Vol. 95, No. 1, 1980 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
July 16, 1980 Pages 334-341

THE PRIMARY STRUCTURE OF BOVINE LENS LEUCINE AMINOPEPTIDASE

Complete amino acid sequence of the N-terminal cyanogen bromide fragment and site of limited tryptic digestion

L.A.H. van Loon-Klaassen, H.Th Cuypers, H. van Westreenen, W.W. de Jong and H. Bloemendal

Department of Biochemistry, University of Nijmegen, Geert Grooteplein Noord 21, 6525 EZ Nijmegen, The Netherlands

Received May 15,1980

SUMMARY

The amino acid sequence of the N-terminal cyanogen bromide fragment of bovine lens leucine aminopeptidase has been determined. This fragment contains a total of 171 amino acid residues and has a calculated molecular weight of 18,637. The sequence data presented here represent the first report of primary structure determination of a member of the class of aminopeptidases.
The single cleavage site produced by limited tryptic digestion of native leucine aminopeptidase was determined to be between arginine-137 and lysine-138 of the total amino acid sequence. The possible existence of distinct structural domains in leucine aminopeptidase is discussed.

Leucine aminopeptidase (EC 3.4.11.1) is an exopeptidase catalyzing the hydrolysis of amino acids from the N-terminus of peptides and proteins (1). It has a molecular weight of 326,000 (2,3) and consists of six identical subunits of molecular weight 54,000 (4,5). Leucine aminopeptidase (LAP)* belongs to the class of aminopeptidases which are widely distributed in nature. At present no primary structure of any of the aminopeptidases has been published. We have therefore undertaken an investigation of the amino acid sequence of LAP from bovine lens. The present study established the sequence of the N-terminal cyanogen bromide fragment of LAP.

Limited tryptic digestion of native LAP resulted in the cleavage of a very specific bond (6). Despite splitting of this bond the enzyme retained all of its catalytic properties due to the fact that the LAP aggregate remained intact. The results of the sequence determination of the N-terminal cyanogen bromide fragment of LAP enabled us to identify the exact cleavage site of limited tryptic digestion of native LAP.

MATERIALS AND METHODS

Leucine aminopeptidase was isolated from bovine eye lenses according to Hanson et al. (7). After reduction of possible disulfide bridges the

* Abbreviations: LAP, leucine aminopeptidase; MW, molecular weight; PTH, phenylthiohydantoin; SDS, sodium dodecyl sulfate; Tris, tris(hydroxymethyl)aminomethane.

