)	0.6 (170)

[a] The hinge-bending angle was determined by first superimposing the C-terminal domains of the two structures being compared and then determining the angle of rotation required to bring the N-terminal domain of one protein into register with the other. In the right-hand columns the numbers in parentheses indicate the number of α-carbon atoms included in the comparison, determined according to Rossmann & Argos (1975).

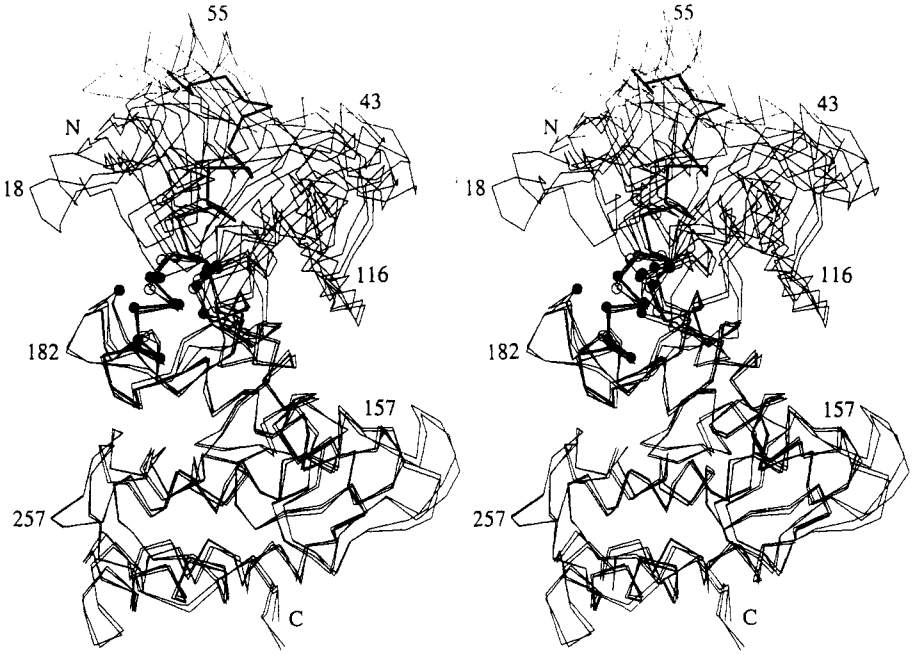

FIGURE 2: α-Carbon backbones of PAE, PAE-I, and NEU superimposed on that of TLN so that the C-terminal domains correspond as well as possible. Solid circles indicate residues that are close to the axis of hinge bending for the PAE and PAE-I pair. Open circles indicate analogous residues for the TLN–NEU pair (see text). The α-helix (65–93 of TLN) that connects the two domains is highlighted in a bold line.

