

# Structure of neurolysin reveals a deep channel that limits substrate access

C. Kent Brown\*<sup>†</sup>, Kevin Madauss\*, Wei Lian<sup>‡</sup>, Moriah R. Beck<sup>§</sup>, W. David Tolbert<sup>¶</sup>, and David W. Rodgers<sup>||</sup>

Department of Molecular and Cellular Biochemistry and Center for Structural Biology, University of Kentucky, Lexington, KY 40536

Communicated by Stephen C. Harrison, Harvard University, Cambridge, MA, December 29, 2000 (received for review November 14, 2000)

**The zinc metallopeptidase neurolysin is shown by x-ray crystallography to have large structural elements erected over the active site region that allow substrate access only through a deep narrow channel. This architecture accounts for specialization of this neuropeptidase to small bioactive peptide substrates without bulky secondary and tertiary structures. In addition, modeling studies indicate that the length of a substrate N-terminal to the site of hydrolysis is restricted to approximately 10 residues by the limited size of the active site cavity. Some structural elements of neurolysin, including a five-stranded  $\beta$ -sheet and the two active site helices, are conserved with other metallopeptidases. The connecting loop regions of these elements, however, are much extended in neurolysin, and they, together with other open coil elements, line the active site cavity. These potentially flexible elements may account for the ability of the enzyme to cleave a variety of sequences.**

**M**ore than 100 biologically active peptides, ranging in size from 2 to 40 residues, have been identified (1). They play a variety of roles in the nervous and endocrine systems through their interaction with cell surface receptors and subsequent modulation of signaling pathways. The signals from neuropeptides and peptide hormones are terminated by extracellular proteolysis or by internalization of the peptide or peptide-receptor complex followed by intracellular degradation (2, 3). A number of peptidases that participate in neuropeptide degradation or modification have been identified and shown to be either intracellular, associated with the plasma membrane, secreted, or present at more than one of these locations. These neuropeptidases can be either exo- or endopeptidases, and nearly all are metallopeptidases that use a zinc ion cofactor. Inhibitors of neuropeptidases have been used in experimental studies to modulate neuropeptide levels (4, 5), and some therapeutic approaches are based on inhibition of neuropeptidases. One well-known example is treatment of hypertension with inhibitors of angiotensin-converting enzyme (6).

A number of neuropeptidases share two unusual properties: they are strict oligopeptidases—that is they hydrolyze only short peptides—and they cleave at a limited set of sites that are nonetheless diverse in sequence (1, 2). One neuropeptidase that exemplifies these properties is neurolysin (EC 3.4.24.16), a zinc metalloendopeptidase that functions as a monomer of molecular mass 78 kDa (7, 8). *In vitro*, neurolysin cleaves a number of bioactive peptides at sequences that vary widely, and its longest known substrate is only 17 residues in length. The enzyme belongs to the M3 family of metallopeptidases (9) along with eight other known peptidases that share extensive sequence homology, including the closely related (60% sequence identity) thimet oligopeptidase (EC 3.4.24.15). Enzymes in the M3 family share with several other metallopeptidase families a common active site sequence motif, His-Glu-Xaa-Xaa-His (HEXXH), that forms part of the binding site for the metal cofactor (10). The two histidines of the motif coordinate the zinc ion, and the glutamate orients and polarizes a water molecule that is believed to act as the attacking nucleophile. Neurolysin is widely distributed in mammalian tissues (7) and is found in different subcellular locations that vary with cell type. Much of the enzyme is

cytosolic, but it also can be secreted or associated with the plasma membrane (11), and some of the enzyme is made with a mitochondrial targeting sequence by initiation at an alternative transcription start site (12).

Although neurolysin cleaves a number of neuropeptides *in vitro*, its most established (5, 13, 14) role *in vivo* (along with thimet oligopeptidase) is in metabolism of neurotensin, a 13-residue neuropeptide. It hydrolyzes this peptide between residues 10 and 11, creating shorter fragments that are believed to be inactive.

Neurotensin (pGlu-Leu-Tyr-Gln-Asn-Lys-Pro-Arg-Arg-Pro  $\downarrow$  Tyr-Ile-Leu) is found in a variety of peripheral and central tissues where it is involved in a number of effects, including modulation of central dopaminergic and cholinergic circuits, thermoregulation, intestinal motility, and blood pressure regulation (15). Neurotensin is also one of the most potent antinociceptive substances known (16), and an inhibitor of neurolysin has been shown to produce neurotensin-induced analgesia in mice (17).

The unusual properties of oligopeptidases are not well understood. Our goal is to define mechanisms of substrate selectivity in these enzymes and to use structure-based analysis to improve specific inhibitors.

## Materials and Methods

**Protein Production and Crystallization.** Production of recombinant neurolysin and growth of crystals have been reported (18). Selenomethionine-substituted protein was prepared according to Doublé (19) and purified in the same way as the native enzyme except for use of 15 mM 2-mercaptoethanol throughout the procedure. Crystallization of selenomethionine-substituted enzyme was done under anaerobic conditions.

