A three-dimensional construction of the active site (region 507–749) of human neutral endopeptidase (EC.3.4.24.11)

Gilles Tiraboschi, Nathalie Jullian¹, Vincent Thery², Serge Antoneczak³, Marie-Claude Fournie-Zaluski and Bernard F.Roques⁴

Département de Pharmacochimie Moléculaire et Structurale, INSERM U266, CNRS URA D1500, UFR des Sciences Pharmaceutiques et Biologiques, 4 Avenue de l’Observatoire, 75270 Paris, Cedex 06, ¹Molecular Simulations SARL, Parc Club Orsay Université, 20 Rue Jean Rostand, 91893 Orsay Cedex, ²Laboratoire de Synthèse Organique, Faculté de Sciences, Les Cézeaux, 63177 Aubiere and ³Laboratoire de Chimie Théorique, CNRS URA 510, Université Nancy 1, BP 239, Domaine Universitaire Victor-Grignard, 54506 Vandoeuvre-Les-Nancy Cedex, France

To whom correspondence should be addressed

A three-dimensional model of the 507–749 region of neutral endopeptidase-24.11 (NEP; EC.3.4.24.11) was constructed integrating the results of secondary structure predictions and sequence homologies with the bacterial endopeptidase thermolysin. Additional data were extracted from the structure of two other metalloproteases, astacin and stromelysin. The resulting model accounts for the main biological properties of NEP and has been used to describe the environment close to the zinc atom defining the catalytic site. The analysis of several thiol inhibitors, complexed in the model active site, revealed the presence of a large hydrophobic pocket at the S₁ subsite level. This is supported by the nature of the constitutive amino acids. The computed energies of bound inhibitors correspond with the relative affinities of the stereoisomers of benzofused macrocycle derivatives of thiorphan. The model could be used to facilitate the design of new NEP inhibitors, as illustrated in the paper.

Keywords: zinc metallopeptidase/homology modelling/secondary structure prediction/sequence alignment/enzyme active-site/S₁ subsite modelling/enzyme inhibitor complex/thiorphan/disulfide bridge/structure refinement

Introduction

Neutral endopeptidase 24.11 (NEP; EC.3.4.24.11, neprilysin) is a zinc-endoenzyme involved in the metabolism of some important regulatory peptides in both the central nervous system and the periphery. NEP is implicated in the physiological degradation of the endogenous opioid peptides Met- and Leu-enkephalins (Roques et al., 1980) and the peptides modulating blood pressure, such as the atrial natriuretic peptide (Kenny and Stephenson, 1988), bradykinin (Pham et al., 1993) and endothelin (Abassi et al., 1992). NEP has also been shown to cleave, in vitro, a variety of other peptides such as neuotensin and substance P (Skidgel et al., 1984). Due to the physiological importance of NEP in the modulation of nociceptive and pressor responses, a variety of inhibitors have been designed in order to develop new analgesic or antihypertensive agents (review in Roques et al., 1993).

NEP, whose primary sequence has been determined from various species (Devault et al., 1987; Malfroy et al., 1987; Letarte et al., 1988; Malfroy et al., 1988; Chen et al., 1992), is a member of the nephrilysin family of zinc metalloproteases, which also contains endothelin converting enzymes ECE-1 and ECE-2, and the putative proteases Kell and PEX (Shimada et al., 1994; Emoto and Yanaginawa, 1995; Lee et al., 1995; Grief et al., 1997). All these enzymes are type II integral membrane proteins with a short intracytoplasmic region, a transmembranous hydrophobic domain, and a bulky extracellular domain that contains the active site. In the case of NEP, these sequences possess 27, 22 and 700 residues, respectively. The extracellular domain of NEP contains 12 cysteine residues involved in disulphide bridges (Tam et al., 1985), and 10 are conserved in ECE-1, ECE-2, Kell and PEX, suggesting that these enzymes have a similar folding (Turner and Tanzawa, 1997). Some other reported neprilysin family members, such as pepO from Lactococcus lactis (Mierau et al., 1993), do not have these conserved cysteine residues. In addition, all these enzymes show sequence homologies in their C-terminal moieties. The similarities between NEP and ECE-1, ECE-2, KELL, PEX and PEPO are 35, 30, 17, 31, 25 and 19% respectively, but increase to 51, 44, 27, 43 and 31% respectively when only the 250 C-terminal residues are considered. In this region the catalytic site of NEP has been identified by the presence of the two zinc binding consensus sequences 5⁵HExxH₅⁸⁷ (Devault et al., 1987) and 6₄⁴ExxxD₆₅⁰ (Roques, 1993), which are the signatures of the thermolysin (TLN), neprilysin and angiotensin-converting enzyme families (Hooper, 1994).

