N-TERMINAL SEQUENCE ANALYSIS OF CHICKEN PEPSINOGEN AND PEPSIN

V.KOSTKA, H.KEILOVÁ, K.GRÜNER and J.ZBROŽEK

Institute of Organic Chemistry and Biochemistry, Czechoslovak Academy of Sciences, 166 10 Prague 6

Received April 6th, 1977

A homogeneous preparation of chicken pepsinogen, characterized by molecular weight determination (42800 ± 1300), and the mixture of peptides liberated during its activation to chicken pepsin were subjected to sequential analysis. Automatic sequential degradation of native chicken pepsinogen or its derivatives and the analyses of the activation peptides permitted the N-terminal 26-residue amino acid sequence of chicken pepsinogen to be defined as follows: Ser-Ile-His-Arg-Val-Pro-Leu-Leu-Lys-Gly-Lys-Ser-Leu-Arg-Lys-Gln-Leu-Lys-Asp-His-Gly-Leu-Leu-Glu-Asp-Phe. This sequence is homologous with the N-terminal sequences of several other zymogens of carboxyl proteases. A similar sequential investigation was carried out also with chicken pepsin (mol.wt. 33400 ± 1000). Its N-terminal amino acid sequence, Thr-Ala-(Thr)-Glu-Glu-Tyr-Asp-... does not show any homology with the N-terminal sequences of the corresponding proteases.

In spite of the fact that acid proteases are enzymes of very frequent occurrence in nature and that some of them were isolated as the first ones from various organisms, their sequential investigation has not started until recently. The elucidation of the covalent structure of hog pepsin was completed not long ago[1,2]. Only partial structures are known of all other acid proteases and their zymogens (bovine pepsinogen A and B[3-8], human pepsin[9,10] *etc.*) with the exception of prochymosin[11-14] and penicillopepsin[15] whose sequential studies have been concluded. Chicken pepsin differs from all acid proteases reported so far in its content of seven half-cystine residues; of the latter 6 are joined together by 3 disulfide bonds and one is present as a cysteine residue with a free thiol group. This free thiol group is buried inside the zymogen molecule and is unmasked after its activation; in pepsin this thiol group can be titrated and chemically modified[16]. Since thiol groups can easily be blocked by heavy metal atoms, chicken pepsin is a perspective object of X-ray diffraction analysis. Its necessary prerequisite is the knowledge of the primary structure of the protein.

In preceding papers from this Laboratory we reported on the isolation and purification of active pepsin from chicken forestomachs[17,18]. We observed that chicken pepsin isolated in active form is not suitable for sequential studies; we decided therefore to isolate its inactive zymogen. In this paper experiments are described which led to homogeneous preparations of chicken pepsinogen and chicken pepsin obtained by activation of the zymogen. The results of sedimentation analyses of both proteins as well as the sequential investigation of the N-terminal parts of their molecules are reported.

EXPERIMENTAL

Material

Chicken forestomaches were obtained from chickens within one hour after their slaughter. Adhering fat was removed and the forestomachs were washed with cold physiological saline, then frozen to −20°C, and stored at this temperature until treated further. Carboxypeptidase A and B were DFP-treated commercial preparations of Worthington Bioch. Corp., Freehold, N. J., U.S.A. DEAE-cellulose (Cellex D, std. cap.) was from Calbiochem, San Diego, Cal., U.S.A., cellulose was purchased from Whatman Biochemicals, Maidstone, England. Sephadex G-25 (particle size 20—80 μ) and Sephadex G-100 (particle size 40—120 μ) were products of Pharmacia Fine Chemicals, Uppsala, Sweden. The chemicals used for automatic sequential degradation were supplied by the sequencer manufacturer (Beckman Inst., Spinco Div., Palo Alto, Cal., U.S.A.). Silufol, silica gel thin layer sheets for chromatography of 3-phenyl-2-thiohydantoins of amino acids, were purchased from Kavalier, Czechoslovakia. All the remaining chemicals used in this study were of analytical purity.

Methods

The proteolytic activity of pepsin or the potential proteolytic activity of pepsinogen were determined with denatured hemoglobin as substrate as described elsewhere[17].

Nitrogen content was determined by the method of Kjeldahl in protein samples of known moisture and ash content.