**Crystallography.** Crystals grew to full size (0.15  $\times$  0.1  $\times$  0.03 mm) within 3 days. Their space group is P2<sub>1</sub>2<sub>1</sub>2 (a = 157.8 Å, b = 88.0 Å, c = 58.4 Å) with one molecule in the asymmetric unit. A methyl mercury chloride derivative was prepared by adding MeHgCl in solution to the crystallization drop (final concentration 0.2 mM) for 6 h before harvesting the crystals. All data sets (Table 1) were collected from flash-cooled crystals, which were exposed to 20% polyethylene glycol 400 in the normal crystal-

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.rcsb.org (code 111).

\*C.K.B. and K.M. contributed equally to this work.

<sup>†</sup>Present address: Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota, Minneapolis, MN 55455.

<sup>‡</sup>Present address: Department of Medicinal Chemistry, University of Florida, Gainesville, FL 32610.

<sup>§</sup>Present address: Department of Biochemistry, Washington University, St. Louis, MO 63130.

<sup>¶</sup>Present address: Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853.

<sup>||</sup>To whom reprint requests should be addressed. E-mail: rogers@focus.gws.uky.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

**Table 1. Summary of crystallographic data**

	MeHgCl*	Se-Met†	Native
Wavelength (Å)	1.5418	0.9793	1.0000
Resolution (Å)	20.0–3.5	20.0–3.5	20.0–2.3
Last shell (Å)	3.66–3.5	3.66–3.5	2.38–2.3
Avg. redundancy	6.55	6.67	3.90
$R_{\text{sym}}^{\ddagger}$ (last shell) (%)	11.0 (21.3)	17.4 (29.1)	4.2 (15.6)
$I/\sigma I$ (last shell)	16.7 (7.7)	11.6 (7.6)	23.9 (7.8)
Completeness (%)	99.1	98.0	99.0
Phasing power <sup>§</sup> (acent./cent.)	2.34/1.61	0.76/0.68	
$R_{\text{Cullis}}^{\parallel}$ (acent./cent.)	0.69/0.68	0.89/0.86	
$R_{\text{Cullis}}$ (anomalous)	0.79	0.91	

\*Methyl mercury chloride derivative (6 sites).

†Selenomethionine-substituted protein (20 sites).

$R_{\text{sym}} = \sum |I_j| - \langle I \rangle / \sum I_j$ .

<sup>§</sup>Phasing power =  $\langle F_h \rangle / E$ , where  $\langle F_h \rangle$  is the rms heavy atom structure factor and  $E$  is the residual lack of closure error.

$R_{\text{Cullis}} = \sum |F_{\text{PH}} \pm F_{\text{P}}| - F_{\text{H}} / \sum |F_{\text{PH}} \pm F_{\text{P}}|$ , where  $F_{\text{P}}$  and  $F_{\text{PH}}$  are the structure amplitudes of the parent and the heavy-atom derivative and  $F_{\text{H}}$  is the calculated heavy-atom structure factor.

lization solution for approximately 10 sec, mounted in a nylon loop, and plunged into liquid nitrogen.

Data for the MeHgCl derivative were collected on an R-AXIS IV image plate detector with a CuK $\alpha$  source. Beamline B of the BioCARS facility at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL) was used to collect data from a crystal of selenomethionine-substituted neurolysin. Only the anomalous peak wavelength ( $\lambda = 0.9793$  Å) was collected because of time considerations. The data set for the native protein was collected at beamline C of the BioCARS facility on an Area Detector Systems Corporation (1) Q4 charge-coupled device detector. All data were reduced with the HKL package (20).

One mercury heavy atom site was identified from isomorphous difference Patterson maps produced by using the CCP4 package (21) and refined in MLPHARE (22). Subsequent sites (total of six) then were located on isomorphous and anomalous difference electron density maps and refined in MLPHARE. Selenium sites (total of 20 of 24 methionines in the molecule) were located on isomorphous difference maps by using density-modified (23) mercury phases and refined in MLPHARE. The overall figure of merit was 0.55- to 3.5-Å resolution and was improved to 0.76 with density modification. Maps made with these phases were readily interpreted, allowing an initial model to be built in O (24). Subsequent cycles of dynamics and minimization refinement in CNS (25), manual rebuilding, and addition of metal ions and water molecules gave an  $R_{\text{work}}$  of 22.4% and an  $R_{\text{free}}$  of 26.8% with good geometry (Table 2). The final model contains residues 14–678 of the 681 residues in the

expressed protein, 172 water molecules, and two zinc ions (one at the active site and one at a crystal packing interface).

## Results and Discussion

**Structure of Neurolysin.** We have determined the crystal structure to 2.3-Å resolution (Table 2) of recombinant rat neurolysin, a normally soluble neuropeptidase from the M3 family of metallopeptidases. The structure of the extracellular domain of neprilysin (EC 3.4.24.11), an integral membrane metallopeptidase (M13 family) that also functions as a neuropeptidase, was reported recently (26), and prolyl oligopeptidase (EC 3.4.21.26), a serine peptidase that processes neuropeptides, also has been described (27).

Neurolysin adopts an overall prolate ellipsoid shape with a dramatic deep cleft or channel running the length of the molecule (Fig. 1). The channel extends for about 60 Å and plunges to almost 40 Å below the top of the molecule at its deepest point. It effectively divides the enzyme into two large domains (labeled domain I and domain II) that are connected only at the floor of the cleft by a small number of secondary structural elements. The active site of the enzyme lies at the base of domain II near the bottom of the channel and about midway along its length. It is buried by the superstructure of the channel walls with only a narrow pathway for access, as illustrated in the cross-sectional view of the channel shown in Fig. 1C.