The 316 amino acid TLN has been co-crystalized with various inhibitors including NEP inhibitors such as thiorphan and retrothiorphan (Roderick et al., 1989), and crystallographic analyses of these structures has allowed the residues involved in its mechanism of peptide hydrolysis to be defined (Matthews, 1988). The zinc atom is coordinated to two histidine residues, H₁⁴₂ and H₁⁴₆, which are found in the 1⁴₂HExxH₁₄₆ consensus sequence, whilst the glutamate of this sequence, E₁⁴³, transfers a hydrogen atom and activates the nucleophilic attack of the polarized zinc-bound water on the scissile peptide bond. The third zinc binding ligand is the glutamate, E₁⁶⁶, that is found in the second consensus sequence 1⁶⁶ExxxD₁⁷⁰. Furthermore, the aspartate residue D₁⁷⁰ is involved in a functionally important network of hydrogen bonds (Marie-Claire et al., 1997). Other important residues are H₂¹¹ and Y₁⁵⁷, which stabilize the transition state, and N₁¹², A₁¹³ and R₂⁰⁰, which are involved in substrate binding (Matthews, 1988).

Despite the low homology between the primary sequences of NEP and thermolysin, it has been shown that these two enzymes recognize the same substrates by cleaving the peptide bond at the N-terminus of hydrophobic residues, are inactivated by the same type of inhibitors, and have the same stereochemical dependence (Roques et al., 1993). Moreover, structural analogies have been proposed between TLN and the C-terminal region of NEP, using the hydrophobic cluster analysis method (Benchetrit et al., 1988). This was the case of helices bearing the two consensus sequences, HExxH and ExxxD. Site-directed
mutagenesis experiments have allowed several important residues of the NEP active site to be identified, suggesting a quasi identical organization of the active sites of NEP and TLN. Thus, in addition to the residues included in the two consensus sequences, i.e. the three zinc-binding ligands H583, H587 and E646 (corresponding to H142, H146 and E166 in TLN), the catalytic E584 and D585 (Devault et al., 1988a; Le Moual et al., 1991, 1994), (corresponding to E143 and D170), residues N542 (Dion et al., 1995), H711 (Dion et al., 1993), R717 (Marie-Claire et al., 1997) have been proposed to be the counterparts of N112, H231 and R203 in TLN, respectively.

In the absence of crystallographic data, the strategy used for the design of NEP inhibitors has been based on the analogies between the active site of this physiological enzyme with that of TLN (for a review see Roques et al., 1993). However, a more detailed three-dimensional model, using a molecular modeling approach, could greatly facilitate the rational design of selective inhibitors. We report in this paper a proposed model for the tertiary structure of NEP restricted with that of TLN (for a review see Roques et al., 1993). A more detailed three-dimensional model, using a molecular modeling approach, could greatly facilitate the rational design of selective inhibitors. We report in this paper a proposed model for the tertiary structure of NEP restricted with that of TLN (for a review see Roques et al., 1993).

Materials and methods

The construction of the model of NEP was achieved by the following steps.

