Volume 305, number 1, 55–61 FEBS 11182 June 1992

A molecular model for the tumour-associated antigen, p97, suggests a Zn-binding function

Richard C. Garratt[a] and Harren Jhoti[b,*]

[a]*Instituto de Fisica y Quimica de São Carlos, Departamento de Fisica y Ciência dos Materials, Universidade de São Paulo, 13560 São Carlos SP, Brazil* and [b]*Laboratory of Molecular Biophysics, Rex Richards Building, South Parks Road, Oxford, UK*

Received 7 April 1992; revised version received 30 April 1992

The primary structure of p97 (melanotransferrin) has been compared with other members of the transferrin superfamily. A molecular structure of p97 has been modelled based on the crystal structure of diferric rabbit serum transferrin. The most significant amino acid substitutions in p97 are almost exclusively limited to only two regions; the C-lobe iron-binding cleft and the interlobe contact region. The latter includes within the N-terminal lobe a Zn-binding consensus sequence found in metallopeptidases, and in the C-terminal lobe a glutamic acid residue (Glu-394) capable of completing a potential thermolysin-like Zn-binding site. Thus, p97 may have a Zn-binding potential, unique amongst the transferrin superfamily.

Transferrin; Zn-binding; Molecular modelling; Tumour antigen; Metallopeptidase; Thermolysin

1. INTRODUCTION

p97 (melanotransferin) is a tumour-associated differentiation antigen expressed at high levels in most human melanomas but present in only trace amounts in normal adult tissues [1]. Whilst most melanomas express between 50,000–500,000 molecules/cell the normal adult tissues with the most abundant p97 expression (smooth muscle) possess only about 8,000/cell. The molecule is a membrane-bound protein whose extracellular domain shares approximately 40% sequence identity with members of the transferrin superfamily, hence its pseudonym [2]. The transferrins are bi-lobal iron-binding glycoproteins capable of sequestering two ferric irons together with synergistic anions at specific sites within each lobe. p97 has similarly been shown to bind iron [3].

The exact role of p97 is unclear, but it may represent a descendent from an ancestral protein which existed prior to the emergence of the soluble transferrins (ovotransferrin, serum transferrin and lactoferrin) [4]. It has recently been suggested that an iron cell-surface binding component may be p97, implying its participation in a transmembrane iron transport system alternative to that of the transferrin receptor-mediated system [5].

Interest in p97 has arisen as a consequence of its potential application in various forms of active and passive antitumour immunotherapy. For example, recombinant vaccinia virus expressing p97 has been shown to induce an immune response in both mice and monkeys [6,7]. In an alternative approach, 3 anti-idiotypic antibodies against 96.5 (an antibody specific for p97[a] epitope) were demonstrated to completely inhibit the binding of p97 to 96.5, however, they were unable to protect mice against p97-positive mouse melanoma cells [8].

In the absence of a crystal structure for p97 little is known of its 3D structure and, in particular, how it differs from the soluble members of the transferrin superfamily for which there are two X-ray crystal structures known; rabbit serum transferrin [9] and human lactoferrin [10]. This has been the stimulus for this investigation.

2. MATERIALS AND METHODS

With the exception of Manduca transferrin [11], whose precise iron-binding characteristics have yet to be determined, all transferrins for which complete amino acid sequences have been reported in the literature were aligned in the following manner. Sequences were divided into individual N- and C-terminal lobes based on the position of the 8/9 exon junction observed in chicken ovotransferrin and human serum transferrin [12–14], and which represents the division of the two halves of the duplicated transferrin gene. For those sequences for which such information was not available the division was based on an initial alignment of the complete sequences.

Lobe sequences were aligned using the program suite AMPS [15] based on a pseudo-phylogenetic tree generated by the method of Feng and Doolittle [16]. This alignment was subject to slight manual adjustment in order that it was consistent with an alignment of human serum transferrin [17,18] and human lactoferrin [10,19], based on the superposition of the crystal structures of rabbit serum transferrin [9,20] and human lactoferrin [10], respectively.

Using the above alignment all positions were identified which have identical or conservatively substituted residues within a given lobe of

Present address: Protein Structure Group, Glaxo Group Research, Greenford Road, Greenford, Middlesex, UB6 0HE, UK.

