Lockard, R. E., & RajBhandary, U. L. (1976) *Cell (Cambridge, Mass.) 9*, 747-760.
Lockard, R. E., Alzner-Deweerd, B., Heckman, J. E., MacGee, J., Tabor, M. W., & RajBhandary, U. L. (1978) *Nucleic Acids Res. 5*, 37-56.
Lockard, R. E., Connaughton, J. F., & Kumar, A. (1982) *Nucleic Acids Res. 10*, 3445-3457.
Maly, P., & Brimacombe, R. (1983) *Nucleic Acids Res. 11*, 7263-7286.
Mandal, C., Kallenbach, N. R., & Englander, S. W. (1979) *J. Mol. Biol. 135*, 391-411.
Mizushima, S., & Nomura, M. (1970) *Nature (London) 226*, 1214-1217.
Moazed, D., Stern, S., & Noller, H. F. (1986a) *J. Mol. Biol. 187*, 399-416.
Moazed, D., Van Stolk, B. J., Douthwaite, S., & Noller, H. F. (1986b) *J. Mol. Biol. 191*, 483-493.
Peattie, D. A., & Gilbert, W. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 4679-4682.
Pleij, C. W. A., Rietveld, K., & Bosch, L. (1985) *Nucleic Acids Res. 13*, 1717-1731.
Pörschke, D. (1974) *Biophys. Chem. 2*, 97-101.
Qu, L.-H., Michot, B., & Bachellerie, J. P. (1983) *Nucleic Acids Res. 11*, 5903-5920.
Seeman, N. C., Rosenberg, J. M., & Rich, A. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 804-808.
Sigmund, C. D., Ettayebi, M., & Morgan, E. A. (1984) *Nucleic Acids Res. 12*, 4653-4663.
Spangler, E. A., & Blackburn, E. H. (1985) *J. Biol. Chem. 260*, 6334-6340.
Stern, S., Wilson, R. C., & Noller, H. F. (1986) *J. Mol. Biol. 192*, 101-110.
Tinoco, I., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M., & Gralla, J. (1973) *Nature (London), New Biol. 246*, 40-41.
Van Stolk, B. J., & Noller, H. F. (1984) *J. Mol. Biol. 180*, 151-177.
Vincze, A., Henderson, R. E. L., McDonald, J. J., & Leonard, N. J. (1973) *J. Am. Chem. Soc. 95 (8)*, 2677-2682.
Wittman, H. G. (1982) *Annu. Rev. Biochem. 51*, 35-65.
Wool, I. G. (1980) in *Ribosomes, Structure, Function, and Genetics* (Chambliss, G., Ed.) pp 797-824, University Park Press, Baltimore, MD.

Primary Structure Homologies between Two Zinc Metallopeptidases, the Neutral Endopeptidase 24.11 ("Enkephalinase") and Thermolysin, through Clustering Analysis†

T. Benchetrit,‡ V. Bissery,§ J. P. Mornon,§ A. Devault,∥ P. Crine,∥ and B. P. Roques*,‡

Département de Chimie Organique, U 266 INSERM, UA 498 CNRS, UER des Sciences Pharmaceutiques et Biologiques, 4 avenue de l'Observatoire, 75006 Paris, France, Laboratoire de Minéralogie et Cristallographie, Université Paris VI, 4 Place Jussieu, 75005 Paris, France, and Département de Biochimie, Université de Montréal, Montréal, Québec, Canada H3C 3J7

Received April 28, 1987; Revised Manuscript Received August 20, 1987

ABSTRACT: Analogies in the sequences of two related zinc metallopeptidases, the bacterial thermolysin (316 amino acids) and the recently cloned neutral endopeptidase 24.11 ("enkephalinase", 749 amino acids), have been demonstrated by a hydrophobic cluster analysis method derived from the Lim theory. Two sequence alignments are proposed for the entire primary structure of thermolysin and the C-terminal part of endopeptidase 24.11. Except for an arginine residue, all the amino acids involved in the active site of thermolysin have been retrieved in both models of endopeptidase 24.11 within conserved clustered structures. The first model is characterized by a deletion of the Ca^{2+}-binding coil present in thermolysin and the second by replacement of this coil by two α-helices. In both models an Arg residue can be located in the active site of the neutral endopeptidase.