Definition of homologous regions between NEP and reference proteins

Global multialignment sequences studies, using CLUSTALW 1.7 (Thompson et al., 1994), local alignment methods using the BLAST algorithm (Alschul et al., 1990) and threading methods [TOPIT (Rost, 1995); FRSVR (Fischer and Eisenberg, 1996); SWISSMODEL (Peitsch, 1996)] were first performed. Insight II (Insight II user guide 1995) was used for the definition of the structurally conserved regions (SCRS) between TLN, astacin (AST) (Gomis-Rüth et al., 1993) and the C-terminal moiety of stromelysin-1 (STR) (Becker et al., 1995). PHD (Rost and Sander, 1994) and PREDATOR (Frishman and Argos, 1995) multialignment secondary structure prediction methods were performed using the human NEP, human ECE1, bovine ECE2, KELL and PEX sequences. Pairwise alignment between NEP regions and sequences of SCRS were achieved using the segment pair overlap algorithm (Schuler et al., 1991), as implemented in the Homology module of InsightII (Homology user guide, 1995). The same module was used for the assignment of coordinates of homologous regions and the generation of loop regions described below.

Assignment of the coordinates of homologous regions

Assignment of the coordinates of homologous regions was achieved using the crystallographic TLN coordinates set [ref 1LNF in Protein Data Bank (PDB; Bernstein et al., 1977)] for the template structure. The backbone coordinates were transferred from TLN to NEP, as well as the values of the dihedral angle, in common with the side chain of TLN and NEP residues. If a side chain of a NEP residue had more atoms than that of TLN, the more distal atoms were given in an extended conformation and bad van der Waals (VDW) contacts removed manually.

Loops between homologous regions

Loops between homologous regions were generated as a function of the increasing order of their size by searching related three-dimensional structures in PDB, with a distance geometry loop-searching algorithm originally described by Jones and Thirrup (1986), or by randomizing dihedral angle values. From the 10 loops generated in each case, selection was carried out on the basis of non-repulsive VDW contacts with the rest of the model. Local bad VDW contacts were corrected by rotating the dihedral angles of the amino acids involved. The zinc atom was positioned in a manner similar to that observed in the TLN structure (1LNF).

Refinement of the model and inhibitor docking

All the refinement steps were performed, in vacuo, using the Discover program of Insight II (Insight II user guide, 1995). The AMBER forcefield was modified in order to take into account the zinc dication ligation (Jacob, 1990). A dielectric constant of 4*r was used for all calculations with a cut-off of
12 Å. Refinement steps included minimization combined with simulated annealing protocols. The choice among the cooled structures was made according to both the energy criteria, and the geometry residue analysis achieved with the PROCHECK software (Laskowski et al., 1993).

The non-aligned regions were first minimized and dynamized at 500 K during 50 ps after 5 ps of equilibration. One structure was saved every 10 ps and submitted to a simulated annealing procedure, generating a total of 50 cooled structures. Each structure was cooled from 500 to 10 K in 15 ps and submitted to a steepest descent minimization until the maximum derivative was lower than 0.05 kcal mol⁻¹ Å⁻¹.

The model was further refined under its complexed form with the thiol inhibitor, thiorphan. The thiorphan molecule was docked in the NEP model, using the crystallographic TLN–thiorphan complex (Roderick et al., 1989), after superimposition of the backbone of the aligned residues between TLN and NEP. As the conformation of the TLN moiety in the crystal complex did not differ from the template structure (1LN), the thiorphan molecule could be docked in the NEP model, directly by transferring coordinates, without generating bad VDW contacts. This complex was submitted to a simulating annealing procedure at 300 K (with the procedure described above) where only loops and side chains of the aligned residues were allowed to move. The complex was then fully minimized with a conjugate gradient algorithm (including the thiorphan moiety) until the maximum derivative energy was lower than 0.001 kcal mol⁻¹ Å⁻¹. Finally, a constrained dynamic (the backbone of the aligned residues of NEP were held fixed) was performed at 300 K for 100 ps after 5 ps of equilibration.