Neurolysin is predominately (53%)  $\alpha$ -helical in nature. The only  $\beta$  secondary structure, which accounts for just 5.9% of all residues, occurs as a five-stranded sheet midway along domain II and a small two-strand element at the end of domain I (Figs. 1A and 2A). The N-terminal 13 residues are disordered in the crystal structure, and possibly a portion of this flexible sequence mediates the reported (11, 28, 29) transport of neurolysin and its close homologue, thimet oligopeptidase, through secretory pathways in certain cell types. Residues 31–114 make up the first tertiary structural element, folding into a three-helix bundle ( $\alpha 1$ – $\alpha 3$ ) similar to those found in a number of other proteins. The neurolysin helical bundle, together with two subsequent helices, forms approximately half the channel wall in domain II. The polypeptide chain then crosses the base of the channel to fold into a series of long, roughly parallel helices (and two small  $\beta$ -strands) that make up the bulk of domain I. It crosses once again into domain II, forming the five-stranded sheet and its flanking helices, including the two active site helices ( $\alpha 16$  and  $\alpha 17$ ). The first of these active site helices contains the HEXXH motif, with its two zinc-coordinating histidines (H474, H478), and the second contains a glutamate residue (E503) that is the

**Table 2. Summary of refinement**

Resolution (Å)	20.0–2.3
$R_{\text{work}}/R_{\text{free}}$ (%) <sup>*</sup>	22.4/26.8
rmsd bond lengths	0.009
rmsd bond angles	1.2°
rmsd improper angles	0.8°
rmsd dihedral angles	20.5°
B value rmsd (main/side)	1.2/1.9
# metal ions	2
# water molecules	172

rmsd, rms deviation.

<sup>\*</sup> $R_{\text{work, free}} = \sum |F_{\text{obs}}| - |F_{\text{calc}}| / F_{\text{obs}}$  for the reflections used in refinement (work) and the 10% of reflections held aside (free).

third zinc ligand. C terminal to the active site sequence, the chain passes across the channel bottom ( $\alpha 18$  and  $\alpha 19$ ) to form two helices ( $\alpha 20$  and  $\alpha 21$ ) that pack in the core of domain I and then back again to end in a series of short helices below the active site ( $\alpha 22$ – $25$ ).

The nature and tight packing of secondary elements found in the two domains suggests that each is likely to be rigid; however, looser packing in the base of the channel may allow some motion of the two domains relative to each other. A higher average main-chain *B* value for domain I (38.6 vs. 29.9 for domain II) indicates some flexibility relative to domain II, which is held more tightly by crystal contacts. The extent of this flexibility is probably limited, however, by rigid helices ( $\alpha 13$ ,  $\alpha 14$ ,  $\alpha 19$ ,  $\alpha 22$ ,  $\alpha 25$ ) that cross the channel at one end of the molecule (Figs. 1*A* and 2*A*), and we would not expect the width of the cleft to vary greatly as a result of this motion.

**Relation to Other Metallopeptidases.** Neurolysin and other members of the M3 family belong to a supergroup of metallopeptidases, the MA clan, that includes enzymes bearing the HEXXH active site motif (9). Outside of this motif, there is no additional sequence similarity between the M3 family and others in this clan. The structure of neurolysin, however, demonstrates an evolutionary relationship beyond the active site motif. A large portion of the active site region is structurally related to the other members of the supergroup for which three-dimensional information is available. For example, the five strands ( $\beta 3$ – $\beta 7$ ) of the active site  $\beta$ -sheet of neurolysin, a helix ( $\alpha 15$ ) N-terminal to the sheet, and the two active site helices ( $\alpha 16$  and  $\alpha 17$ ) all have corresponding elements in thermolysin (EC 3.4.24.27), a well-studied (10) member of the M4 family (see Figs. 1*A* and 2). In addition, the C-terminal four helices in neurolysin ( $\alpha 22$ – $\alpha 25$ ) match those in thermolysin (with a small two  $\beta$ -strand insertion). The topology of the sheet differs in the two enzymes, as do the sizes of individual elements, but the overall structural similarity and like function imply a common origin. In neurolysin, the large N-terminal insertion relative to thermolysin provides the bulk of the superstructure that shields the active site, with only two short elements coming from a second, smaller insertion C terminal to the active site helices (Fig. 2*A*).

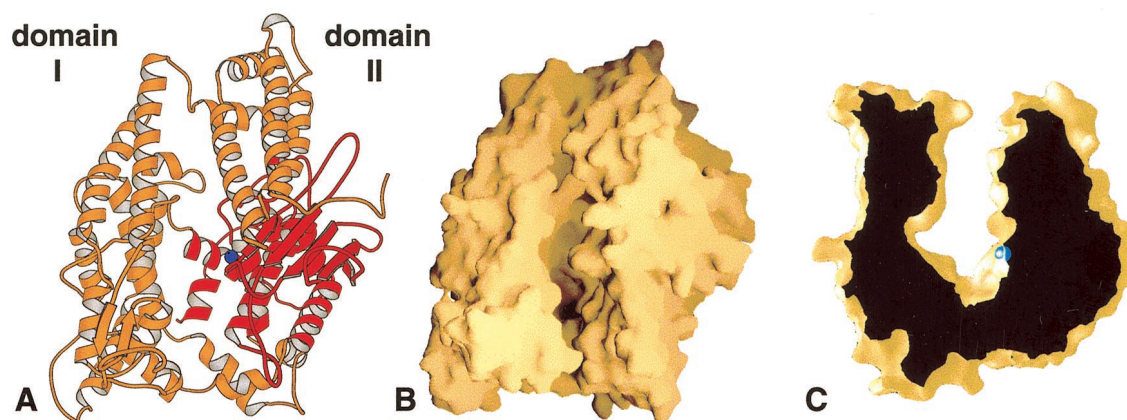
Metallopeptidase family M13, for which the structure of the neuropeptidase neprilysin has been reported (26), is the only other clan MA family with a member of known three-dimensional structure. Neprilysin also has some of the active site elements found in the M3 and M4 families, but they are less well