Neutral endopeptidase, NEP (EC 3.4.24.11), initially characterized in the kidney brush border (Kerr & Kenny, 1974) was also subsequently shown to be present in the brain (Malfroy et al., 1978). The enzyme, localized in the vicinity of opioid receptors (Waksman et al., 1986), is involved in the physiological inactivation of the endogenous opioid peptides methionine- and leucine-enkephalin (Roques et al., 1980), and is therefore often called "enkephalinase".

NEP belongs to an important physiological group of Zn metallopeptidases (Vallee & Galdes, 1984), such as angiotensin converting enzyme, ACE (Das & Soffer, 1975), and collagenase (Seltzer et al., 1977). Two of the enzymes from this group, carboxypeptidase A (Rees et al., 1983) and thermolysin, TLN (Holmes & Matthews, 1982), have been crystallized, alone or in the presence of various substrates or inhibitors. A simplified model of the active site of Zn metallopeptidases, derived from these crystallographic studies, was used by Ondetti et al. (1977) for the rational design of ACE inhibitors such as captopril, which act as potent antihypertensive drugs. Taking into account the similarities in the enzymatic activity of NEP and TLN (Kerr & Kenny, 1974; Llorens et al., 1980; Benchetrit et al., 1987), potent inhibitors of NEP have also been developed, leading to a new class of potential analgesic drugs (Chipkin, 1986; Roques & Fournié-Zaluski, 1986).

Recently, the amino acid sequence of NEP (749 amino acids), deduced from its complementary DNA (Devault et al.,

† This research is supported by grants from Université René Descartes, the Centre National de la Recherche Scientifique, and the Institut National de la Santé et de la Recherche Médicale.
‡ UER des Sciences Pharmaceutiques et Biologiques.
§ Université Paris VI.
∥ Université de Montréal.

0006-2960/88/0427-0592$01.50/0

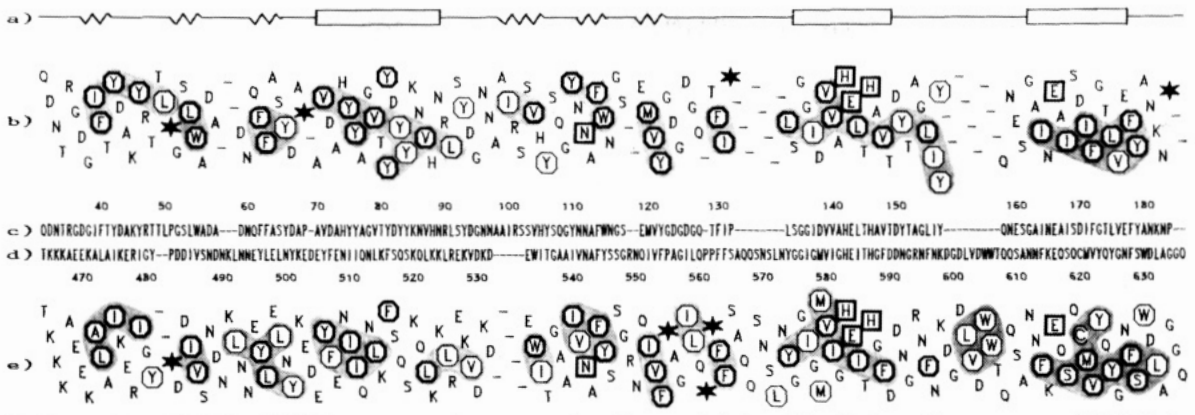

FIGURE 1: Comparison of TLN and NEP sequences: the common domain to models I and II. (a) Secondary structure of TLN derived from crystallography. Zigzags represent β-strands and boxes α-helices. (b) HCA plot of TLN sequence; (c) linear representation of TLN sequence; (d) linear representation of NEP sequence; (e) HCA plot of NEP sequence. Amino acids are given in the single-letter code except for prolines (black star), and insertions are shown by (–). Active site residues are squared, and hydrophobic amino acids are surrounded by a heavy octagon when they are present in both proteins and in other cases by a thin octagon. Hydrophobic clusters are circled and filled in gray; to allow continuity of clusters, some residues were drawn twice.