0006-291X/80/130334-08$01.00/0

cysteine residues of LAP were modified by S-carboxymethylation (3) or S-aminoethylation (8) prior to cyanogen bromide cleavage. The reduction time was extended to 4 h and larger excess of alkylating reagent was used. Cyanogen bromide cleavage was performed in 70% formic acid for 18 h at room temperature in the dark, at a protein concentration of 10 mg/ml using an equal weight of CNBr. The fragments were separated by gel filtration on a column (200 x 3.5 cm) of Sephadex G-100 fine, eluted with 10% acetic acid containing 6 M urea, at a flow rate of 10 ml/h. The fragments were desalted on a column of Sephadex G-25 coarse (30 x 4.0 cm) in 10% acetic acid. SDS polyacrylamide gel electrophoresis was carried out according to de Jong et al. (9). Citraconylation was performed according to Atassi and Habeeb (10). Digestions of CB1 by trypsin (Worthington TRTPCK), chymotrypsin (Worthington CBI) and thermolysin (Calbiochem A grade) were carried out in 0.1 M NH_4HCO_3, pH 8.9, at 37 °C. Tryptic and chymotryptic digestion of CB1 was performed for 2 h at a protein concentration of 10 mg/ml using 1% (w/w) of enzyme initially, and an additional 1% after 1 h. Thermolytic digestion of CB1 was carried out for 0.5 h at a protein concentration of 15 mg/ml using 1% (w/w) of enzyme. Digestion of tryptic peptide T14 by Staphylococcus protease (Miles, S. aureus V8) was carried out for 6 h at 40 °C in 0.1 M NH_4HCO_3, pH 8.0, at a peptide concentration of 0.5 mg/ml using 5% (w/w) of enzyme (11). A first separation of peptides was obtained by gel filtration on a column (120 x 1.5 cm) of Sephadex G-50 superfine, eluted at room temperature with 0.05 M NH_4HCO_3 or 0.1 M ammonia at a flow rate of 5 ml/h. The effluent was monitored at 280 nm and 206 nm. Large peptides were further purified by ion-exchange chromatography on DEAE-cellulose (Whatman DE-52, 11.0 x 1.0 cm) according to Howard et al. (12). Small peptides were purified by high voltage paper electrophoresis at pH 6.5 followed by descending chromatography in the second dimension in n-butanol/glacial acetic acid/water/pyridine (15:3:12:10, by volume). Amino acid analysis was carried out with a Chromaspek amino acid analyzer (Rank Hilger). The dansyl-Edman technique was performed according to Gray and Smith (13). Dansylated amino acids were identified by thin-layer chromatography on polyamide sheets (14). For direct manual Edman degradation the accelerated version of Niall and Potts (15) was used. Identification of PTH-amino acids was performed by thin-layer chromatography and by high-performance liquid chromatography (16). Automated Edman degradation was performed in a Beckman spinning cup sequenator, model 890 C, using the quadrol double cleavage program. Amide groups were assigned on the basis of the charge of the peptides calculated according to Offord (17), or by direct identification as PTH-derivative in the Edman degradation.

RESULTS AND DISCUSSION

Essentially pure LAP was isolated from calf lenses by the method of Hanson et al. (7). About 200 mg of enzyme was obtained routinely from 800 lenses. The amino acid composition of LAP is shown in table I. The values agree reasonably well with those reported earlier (5). The subunit of LAP thus contains approximately 490 residues. The presence of 11 methionine residues in LAP provided a good basis for cyanogen bromide cleavage. The N-terminal amino acid sequence of LAP was determined up to threonine-21 by automated Edman degradation.

The cyanogen bromide cleavage mixture of CM-LAP was chromatographed on Sephadex G-100 (fig. 1A). Dissociating conditions were necessary to prevent aggregation of the fragments. Analysis of the different pools by SDS gel electrophoresis (fig. 1B) revealed that pool II contained a fragment of

Table I Amino acid composition of bovine lens leucine aminopeptidase and of some fragments

LAP: (a) present results, (b) values taken from Carpenter and Vahl (5); CB1: (a) values obtained from amino acid analysis, (b) values obtained from sequence results; LT-I: (a) values obtained from amino acid analysis, (b) sum of residues 1 to 137.
Values are expressed as residues per subunit (for LAP) or as residues per fragment (for CB1 and LT-I). Values for threonine and serine were extrapolated to zero time hydrolysis, and values for valine and isoleucine were taken from the 72 h hydrolysate.

	LAP		CB1		LT-I	
Amino acid	(a)	(b)	(a)	(b)	(a)	(b)
Aspartic acid	46.5	47	14.8	15	13.2	13
Threonine	24.1	25	5.2	5	4.8	5
Serine	29.2	28	8.4	9	5.6	6
Homoserine			0.7	1		
Glutamic acid	52.3	53	25.0	26	22.0	21
Proline	23.8	26	6.3	6	5.7	6
Glycine	42.3	42	17.7	18	14.9	15
Alanine	52.5	53	17.0	16	13.0	13
Cysteine*	6.8	8	1.8	2	1.9	2
Valine	33.3	33	13.4	14	9.3	9
Methionine	11.2	11			0.0	0
Isoleucine	27.4	29	7.1	7	7.1	7
Leucine	38.3	39	15.7	16	12.0	12
Tyrosine	8.9	10	3.7	4	3.9	4
Phenylalanine	19.2	20	5.3	5	4.1	4
Histidine	8.9	8	3.2	3	2.2	2
Lysine	36.1	34	13.8	14	12.7	12
Arginine	22.4	21	8.0	8	5.4	5
Tryptophan	3.8**	10	1.8	2	n.d.	1
Total	487.0	496	168.9	171	137.8	137
MW	52,799	54,060	18,414	18,637	15,117	15,189