Amino acid analysis of protein or peptide samples was carried out after hydrolysis of the samples (20, 70, or 120 h, in 6M-HCl at 110°C) on Beckman-Spinco Model 120 B Amino Acid Analyzer or on an analyzer of Czechoslovak make (Instrument Development Workshops, Czechoslovak Academy of Sciences, Prague); the method of Spackman and coworkers[19] as modified by Benson and Patterson[20] was used. The oxidation of protein for the determination of half-cystine as cysteic acid was carried out according to Moore[21]. Tryptophan content was estimated spectrophotometrically[22]. Amino sugars were determined after 12 h hydrolysis of the protein in 2M-HCl at 110°C according to Plummer[25].

Determination of sedimentation constant and molecular weight. The sedimentation coefficient was determined in Model E Beckman-Spinco Ultracentrifuge equipped with schlieren optics. The molecular weight was determined in the same instrument by the method of sedimentation equilibrium in the long column according to Chervenka[23].

Manual sequential degradation of peptides was carried out according to Niall and Edman[24] as described in detail elsewhere[25].

Automatic sequential degradation of pepsin, pepsinogen, and their derivatives was effected in Model 890 C Beckman Amino Acid Sequencer. The degradation was allowed to proceed according to our modification of fast Quadrol program No 07172 recommended by the instrument manufacturer. Phenylthiohydantoins of amino acids, obtained by conversion of thiazolinones, were identified by thin layer chromatography[26,27] (Silufol) or by gas chromatography in Model GC-65 Beckman Gas Chromatograph, either as such or after silylation[28].

The C-terminal end groups of the proteins were estimated from the kinetics of cleavage with carboxypeptidase A and B. The digestion was carried out at a 1 : 100 enzyme to substrate molar ratio, 5 min to 24 h at 37°C. The technique of treatment of the digests and their analysis were described in detail elsewhere[25].

The procedures used for the separation of peptide mixtures by paper techniques and the preparation of peptide maps were described in detail earlier[29].

The preparation of chicken pepsinogen was carried out according to Bohak[16]. Chicken forestomachs (100 g) were homogenized in 0·4M-Tris acetate buffer, pH 8·6. Proteins were precipitated by the addition of acetone (−15°C, final concentration 85%) and crude pepsinogen extracted from the precipitate by the same buffer. The extract was centrifuged (50000 *g*) and pepsinogen precipitated by the addition of acetone (−15°C, final concentration 70%) to the clear supernatant. Pepsinogen was extracted from the precipitate with 0·05M-Tris-acetate buffer, pH 8·6, the extract cleared by centrifugation, equilibrated with 0·05M borate buffer, pH 7·6, by dialysis, and chromatographed on a 2·5 × 20 cm DEAE-cellulose column equilibrated with the same buffer. Protein contaminants were washed off with the same buffer made 0·1M in NaCl. Pepsinogen was subsequently displaced by a linear elution gradient of increasing NaCl concentration (0·1−0·5M, volume of mixing device and reservoir 300 + 300 ml). The absorbance of the effluent was measured at 280 nm and the proteolytic activity was determined on aliquots of individual fractions. Pepsinogen-containing fractions were pooled and dialyzed against ammonia water (pH 8·0). The dialyzed material was lyophilized and rechromatographed on DEAE-cellulose under identical conditions. The rechromatographed material after dialysis was lyophilized, dissolved to a 2% solution in 0·4M-NH_4HCO_3 and passed over Sephadex G-100 equilibrated with the same solution. The effluent was tested as described above. Fractions showing identical specific activity were pooled and lyophilized.

Preparation of pepsin from pepsinogen. Pepsinogen (500 mg) was dissolved in 25 ml of 1 mM-HCl and the pH of the solution adjusted to 2·5 by 1M-HCl. The solution was centrifuged 20 min later and the clear solution was passed over a Sephadex G-100 column, equilibrated with 1 mM-HCl, to separate the enzyme from peptides formed in the process of activation. Effluent fractions were subjected to absorbance measurement at 280 nm and to the proteolytic activity assay. Enzyme-containing fractions were pooled and lyophilized. Fractions containing the activation peptides were also pooled and lyophilized. All operations were carried out at +4°C.