1987; Malfroy et al., 1987), was compared to the primary structure of TLN (316 amino acids) showing that, except for two highly homologous short regions, the sequences of these two enzymes are different (Devault et al., 1987). However, most of the residues involved in the catalytic sites of both enzymes appear to be conserved.

The knowledge of the 3D structure of NEP should facilitate the design of even more potent and specific inhibitors. Such a strategy has been recently illustrated for ACE (Hangauer et al., 1984; Mayer et al., 1987). However, the computer prediction of the tertiary structure of a protein from the 3D structure of a related protein depends on identification of analogies in their secondary structures, motifs, domains, or ligand interactions (Blundell et al., 1987).

Traditional alignment techniques (Waterman, 1984) appear to be well adapted to sequences that show more than 30% strict homology, while more sophisticated methods have been proposed for sequences with the case of poor analogies (Lewis & Scheraga, 1971; Cid et al., 1982; Novotny & Auffray, 1984; Bryant & Sternberg, 1987).

Thus in order to reveal the apparently hidden similarities between NEP and TLN, a simpler method (Gaboriaud et al., 1987) based on hydrophobic cluster analysis (HCA), derived from the theory of Lim (1974), has been used. Briefly, the sequences of isofunctional proteins such as TLN and NEP were drawn on a theoretical α-helix where most of the hydrophobic residues display a tendency to cluster. The characteristics of the clusters, including their size and their shape, were then compared, allowing an alignment of the sequences.

This method shows the occurrence of large similarities in the clustered sequences of TLN and NEP. By use of these results and by taking into account the lack of calcium cations in NEP (Kerr & Kenny, 1974), two possible sequence alignments have been proposed. The first model is characterized by a deletion of the Ca^{2+}-binding coil present in TLN and the second model by replacement of this coil by two α-helices. Both models suggest a significant structural analogy between TLN and NEP at their active sites.

Materials and Methods

The present study was mainly performed by a detailed hydrophobic cluster analysis, HCA (Gaboriaud et al., 1987), derived from the method of Lim (1974). Lim's theory postulates that before folding into its native structure a globular protein roughly takes on an overall α-helical conformation. At the surface of this native structure, most of the hydrophobic residues cluster in a limited number of areas, leading to a final regular secondary structure, characterized by α-helices and β-strands or combinations of both.

It has often been observed that the secondary and tertiary structures of related proteins are far more conserved than the sequences, e.g., the aspartic (Carlson et al., 1984) and serine proteases (Greer, 1981). A major advantage of the HCA is that it allows the alignment of two related proteins even if the sequences show very little homology.

To simplify the method, one main program, HCAplot, has been written in Fortran 77. The sequence was drawn on a classical α-helix (3.6 amino acids per turn), and the helix was then cut parallel to its axis and unrolled.

To help the visual inspection of the pattern, hydrophobic residues (I, L, F, W, M, Y, V, and sometimes A and S) were circled and prolines symbolized by a black star. Alanine accepts a hydrophilic environment and therefore may or may not participate in hydrophobic clusters. Serine could be inserted in a hydrophobic environment as it could be linked by a H bond inside the main chain of α-helices, therefore masking its OH group. Finally hydrophobic clusters, composed of adjacent hydrophobic residues not separated by prolines, were visualized.

To allow an easier amino acid by amino acid comparison of the aligned sequences, the linear primary structures of the proteins were drawn alongside their HCA representations. A quantitative index to evaluate the correspondence between two HCA patterns (matching score) was defined as follows: for the hydrophobic clusters that can be aligned, the quantity $(2CR \times 100)/(RC1 + RC2)$ is calculated, where RC1 and RC2 are the number of hydrophobic residues in the clusters of sequence 1 and sequence 2, respectively, and CR is the number of residues in the clusters that are similarly positioned.