Correspondence address: R.C. Garratt, Instituto de Fisica y Quimica de São Carlos, Departamento de Física y Ciência dos Materials, Universidade de São Paulo, 13560 São Carlos SP, Brazil. Fax: (55) (162) 722 218.

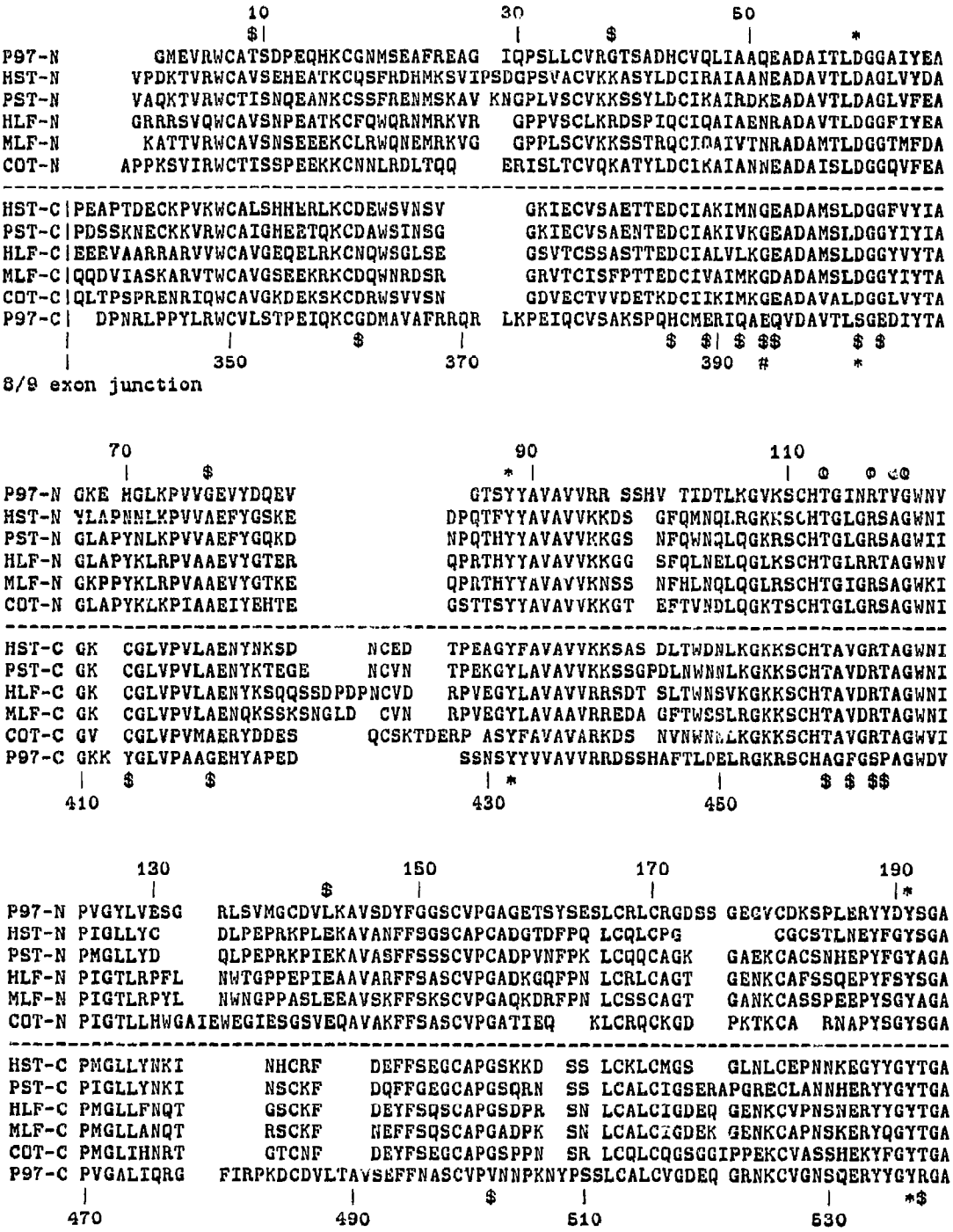

Fig. 1. Alignment of the N- and C-terminal lobes of transferrin sequences. p97, HST, PST, HLF, MLF and COT refer to p97, human serum transferrin, pig serum transferrin, human lactoferrin, mouse lactoferrin and chicken ovotransferrin, respectively. Iron-binding residues are indicated by *, anion-binding