conserved. The  $\beta$ -sheet and the two active site helices are present, but the sheet has only three strands instead of five. The small C-terminal helices also are found in neprilysin, but there are three rather than the four found in neurolysin and thermolysin. Beyond the active site region, neprilysin bears no structural similarity to neurolysin other than its predominately  $\alpha$ -helical nature (26).

Members of a second supergroup, the MB clan, also contain an active site HEXXH motif, but the third zinc ligand is either a histidine or aspartate rather than the glutamate found in clan MA. The five-stranded  $\beta$ -sheet (again with a different topology) and the preceding helix are conserved, for example, in the clan MB enzyme astacin (EC 3.4.24.21; M12 family), but only the N terminal of the two active site helices is present. Instead of arising from the second of these helices, the third zinc ligand is located in an open coil segment.

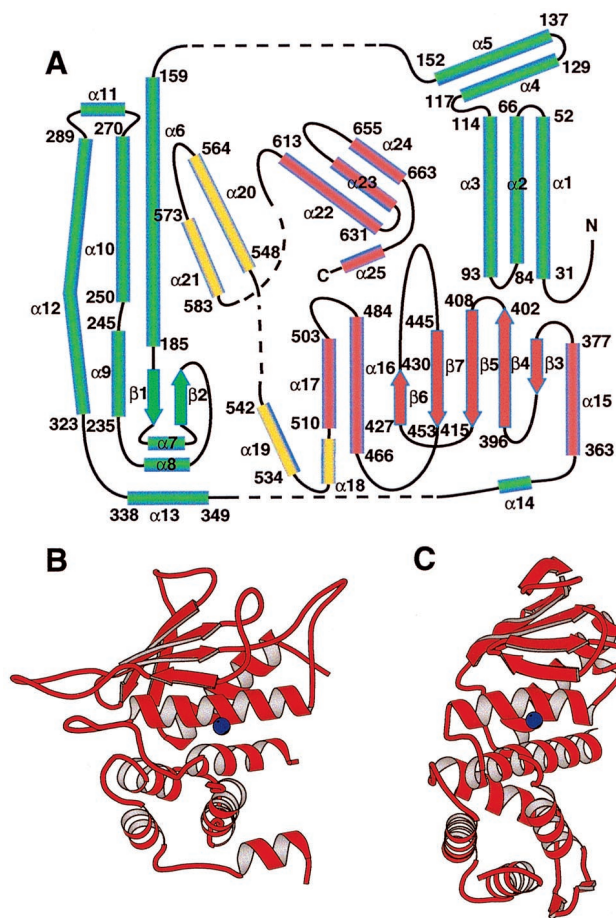
**Model for Substrate Length Restriction.** The extensive shielding elements erected over the neurolysin active site clearly account for its inability to hydrolyze large, folded substrates. Even the longest surface loops of proteins could not extend far enough into the narrow channel to bind at the active site. Oligopeptides of sufficient length to form  $\alpha$ -helices or bulky tertiary structure also would be excluded by the narrow upper regions of the channel. Neurotensin, a 13-residue substrate of neurolysin, is known to be unstructured in solution (30), as are most mixed sequence peptides shorter than 20 residues (31).

But is the requirement for no bulky secondary or tertiary structure the only restriction on substrate size? Can any length of peptide C or N terminal to the cleavage site be accommodated as long as the sequence adopts an extended conformation? The neurolysin structure suggests that there may be additional restrictions. Taking advantage of neurolysin's structural homology with other metallopeptidases, we have modeled a 13-residue peptide having the neurotensin sequence into the active site channel. The orientation of the peptide and its conformation near the active site are based on structures of thermolysin and astacin with bound peptide analogs (32, 33). The known single cleavage site in neurotensin, between residues 10 and 11, was positioned so that the adjacent carbonyl oxygen interacts with the catalytic zinc ion, and the N-terminal portion of the peptide was extended along the nearby  $\beta$ -strand by analogy with other metallopeptidase complexes. After fitting all residues, the modeled complex was subjected to energy minimization using a CHARMM force field (34).