Results

Thermolysin HCA Pattern. The thermolysin HCA pattern was made up of 21 clusters, approximately one for every 15 residues (Figures 1b and 2b). Their shapes and dimensions were quite different, but two types predominated. Type I clusters were rather horizontal and were composed of more than 10 residues, such as clusters V71–L91, I164–Y179, or L263–L275. Type II clusters are rather vertical and comprise less than 10 residues, such as Y110–W115 or M120–Y122. A correlation of the cluster shapes with their known secondary structure, derived from crystallographic analysis (Holmes & Matthews, 1982), showed that type I's corresponded essentially

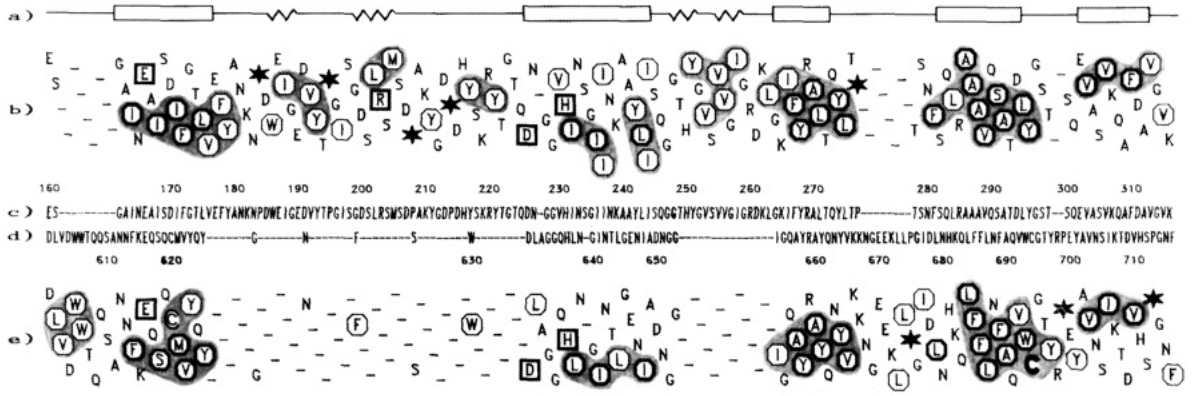

FIGURE 2: Comparison of TLN and NEP sequences: the C-terminal part of model I. (a) Secondary structure of TLN derived from crystallography; (b) HCA plot of TLN sequence; (c) linear representation of TLN sequence; (d) linear representation of NEP sequence; (e) HCA plot of NEP sequence. For details of symbols, see Figure 1.

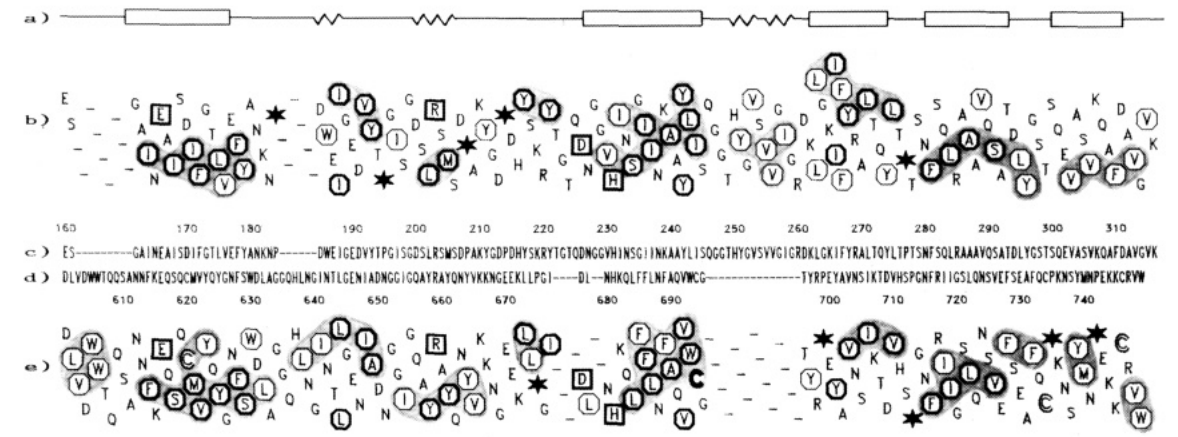

FIGURE 3: Comparison of TLN and NEP sequences: the C-terminal part of model II. (a) Secondary structure of TLN derived from crystallography; (b) HCA plot of TLN sequence; (c) linear representation of TLN sequence; (d) linear representation of NEP sequence; (e) HCA plot of NEP sequence. For symbols, see Figure 1.

to α-helices and type II's to β-strands.