Reduction and aminoethylation of pepsinogen. The procedure[30] developed by Dopheide and coworkers[31] for hog pepsin was used. The protein (200 mg, 4·56 µmol) was dissolved in 8M solution of deionized urea containing 5 mg of ethylenediamine tetraacetic acid and 340 mg of amediol (2-amino-2-methyl-1,3-propanediol). The total volume of the reaction mixture was made up to 12 ml with water; the resulting pH was 10·6. The solution was flushed with nitrogen, 1·2 ml (1·38 g, 18 mmol) of 2-mercaptoethanol was added, the solution was repeatedly flushed with nitrogen, and incubated in a sealed-off vessel 4 h at 37°C. The molar excess of 2-mercaptoethanol was 563-fold (on the assumption of the content of seven half-cystine residue per mol). At the end of the reduction the reaction mixture was mixed with an equal volume of 5M amediol, pH 8·6. The SH-groups were labeled by the addition of ethylene imine (4 × 0·6 ml added at 30 min intervals). The absence of SH-groups was tested by the nitroprusside reagent. The molar excess of ethylene imine with respect to 2-mercaptoethanol was 2·6-fold. After completion of the substitution the reaction mixture was desalted either by gel filtration on Sephadex G-25 equilibrated with 0·1% NH_4HCO_3 or by dialysis against several changes of water adjusted to pH 9·0 by concentrated ammonia. The desalted product was lyophilized.

Aminoethylation of native pepsin. Native chicken pepsin containing 0·88 mol of free SH-group per mol of protein according to the results of the titration assay[32] was used to start with. Pepsin (198 mg, 5·92 µmol) was dissolved in 20 ml of 1M-Tris-HCl buffer, pH 8·6, which was 0·45 mM in dithiothreitol and the solution was stirred 10 min at room temperature. Subsequently the solution was treated with 4 × 10 µl (8·3 mg, 193 µmol) of ethylene imine in 30-min intervals (a total of 77 µmol of ethylene imine, *i.e.* a 150-fold excess per one SH-group). The total reaction time was 120 min. The reaction product was desalted on a column (2·5 × 34 cm) of Sephadex G-25 equilibrated with 1% NH_4HCO_3 and lyophilized.

Reduction and S-sulfonation of pepsinogen and pepsin. The original procedure of Pechére and coworkers[33] was used for the preparation of both derivatives. The reaction mixture was desalted on a column of Sephadex G-25 equilibrated with 0·2M-$(NH_4)_2CO_3$.

The oxidation of pepsin was effected by the method of Hirs[34].

RESULTS

Chromatography of crude pepsinogen. Crude pepsinogen was isolated by procedures carried out at alkaline pH in order to prevent its spontaneous activation to pepsin. When subjected to chromatographic separation of DEAE-cellulose one main fraction and three additional fractions showing weak potential proteolytic activity were obtained (Fig. 1). The latter zymogens or enzymes were not studied in detail. The active fraction was rechromatographed on DEAE-cellulose and subjected to gel filtration of Sephadex G-100 to separate the remaining traces of high molecular weight contaminants. Effluent fractions were pooled according to their specific activity; hence, essentially fractions presenting the middle part of the peak were pooled. By this approach we were able to obtain 10·5 g of homogeneous pepsinogen from 1 500 g of tissue.

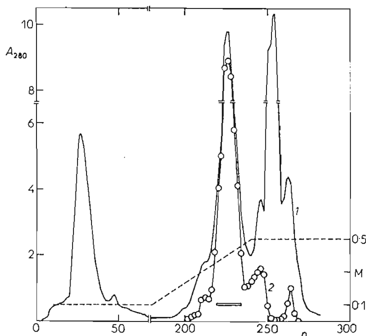

FIG. 1

Chromatography of Crude Chicken Pepsinogen on DEAE-Cellulose

The fraction obtained by precipitation with acetone (170 ml, corresponding to 100 g of material) was placed on a 2·5 × 20 cm column. Flow rate 38·4 ml/15 min (1 fraction). First buffer 0·05M sodium borate, pH 7·6, 0·1M in NaCl. The column was eluted by a gradient of increasing concentration of sodium chloride (0·1—0·5M-NaCl, 300 + 300 ml). *1* Absorbance at 280 nm, *2* proteolytic activity, *n* tube number, ----- molarity of sodium chloride (0·1—0·5M).