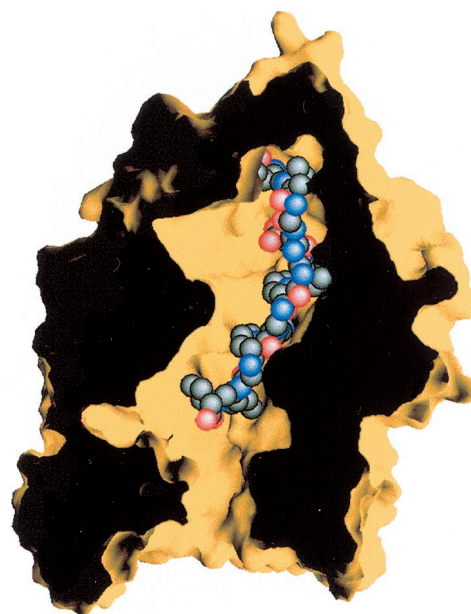
**Fig. 1.** Overview of the rat neurolysin structure. (A) Ribbon view of the molecule looking down on the substrate-binding channel. The active site zinc ion is shown as a blue sphere, and the region structurally similar to other metallopeptidases is shown in red. (B) The same view shown as a molecular surface representation. (C) Cross-sectional view of the neurolysin substrate-binding channel. The depth of the channel at the active site (zinc ion in blue) is approximately 35 Å. This figure and all other ribbon and surface figures were prepared with MOLSCRIPT (50) and GRASP (51).





**Fig. 2.** Topology and the structurally conserved region. (A) Topology diagram of neurolysin with  $\alpha$ -helices and  $\beta$ -strands numbered sequentially from the N terminus. Residue numbers for the beginning and end of larger secondary elements are given whenever possible. Dashed lines indicate points where regions have been separated for clarity. Portions shown in red are conserved in some other families of metallopeptidases. Ribbon diagrams for conserved regions of (B) neurolysin and (C) thermolysin also are shown. The active site zinc ion is in blue.

In the minimized complex (Fig. 3), the N terminus of the modeled peptide extends into a cavity at the bottom of the active site channel. This cavity runs beneath the three-helix bundle formed by residues near the protein N terminus, and one of its walls is formed by loops and turns extending from the conserved active site elements. Because the active site channel narrows at this end of the molecule, the cavity is sealed off, and a peptide extending into it would be trapped, unable to make its way back to wider sections of the channel. The cavity is only just long enough to accommodate the 10 residues N-terminal to the neurotensin cleavage site in the model. Although an alternative peptide conformation might permit a few additional residues, it seems likely that the cavity acts as a length filter or molecular yardstick, restricting to about 10 the maximum number of residues N-terminal to the site of hydrolysis. Longer N-terminal sequences would not permit the cleavage site to be positioned correctly relative to the catalytic machinery. As predicted by this model, neurotensin has one of the longest N-terminal sequences of known neurolysin substrates, whose N-terminal lengths range from 4 to 11 residues for peptides with single cleavage sites (8, 35). In contrast to the predicted restriction on the N terminus, the neurolysin structure suggests that the C terminus of a substrate may be able to exit from the channel where it widens

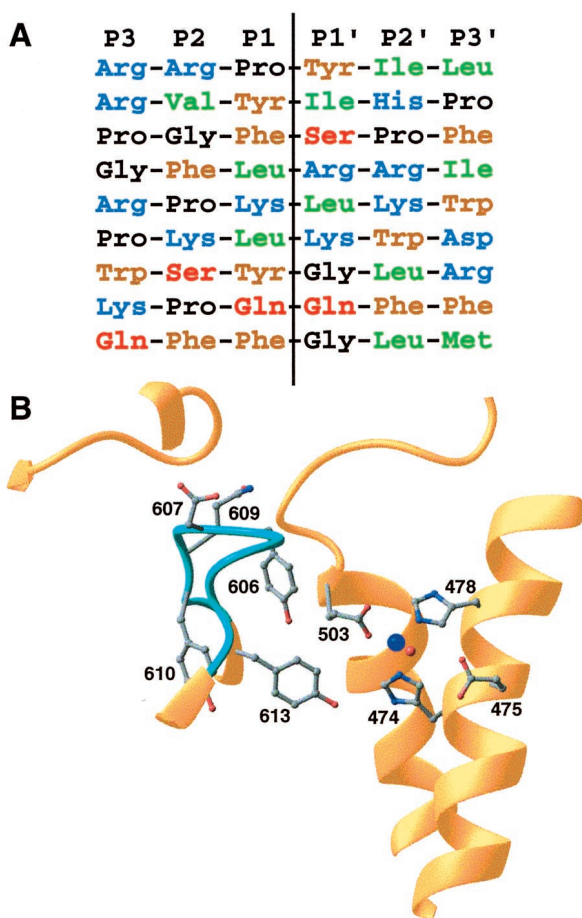


**Fig. 3.** Model of substrate peptide binding. A molecular surface representation of neurolysin sectioned to show the large cavity at the bottom of the active-site channel is shown with the 13-residue substrate neurotensin modeled as described in the text. The N terminus of neurotensin is at the top.

at the other end of the molecule. The activity of the highly homologous enzyme mitochondrial intermediate peptidase (E.C. 3.4.24.59), another M3 family member, supports this possibility. Mitochondrial intermediate peptidase cleaves an N-terminal octapeptide from the targeting sequence of proteins imported into the mitochondria (36). It seems a substrate peptide may even have a bulky molecule at the C terminus as long as a sufficient length of unstructured polymer can thread into the channel. Significantly, mitochondrial intermediate peptidase activity depends on prior removal of the first nine residues of the targeting sequence (37), consistent with our prediction that 17 residues would not fit in the cavity that receives the substrate N terminus. We suggest, then, that residues near a site of hydrolysis (see below) interact with the enzyme to position the peptide for catalysis. This positioning cannot occur, however, if bulky structure is present or if the number of residues N-terminal to a potential site exceeds the capacity of the active-site cavity.