Nevertheless, some exceptions to this rule appeared, especially in horizontal clusters where Tyr residues are surrounded in the sequence by hydrophilic moieties. For instance, the I40–L50 or Y106–W115 sequence (Figure 1b) does not correspond to α-helices but contains β-strands followed or preceeded by turns or ω-loops. Likewise, L133–Y157 (Figure 1b) exists as an α-helix and an extended strand due to the presence of Tyr-151 surrounded by Asp-150 and Thr-152.

Conversely, some helices, such as V230–I244 or F281–Y296 (Figure 2b), were not detected by the HCA plot. However, in these cases, independent clusters could be linked with alanines which are surrounded by hydrophobic residues in the HCA plot as occurs for Ala-240 in the sequence V230–I244.

In addition, Ser-234 and Ser-291 (Figure 2b), which are included in α-helices, have to be considered as hydrophobic residues. The hydroxyl groups of these amino acids form H bonds with the oxygen of the carbonyl of His-231 (d = 3.3 Å) and Ala-287 (d = 2.8 Å), respectively, avoiding the formation of hydrophilic interaction outside the helix core. These results were taken into account when the clusters of TLN and NEP were aligned.

Neutral Endopeptidase 24.11 HCA Pattern. The HCA pattern of the entire sequence of NEP was drawn and analyzed although only the plot of the last 300 residues is presented here (Figures 1e, 2e, and 3e).

The N-terminal domain, G1–T464, contains a large and compact hydrophobic cluster, L26–C56, previously proposed as a membrane spanning domain (Fulcher et al., 1986; Devault et al., 1987; Malfroy et al., 1987). The HCA plot of the N-terminal part of NEP suggested the occurrence of the α-helices C80–I96, Y103–L121, I269–I283, L305–I316, P367–Y386, and V439–M461, whereas L156-L160, V163–W166, V198-V205, V344-Y347, Y351-L355, and I358-F363 could correspond to β-strands. It should be noticed that these predictions are in good agreement with the secondary structure predictions obtained from the GOR method (Garnier et al., 1978).

The C-terminal domain of NEP, T464–W749 (Figures 1e, 2e, and 3e), showed clusters and sequences with significant homologies with thermolysin. Therefore, the results are presented as a comparison with the primary structure of TLN.

Comparison of HCA Plots for Thermolysin and the C-Terminal Domain of Neutral Endopeptidase. The sequences G575–H587 and G634–N642 in NEP were initially suggested to be homologous to the sequences G135–H146 and G228–N238 in TLN (Devault et al., 1987). The shape of the cluster G575–H587 of NEP shows a very large homology with the G135–H146 segment (Figure 1) of TLN. Moreover, these sequences contain the three residues thought to be conserved in the active sites: H142, E143, and H146 in TLN and H583, E584, and H587 in NEP. Taking into account the presence of this highly conservative segment in both proteins, the alignment was performed cluster by cluster starting from this segment to both the N- and C-terminal parts of the T464–W749 domain. By use of this approach, two models (I and II) can be proposed.

In both models (Figures 1 and 2 and Figures 1 and 3), one TLN segment, Q246–R260, has been deleted to improve the alignment between TLN and NEP clusters. Moreover, the

NEP D599–T607 sequence needs to be inserted near Y157 in TLN. Given this arrangement, the essential amino acids present in the active site of TLN, E166 and N112, probably correspond to E616 and N542 in NEP. In addition, the asparagines (N112 and N542) appear in the enzymes at the center of a well-conserved local sequence.

In TLN, the A180–Q225 sequence is stabilized by three Ca^{2+} ions. In NEP devoid of Ca^{2+} ions, the corresponding sequence was either deleted in model I (Figure 2) or structured by two large hydrophobic clusters, very likely α-helical in nature, in model II (Figure 3). Therefore, in the first model of NEP, there is no evident equivalent to the residue Arg-203, present in the A180–Q225 sequence of TLN, while in model II this residue could correspond to Arg-659.

In addition, in model I, the sequence K667–P675 of NEP was inserted in TLN near the residue P277 to improve the cluster alignment between both proteins.