* determined as S-carboxymethylcysteine
** unreliable value
n.d. = not determined

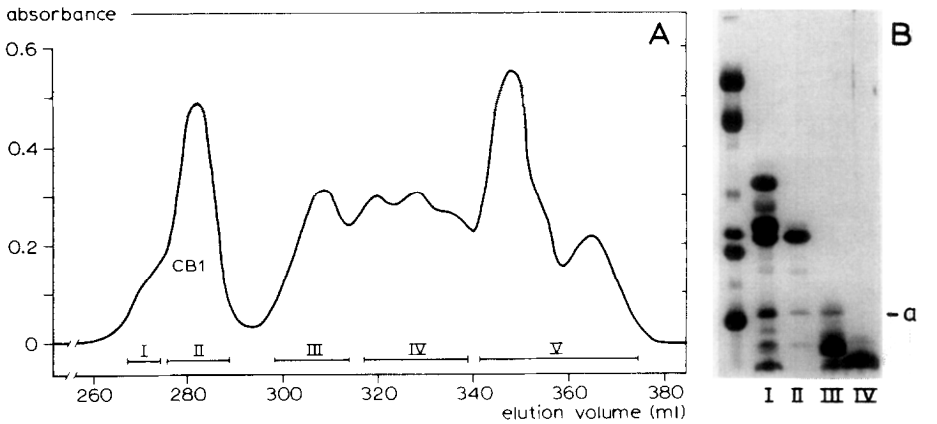


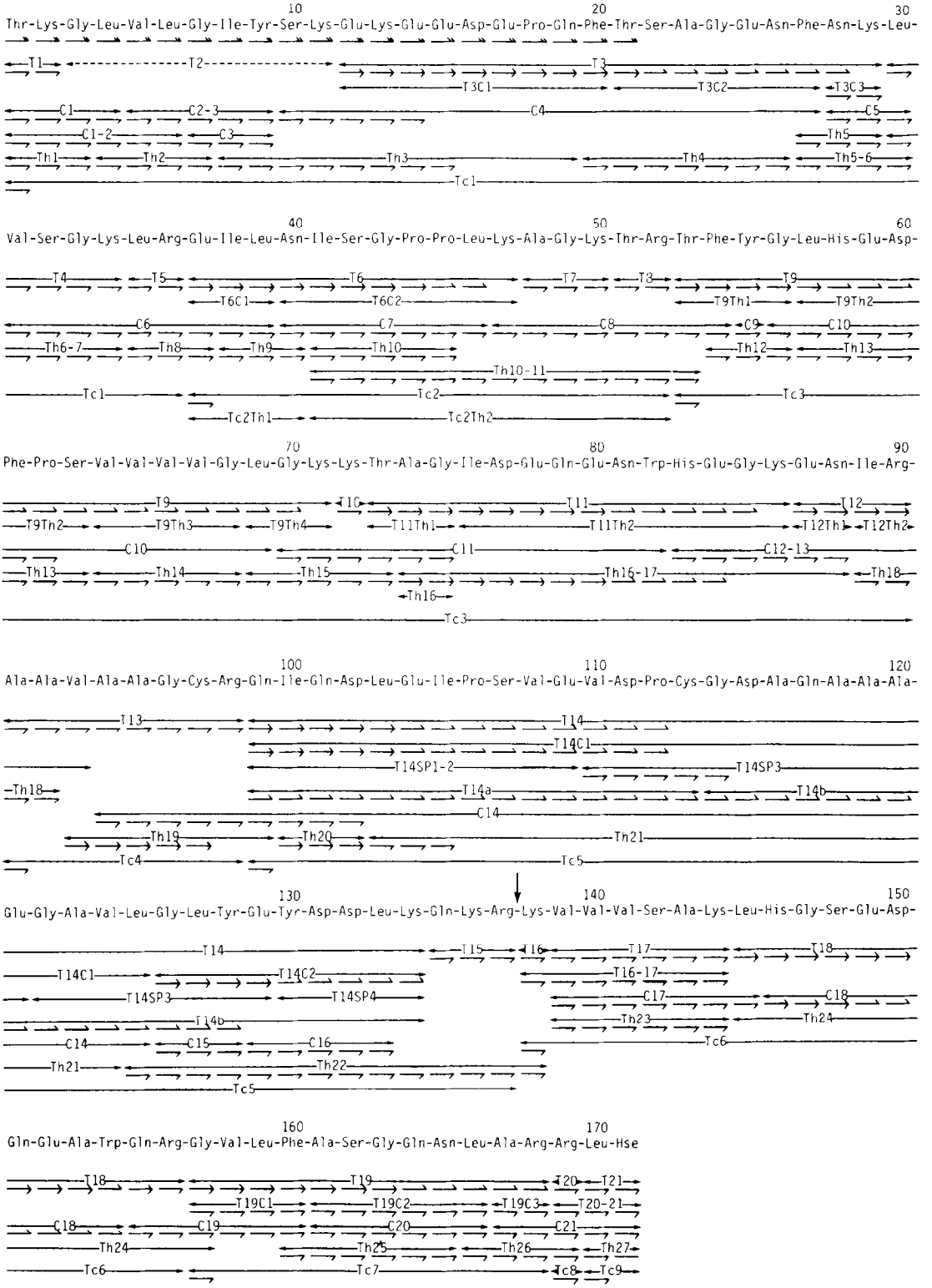
Fig. 1 A Chromatography of cyanogen bromide fragments of reduced and S-carboxymethylated LAP on Sephadex G-100.
B SDS polyacrylamide gel electrophoresis of pools I to IV.

molecular weight 21,000. This fragment was essentially pure after re-chromatography on the same column. A sequenator run of this fragment up to residue 21 gave the same result as found for native LAP. Hence we concluded that this fragment was the N-terminal cyanogen bromide fragment, and it was designated CB1. Occasionally a shortened fragment (a in fig. 1B) of molecular weight 12,500 was found, resulting from the cleavage of an Asp-Pro bond in CB1 under the acidic conditions of the reaction. No further purification was performed when CB1 contained small amounts of this fragment. The amino acid composition of CB1 is listed in table I.

The accumulated evidence for the sequence determination of CB1 is given in fig. 2. The amino acid sequence of the tryptic peptides of CB1 was determined by direct Edman degradation and by the dansyl-Edman method. In some cases subdigestion of tryptic peptides by thermolysin, chymotrypsin or Staphylococcus protease was performed in order to confirm or complete the sequence. In peptide T3 a lysine bond was not cleaved by trypsin, probably because of several adjacent acidic residues. The sequence of the large peptide T14 was difficult to determine because of the presence of two glutamine residues which blocked the Edman degradation due to cyclization, causing a rapid decrease in yield of the next PTH-amino acids. The presence of a cysteine residue in the middle of T14 enabled us to solve the problem by isolating peptides T14a and T14b from a tryptic hydrolysate of CB1 from reduced and S-aminoethylated LAP. Direct Edman degradation of these peptides completed the sequence of T14. Peptide T15 contained a Lys-Arg bond that was not cleaved by trypsin. Two peptides containing a homoserine residue were found as a result of the partial splitting of an Arg-bond by trypsin.