Other oligopeptidases may have functionally similar but structurally distinct mechanisms of substrate length restriction. Prolyl oligopeptidase has a  $\beta$  propeller motif that shields a large active site cavity (27). It appears that substrate peptides must insert through the narrow center of the propeller to reach the catalytic residues, and in this way, bulky substrates would be excluded. In neprilysin, the active site is located in a spherical cavity that has a narrowed opening to solvent (26). Again, the restricted opening may prevent access by structured peptides and proteins. In both of these enzymes, however, the path of the peptide may be less constrained than in neurolysin because of their broader active site cavities. Length restriction therefore may not be directional as we propose for neurolysin.

**Sequence Selectivity and the Active Site.** A peptidase cleavage site is generally determined by the identity or type of residues within four or five positions of the scissile bond (38). For many peptidases, the two flanking residues, particularly the adjacent N-terminal (P1) residue, are most important in determining the hydrolysis rate. There are, however, enzymes, such as pepsin, where the effect of sequence substitutions is more evenly dis-



**Fig. 4.** Sequence selectivity. (A) Aligned sequences for known neurolysin cleavage sites. The sites listed are, top to bottom, from the neuropeptides neurotensin, angiotensin II, bradykinin, dynorphin A (residues 1–8), dynorphin A (residues 1–17), a second site in dynorphin A (residues 1–17), lutenizing hormone-releasing hormone, substance P, and a second site in substance P (8, 35). Residue types are indicated by different colors (blue = basic, brown = aromatic, green = aliphatic, black = proline or glycine, red = polar), and the hydrolysis position is indicated by the vertical line. (B) Details of the active site and nearby disordered loop (light blue; residues 600–612). The zinc cofactor is shown in dark blue and the catalytic water in red. Some side chains from residues in the mobile loop and active site are shown.

tributed over the nearby positions (39). Based on sequences of its known cleavage sites (Fig. 4A), neurolysin, like several other neuropeptidases, most likely belongs to the latter category. Various classes of residues occur at all positions near the scissile bond, and the only common features are a predominance of hydrophobic, aromatic, and basic residues. Cleavage sites therefore must be specified by the presence of favorable residues at different positions or by different favorable combinations of residues. Systematic variation of two substrate peptides cleaved by the closely related thimet oligopeptidase confirms the complex nature of specificity in these enzymes (40).

Even for enzymes that show consistent preferences at particular positions, various amino acids at nearby sites must be accommodated without compromising binding affinity or positioning of the substrate. The need to accommodate different residues is even more severe when all positions are subject to extensive variation as in neurolysin. There appear to be two mechanisms for tolerating sequence variation (41). In some cases, the bound substrate changes conformation, interacting with alternative subsites or portions of subsites on the enzyme. Examples of changes in substrate conformation include differ-

ences in the manner of P1-Phe and P1-Lys binding to *Bacillus amyloliquefaciens* subtilisin (42) and changes in the binding of P1 basic and hydrophobic residues to crab collagenase (43). A second mechanism for accommodating sequence variation involves plasticity of individual active site residues or larger structural elements. Here, local conformational changes in the enzyme allow it to adapt to sequence changes in the substrate. Accommodation of different-sized P1 hydrophobic groups by changes in the binding pocket volume in mutants of  $\alpha$ -lytic protease (44) and changes in the binding site for the P4 residue in *B. amyloliquefaciens* subtilisin (45) are examples of this mechanism. The regions that exhibit plasticity correlate at least to some extent with sites of high mobility in the free enzyme, and these regions become more fixed on binding substrate (46, 47).

In neurolysin, the channel near the active site is lined with potentially flexible loops and open coil regions (residues 84–93, 281–298, 415–427, 430–445, 453–466, 484–503, 542–548, 600–612, 645–655, 663–670, and 675–678). Loops in domain II arise primarily from the conserved structural elements and are extended relative to corresponding coil regions in thermolysin (compare Fig. 2B and C). They fan out along one wall of the channel covering much of the potential substrate interaction surface around the two active-site helices. On the domain I side of the channel, an additional two large loops sit just opposite the catalytic site, well positioned to interact with the substrate peptide. It is possible that these and other loop elements in the neurolysin channel help to accommodate different cleavage site sequences by altering conformation. Other elements lining portions of the channel more distant from the active site are primarily large helices less likely to undergo local rearrangements involving main-chain shifts.

One of the active site loops arising from domain I (residues 600–612) has particularly high thermal factors (average main-chain *B* value of 45.1) over nearly the entire loop, indicating an inherent flexibility. Residues in this loop are well positioned to interact with the substrate peptide near the scissile bond (Fig. 4B), and rearrangements of the loop might be important in allowing the enzyme to adapt to different substrates. In particular, Tyr-606 sits over the active site region in a way that would allow interaction with the side chain of the P1 residue. Appropriately, tyrosine could interact with different types of substrate residues through either its aromatic or polar side-chain groups. Tyr-606 and other residues in this loop are conserved between closely related neurolysin and thimet oligopeptidase. These two enzymes share many of the same cleavage site sequences, but on the peptide neurotensin, neurolysin cuts with a tyrosine in the P1 position whereas thimet oligopeptidase cuts with an arginine at that position. Perhaps Tyr-606 or its equivalent in thimet oligopeptidase aids in accommodating both residues, with other elements mediating the different specificities of the two enzymes. Other side chains in this loop also could interact with the substrate if some rearrangement occurs. In that regard, the presence of five glycine residues (600, 604, 605, 608, and 612) provides considerable potential for main-chain rearrangement in this loop. Structures of complexes between neurolysin and peptide analogues having different sequences will allow the model of sequence selectivity to be tested.