The minimized NEP/thiorphan model (coordinates will be submitted to the PDB) was used for the docking of compounds II, III IV and V (see Table I) in NEP, superimposing the common chemical groups between the inhibitor moieties. Compounds II(S,S) and III(S,S) were constructed with ring conformations similar to those proposed for their complexed forms with TLN [see figure 4 of Ksander et al., 1997a and figure 8 of Ksander et al., 1997b for II(S,S) and III(S,S) respectively]. Stereoisomers II(R,S) and III(R,S) were constructed by inversion of the appropriate carbon configuration. All the stereoisomers of II and III were then minimized. During minimization the sulfur atom, atoms of the lactam and carboxylate functions were held fixed in the positions adopted for these groups in the NEP–thiorphan complex. In this manner the four resulting structures could be docked in the NEP model without major bad VDW contacts. All the complexes were then fully minimized as previously described for the NEP–thiorphan complex. A quantitative evaluation of the inhibitor potencies was then carried out, computing three energy terms:

The strain energy of inhibitors were defined as the energy difference between the bounded conformation and the lowest energy conformation of the isolated compound. The lowest energy conformation was chosen among a set of 100 structures obtained by a dynamic/minimization protocol. During the dynamic run (900 K; 1 ns) one structure was saved every 0.01 ns and minimized.

The total energy of a given complex corresponds to the energy of the enzyme and the inhibitor, including that provided by the covalent binding to the zinc atom. The relative energy of complexation of the studied inhibitors were thus estimated as:

\[ \Delta E = [E_{complex1} - (E_{II1} + E_d)] - [E_{complex2} - (E_{II2} + E_d)] \]

where \( E_{complex} \) corresponds to the total energy of the complex

Table II. Secondary structure prediction of the 507–749 region of the human NEP sequence resulting from PREDATOR (PRE) and PHD methods. Consensus sequences are in bold

<table>
<thead>
<tr>
<th>NEP</th>
<th>PRE</th>
<th>PHD</th>
</tr>
</thead>
<tbody>
<tr>
<td>507</td>
<td>FENVIQNLKFSQKQLRKBVRKDKDEWISGAAYV</td>
<td></td>
</tr>
<tr>
<td>512</td>
<td></td>
<td></td>
</tr>
<tr>
<td>517</td>
<td>HAFYSSGRQNOIVPPAGILQPPFSSAQQSNLSNYGG</td>
<td></td>
</tr>
<tr>
<td>522</td>
<td></td>
<td></td>
</tr>
<tr>
<td>527</td>
<td>GGMAIGERITHGFDNHRNFHGDGLVDWDTQQSA</td>
<td></td>
</tr>
<tr>
<td>532</td>
<td></td>
<td></td>
</tr>
<tr>
<td>537</td>
<td>SNFKEQQSCMVYQGMFSDLLAGQHNLGINTLGH</td>
<td></td>
</tr>
<tr>
<td>542</td>
<td></td>
<td></td>
</tr>
<tr>
<td>547</td>
<td>HIAANDGQLQAYRAYQNYIKKGNEKKLPGLDLNH</td>
<td></td>
</tr>
<tr>
<td>552</td>
<td></td>
<td></td>
</tr>
<tr>
<td>557</td>
<td>KQLFLLNFAMQVGCYRPEAYAVSIKTVDVNPSQH</td>
<td></td>
</tr>
<tr>
<td>562</td>
<td></td>
<td></td>
</tr>
<tr>
<td>567</td>
<td>RISTLQMEASEFSEASFCRNSYNRRPEKCRVW</td>
<td></td>
</tr>
</tbody>
</table>

Black and shaded regions correspond to respectively \( \alpha \) helical and \( \beta \) strand domains.

between NEP and inhibitor \( n \), and \( E_{I_0} \) to the lowest energy conformation of the free inhibitor \( n \). The energy of the free enzyme (\( E_0 \)) was not calculated since this term disappears in the equation.

Results

‘Automated’ sequence alignment

Global multialignment studies (CLUSTALW 1.7) failed because the two consensus sequences are separated by 54 residues in NEP (583HExxH587, 646ExxxD650) and only 20 residues (142HExxH146, 166ExxxD170) in TLN; and because the reported equivalent residues to R203 and H231 of TLN are in an inverted sequence order in NEP (R717 and H711 respectively). Due to the very low sequence similarity, even in regions where the equivalent residues have been experimentally identified in NEP and TLN, the BLAST algorithm, using standard parameters with BLOSUM62 (Henikoff and Henikoff, 1992) or PAM250 (Dayhoff et al., 1983) substitution matrices, did not find usable local alignments. For the same reasons, and because of the presence of disulfide bridges, threading methods (TOPIT, FR3VR, SWISSMODEL) did not suggest reliable folds (low z score, incompatible fold). In the absence of similarity to a known three-dimensional structure, construction of the conserved C-terminal region (507–749) of the human peptidase was carried out by using secondary structure prediction methods combined with local sequence comparisons and reported equivalent residues between TLN and NEP.