residues by @, the proposed zinc-binding residues by #, and the most significant residue substitutions in p97 by $. The residues are numbered according to the sequence of p97. Symbols at the upper and lower halves of each section refer to the N- and C-lobes, respectively. The zinc-binding consensus sequence (around residue 330) is shown boxed, as are the C-terminal helices of lactoferrin and human serum transferrin (based on that of rabbit transferrin) and the proposed membrane anchor of p97.

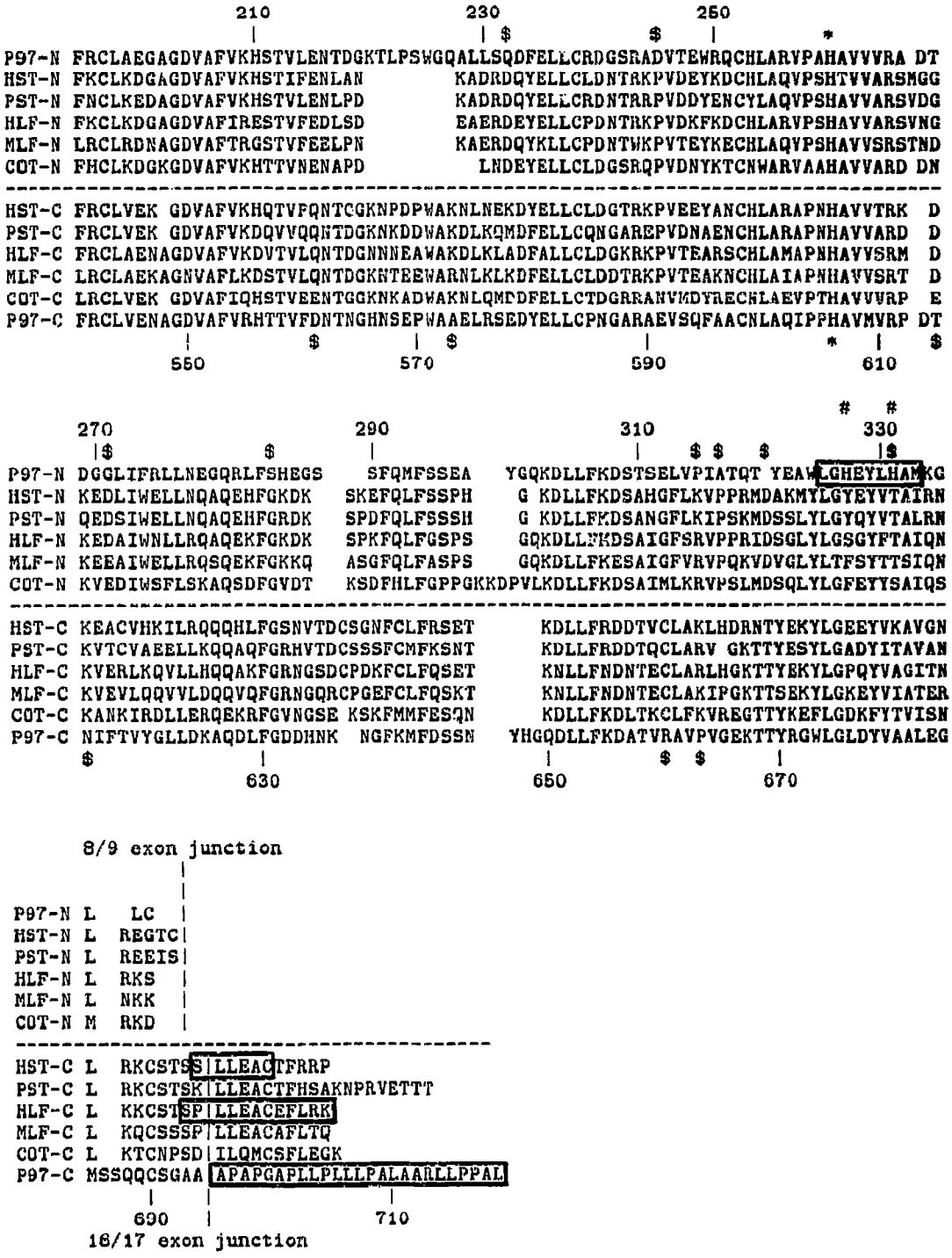

Fig. 1 continued.