As noted, the catalytic residues are positioned directly across the channel from this flexible loop. The three zinc-coordinating residues (His-474, His-478, and Glu-503) adopt conformations very similar to those found in other zinc metallopeptidases (Fig. 4B). This similarity holds true for the position of the catalytic water, which also coordinates the zinc, and the hydrogen-bonded glutamate residue (Glu-475) that is thought to facilitate catalysis by polarizing the water and transferring a proton to the leaving main-chain nitrogen (48, 49). Absent, however, are the nearby histidine and tyrosine residues found in thermolysin (residues 231 and 157, respectively) that together with the zinc ion stabilize



the oxyanion intermediate formed after nucleophilic attack by the water. Instead, Tyr-613 in helix  $\alpha 22$  may play this role. Its side-chain hydroxyl group is located near the zinc ion, where it would be in position to interact with the substrate carbonyl oxygen from the same direction as His-231 in thermolysin. In this respect, neurolysin may be more like astacin, which also has a tyrosine (residue 149) in a similar position.

**Conclusions.** Large structural elements prevent access of all potential substrates except unstructured peptides to the neurolysin active site. Shielding elements also are found in two other oligopeptidases, but the precise nature of these structures varies considerably and so, therefore, may the details of size restriction and site recognition. The narrow channel in neurolysin, with one end effectively sealed, probably limits the size of a substrate N terminal to the site of cleavage. In addition, the extensive loop elements lining the neurolysin channel suggest that the wide

sequence variation in known cleavage sites may be explained by plasticity in the regions of the enzyme responsible for substrate binding. Given the close sequence homology between neurolysin and other M3 family members, it is likely that there are only small variations from the neurolysin structure, accounting for the predominance of oligopeptidase function within this group. Interestingly, one family member, peptidyl-dipeptidase Dcp, removes only C-terminal dipeptides, an exopeptidase-like activity, suggesting that additional restrictions are placed on substrate positioning in this enzyme.

We are grateful to Wilfried Schildkamp, Vukica Srajer, and Reinhard Pahl of the BioCARS sector at the Advanced Photon Source for help with data collection and Trevor Creamer, Chad Haynes, Christina Hines, Louis Hersh, Kallol Ray, and Karin Reinisch for comments on the manuscript. The work was supported by the National Science Foundation.