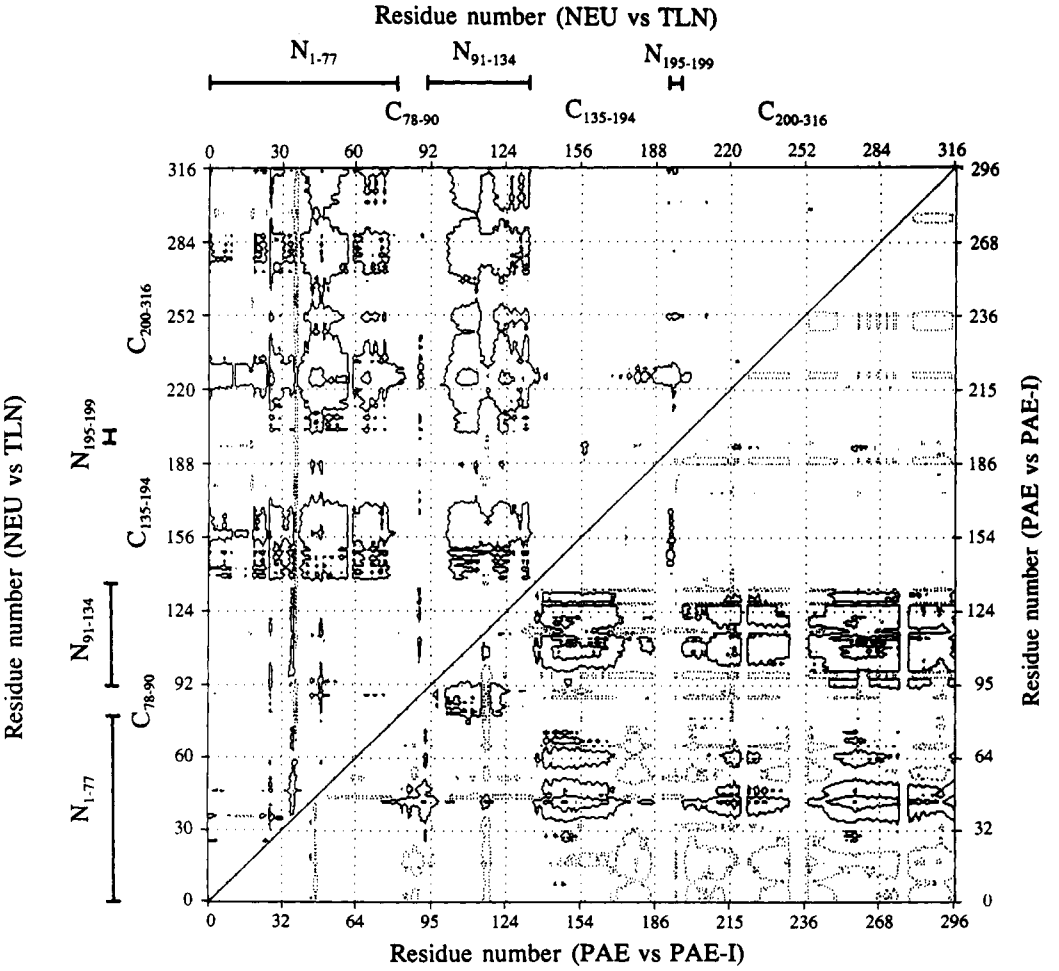

FIGURE 3: Double-difference plot comparing C^{α}–C^{α} distances in TLN and NEU (upper left) and PAE and PAE-I (lower right). If the distance between α-carbons i and j in a given structure is r_{ij}, then the values plotted in the upper left are $(r_{ij,\mathrm{NEU}} - r_{ij,\mathrm{TLN}})$ and in the lower right are $(r_{ij,\mathrm{PAE}} - r_{ij,\mathrm{PAE\text{-}I}})$. Contours are drawn at 1 Å, 3 Å, 5 Å, ... (solid contours) and −1 Å, −3 Å, −5 Å, ... (broken contours). The bars indicate residues that are within the N-terminal domains (e.g., N_{1-77}). The remaining residues (e.g., C_{78-90}) are in the C-terminal domains. The residue identification corresponds to the amino acid sequences numbering given in Figure 1. In order to keep the PAE structure in register with TLN, residues that are present in TLN but not in PAE appear as "gaps" in the PAE sequence.