Preparation of pepsin from pepsinogen. The separation of active pepsin on a column of Sephadex G-100 after the activation of pepsinogen is shown in Fig. 2. The first peak represents active pepsin, the third peak contains peptides formed in the process of activation. In a typical experiment 500 mg of pepsinogen yielded about 315 mg of the active enzyme and 170 mg of the peptide mixture.

Sedimentation analysis of pepsinogen and pepsin. Both proteins gave one peak only in the ultracentrifuge. The sedimentation coefficient of pepsinogen was $s_{20,w} = 3{\cdot}5S$ and the molecular weight 42800 ± 1300. The sedimentation coefficient of pepsin is $s_{20,w} = 3{\cdot}7S$ and the molecular weight 33400 ± 1000.

Amino acid composition of pepsinogen and pepsin. The results of amino acid analyses of both proteins are given in Table I. Pepsinogen contains 379 amino acid residues and its molecular weight calculated from this analysis is 43879. Analytical data on pepsin show 304 residues and a molecular weight of 34 220. Both the molecule of pepsinogen and of pepsin contains glucosamine (6 residues in both cases). The presence of other carbohydrates was not examined in our Laboratory. From the difference between the two analyses 75 residues should be split off in the process of activation. Amino acid analysis of the activation peptides indicates the presence of 71 residues.

N-*terminal end group analysis.* The determination of the N-terminal end groups of pepsinogen and pepsin was effected by manual Edman degradation. Pepsinogen was N-terminated by serine, pepsin by threonine.

C-*terminal analysis.* The quantities of amino acids liberated by carboxypeptidase digestion of the native proteins were very small (about 0·1 mol per mol of protein).

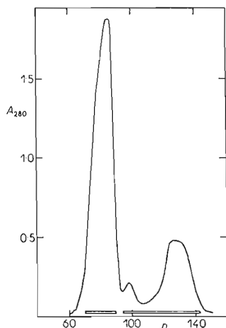

FIG. 2

Activation of Chicken Pepsinogen to Pepsin. Gel Filtration of Activation Mixture on Sephadex G-100

Pepsinogen (500 mg) activated 20 min at pH 2·5 was placed on a 4·5 × 114 cm column. Flow rate, 30 ml/h, fractions 15 ml/ /30 min. The column was eluted by 1 mM-HCl. *n* tube number.

TABLE I

Amino Acid Composition of Chicken Pepsinogen and Chicken Pepsin

Duplicate analyses[19,20] were performed on samples hydrolyzed 20, 70, and 120 h. Unless otherwise stated the figures shown below are averages of all these values. Replicate analyses always agreed to $\pm 3\%$. For the calculation of molar amino acid ratios the average micromoles of alanine were taken to represent 18 residues per molecule for pepsinogen and 15 residues per molecule for pepsin.

Amino acid	Amino acid residues						
	pepsinogen		pepsin		difference	mixture of activation peptides	
	found	to nearest integer	found	to nearest integer		found	to nearest integer
Lysine	18·78	19	8·74	9	10	10·30	10
Histidine	7·84	8	3·87	4	4	4·15	4
Arginine	7·37	7	4·47	4	3	3·00	3
Half-cystine[a]	6·80	7	6·85	7	0	0·52	1
Aspartic acid	41·13	41	34·16	34	7	7·50	8
Threonine[b]	26·85	27	24·83	25	2	3·63	4
Serine[b]	38·59	39	34·76	35	4	4·87	5
Glutamic acid	27·23	27	21·05	21	6	5·60	6
Proline	18·46	19	13·90	14	5	3·94	4
Glycine	31·64	32	27·41	27	5	4·80	5
Alanine	18·00	18	15·00	15	3	3·00	3
Valine	25·73[c]	26	21·95[d]	22	4	3·41	3
Methionine	9·25	9	7·94	8	1	1·14	1
Isoleucine	22·95[c]	23	19·66[d]	20	3	2·72	3
Leucine	29·81	30	19·96	20	10	5·83	6
Tyrosine	21·69	22	17·58	18	4	2·28	6
Phenylalanine	21·32	21	17·18	17	4	2·69	3
Tryptophan[e]	4·24	4	3·86	4	0		
Ammonia	36·67	(37)[g]	33·86	(34)[g]	(3)	—[i]	—
Total		379		304			
Glucosamine[f]	6·27	6	5·68	6	0	—[i]	
Percentage of nitrogen calculated	96·21		96·15				
Molecular weight		43 879		34 220			