The validity of the proposed alignment can be evaluated through the cluster matching score. For instance, the score of cluster 133–148 (TLN numeration) containing a part of the catalytic site of the enzymes is 77%. In the calculation, the deleted and/or inserted segments for each part of the alignment were not taken into account. Under these conditions, the score of the common domain (Figure 1) reached 70%, while for the uncommon domains the values are 75% and 67% for model I (Figure 2) and model II (Figure 3), respectively. The slightly lower percentage obtained for model II is essentially due to the Ca^{2+}-binding coil, since the exclusion of this sequence increased the score to 75%.

Discussion

The putative knowledge of the tertiary structure of the active site of proteases, such as ACE (Hangauer et al., 1984; Mayer et al., 1987) or renin (Carlson et al., 1984), two enzymes involved in the contraction of arterial tissues, represents for many pharmaceutical companies an attractive challenge to develop highly potent and selective inhibitors. Likewise, the implication of the neutral endopeptidase in the control of natural analgesia explains the intensive effort devoted to the development of NEP inhibitors (Chipkin, 1986; Roques & Fournié-Zaluski, 1986). Until now, NEP has not been crystallized, and its 3D structure remains unknown, although recently its sequence has been determined by cDNA sequencing (Devault et al., 1987; Malfroy et al., 1987). As a number of biological studies have emphasized the similarities between NEP and thermolysin (Kerr & Kenny, 1974; Beaumont & Roques, 1986; Benchetrit et al., 1987), a bacterial metallopeptidase whose 3D structure is known (Holmes & Matthews, 1982), the modeling of the tertiary structure of NEP could now be possible. However, such studies are subject to error, even if the sequences of both proteins show large homologies (Delbaere et al., 1979). In the case of TLN and NEP, the linear sequence homologies are low (Devault et al., 1987), and improved methods, taking into account secondary structures and hydrophobic patterns (Lewis & Scheraga, 1971; Cid et al., 1982; Novotny & Auffray, 1984; Bryant & Sternberg, 1987) such as HCA, were required to demonstrate the similarities in the 3D structures of both enzymes.

The folding of a protein leading to its tertiary structure depends both on its secondary structure and on interactions between structured sequences. The HCA method, based on Lim's theory (Lim, 1974), yields the structured sequences and the location of the hydrophobic residues, which induce interactions between structural elements, in these sequences. Obviously, the HCA method is unable to directly predict the folding of a protein, but it allows the sequences of proteins to be compared and the tertiary structure of a protein to be predicted from the 3D structure of another related one. The efficiency of the HCA method has been checked with structures obtained from the Protein Data Bank or the literature (Gaboriaud et al., 1987).

The HCA plots of TLN and NEP were drawn and visually aligned. Two models have been proposed, differing in the manner in which the stabilization of TLN by four Ca^{2+} cations was taken into account. In TLN, one segment binds three of these ions. In NEP, this segment appears not to be conserved, and in model I, it was simply deleted. Interestingly in a comparison of TLN with another TLN-like enzyme, *Bacillus subtilis* neutral endopeptidase, which contains only three calcium atoms, the deletion of a part of the corresponding coil from this latter enzyme was also proposed (Sidler et al., 1986). The sequence D185–Q225 forms a large coil of roughly 40 residues. Its deletion would induce large modifications, such as a replacement of the arginine residue present in the active site, changes in the hydrophobic environment of clusters 164–179 and 232–244 (TLN numeration), and the joining of these segments. In NEP, the cluster F614–Y625, probably helical, could be cut after residue Y625, corresponding to residue L175 in TLN. Thus, the sequence Y625–D631 would be long enough to join the clusters, since a distance of 20 Å separates residues L175 and D226 in TLN. The residue C620 of NEP, corresponding to residue D170 of TLN, is located in the proximity of the glutamate residue E616 present in the active site. The investigation of the structure of TLN indicates that after deletion of the coil the side chain of D170 points outside the core of the protein. Assuming the same type of structural arrangement in NEP, the C620 residue could form a disulfide bridge, bringing an essential arginine amino acid into the active site. Indeed, disulfide bridges have already been shown to play a important structural role in the endopeptidase 24.11 (Tam et al., 1985), confirmed by the even number of cysteines in the primary structure of NEP (Devault et al., 1987; Malfroy et al., 1987). In these conditions, three arginines with a vicinal cysteine moiety (Arg-408, Arg-409, and Arg-747) could assume the role of TLN Arg-203.