The order of the tryptic peptides in the primary structure was determined by isolating overlapping peptides obtained after digestion of CB1 by thermolysin and chymotrypsin, and tryptic digestion of citraconylated CB1. Although peptide T2 has never been isolated, its sequence was not only proved by the sequenator run on total CB1, but also by the chymotryptic and thermolytic peptides. Most overlaps between the tryptic peptides involved two or more residues in the chymotryptic or thermolytic peptides. When an overlap only involved one residue, the order of the tryptic peptides was established by the citraconyl peptides. In this way all residues in CB1 were identified and all overlaps rigidly established. As can be seen from fig. 2, CB1 contains a total of 171 amino acid residues. The sequence is in good agreement with the amino acid composition obtained from hydrolysis (table I).

The molecular weight calculated from the sequence is 18,637. However, polyacrylamide gel electrophoresis in the presence of SDS resulted in an estimated molecular weight of 21,000. The discrepancy between the two values


10
20
30
Thr-Lys-Gly-Leu-Val-Leu-Gly-Ile-Tyr-Ser-Lys-Glu-Lys-Glu-Glu-Asp-Glu-Pro-Gln-Phe-Thr-Ser-Ala-Gly-Glu-Asn-Phe-Asn-Lys-Leu-
T1
T2
T3
T3C1
T3C2
T3C3
C1
C2-3
C4
C5
C1-2
C3
Th5
Th1
Th2
Th3
Th4
Th5-6
Tc1
40
50
60
Val-Ser-Gly-Lys-Leu-Arg-Glu-Ile-Leu-Asn-Ile-Ser-Gly-Pro-Pro-Leu-Lys-Ala-Gly-Lys-Thr-Arg-Thr-Phe-Tyr-Gly-Leu-His-Glu-Asp-
T4
T5
T6
T7
T8
T9
T6C1
T6C2
T9Th1
T9Th2
C6
C7
C8
C9
C10
Th6-7
Th8
Th9
Th10
Th12
Th13
Th10-11
Tc1
Tc2
Tc3
Tc2Th1
Tc2Th2
70
80
90
Phe-Pro-Ser-Val-Val-Val-Val-Gly-Leu-Gly-Lys-Lys-Thr-Ala-Gly-Ile-Asp-Glu-Gln-Glu-Asn-Trp-His-Glu-Gly-Lys-Glu-Asn-Ile-Arg-
T9
T10
T11
T12
T9Th2
T9Th3
T9Th4
T11Th1
T11Th2
T12Th1
T12Th2
C10
C11
C12-13
Th13
Th14
Th15
Th16-17
Th18
Th16
Tc3
100
110
120
Ala-Ala-Val-Ala-Ala-Gly-Cys-Arg-Gln-Ile-Gln-Asp-Leu-Glu-Ile-Pro-Ser-Val-Glu-Val-Asp-Pro-Cys-Gly-Asp-Ala-Gln-Ala-Ala-Ala-
T13
T14
T14C1
T14SP1-2
T14SP3
Th18
T14a
T14b
C14
Th19
Th20
Th21
Tc4
Tc5
130
140
150
Glu-Gly-Ala-Val-Leu-Gly-Leu-Tyr-Glu-Tyr-Asp-Asp-Leu-Lys-Gln-Lys-Arg-Lys-Val-Val-Val-Ser-Ala-Lys-Leu-His-Gly-Ser-Glu-Asp-
T14
T15
T16
T17
T18
T14C1
T14C2
T16-17
T14SP3
T14SP4
C17
C18
T14b
Th23
Th24
C14
C15
C16
Tc6
Th21
Th22
Tc5
160
170
Gln-Glu-Ala-Trp-Gln-Arg-Gly-Val-Leu-Phe-Ala-Ser-Gly-Gln-Asn-Leu-Ala-Arg-Arg-Leu-Hse
T18
T19
T20
T21
T19C1
T19C2
T19C3
T20-21
C18
C19
C20
C21
Th24
Th25
Th26
Th27
Tc6
Tc7
Tc8
Tc9

is another example of the unreliability of this method for molecular weight determination (9).