[a] Determined as cysteic acid on a sample of the oxidized protein[21]; [b] Extrapolated to zero time of hydrolysis; [c] 70-h hydrolysis; [d] 120-h hydrolysis; [e] Determined spectrophotometrically[22]; [f] Determined according to[38]; [g] Not included in total number of amino acid residues; [h] Determined according to[35]; [i] Not determined.

Therefore derivatives were subjected to digestion whose chain had been unfolded by S-sulfonation or aminoethylation. The results of these experiments are summarized in Table II and show that pepsinogen and pepsin share a common C-terminal sequence, -Leu-Ser-COOH.

Automatic sequential degradation. The stepwise degradation of native pepsinogen (0·4 μmol) proceeded with difficulties. With the increasing number of steps the sample became denaturated and could be dissolved only with difficulties. Therefore the degradation was performed on aminoethylated pepsinogen (0·35 μmol). Twenty-four steps were carried out yet 17 residues only could be determined, of their number 13 unambiguously as follows: H_2N-Ser-Ile-His-Arg-Val-Pro-Leu-Leu-Lys-Gly-Lys--X-Leu-Arg-(Lys)-(Gln)-Leu. The residues given in brackets gave low yields only of the corresponding phenylthiohydantoins. The amino acids degraded in the 17th to 24th step could not be identified beyond doubt. Native pepsin (0·5 μmol) was degraded to the 16th step, yet only the first six amino acids were determined with certainty.

TABLE II

Kinetics of Digestion of Chicken Pepsinogen and Pepsin by Carboxypeptidase A

The protein samples (140–180 nmol) in 1 ml of 0·2M-Tris-HCl buffer, pH 8·5, were digested with carboxypeptidase A at a molar enzyme to substrate ratio of 1 : 100 at 37°C. The technique of treatment of the digests and of their analysis were described elsewhere[25].

Protein	Reaction time min	Amino acids liberated mol/mol protein		
		serine[a]	leucine	other
S-sulfo-pepsinogen	5	0·09	0·06	Tyr 0·06
	15	0·09	0·08	Tyr 0·06
	30	0·13	0·11	Tyr 0·06
	60	0·25	0·19	Tyr 0·06
	300	0·54	0·50	Tyr 0·13
	1 440	0·71	0·71	Tyr 0·23; Ala 0·19
S-sulfo-pepsin	5	0·08	0·03	
	15	0·16	0·06	
	30	0·24	0·13	
	60	0·34	0·21	
	300	0·51	0·38	
	1440	0·85	0·71	

[a] The absence of asparagine or glutamine in the digests was shown in specialized experiments. No basic amino acids were liberated by carboxypeptidase B.

The results of the degradation experiments carried out with oxidized pepsin or S-sulfo-pepsin (0·5 μmol) gave essentially the same results which permit the sequence of seven residues at the terminus to be defined as H_2N-Thr-Ala-(Thr)-Glu-Glu-Tyr-Asp. The yield of threonine in step 3 was very low; traces of serine were also detected in step 5, and traces of serine and tyrosine in step 7.

Analysis of peptide mixture liberated during activation of pepsinogen to pepsin. The first fractionation of the mixture was effected by gel filtration. The material (433 mg) was dissolved in 1% $(NH_4)_2CO_3$ to a 2% solution and placed on a 4·5 × × 120 cm column of Sephadex H-25 equilibrated with the same buffer. The fractionation was monitored by absorbance measurement of effluents at 230 nm and by paper chromatography[29] of aliquots corresponding to 2% of each fraction. According to the results of these analyses effluent fractions containing similar material were pooled. The pools were lyophilized and analyzed by the technique of peptide maps[29]. Individual peptides were isolated from pooled fractions by paper techniques. Only those peptides which according to their amino acid composition (Table III) were likely to provide additional information extending the sequence obtained in the sequencer were chosen for further treatment. These peptides were sequenced manually[24,25].