the soluble transferrins, but which differ in p97. In order to be more selective in highlighting the most significant of these substitutions these residues were subjected to a filtering process which removed those which were: (i) ambiguous due to ambiguities in the alignment; (ii) those which were only conservative substitutions; and (iii) those for which at least one example of the substituted amino acid occurred in the other lobe.

In order to examine the positions of these substituted residues in the tertiary structure of the molecule a model of p97 was built based primarily on the 3.3 Å crystal structure of diferric rabbit serum transferrin (RST) [9,20]. Using the above determined alignment substitutions were made where necessary using TOM [21] running on a CYBER 910-400 graphics workstation. Loops were modelled by selecting the most appropriate from the four lobes of known crystal structure (N- and C-lobes of RST and human lactoferrin) based primarily on a criterion of length, and secondarily on that of sequence similarity. Loops were excised from the appropriate structure after superposition of the elements of secondary structure immediately preceding and following the loop, using the graphics program WHATIF [22], and the geometry of the loop junctions was corrected using the REFI option within TOM. If no loop of identical length existed insertions or deletions were made, and adjusted manually and using REFI to give good stereochemistry. Despite the presence of seven prolines in this region the membrane-anchoring domain was modelled in a helical conformation (see below). The prolines were modelled into an idealised helix by adjusting the main-chain torsion angles in accord with the average values observed in the study of Barlow and Thornton [23] for up to five residues either side of the proline. Since the number of prolines result in many overlapping segments which require contradictory torsion angles we were guided in these cases by the phi angles at i-1 and i-3 which are observed to differ most from ideal values [23]. The helix was subsequently energy minimised using the program WHATIF [22]. The residues identified as being the most significant substitutions in p97 were examined with respect to this modelled 3D structure.

p97 has been shown by Jongeneel et al. [24] to possess the consensus sequence found in Zn-dependent metallopeptidases such as thermolysin, and which forms the core of their active sites [24–26]. The consensus sequence proposed by Jongeneel et al. [24] is **HE**H*Φ where * and Φ represent uncharged and hydrophobic residues, respectively, and where both histidines form ligands to the zinc ion. In both thermolysin and the modelled structure of p97 this region adopts an alpha-helical conformation. In order to examine further the feasibility of the existence of a possible Zn-binding site in p97 the structures of the consensus regions of the two molecules were superposed using the program WHATIF [22]. The coordinate set used for the thermolysin structure was P3TLN, as deposited in the Brookhaven databank [26]. The resulting positions of the zinc ion and its remaining ligands (Glu-166 and a water molecule) were examined for clashes and favourable interactions with the remainder of the p97 molecule.

3. RESULTS AND DISCUSSION

The alignment of sequences of six members of the transferrin superfamily is shown in Fig. 1. The iron- and anion-binding residues, as identified in the crystal structures of human lactoferrin, rabbit serum transferrin and a half-molecule fragment thereof [27], are indicated (* and @, respectively). Also shown are residues identified as being the most significant residue changes in p97 in comparison with the remaining (soluble transferrin) sequences, indicated by $.

Fig. 2 shows the distribution of these residues in the modelled 3D structure. The figure shows C-α positions only, and clearly highlights that the majority of the amino acid substitutions considered the most dramatic have occurred in one of the two localities; the C-terminal iron-binding cleft (including the iron- and anion-binding sites) and the interlobe contact region. Some of the changes at the C-lobe iron site have been previously noted [29], and this lobe, as expected, does not bind iron [28]. Specifically, the substitution, D-to-S402, affects one of the otherwise conserved iron-binding residues; R-to-S463 and T-to-A459 involve changes to the hydrogen-bonding network which binds the carbonate anion; and T/S-to-P464 and G-to-E404 involve large changes to residues involved in stabilising the orientation of the aspartic acid ligand, as found in the crystal structure of rabbit serum transferrin half molecule [27].