Secondary structure predictions

The results of secondary structure predictions for the 507–749 region of NEP are reported in Table II with \( \alpha \) helices and
β-strand regions being predicted by the two methods used here (PHD and PREDATOR). The regions 541–545, 551–553, 690–693 and 716–720 correspond to β-strands. The regions 507–525, 611–624, 645–666, including the 646ExxH650 consensus sequence and 725–731, are predicted to be helical. Surprisingly, the two methods predict the zinc-binding HExxH motif to be involved in a β-strand, whereas, in all crystallized zinc peptidases this consensus region is included in a helical segment. Consequently, this result has not been considered for the construction of the model.

Definition of the structurally conserved regions between thermolysin, astacin and stromelysin

AST and STR enzymes belong to the metzincin metallopeptidase family. They have a common organization around the active site: a sheet consisting of four parallel (sI, sII, sIII, sV) and one antiparallel (sIV) β-strand and three α helices (hA, hB and hC) where hB contains the HExxH zinc-binding motif and sIV interacts with bound substrate (Stöcker and Bode, 1995). Topological similarities with many secondary structure elements have been described between metzincins and TLN, despite their low sequence homology (Gomis-Rüth et al., 1993). After superimposition of the common residues of the sequence consensus of TLN, AST and STR (142HExxH146, 92HExxH96 and 201HExxH205 respectively) (Figure 2), SCRs were defined (Table IIIa) involving H1, H2, and H6 α helices of TLN and hA, hB and hC of AST and STR respectively, S9, S10 and S11 strands of TLN and the sIII, sIV and sV strands of AST and STR, respectively. The r.m.s.d. values after superimposition of the α carbons of the TLN/AST, TLN/STR and AST/STR SCRs pairwises are 3.5, 2.9 and 2.9 Å respectively.

‘Directed’ sequence alignments

The predicted regions of NEP discussed above, and the region including the HExxH sequence, were first aligned with the secondary structure elements of TLN. All the predicted regions for which no TLN homologous regions could be found were then aligned with the defined SCRs of AST and STR (Table IIIa). As the length of the predicted regions varied, depending on the method used, they were considered as ‘test-areas’ for alignments rather than strictly defined segments for construction. Table IIIb shows the sequence alignments of NEP with secondary structural elements of the three reference proteins. The numbering of the secondary structure elements found in TLN (helices H and sheet S) are used to designate the corresponding homologous regions of NEP (e.g. H2 of TLN and H1 of NEP). The 507–525 region (H1) can be aligned with the C-terminal region (71–90) of the H1 helix of TLN. The region 570–591 (H2), which includes the 580HExxH587 consensus sequence and Y580, proposed to be the structural homolog of V139 in TLN (Vijayaraghavan et al., 1990) can be aligned with the 130–150 helical region of TLN (H2) and the 639–659 (H4) region, including the second consensus sequence, 646ExxD650, with the 159–179 (H4) helical region of TLN. The 542–547 region (S11) shows a high sequence homology with the 112–117 (S11) β-strand of TLN. This homology is reinforced by a N542 mutation study, which suggested a functional equivalence between residues N542 and A543 of NEP and residues N112 and A113 of TLN, respectively (Dion et al., 1995). The 551–555 (S11) region can be aligned with the 119–123 (S11) β-strand region of TLN. On the basis of the functional equivalence between R717 of NEP and R703 of TLN (Marie-Claire et al., 1997), the β-strand predicted region (716–720) (S13) can be aligned with the 202–206 (S13) β-strand region of TLN. The sequence 611–619 (H6) of NEP can be aligned with the 235–243 region of STR, structurally homologous to the 263–270 (H6) helical region of TLN.