N-*terminal* 26-*residue sequence of pepsinogen.* The sequences obtained by manual Edman degradation of some of the activation peptides were aligned with the 16-residue sequence obtained by automatic degradation; thus the sequence of 26-residues

TABLE III

Amino Acid Composition of Peptides AP-1 — AP-10 Derived from N-Terminal Part of Molecule of Chicken Pepsinogen

Amino acid	Peptide, mol/mol									
	AP-1	AP-2	AP-3	AP-4	AP-5	AP-6	AP-7	AP-8	AP-9	AP-10
Lysine	1·70	1·07		1·10	1·00	1·00	0·97	0·86		
Histidine	0·90	0·93		0·98	0·93					
Arginine	0·90						1·00			
Aspartic acid	1·00	1·00	0·68	1·00	1·00				0·86	1·00
Serine						0·87				
Glutamic acid	1·90	0·98	1·00				0·90	0·71		1·00
Glycine	1·30	1·00		1·02	1·21					
Leucine	2·90	2·00		1·98	1·00	1·00	0·93	1·00		
Phenylalanine									1·00	1·05

at the N-terminus of chicken pepsinogen was derived as follows: H_2N-Ser-Ile-His--Arg-Val-Pro-Leu-Leu-Lys-Gly-Lys-X-Leu-Arg-Lys-Gln-Leu-Lys-Asp-His-Gly-Leu--Leu-Glu-Asp-Phe. The alignment of the peptides is shown in Fig. 4.

DISCUSSION

Chicken pepsin like hog pepsin[35] which has been isolated from the stomach mucosa directly, *i.e.* in the form of active enzyme omitting the zymogen step, is not homogeneous enough to serve as starting material for sequential work. We decided therefore to prepare first its zymogen. The pepsinogen preparation obtained by us was homogeneous as regarded its sedimentation behaviour. N-Terminal end group analysis of this preparation carried out by manual Edman degradation showed the presence of the one N-terminal group only, of serine. This finding is in disagreement with earlier data on chicken pepsinogen in this respect: Bohak[16], who was the first one to analyze chicken pepsinogen from this viewpoint, did not find any N-terminal end group, Green and Llewellin[36] did find N-terminal serine yet only after reduction and carboxymethylation of the protein. For C-terminal end group analysis we converted pepsinogen into its S-sulfo derivative because the native zymogen was resistant to carboxypeptidase digestion. Our results like those of other authors[16,36] provide evidence of C-terminal sequence -Leu-Ser-COOH (Table II).

The amino acid composition of chicken pepsinogen shown in Table I is in accordance with the data reported by other authors[36,39], especially with the analysis made by Bohak[16], except for the amide concent of pepsinogen where we arrive at a considerably higher number. The molecule of pepsinogen (and pepsin) contains also a carbohydrate moiety. Our analysis of aminohexose content according to Plummer[38] showed 6 glucosamines; Kay[37] reports for our preparation 4 residues of N-acetyl-glucosamine and 2 residues of mannose. Bohak[16] found in its preparation a carbohydrate moiety corresponding to 6–7 "glucose equivalents" and 2 glucosamine residues by amino acid analysis. Green and Llwellin report 7 galactosamine residues in pepsinogen and 6 residues in pepsin. The molecular weight of pepsinogen calculated from its amino acid analysis, *i.e.* 43879 is in good agreement with the data obtained by sedimentation analysis (42800 $\pm$ 1300).

Like hog pepsin[35] chicken pepsin obtained by activation of purified pepsinogen is homogeneous and thus differs from the enzyme obtained directly from stomach mucosa[40]. Our preparation was homogeneous on ultracentrifugation. One N-terminal end group only, threonine, was found. The carboxyl terminal sequence of chicken pepsin was the same as the sequence of the zymogen, *i.e.* -Leu-Ser-COOH. The results of our terminal analyses, similarly to the data obtained by other authors, demonstrate that chicken pepsin is formed from its zymogen by cleavage in the N-terminal part of the zymogen molecule liberating a certain peptide segment of its polypeptide chain. A similar observation has been made with all animal acid (carboxyl)

N-Terminal 26-residue Sequence of Chicken Pepsinogen Derived from Automatic Edman Degradation of Protein and Analysis of Activation Peptides

H_2N - Ser - Ile - His - Arg - Val - Pro - Leu - Leu - Lys - Gly - Lys - Ser - Leu - Arg - Lys - Gln - Leu - Lys - Asp - His - Gly - Leu - Leu - Glu - Asp - Phe - ...