The residue changes at the interlobe interface are partly due to the addition of a hydrophobic tail at the C-terminus of the p97 sequence. This region has been predicted to form the membrane-anchoring domain [2], and is indicated on Fig. 1, along with the region which forms the C-terminal helix in soluble transferrins, and the exon, 8/9, and 16/17 splice junctions. It can be seen that the C-terminal helix of the soluble transferrins (an element of secondary structure unique to the C-terminal lobe) is encoded by exon 17, which is similarly unique to the C-terminal half of the transferrin gene. The equivalent structure to this region in p97 is the membrane-anchoring helix. If it is the case that soluble transferrins were originally membrane proteins which only lost their membrane anchors after the gene duplication event (as has been suggested [4]) then the implication is that the final exon originally encoded a specific functional domain [30], the membrane anchor, which became shortened and folded between the N- and C-terminal lobes. In the model shown (Fig. 2), the membrane anchor has been built such that it extends away from the protein surface. Only in such a way is its length sufficient to fully span a lipid bilayer.

Although it cannot be asserted with confidence that this anchoring region is alpha-helical in nature it has been built as such despite the presence of seven prolines in this region because such short anchoring domains are generally believed to be helical in nature. This is in order to satisfy the hydrogen bonding functions of the backbone amide and carbonyl groups within the lipid membrane. This can only otherwise be achieved through the formation of a beta-barrel structure, for which the p97 anchor is not of sufficient length. There is no experimental evidence which sheds light on this matter, nor does the model suggest any alternative way in which the molecule may interact with the membrane. The result of the modelling is a helix which is substantially curved due to the radially asymmetric distribution of prolines about the helix axis but which avoids bad contacts. The conformation of this region does not affect the conclusions to be drawn in the present study and has only been included for the sake of completeness. The direction of this helix was chosen such that it has the same orienta-

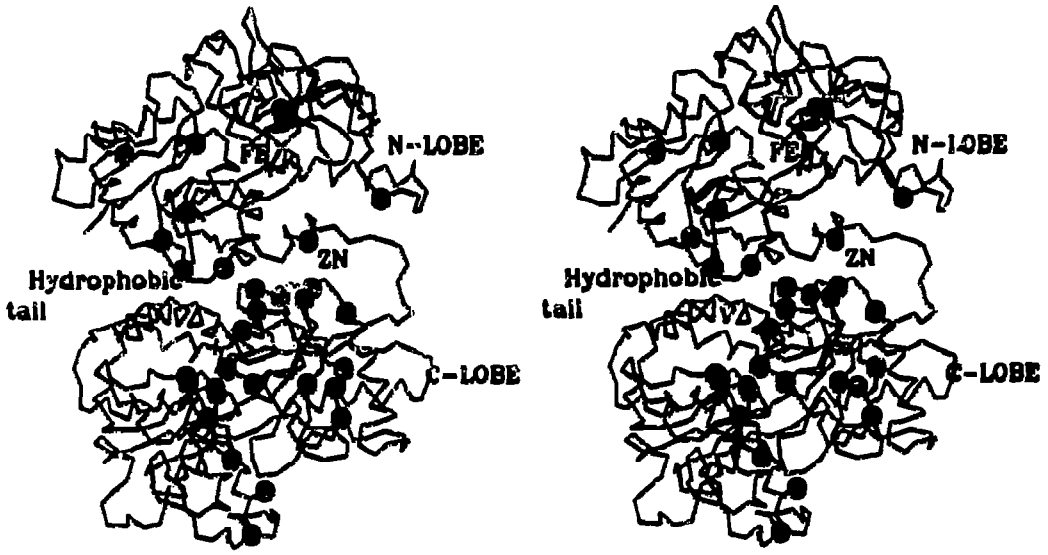

Fig. 2. Stereo C-alpha trace of modelled p97. The most significant residue substitutions in comparison with soluble transferrins are indicated by solid spheres. They can be seen to cluster in the C-lobe iron-binding cleft and at the interlobe interface. The N-terminal iron binding site and proposed zinc site are indicated as is the hydrophobic C-terminal tail.

tion with respect to the C-terminal lobe as does the interlobe helix to the N-terminal lobe of lactoferrin. Although this orientation cannot be asserted with absolute confidence it seems clear that in order to traverse the membrane the helix cannot adopt the same position as the C-terminal helix of soluble transferrins, a hypothesis supported by the absence of a disulphide bridge in this region which presumably allows the tail to adopt a different orientation [29]. The concentration of substitutions at the lobe interface suggests that the relative orientations of the two lobes may be different in p97.