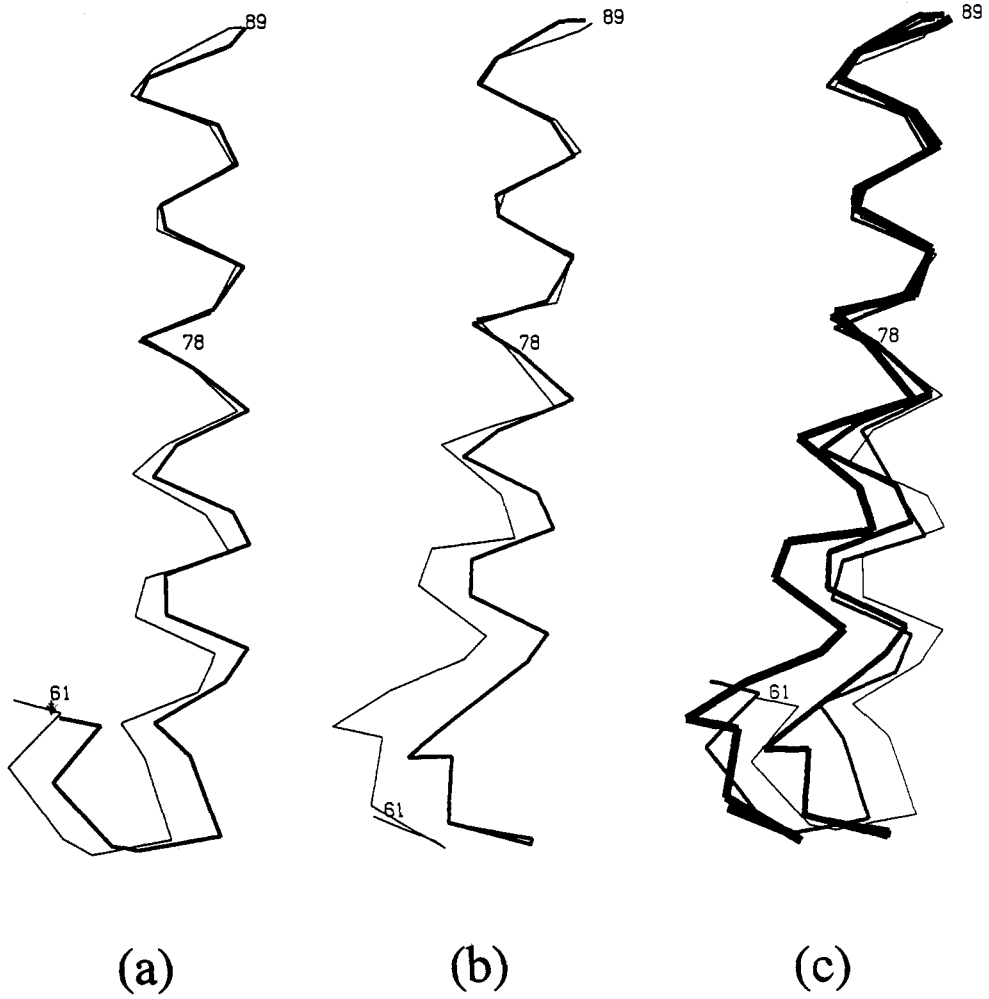

FIGURE 4: Comparison of α-helix 65–89 in the four metalloprotease structures. (a) NEU (light) and TLN (bold); (b) PAE (light) and PAE-I (bold); (c) comparison of the helix in all four structures. In all cases the helices are aligned on the basis of optimal superposition of the full C-terminal domains of the respective proteins.

for NEU vs TLN and PAE vs PAE-I are easiest to interpret since the proteins being compared have the highest sequence identity. Comparisons of PAE vs TLN and PAE vs NEU (not shown) show similar domain motion as well as localized shifts, especially in loop regions adjacent to insertions and deletions (Figure 1).

Relative to TLN, which is the most "closed" of the four structures, the native form of PAE is most "open" with a hinge-bending displacement of 16° (Table II). The NEU and PAE-I structures have intermediate hinge-bending displacements. The axes describing the respective hinge-bending displacements are roughly parallel to each other and pass in the vicinity of residues 77–82, 176–182, 137–138, 198–200, 191–193, and 31–34. As highlighted in Figure 2, these residues occupy a small region between the N-terminal and C-terminal domains.

It might be noted that residues 77–82 are located in the long helix that connects the two domains (Figure 2). Superpositions of this region (Figure 4a–c) suggest that this helix is bent in PAE and NEU relative to PAE-I and TLN. If the C-terminal domains of the respective proteins are superimposed the carboxyl terminal halves of the helices agree well, but there is a gradual deviation as one progresses toward the amino terminus of the helix.

Comparison of Active Sites. In the case of the PAE and NEU crystal structures, the active site region is presumed to be unoccupied, except for solvent. PAE-I is a complex with the covalent inhibitor *S*-homoPhe[N$\alpha\alpha$]Phe-IsoAsn (McKay et al., 1992; unpublished results of H. W. Pley, K. M. Flaherty, and D. B. McKay), and the thermolysin structure is a complex with the presumed transition-state analogue Cbz-Phe-(PO_2)-Leu-Ala (Holden et al., 1987). Residues presumed to be critical for catalysis (Kester & Matthews, 1984; Matthews, 1988) are conserved in all of the structures. In each case the "upper" part of the active site cleft is formed by residues from the N-terminal domain while residues from the C-terminal domain constitute the "lower" half of the active site region.