Limited proteolysis of native proteins has proved to be an important tool in the investigation of protein structure (18). The scissile bond (or bonds) may be located in terminal parts or in exposed loops. Alternatively, limited proteolysis may occur in a region which connects preformed domains of the molecule. When native LAP in 0.1 M Tris-HCl buffer, pH 8.0, was treated with trypsin (0.5%, w/w), two fragments were obtained which could only be separated under dissociating conditions (6). The isolated fragments had molecular weights of 17,000 and 37,000. The 17,000 MW fragment with N-terminal threonine was found to be the N-terminal fragment, and the 37,000 MW fragment with N-terminal lysine was the C-terminal fragment. They were denoted LT-I and LT-II, respectively. In order to determine the exact cleavage site of trypsin, fragment LT-I was further investigated. The amino acid analysis of LT-I is listed in table I. Analysis of the tryptic hydrolysate of reduced and S-carboxymethylated LT-I revealed that peptides T1 to T15 were present in LT-I. These results, and the fact that lysine was the N-terminal residue of LT-II, indicated that the exact cleavage site of trypsin in native LAP is between Arg-137 and Lys-138 in the primary structure. This position is indicated by a vertical arrow in fig. 2. The bond is contained in a cluster of basic residues, namely Lys-Gln-Lys-Arg-Lys (res. 134-138).

The fact that the splitting of the Arg-Lys bond is complete after approximately 3 h, indicates that this bond is split in all six subunits. Consequently the surface topology around the susceptible bond probably is identical in all subunits. Investigations of the quaternary structure of LAP have shown that to each of the subunits an additional, elongated "tail" is attached (19). These tails, extending from the molecule, may be especially susceptible to attack by proteolytic enzymes. Alternatively, the trypsin-susceptible bond may be located in a region connecting distinct domains. The N-terminal part of LAP thus may represent a distinct structural domain comprising about a quarter of the molecule. The size of LT-I (137 residues) corresponds well with that of separate domains found in many proteins (between

---

Fig. 2 Proposed amino acid sequence of the N-terminal cyanogen bromide fragment of bovine lens leucine aminopeptidase

CB1 was digested by trypsin (T), chymotrypsin (C), thermolysin (Th) and by trypsin after citraconylation (Tc). Some tryptic peptides were subdigested by chymotrypsin, thermolysin or Staphylococcus protease (SP). Peptide T2 was not isolated. Peptides T14a and T14b resulted from tryptic digestion of CB1 of reduced and S-aminoethylated LAP. Residues were identified by automated Edman degradation (⟶), manual Edman degradation (⟶), dansyl-Edman method (⟶), or by both manual methods (⟶). The heavy vertical arrow indicates the single cleavage site of limited tryptic digestion of native LAP.

100 and 200 residues). The remaining C-terminal part of LAP may comprise one or more domains. In view of this hypothesis the finding of Taylor et al. (20) that the subunits of LAP are bilobal with a principal and a minor lobe, is of special interest. However, the existence of distinct domains in LAP can only be verified by detailed X-ray structural analysis.

The finding that the LAP aggregate remains intact after limited tryptic digestion indicates that the fragments are held together either by disulfide bonds or by strong secondary interactions. The fact however, that LT-I and LT-II can be separated without prior reduction, excludes the possibility of disulfide bonds holding the fragments together. As can be seen in fig. 2, two cysteines are present in LT-I, namely Cys-97 and Cys-133. Neither of these is bound by an S-S bridge to a cysteine residue farther away in the primary structure. Carpenter and Vahl (5) and Frohne (21) found that of the 8 cysteines thought to be present in each subunit of LAP, six were in the sulfhydryl form, and hence they concluded that one S-S bridge is present. Our results of limited tryptic digestion of native LAP point to two possibilities: either both cysteines in LT-I are in the sulfhydryl form, or these residues form the disulfide bridge. Future investigations using [^{14}C]-monoiodoacetic acid will discriminate between these two possibilities.