Further alignments were achieved by considering isolated corresponding residues between NEP and TLN: site-directed mutagenesis studies have shown an equivalence between residues H113 of NEP and H121 of TLN (Dion et al., 1993). In addition, the residue Y525 has been also proposed to be analogous to Y137 of TLN (Sansom et al., 1995) and a recent mutation in a thermolysin-like enzyme suggests that the salt bridge between H231 and D226 in TLN is important for stability.
Molecular modelling of the active site of NEP

Table III. (a) Structurally conserved regions (in boxes) between TLN, AST and STR defined after superimposition of the residues H, E and H of their respective HExxH consensus sequences (in bold). (b) Local alignment between NEP and the three reference proteins (TLN or AST or STR)

Table IV. Comparison between (A) the defined homologous region of NEP used for the construction. (B) Secondary structures of the 61–316 region of TLN. Consensus sequences are in bold

After loop generation, all the non-aligned residues (526–533, 538–541, 548–550, 556–569, 592–610, 620–624, 626–638, 660–708, 710, 712–716 and 721–749) and the 570–580 H2 fragment, including the G578 insertion, were annealed. Even though modeling of a large loop (660–708) is risky, it preserved the connectivity between the aligned regions so that the model could be further refined.

Refinement of the model: the NEP–thiorphan complex

TLN and NEP exhibit equivalent stereochemical dependence towards thiol inhibitors such as thiorphan and retro-thiorphan (Benchetrit et al., 1987). The crystallographic TLN–thiorphan complex shows interactions involving N112, H231 and R203, which, as suggested by mutagenesis experiments, could be equivalent to N542, H711 and R717, respectively (Figure 2). Therefore, the 507–749 region of NEP was further refined under its complexed form with this inhibitor, providing additional structural information. Recent mutagenesis experiments (Marie-Claire et al., 1997) suggest that residues H583, D650 and R717, homologous to H142, D170 and R203 of TLN, are probably involved in a hydrogen bond network very similar to that observed in all the crystallographic structures of TLN, between residues H142 and D170, D170 and R203, and between D170 and N238. Analysis of the R203M and D170A TLN mutations as well as R717M and D650A NEP mutations has shown the importance of this network for the topology of the active sites.
Fig. 3. Comparison between the three-dimensional organization of the 61–316 region of TLN (A) and that proposed in this study for the 507–749 region of NEP (B) following refinement. For clarity, only the aligned secondary structures of TLN used for the construction are drawn in MOLSCRIPT style.

and the activities of both enzymes (Marie-Claire et al., 1997). During the annealing of the NEP–thiorphan complex, the hydrogen bonds between H583 and D650 and between D650 and R717 were not maintained. This is probably due to the absence of a fourth partner, equivalent to N238 of TLN, in the model of NEP. We therefore restarted this refinement step, applying distance constraints between pairwise OD1 : D650/ND1 : H583, OD1 : D650/NH2 : R717 and OD2 : D650/NE1 : R717 of NEP atoms. After the full minimization of the complex (the refined structure is reported in Figure 3), a final constrained dynamics (300 K; 100 ps) was performed for analysis of the dynamic behaviour of the model.

In the fully optimized model of the NEP–thiorphan complex, the interactions and the conformation of the inhibitor remain similar to those proposed for the crystallographic TLN–thiorphan complex: the amide carbonyl oxygen of the inhibitor lies within hydrogen bonding distances of the NE1 and NE2 atoms of R717. However, as previously proposed (Dion et al., 1995), the C-terminal glycine moiety displays a unique hydrogen bond between its nitrogen and the OD1 atom of the side chain of N542 (Figure 2). The phenyl ring, which differs by a rotation of 40 Å relative to the TLN complexed conformation, is docked in the S1′ subsite.