Anderson et al. [10] have identified residues believed to be involved in the interaction between the two lobes of lactoferrin. These residues, which are responsible for the formation of two salt bridges and a hydrophobic patch, are shown in Table I together with their counterparts in p97. Unlike the other soluble transferrins, in which at least one and normally both salt bridges identified previously are at least theoretically possible, p97 is the only sequence to possess neither. A similar argument can be applied to those residues involved in hydrophobic interactions where it can be seen that the nature of the C-terminal hydrophobic surface has been completely altered by virtue of the presence of the membrane anchor and amino acid substitutions. These combined alterations have resulted in the loss of the normal lobe interface from p97 which, together with the observed concentration of the most significant residue changes in this region, suggests that it may either possess a significantly different relative lobe orientation and/or possess a unique function.

Fig. 3 shows the consensus proposed by Jongeneel et al. [24] for Zn-metallopeptidases together with the thermolysin sequence and the region of p97, which they observed to be compatible with the consensus. Also given are the assigned secondary structures for thermolysin and p97, as given by the program of Kabsch and Sander [31]. As can be seen the consensus region adopts, or is predicted to adopt, a helical structure in both

Table I

Residues reported to be directly involved or possibly important in the interlobe interactions of lactoferrin and their counterparts in p97

Lactoferrin		p97		Comments
N-lobe	C-lobe	N-lobe	C-lobe	
Salt bridges				
R313 with D379		Q319	H386	At least one and normally both present in other transferrins
D315 with K386		–	A393	
Hydrophobics				
V310		I316		The hydrophobic character of these residues is largely conserved in other transferrin sequences
P311		A317		
I314		T320		
L318		A323		
Y319		W324		
F325		L330		
	A382		E389	
	L385		Q392	
	L682		No equivalent due to membrane anchor	
	A685			
	F688			
	L689			

CONSENSUS		* * H E * * H * Φ	
Thermolysin	140	V G H E L T H A V	148
p97	325	L G H E Y L H A M	333
Thermolysin		α α α α α α α α α	
p97		C α α α α α α α α	

Fig. 3. The Zn-metallopeptidase consensus sequence is shown together with that of thermolysin (a member of the metallopeptidase family) and the region of p97 between residues 325 and 333 which matches the consensus. * and Φ refer to uncharged and hydrophobic residues, respectively. Given below the sequences is the secondary structure assignment for both proteins, where α and C designate helix and coil, respectively [31].

molecules. Based on this observation it was decided to test the hypothesis that p97 might contain a Zn-binding site.

After least-squares superposition of the two molecules using the C-α positions of the consensus sequence no clashes were found between the Zn ion, its water ligand and the remainder of the p97 molecule. Furthermore, the superposition revealed that the closest equivalent C-α position to that of Glu-166 in thermolysin (which together with the two histidines of the consensus sequence completes the Zn coordination sphere) is also that of a glutamic acid residue in p97 (Glu-394). This strongly suggests that a Zn-binding site might easily be formed by this constellation of residues in p97, and the presence of Glu-328 within the consensus sequence indicates that it may also have peptidase activity. The equivalent residue in thermolysin (Glu-143) is believed to be the critical residue in catalysis [32,33]. More specifically the role of Glu-143 in thermolysin is probably to act as a general base and as a proton shuttle, transfering the abstracted proton from the Zn-bound water molecule and adding it to the scissile peptide nitrogen. Furthermore, the fact that this residue is the most important component of the catalytic apparatus of Zn-metalloproteinases is demonstrated by the observation that the only elements found in a truly equivalent spatial arrangement in thermolysin and the exopeptidase, carboxypeptidase, are the zinc and Glu-143 [32,34]. The amino acid sequence and the remainder of the 3D structure of carboxypeptidase are completely unrelated to those of thermolysin and yet they are believed to share a common catalytic mechanism. The modelled p97 Zn-binding site is shown together with that of thermolysin in Fig. 4.