The partial amino acid sequence of bovine lens leucine aminopeptidase proposed here represents the first report of primary structure determination of a member of the class of aminopeptidases. The sequence can provide a basis for the interpretation of X-ray crystallographic studies, of which a preliminary account has been presented (22). Amino acid sequence and three-dimensional structure could provide a better molecular insight into the mechanism of action of the aminopeptidases.

ACKNOWLEDGEMENTS

We thank Marlies Versteeg for amino acid analyses and Gerrit Groenewoud for isolation of leucine aminopeptidase. Automated Edman degradation was performed at the Department of Biochemistry, University of Groningen, The Netherlands (Prof. J.J. Beintema). This work was supported in part by the Netherlands Foundation for Chemical Research (S.O.N.) with financial aid from the Netherlands Organization for Pure Research (Z.W.O.).

REFERENCES

1. Hanson, H. and Frohne, M. (1977) Methods Enzymol. 45, 504-521.
2. Kretschmer, K. and Hanson, H. (1965) Z. Physiol. Chemie 340, 126-137.
3. Melbey, S.W. and Carpenter, F.H. (1971) J. Biol. Chem. 246, 2459-2463.
4. Carpenter, F.H. and Harrington, K.T. (1972) J. Biol. Chem. 247, 5580-5586.
5. Carpenter, F.H. and Vahl, J.M. (1973) J. Biol. Chem. 248, 294-304.

6. Van Loon-Klaassen, L., Cuypers, Th. and Bloemendal, H. (1979) FEBS Lett. 107, 366-370.
7. Hanson, H., Glässer, D. and Kirschke, H. (1965) Z. Physiol. Chemie 340, 107-125.
8. Raftery, M.A. and Cole, R.D. (1966) J. Biol. Chem. 241, 3457-3461.
9. De Jong, W.W., Zweers, A. and Cohen, L.H. (1978) Biochem. Biophys. Res. Commun. 82, 532-539.
10. Atassi, M.Z. and Habeeb, A.F.S.A. (1972) Methods Enzymol. 25, 546-553.
11. Austen, B.M. and Smith, E.L. (1976) Biochem. Biophys. Res. Commun. 72, 411-417.
12. Howard, N.L., Joubert, F.J. and Strydom, D.J. (1974) Comp. Biochem. Physiol. 48B, 75-85.
13. Gray, W.R. and Smith, J.F. (1970) Anal. Biochem. 33,36-42.
14. Woods,K.R. and Wang, K.T. (1967) Biochim. Biophys. Acta 113, 369-370.
15. Niall, H.P. and Potts, J.T. (1970) in Peptides, Chemistry and Biochemistry (Weinstein, B. and Lande, S., eds.) pp. 215-241, Dekker, New York.
16. Frank, G. and Strubert, W. (1973) Chromatographia 6, 522-524.
17. Offord, R.E. (1966) Nature 211, 591-593.
18. Mihalyi, E. (1978) Application of Proteolytic Enzymes to Protein Structure Studies (Weast, E.D., ed.) 2nd edition, CRC-Press, Florida.
19. Ludewig, M., Hanson, H., Kiselev, N.A., Stel'mashchuk, V.Ya. and Tsuprun, V.L. (1977) Acta Biol. Med. Germ. 36, 157-165.
20. Taylor, A., Carpenter, F.H. and Wlodawer, A. (1979) J. Ultrastruct. Res. 68, 92-100.
21. Frohne, M. (1977) Chemische Modifizierung von Leucinaminopeptidase, Thesis, University of Halle, G.D.R.
22. Jurnak, F., Rich, A., van Loon-Klaassen, L., Bloemendal, H., Taylor, A. and Carpenter, F.H. (1977) J. Mol. Biol. 112, 149-153.