Pairwise comparisons of the active site regions, based on superposition of the C-terminal domains, are shown in Figure 5. As shown in Figure 5a (top) the conformations of the two ligand-bound structures, PAE-I and TLN, are very similar. The main-chain conformations in the active sites of the free enzymes NEU and PAE are also similar (Figure 5b, middle) except for a change in the vicinity of Phe 114.

A comparison of the bound (TLN) and unbound (NEU) enzymes, however (Figure 5c, bottom) shows that the β-sheet strand (residues 112–116 of TLN) that extends across the "top" of the cleft and interacts with substrate moves ~2 Å "upward". This means that the "jaws" of the enzyme that hold a bound substrate are approximately 2 Å further apart in the unbound than in the bound conformations. A very similar conformational change was noted previously in comparison of the bound (PAE-I) and unbound forms of PAE (Thayer et al., 1991; McKay et al., 1992; unpublished results of H. W. Pley, K. M. Flaherty, and D. B. McKay). Comparison of PAE with TLN (not shown) reveals a similar

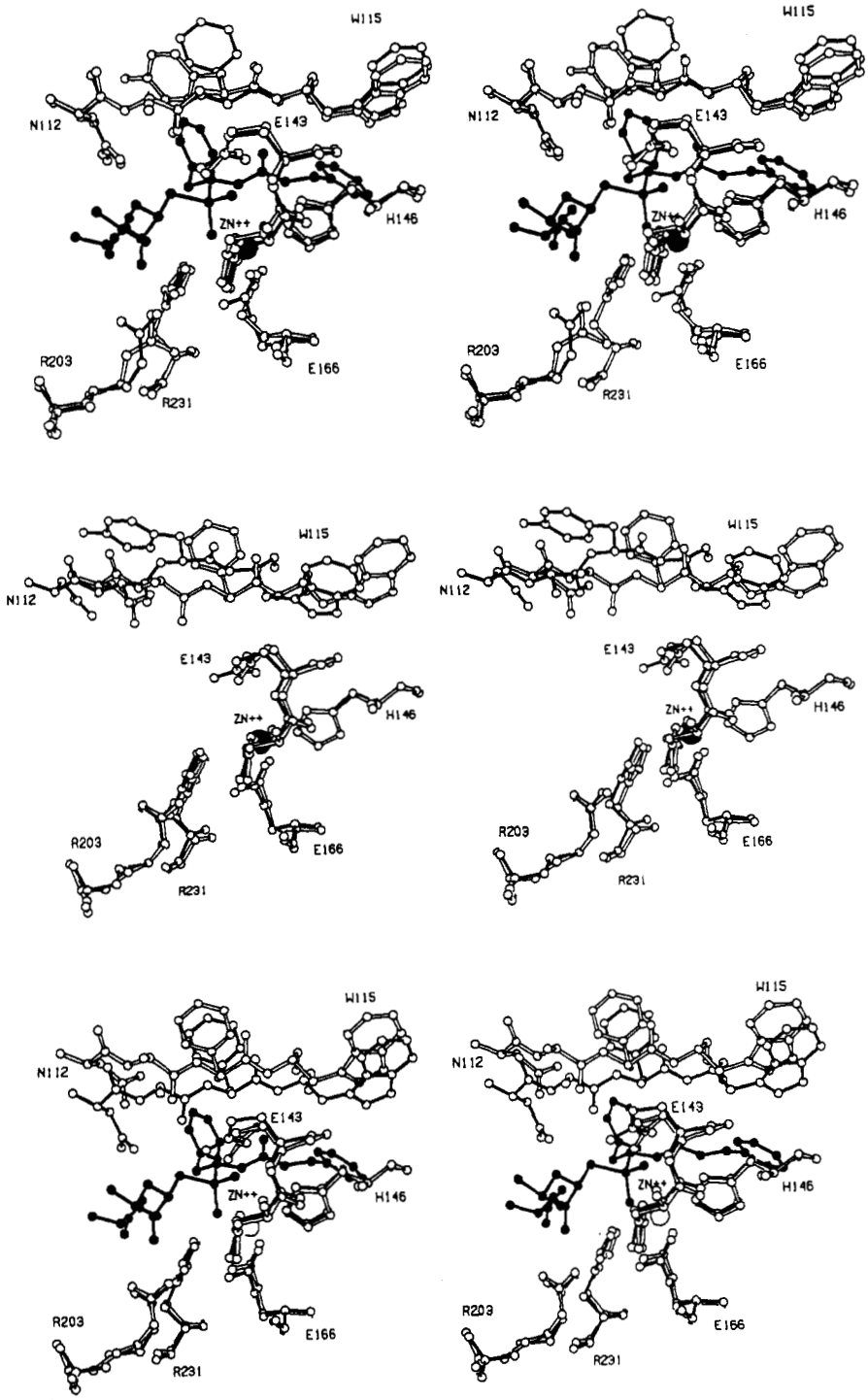

FIGURE 5: Superposition of the active site regions of different metalloprotease structures. (a, top) Comparison of the two inhibitor-bound structures TLN (solid bonds) and PAE-I (open bonds). The inhibitor included in the figure, shown completely solid, is Cbz-Phe-(PO_2)-Leu-Ala bound to TLN. (b, middle) Comparison of the two inhibitor-free structures, PAE (solid bonds) and NEU (open bonds). (c, bottom) Comparison of the bound/free pair TLN (solid bonds) and NEU (open bonds). The TLN inhibitor is again shown in all solid. Numbering for all three comparisons corresponds to that of TLN.

opening of the mouth of the active site cleft by an amount even larger than 2 Å.

DISCUSSION

The results presented above suggest that the neutral metallopeptidases may undergo a conformational change on binding inhibitors and presumably substrates as well. The observed conformational change includes an increase in the hinge-bending angle between the N-terminal and the C-terminal domain as well as an opening of the jaws of the active site cleft by about 2 Å.

An apparent difficulty with this scenario, however, is that the crystal structure of thermolysin has never been observed to change significantly on the binding of inhibitors [e.g., Kester and Matthews (1974), Holden et al. (1987), and Matthews (1988)]. However, recent refinement of the structure of native thermolysin using data recollected to 1.6 Å has suggested that there is electron density in the active site that can be interpreted as corresponding to the bound dipeptide Val-Lys or a close homologue (panels a and b (top, middle) of Figure 6). Val-Lys corresponds to the last two residues in the sequence of thermolysin (Table I), and it is possible that the dipeptide