In TLN the S1′ subsite is a selective pocket for hydrophobic residues, well delineated by the side chains of residues F130, L133, V139, I188, G189, V192 and L202 (Figure 4a). In our model, residues L564, F716, I718, F563, F564, M579 and V580 also define a large lipophilic pocket surrounding the phenyl ring of thiorphan (Figure 4b). Residues V580 and F563 have similar positions to residues V139 and F130 in TLN, respectively, and display equivalent close contact with the phenyl moiety of thiorphan. However, some structural differences can be observed between the two S1′ subsites of TLN and NEP after energy minimization of the latter enzyme. Thus, three aromatic residues (F563, F564, F716) appear to be involved in the NEP subsite (Figure 4b), compared with only one for TLN (F130) (Figure 4a). The side chain of the residue F716 shows close contacts with the meta and para hydrogens of the phenyl ring of thiorphan, thus limiting the bottom of the S1′ pocket. However, analysis of the dynamic simulation reveals large rotations around the χ1 and χ2 (flip–flop) dihedrals angles of F716. Most of the dynamic conformations of NEP, F716 does not evidence interactions with thiorphan. Thus, the S1′ subsite appears deep, with its bottom limited by the residues F718 and L564. The dynamic behavior of the NEP–thiorphan complex agrees with the previously reported ability of NEP to accommodate larger groups than a phenyl ring in the P1′ position (Berger, 1984; Fournié-Zaluski et al., 1984; Ksander et al., 1997a,b).

Position of cysteines

The NEP sequence includes 12 cysteine residues and carboxymethylation studies have suggested the presence of six disulfide bridges (Tam et al., 1985) whose localization is still unknown. In our model, only four cysteines, C620, C694, C734 and C746 are present (Table IV). C694 and C734 display the shortest distances between their respective sulfur atoms during the final dynamic calculation (from 15.7 to 7.5 Å). Moreover, these
two cysteines are buried, whereas C620 and C746 remain accessible on the surface of the model. Thus, a C694–C734 bridge could be hypothesized and residues C620 and C746 could be involved in disulfide bridges with cysteines of the N-terminal region. The formation of the disulfide bond joining C694 and C734 showed that no significant differences occur in our computed model of NEP.

Improvement of the NEP model and inhibitor design

Thiorphan was substituted in our model by benzofused macrocycles thiorphan derivatives (compounds II and III of Table I).

The strain energy of the inhibitors, and the total energy of complexes, have been calculated to test the ability of our model to predict the relative affinities between stereoisomers pairs of compounds II and III (Table V). In agreement with the in vitro potency measurements (Ksander et al., 1997a,b), our results show that the NEP model accommodates the cis (S,S) slightly better than the trans (R,S) isomer of compound II; whereas it highly discriminates between compound III (S,S) and III (R,S) (IC50 are 8 and 1000 nM respectively). The computed relative energy of complexation failed to predict compound II (S,S) to be more potent than the thiorphan inhibitor. This could be due to the neglect of entropic and/or solvation factors in the energy balance.

The ability of the model to accommodate a larger group than a phenyl ring in the P19 position was tested by docking the compound IV, in which the P19 phenyl moiety of thiorphan was replaced by a biphenyl ring (Table I). A comparison of NEP/thiorphan (Figure 3b) and NEP complexed with the biphenyl moiety IV after minimization (Figure 4c) clearly shows that the rearrangement concerns only the residue F716, whereas the interactions between F563 and the proximal ring are similar to those observed in the NEP–thiorphan complex. Thus, the bottom of the S1′ subsite now appears to be limited by residues I718 and L654 (Figure 4c).

As previously reported (Fournié-Zaluski et al., 1994), NEP could accommodate methyl or ethyl substituents on the methylene group of the benzylcic moieties of thiorphan. The relevance of the lipophilicity and size of the S1′ subsite of the NEP model.
model was investigated by using a thiol inhibitor (compound V) with a large and somewhat hindered P1' moiety (Table I). Compound V could be constructed and easily docked in the NEP model (Figure 5). The strain energy of the inhibitor and the relative energy of complexation are similar to those calculated for compound IV (Table V). This is consistent with the experimental results.

Discussion

Sequence homologies between the various zinc metallopeptidase families are very low. Thus, in our model only 38% of the NEP residues could be aligned with TLN. Most of these regions (H2', H3', S10') are very close to the zinc dication, and define, in part, the active site of NEP, which could have a similar organization to that of the TLN active site. The two consensus sequences, HExxH and ExxxD, are included in two α helical regions. They display the same relative orientation, optimizing zinc ligands (H583, H587 and E646) and the catalytic glutamic residue (E584) positions, as well as D650, to allow hydrogen bonding with H583.