It is of note that the proposed Zn-site resides in the cleft between lobes (a common feature of catalytic sites) and that the Zn ion receives ligands from both lobes, as in thermolysin. As described above this particular region of the p97 molecule shows many significant sequence differences in comparison with other transferrins. Included amongst such substitutions are the proposed ligands, Glu-394 and His-331. His-327 is also unique to p97. This accumulated evidence, in conjunction with the loss of iron-binding capacity in the C-terminal lobe, suggests that p97 may have acquired an additional function, that of Zn-binding and, possibly, peptidase activity. Although, to our knowledge, no direct experimental evidence exists concerning Zn-binding or peptidase activity of p97, the evidence presented herein for the proposed site, although entirely theoretically based, is compelling and merits further investigation.

Acknowledgements: The authors would like to thank the members of the Transferrin group at Birkbeck College and also Ted Baker for coordinates of lactoferrin. R.C.G. gratefully acknowledges the support of the Brazilian National Research Council (CNPq) via Grant RHAE/CNPq Proc. 317/88-BIO and São Paulo State Research Council (FAPESP) under the auspices of the Institute of Advanced Studies of the University of São Paulo.

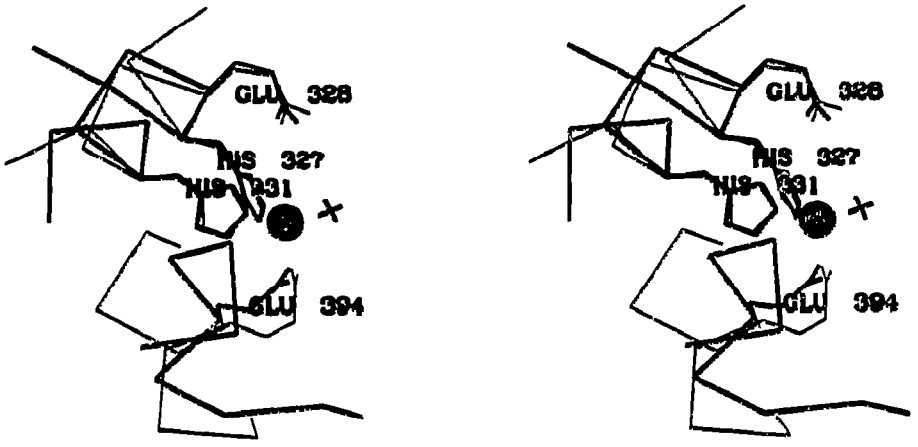

Fig. 4. The proposed Zn-binding site of p97 (bold lines) is shown together with that of thermolysin (thin lines) which was used for the modelling after superposition of the consensus sequences. The structure of the consensus region itself can be modelled without problem into p97. The remaining zinc ligand (Glu-394) adopts a slightly different conformation to that observed in thermolysin but can still coordinate to the Zn. The position of the water molecule is shown by the X, and the Zn by the sphere.