- Konkoy, C. S. & Davis, T. P. (1996) *Trends Pharmacol. Sci.* **17**, 288–294.
- Csuhai, E., Safavi, A., Thompson, M. W. & Hersh, L. B. (1998) in *Proteolytic and Cellular Mechanisms in Prohormone and Neuropeptide Precursor Processing*, ed. Hook, V. (Springer, Heidelberg), pp. 173–182.
- Checler, F. (1993) in *Methods in Neurotransmitter and Neuropeptide Research*, eds. Nagatsu, T., Parvez, H., Naoi, M. & Parvez, S. (Elsevier, Amsterdam), Vol. 2, pp. 375–418.
- Charli, J. L., Mendez, M., Vargas, M. A., Cisneros, M., Assai, M., Joseph-Bravo, P. & Wilk, S. (1989) *Neuropeptides* **14**, 191–196.
- Vincent, B., Dive, V., Yiotakis, A., Smadja, C., Maldonado, R., Vincent, J.-P. & Checler, F. (1995) *Brit. J. Pharmacol.* **115**, 1053–1063.
- Bauer, J. H. (1990) *Am. J. Hypertens.* **3**, 331–337.
- Checler, F., Barelli, H., Dauch, P., Dive, V., Vincent, B. & Vincent, J. P. (1995) *Methods Enzymol.* **248**, 593–614.
- Barrett, A. J., Brown, M. A., Dando, P. M., Knight, C. G., McKie, N., Rawlings, N. D. & Serizawa, A. (1995) *Methods Enzymol.* **248**, 529–556.
- Rawlings, N. D. & Barrett, A. J. (1995) *Methods Enzymol.* **248**, 183–228.
- Matthews, B. W., Weaver, L. H. & Kester, W. R. (1974) *J. Biol. Chem.* **249**, 8030–8044.
- Vincent, B., Beaudet, A., Dauch, P., Vincent, J.-P. & Checler, F. (1996) *J. Neurosci.* **16**, 5049–5059.
- Kato, A., Sugiura, N., Saruta, Y., Hosoiri, T., Yasue, H. & Hirose, S. (1997) *J. Biol. Chem.* **272**, 15313–15322.
- Barelli, H., Fox-Threlkeld, J. E. T., Dive, V., Daniel, E. E., Vincent, J. P. & Checler, F. (1994) *Br. J. Pharmacol.* **112**, 127–132.
- Chabry, J., Checler, F., Vincent, J.-P. & Mazella, J. (1990) *J. Neurosci.* **10**, 3916–3921.
- Goedert, M. (1984) *Trends Neurosci.* **7**, 3–5.
- Clineschmidt, B. V. & McGuffin, J. C. (1977) *Eur. J. Pharmacol.* **46**, 395–396.
- Vincent, B., Jiracek, J., Nobel, F., Loog, M., Roques, B., Dive, V., Vincent, J.-P. & Checler, F. (1997) *Br. J. Pharmacol.* **121**, 705–710.
- Lian, W., Chen, G., Wu, D., Brown, C. K., Madauss, K., Hersh, L. B. & Rodgers, D. W. (2000) *Acta Crystallogr. D* **56**, 1644–1646.
- Doublet, S. (1997) *Methods Enzymol.* **276**, 523–530.
- Otwinowski, Z. & Minor, W. (1997) *Methods Enzymol.* **276**, 307–326.
- CCP4 (1994) *Acta Crystallogr. D* **50**, 760–763.
- Otwinowski, Z. (1991) *Daresbury Study Weekend Proceedings* (Science and Engineering Research Council, Daresbury Laboratory, Warrington, U.K.).
- Cowtan, K. (1994) *Joint CCP4 ESW-EACBM Newl. Protein Crystallogr.* **31**, 34–38.
- Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991) *Acta Crystallogr. A* **47**, 110–119.
- Brünger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J. N., Pannu, N. S., Reed, R. J., et al. (1998) *Acta Crystallogr. D* **54**, 905–921.
- Oefner, C., D'Arcy, A., Hennig, M., Winkler, F. K. & Dale, G. E. (2000) *J. Mol. Biol.* **296**, 341–349.
- Fülöp, V., Böcskei, Z. & Polgár, L. (1998) *Cell* **94**, 161–170.
- Garrido, P. A., Vandenbulcke, F., Ramjaun, A. R., Vincent, B., Checler, F., Ferro, E. & Beaudet, A. (1999) *DNA Cell Biol.* **18**, 323–331.
- Ferro, E. S., Tullai, J. W., Glucksman, M. J. & Roberts, J. L. (1999) *DNA Cell Biol.* **18**, 781–789.
- Xu, G. Y. & Deber, C. M. (1991) *Int. J. Pept. Protein Res.* **37**, 528–535.
- Dyson, H. J. & Write, P. E. (1991) *Annu. Rev. Biophys. Biophys. Chem.* **20**, 519–538.
- Holden, H. M., Tronrud, D. E., Monzingo, A. F., Weaver, L. H. & Matthews, B. W. (1987) *Biochemistry* **26**, 8542–8553.
- Grams, F., Dive, V., Yiotakis, A., Yiallourous, I., Vassiliou, S., Zwilling, R., Bode, W. & Stocker, W. (1996) *Nat. Struct. Biol.* **3**, 671–675.
- Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983) *J. Comput. Chem.* **4**, 187–217.
- Dahms, P. & Mentlein, R. (1992) *Eur. J. Biochem.* **208**, 145–154.
- Kalousek, F., Hendrik, J. P. & Rosenberg, L. E. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7536–7540.
- Isaya, G., Kalousek, F., Fenton, W. A. & Rosenberg, L. E. (1991) *J. Cell. Biol.* **113**, 65–76.
- Keil, B. (1992) *Specificity of Proteolysis* (Springer, Berlin).
- Powers, J. C., Harley, A. D. & Myer, D. V. (1977) *Adv. Exp. Med. Biol.* **95**, 141–157.
- Camargo, A. C. M., Gomes, M. D., Reichl, A. P., Ferro, E. S., Jacchieri, S., Hirata, I. Y. & Juliano, L. (1997) *Biochem. J.* **324**, 517–522.
- Perona, J. J. & Craik, C. S. (1995) *Protein Sci.* **4**, 337–360.
- Poulos, T. L., Alden, R. A., Freer, S. T., Birkhof, J. J. & Kraut, J. (1976) *J. Biol. Chem.* **251**, 1097–1103.
- Tsu, C. A., Perona, J. J., Fletterick, R. J. & Craik, S. S. (1997) *Biochemistry* **36**, 5393–5401.
- Bone, R., Fujishige, A., Kettner, C. A. & Agard, D. A. (1991) *Biochemistry* **30**, 10388–10398.
- Takeuchi, Y., Noguchi, S., Satow, Y., Kojima, S., Kumagai, I., Muira, K., Nakamura, K. T. & Mitsui, Y. (1991) *Protein Eng.* **4**, 501–508.
- Yuan, P., Marshall, V. P., Petzold, G. L., Poorman, R. A. & Stockman, B. J. (1999) *J. Biomol. NMR* **15**, 55–64.
- Davis, J. H. & Agard, D. A. (1998) *Biochemistry* **37**, 7696–7707.
- Rees, D. C. & Lipscomb, W. N. (1982) *J. Mol. Biol.* **160**, 475–498.
- Monzingo, A. F. & Matthews, B. W. (1984) *Biochemistry* **20**, 5724–5729.
- Kraulis, P. (1991) *J. Appl. Crystallogr.* **24**, 946–950.
- Nicholls, A., Sharp, K. A. & Honig, B. (1991) *Proteins Struct. Funct. Genet.* **11**, 281–296.