In model II, it seemed probable that, to compensate for the lack of ion binding in NEP, the sequence G634–I677 would be structured. The analysis of this segment (Figure 3) indeed seems to indicate the probable occurrence of two successive α-helices, one of them containing Arg-659. A preliminary modeling of the structure of neutral endopeptidase indicates that the joining of the segments 164–179 and 232–244 (TLN numeration) by these helices with superimposition of the side chains of TLN Arg-203 and NEP Arg-659 is possible.

Other less essential structural modifications have also been investigated. The residue Y157 of TLN belongs to a loop between α-helices, and the proposed insertion of the sequence D599-T607 in NEP (Figure 1) only raises the problem of the folding of this segment on the protein core.

The Cα's of residues S245 and R260 are separated by 5 Å in TLN, and the deletion of the β-hairpin Q246–R260 of TLN induces no large structural modification. Interestingly, we have noticed that in *B. subtilis* endopeptidase (Sidler et al., 1986) this sequence is also deleted. Finally, in model I, the NEP segment K667–P675 (Figure 2) has been inserted near Pro-277 in TLN. In the 3D structure of TLN, this residue is near Y157, and therefore the inserted segments D599–T607 and K667–P675 could interact to stabilize themselves.

These models could explain some of the differences in biological activity between TLN and NEP. The hydrophobic pocket essential for binding of inhibitors or substrates seems

to be larger in NEP than in TLN. In TLN, this pocket is composed of residues F130, L133, V139, I188, G189, and L202 belonging to segments that are strongly modified in both of our NEP models: TLN segment F130–V139 is larger in NEP (F563–V580), and TLN segment I188–L202, which belongs to the Ca^{2+}-binding coil, is deleted or restructured in NEP. Moreover, TLN functions mainly as an endopeptidase while NEP has been shown to have both endopeptidase and dipeptidylcarboxypeptidase activities (Malfroy & Schwartz, 1985). This behavior of NEP has been attributed to the occurrence, in the active site, of an arginine residue mimicking the Arg-145 of carboxypeptidase A. In the case of NEP, all the essential amino acids of the active site are in close correspondence with those of TLN except for the Arg residue, a finding that may explain the dual enzymatic activity of NEP.

In conclusion, the use of HCA has allowed the sequence alignment of two isofunctional enzymes having low sequence homologies, TLN and NEP. Two models showing the existence of structural similarities between both peptidases have been proposed. Biochemical studies such as the determination of disulfide bridges, labeling of the active site, and/or site-directed mutagenesis are now required to enable a choice to be made between the two. These studies as well as the modeling of the tertiary structure of NEP by computer calculations coupled with graphic analysis are now in progress in our laboratory.

ACKNOWLEDGMENTS

We are grateful to C. Gaboriaud (Université Paris VI) for helpful discussions on HCA method and to Professor M. C. Fournié-Zaluski for valuable comments on the proposed models. We thank Dr. A. Beaumont for stylistic revisions and A. Bouju for typing the manuscript. Calculations were performed at the Université Paris VI.

REFERENCES

Beaumont, A., & Roques, B. P. (1986) *Biochem. Biophys. Res. Commun. 139*, 733–739.

Benchetrit, T., Fournié-Zaluski, M. C., & Roques, B. P. (1987) *Biochem. Biophys. Res. Commun. 147*, 1034–1040.

Blundell, T. L., Sibanda, B. L., Sternerg, M. J. E., & Thornton, J. M. (1987) *Nature (London) 326*, 347–352.

Bryant, S. H., & Sternberg, M. J. E. (1987) *J. Mol. Graphics 5*, 4–7.

Carlson, W., Karplus, M., & Haber, E. (1985) *Hypertension (Dallas) 7*, 13–26.

Chipkin, R. E. (1986) *Drugs Future 11*, 593–606.

Cid, H., Bunster, M., Arriagada, E., & Campos, M. (1982) *FEBS Lett. 150*, 247–254.

Das, M., & Soffer, R. L. (1975) *J. Biol. Chem. 250*, 6762–6768.

Delbaere, L. T., Brayer, G. D., & James, M. N. G. (1979) *Nature (London) 279*, 165–168.