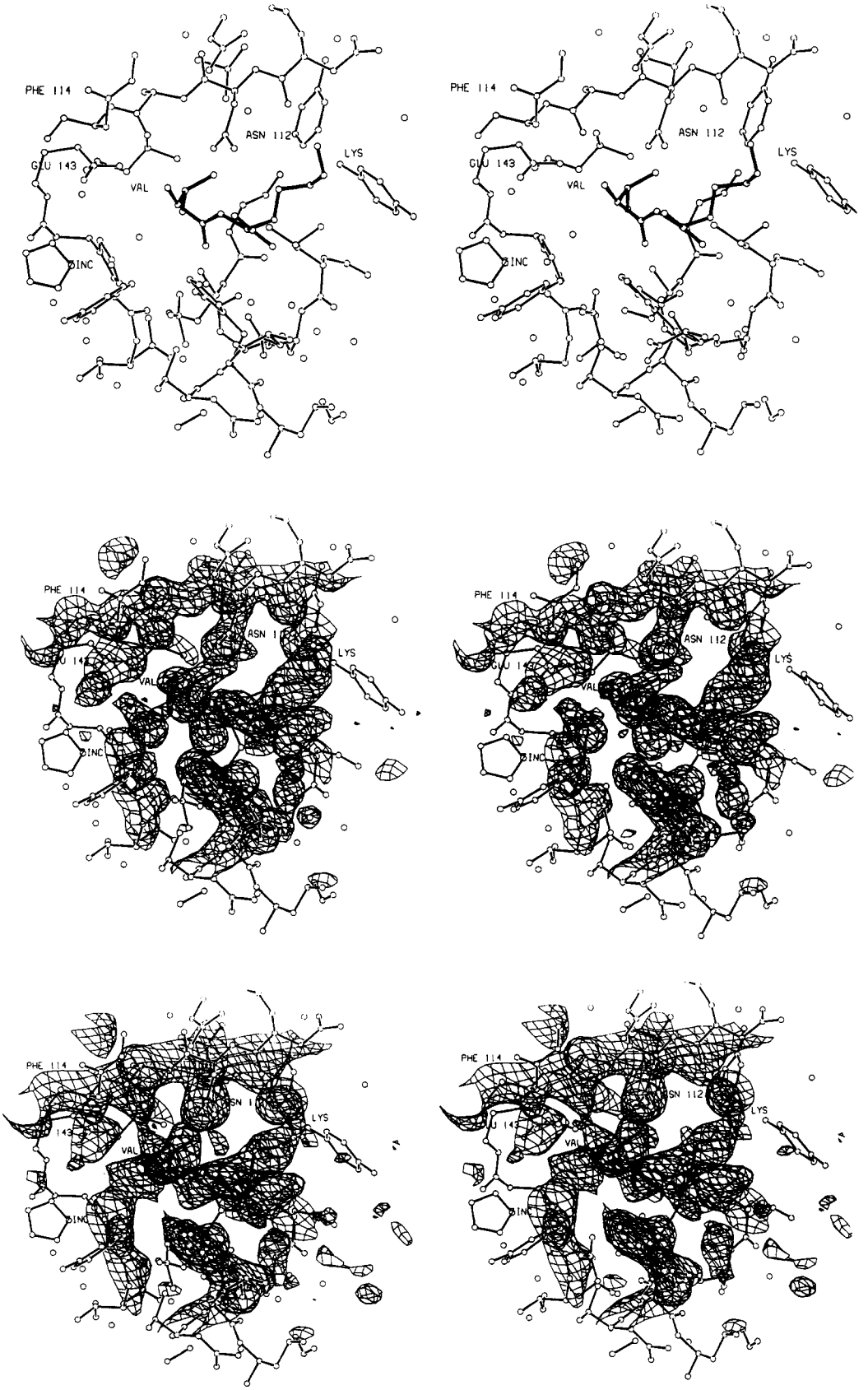

FIGURE 6: Stereo plots showing the evidence for the binding of a dipeptide in the active site of what was previously considered "native" thermolysin. The atoms in all plots are from the most highly refined model (D. E. Tronrud, unpublished results). (a) Refined model of thermolysin in the active-site region with the apparent bound dipeptide now included in the atomic model as Val-Lys and drawn with thicker bonds. (b) Superposition of the refined model on a $(2F_o - F_c)$ density map (20–1.6-Å resolution) contoured at 1 σ, where all atoms were included in the calculation of amplitudes, F_c, and phases. The electron density map indicates the presence of the putative Val-Lys dipeptide, but such maps can be subject to bias because the dipeptide has been included in the phase calculation. (c) To provide evidence without model bias, this stereo drawing shows the multiple isomorphous replacement map (20–2.3-Å resolution) calculated using the original native and heavy-atom data measured from precession films (Matthews et al., 1974). The clear electron density corresponding to an apparent bound dipeptide is strong evidence that the active site of crystalline thermolysin is occupied.

might have been cleaved from the intact protein during protein purification or during crystallization. In the initial structure determination of thermolysin, electron density was observed in the active site of the enzyme and was assumed to be due to bound solvent and a molecule such as Tris, which was present in the crystallizing medium (Kester & Matthews, 1977). Also, in early studies in which crystals of thermolysin were soaked in dipeptide solutions the change in electron density at the active site was minimal (Kester & Matthews, 1977), consistent (in retrospect) with the idea that a dipeptide might have already been present in the crystals.

Prompted by these considerations, we recalculated and reexamined the original isomorphous replacement map (Figure 6c (bottom). The isomorphous replacement map has the advantage that it is completely free of the bias that can enter once one uses crystallographic phases based on a presumed structural model. As can be seen, the electron density in the active site does, in fact, correspond to a bound dipeptide.

Thus, it seems likely that the crystal structure of thermolysin, which was always assumed to be that of the free enzyme, actually represents the structure of an enzyme–inhibitor, or an enzyme–product, complex. Assuming this to be the case, the structure of free thermolysin remains to be determined.

This also leads one to reconsider the implications of the observation that the crystal structure of thermolysin has never been observed to change substantially on the addition of inhibitors. This might simply reflect the possibility that one thermolysin–ligand complex is being replaced with another. It has been observed on a number of occasions that the addition of a ligand to crystalline thermolysin causes the crystals to develop transient surface cracks which gradually disappear with time [e.g., Bolognesi and Matthews (1979), Monzingo and Matthews (1982), and Holden et al. (1987)]. This might now be interpreted as due to transient opening of the active site cleft in association with displacement of the presumed dipeptide already present in the crystals by a tighter-binding ligand.