Thermolysins and metzincins show a common structural element constituted by five β-strand sheets. However, only three of these five strands are superimposable, and these correspond to S6, S10 and S11 in TLN. Due to the homology between the S6, S10 and S11 of TLN, and between the S6', S10' and S11' in NEP, it is likely that a similar sheet exists in NEP. As only three strands can be proposed in our C-terminal model, the two missing strands could be found in the N-terminal region of NEP, as suggested by secondary structure prediction methods (data not shown).

A hydrogen bond network is observed in TLN between H142, D170, R203 and N238. The first three residues have been shown to correspond to H583, D650 and R717 in NEP, respectively, but no equivalent residue for N238 has yet been proposed. In the structurally non-aligned regions of our model, several sequences are conserved with other nephrilysin family members. Among the conserved residues, Q683 or N741 could be proposed as the N238 analog. The choice is difficult because all these residues are included in large non-aligned regions (660–708 and 721–749) which, in addition, contain three of the 12 cysteines involved in disulfide bridges. The results of a recent directed mutagenesis study on human NEP cysteines (P.Crine, personal communication) are consistent with our model, which suggests a possible C694–C734 bridge.

Inhibition as well as kinetic studies with different substrates of TLN and NEP (Fourmí-Zaluski et al., 1984; Pozsgay et al., 1986) have shown evidence of a preference for aromatic residues in the P1' position. Nevertheless, as compared with TLN, the S1' subsite of NEP accepts larger aromatic side chains than the S1' subsite of NEP (review in Roques et al., 1993; Ksander et al., 1977a,b). This is well illustrated by the only fivefold lower IC50 (10 ± 2 nM) as compared with thiopran of a derivative in which the benzyl group of this inhibitor interacting with the S1' subsite of NEP has been replaced by the very large methylene pyrenyl group which contains four condensed phenyl rings (unpublished result). Moreover, quantitative studies of compound III stereoisomers docked in the active site of the NEP model are in good agreement with the measured inhibitory potencies of these molecules. Finally, a comparison with complexes formed between the NEP model and thiopran, or compound IV, shows that residue F716 could reorientate, suggesting a plasticity to the S1' pocket of NEP. Based on the observed large size of the S1' pocket, new potent inhibitors could be designed to perfectly fit the subsite (Figure 5). One such molecule has been designed (compound V), and has displayed a good inhibitor potency toward NEP, as expected.
Molecular modelling of the active site of NEP

Conclusion

In this work, we have attempted to propose a three-dimensional model of the active site of NEP, taking into account all directed mutagenesis results, except for residue R102, which is at the edge of the active site and involved in the dipeptidyl carboxypeptidase activity of NEP (Beaumont et al., 1992). Despite the omission of the N-terminal region of NEP, and of some uncertainties concerning the large non-aligned regions, our model accounts for both the common structural elements of TLN and NEP, which reflect their identical mechanisms of action, and for the differences in their $S_1'$ subsite topology. This model provides a preliminary tool for further characterization of this enzyme, and for other members of the nephrilysin subfamily, such as ECE. We are aware of the possible mistakes in the arrangement of amino acids, but in the absence of crystallographic data for NEP, this model could facilitate the design of potent and selective inhibitors. Moreover, in the case where the 3D structure of NEP would be solved, the accuracy of the modeling strategy could be challenged. Based on our structural hypothesis, mutagenesis studies aimed at characterizing the residues proposed to have a functional role (equivalent to residue N238 and D236 of TLN), and the refining of the $S_1'$ subsite, are now underway. Moreover, potent and selective inhibitors of NEP could be designed through docking studies.

Acknowledgements

The authors wish to thank P.Crime for providing results of the NEP cysteine mutations, C.Marie-Claire for helpful discussions, A.Beaumont for stylistic revisions and C.Dupuis for expert assistance in drafting the manuscript.

References


Received May 18, 1998; revised October 29, 1998; accepted November 9, 1998

Received May 18, 1998; revised October 29, 1998; accepted November 9, 1998