REFERENCES

[1] Hellstrom, K.E. and Hellstrom, I. (1989) FASEB J. 3, 1715–1721.
[2] Rose, T.M., Plowman, G.D., Teplow, D.B., Dreyer, W.J., Hellstrom, K.E. and Brown, J.P. (1986) Proc. Natl. Acad. Sci. USA 83, 1261–1265.
[3] Brown, J.P., Hewick, R.M., Hellstrom, I., Hellstrom, K.E., Doolittle, R.F. and Dreyer, W.J. (1982) Nature 296, 171–173.
[4] Williams, J. (1982) Trends Biochem. Sci. 7, 394–397.
[5] Richardson, D.R. and Baker, E. (1990) Biochim. Biophys. Acta 1053, 1–12.
[6] Hu, S.L., Plowman, G.D., Sridhar, P., Sterenson, U.S., Brown, J.P. and Estin, C.D. (1988) J. Virol. 62, 176–180.
[7] Estin, C.D., Stevenson, U.S., Plowman, G.D., Hu, S.L., Sidhara, P., Helstrom, I., Brown, J.P. and Hellstrom, K.E. (1988) Proc. Natl. Acad. Sci. USA 85, 1052–1056.
[8] Kahn, M., Hellstrom, I., Estrin, C.D. and Hellstrom, K.E. (1989) Cancer Res. 49, 3157–3162.
[9] Bailey, S., Evans, R.W., Garratt, R.C., Gorinsky, B., Hasnain, S.S., Horsburgh, C., Jhoti, H., Lindley, P.F., Mydri, A., Sara, R. and Watson, J. (1988) Biochemistry 27, 5804–5812.
[10] Anderson, B.F., Baker, H.M., Norris, G.E., Rice, D.W. and Baker, E.N. (1989) J. Mol. Biol. 209, 711–734.
[11] Bartfield, N.S. and Law, J.H. (1990) J. Biol. Chem. 265, 21684–21691.
[12] Jeltsch, J.M., Hen, R., Maroteaux, L., Garnier, J.M. and Chambon, P. (1987) Nucleic Acids Res. 15, 7643–7645.
[13] Park, I., Schaeffer, E., Sidoli, A., Baralle, F.E., Cohen, G.N. and Zakin, M.M. (1985) Proc. Natl. Acad. Sci. USA 82, 3149–3153.
[14] Schaetter, F., Lucero, M.A., Jeltsch, J.M., Py, M.C., Levin, M.J., Chambon, P., Cohen, G.N. and Zakin, M.M. (1987) Gene 56, 109–116.
[15] Barton, G. and Sternberg, M.J.E. (1987) J. Mol. Biol. 198, 327–337.
[16] Feng, D.F. and Doolittle, R.F. (1987) J. Mol. Evol. 25, 351–360.
[17] MacGillivray, R.T.A., Mendez, E., Schewale, J.G., Sinha, S., Lineback-Zins, J. and Brew, K. (1983) J. Biol. Chem. 258, 3543–3553.
[18] Yang, F., Lum, J.B., McGill, J.R., Moore, C.M., Naylor, S.L., Van Bragt, P.M., Baldwin, W.D. and Bowman, B.H. (1984) Proc. Natl. Acad. Sci. USA 81, 2752–2756.
[19] Metz-Bartique, M.H., Jolles, J., Mazurier, J., Schoentgen, F., Legrand, D., Spik, G., Martrerill, J. and Jolles, P. (1984) Eur. J. Biochem. 145, 659–676.
[20] Jhoti, H. (1989) Ph.D. Thesis, University of London.
[21] Cambillau, C.C. and Horjoles, E. (1987) J. Mol. Graph. 5, 174–177.
[22] Vriend, G. (1990) J. Mol. Graph. 8, 52–56.
[23] Barlow, D.J. and Thornton, J.M. (1988) J. Mol. Biol. 201, 601–619.
[24] Jongeneel, C.V., Bouvier, J. and Bairoch, A. (1989) FEBS Lett. 242, 211–214.
[25] Holmes, M.A. and Matthews, B.W. (1982) J. Mol. Biol. 160, 623–639.
[26] Holmes, M.A. and Matthews, B.W. (1981) Biochemistry 20, 6912–6920.
[27] Sarra, R., Garratt, R.C., Gorinsky, B., Jhoti, H. and Lindley, P. (1990) Acta Cryst. B46, 763–771.
[28] Baker, E.N., Baker, H.M., Smith, C.A., Stebbins, M.R., Kahn, M., Hellstrom, K.E. and Hellstrom, I. (1992) FEBS Lett. 298, 215–218.
[29] Baker, E.N., Rumball, S.V. and Anderson, B.F. (1987) Trends Biochem. Sci. 12, 350–353.
[30] Gilbert, W. (1978) Nature 271, 501.
[31] Kabsch, W. and Sander, C. (1983) Biopolymers 22, 2577–2637.
[32] Monzingo, A.F. and Matthews, B.W. (1984) Biochemistry 23, 5724–5729.
[33] Hangauer, D.G., Monzingo, A.F. and Matthews, B.W. (1984) Biochemistry 23, 5730–5741.
[34] Kester, W.R. and Matthews, B.W. (1977) J. Biol. Chem. 252, 7704–7710.