Devault, A., Lazure, C., Nault, C., Le Moual, H., Seidah, N. G., Chrétien, M., Kahn, P., Powell, J., Mallet, J., Beaumont, A., Roques, B. P., Crine, P., & Boileau, G. (1987) *EMBO J. 6*, 1317–1322.

Fournié-Zaluski, M. C., Lucas, E., Waksman, G., & Roques, B. P. (1984) *Eur. J. Biochem. 139*, 267–274.

Fulcher, I. S., Pappin, D. J. C., & Kenny, A. J. (1986) *Biochem. J. 240*, 305–308.

Gaboriaud, C., Bissery, V., Benchetrit, T., & Mornon, J. P. (1987) *FEBS Lett. 224*, 149–155.

Garnier, J., Osguthorpe, J., & Robson, B. (1978) *J. Mol. Biol. 120*, 97–120.

Greer, J. (1981) *J. Mol. Biol. 153*, 1027–1042.

Hangauer, D. G., Monzingo, A. F., & Matthews, B. W. (1984) *Biochemistry 23*, 5730–5741.

Holmes, M. A., & Matthews, B. W. (1982) *J. Mol. Biol. 160*, 623–639.

Kerr, M. A., & Kenny, A. J. (1974) *Biochem. J. 137*, 477–488.

Leszozynski, J. F., & Rose, G. D. (1986) *Science (Washington, D.C.) 234*, 849–855.

Lewis, P. N., & Scheraga, H. A. (1971) *Arch. Biochem. Biophys. 144*, 584–588.

Lim, V. I. (1974) *J. Mol. Biol. 88*, 857–872.

Llorens, C., Gacel, G., Swerts, J. P., Perdrisot, R., Fournié-Zaluski, M. C., Schwartz, J. C., & Roques, B. P. (1980) *Biochem. Biophys. Res. Commun. 96*, 1710–1715.

Malfroy, B., & Schwartz, J. C. (1985) *Biochem. Biophys. Res. Commun. 130*, 372–378.

Malfroy, B., Swerts, J. P., Guyon, A., Roques, B. P., & Schwartz, J. C. (1978) *Nature (London) 276*, 523–526.

Malfroy, B., Schofield, P. R., Kuang, W. J., Seeburg, P. H., Mason, A. J., & Henzel, W. J. (1987) *Biochem. Biophys. Res. Commun. 144*, 59–66.

Mayer, D. M., Naylor, C. B., Motoc, I., & Marshall, G. R. (1987) *J. Comput.-Aided Mol. Des. 1*, 3–16.

Novotny, J., & Auffray, C. (1984) *Nucleic Acids Res. 12*, 243–255.

Ondetti, M. A., Rubin, B., & Cushman, D. W. (1977) *Science (Washington, D.C.) 196*, 441–444.

Rees, D. C., Lewis, M., & Lipscomb, W. N. (1983) *J. Mol. Biol. 168*, 367–387.

Roques, B. P., & Fournié-Zaluski, M. C. (1986) *NIDA Res. Monogr. No. 70*, 128–154.

Roques, B. P., Fournié-Zaluski, M. C., Soroca-Luca, E., Lecomte, J. M., Malfroy, B., Llorens, C., & Schwartz, J. C. (1980) *Nature (London) 288*, 286–288.

Seltzer, J. L., Jeffrey, J. J., & Eisen, A. Z. (1977) *Biochim. Biophys. Acta 485*, 179–187.

Sidler, W., Neiderer, E., Suter, F., & Zuber, H. (1986) *Hoppe-Seyler's Z. Physiol. Chem. 367*, 643–657.

Tam, L., Engelbrecht, S., Talent, J. M., Gracy, R. W., & Erdös, E. G. (1985) *Biochem. Biophys. Res. Commun. 133*, 1187–1192.

Vallee, B. L., & Galdes, A. (1984) *Adv. Enzymol. Relat. Areas Mol. Biol. 56*, 283–430.

Vaney, M. C., Surcouf, E., Cherfils, J., Morize, I., & Mornon, J. P. (1985) *J. Mol. Graphics 3*, 123–124.

Waksman, G., Hamel, E., Fournié-Zaluski, M. C., & Roques, B. P. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 1523–1527.

Waterman, M. S. (1984) *Bull. Math. Biol. 46*, 473–500.