Inspection of the most recent electron density maps of the refined structure of the NEU protease does not reveal any density in the active site region that resembles either a bound peptide or another molecule (except for a couple of apparent solvent molecules in the S_1' subsite). The available evidence therefore tends to confirm that this structure is of the free enzyme.

In the case of PAE, both inhibitor-bound and inhibitor-free structures have been determined for the same enzyme, and the change in hinge-bending angle in this case is 14° (Table II). For neither of the other two enzymes (TLN and NEU) is a direct comparison possible of the inhibitor-bound and inhibitor-free form of the same enzyme. However, because these two enzymes are closely related (73% sequence identity), it seems reasonable to compare one with the other. In this case, the difference in hinge-bending angle between the free (NEU) and bound (TLN) forms is only 6° (Table II). Possibly, hinge bending in the case of TLN and/or NEU is more restricted than in the case for PAE. Nishino and Powers (1980) observed that the potent inhibitor phosphoramidon binds with essentially equal affinity to TLN and PAE, but the kinetics of binding to TLN are much slower, consistent with a less flexible active site [cf. McKay et al. (1992)].

There are a number of examples indicating that enzymes are dynamic and that conformational change is an essential component of catalytic activity (Bennett & Huber, 1984; Ringe & Petsko, 1985). The present results add to this body of evidence.

REFERENCES

Bennett, W. S., & Huber, R. (1984) *Crit. Rev. Biochem. 15*, 291–384.

Bever, R. A., & Iglewski, B. H. (1988) *J. Bacteriol. 170*, 4309–4314.

Bolognesi, M. C., & Matthews, B. W. (1979) *J. Biol. Chem. 254*, 634–639.

Holden, H. M., Tronrud, D. E., Monzingo, A. F., Weaver, L. H., & Matthews, B. W. (1987) *Biochemistry 26*, 8542–8553.

Kester, W. R., & Matthews, B. W. (1977) *Biochemistry 16*, 2506–2516.

Matthews, B. W. (1988) *Acc. Chem. Res. 21*, 333–340.

Matthews, B. W., Weaver, L. H., & Kester, W. R. (1974) *J. Biol. Chem. 249*, 8030–8044.

McKay, D. B., Thayer, M. M., Flaherty, K. M., Pley, H., & Benvegnu, D. (1992) in *Matrix Metalloproteinases and Inhibitors* (Birkedal-Hansen, H., Werb, Z., Welgus, H., & Van Wart, H., Eds.) pp 112–115, Gustav Fischer Verlag, New York.

Monzingo, A. F., & Matthews, B. W. (1982) *Biochemistry 21*, 3390–3394.

Nishikawa, K., Ooi, T., Isogai, Y., & Saito, N. (1972) *J. Phys. Soc. (Japan) 32*, 1331–1337.

Nishino, N., & Powers, J. C. (1980) *J. Biol. Chem. 255*, 3482–3486.

Paupit, R. A., Karlsson, R., Picot, D., Jenkins, J. A., Niklaus-Reimer, A. S., & Jansonius, J. N. (1988) *J. Mol. Biol. 199*, 525–537.

Ringe, D., & Petsko, G. A. (1985) *Prog. Biophys. Mol. Biol. 45*, 197–235.

Rossmann, M. G., & Argos, P. (1975) *J. Biol. Chem. 250*, 7525–7532.

Sidler, W., Niederer, E., Suter, F., & Zuber, H. (1986) *Biol. Chem. Hoppe-Seyler 367*, 643.

Stark, W., Pauptit, R. A., Wilson, K., & Jansonius, J. N. (1992) *Eur. J. Biochem. 207*, 781–791.

Thayer, M. M., Flaherty, K. M., & McKay, D. B. (1991) *J. Biol. Chem. 266*, 2864–2871.

Titani, K., Hermodson, M. A., Ericsson, L. H., Walsh, K. A., & Neurath, H. (1972) *Nature, New Biol. 238*, 35–37.