AP-6: (Lys, Ser, Leu)

AP-4: Lys - Asp - His - Gly - Leu - Leu

AP-9: (Asp, Phe)

AP-2: Lys - Asp (His, Gly, Leu, Leu, Glx)

AP-7: Arg - Lys - Gln - Leu

AP-3: (Glx, Asx)

AP-1: Arg - Lys - Glu (Leu, Lys, Asx, His, Gly, Leu, Leu, Glx)

AP-8: (Lys, Gln, Leu)

AP-5: (Lys, Asx, His, Gly, Leu)

AP-10: Glu - Asp - Phe

Edman degradation → automatic, → manual

FIG. 3

N-Terminal 26-Residue Amino Acid Sequence of Chicken Pepsinogen Derived from Results of Automatic Sequential Degradation of Protein and Analysis of Activation Peptides

The peptides isolated from the activation mixture are marked by symbol AP and horizontal bars. The arrows designate the results of automatic (——→) or manual (——→) sequential degradation.

proteases whose zymogens have been known. The cleaved-off activation segment (propart[42]) represents about 10% of the parent zymogen molecule. The corresponding number of amino acid residues is for hog[11] and bovine[4,8] pepsinogen and for prochymosin[11,12] 42–44 residues The difference in the molecular weight of chicken pepsinogen and chicken pepsin is considerably greater. From sedimentation analysis this difference is 9400 whereas the corresponding value derived from molecular weights calculated from amino acid composition is 9659. Chicken pepsin contains therefore 75 residues less than its zymogen. The part of the pepsinogen chain cleaved off under the experimental conditions described in this study represents a mixture of peptides. This is evidenced by a great number of N-terminal amino acids found with this material. Our results thus differ from Bohak's[16] finding of lysine as the only N-terminal amino acid of the peptide mixture. The amino acid composition of the unfractionated peptide mixture is in good agreement with the theoretical value derived from the difference in amino acid composition of chicken pepsinogen and pepsin. We were able to obtain a number of homogeneous peptides from the activation mixture and to determine their sequence. We regard therefore as unlikely that under the experimental conditions described here one peptide only of molecular weight 2000 (ref.[36]) is split off or that 1–2 bonds only are split, as postulated by some authors[16]. The solution of the problem of bonds split in the process of chicken pepsinogen activation calls, as was the case with hog pepsin[41], for specialized experiments which are under progress in our Laboratory[44] at present.

Automatic sequential degradation of chicken pepsinogen derivatives with cleaved disulfide bonds provided information on 17 residues at the N-terminus of the molecule. The characteristic pattern of the last four residues of this sequence together with the knowledge of the amino acid composition of the activation mixture permitted us to align some of the peptides (Fig. 3) with the sequence determined automatically and thus to extend the sequence to 26 residues. The correctness of the one-residue overlap at position 24 and of the asignment of serine to position 12 was demonstrated in recent experiments with automatic sequential degradation (according to program No 102 974 in dimethylallylamine buffer) of the intact N-terminal 26-residue peptide primarily liberated from the molecule of chicken pepsinogen in the process of its activation[44]. Sequential information on the group of acid proteases and their zymogens is meagre: as of now 3 complete (or nearly complete) primary structures are known, namely of hog pepsin[1,2] and its zymogen[11,43], of calf prochymosin[14], and of one microbial acid protease, penicillopepsin[45]. By contrast fragmentary data especially on terminal parts of polypeptide chains of other acid proteases and their zymogens are relatively abundant. An aligment of these sequences including also the 26-residue N-terminal sequence of chicken pepsinogen determined in this study, is shown in Fig. 4. It can be seen that in 6 zymogens whose N-terminal sequence of residue 1–27 (numbered with respect to prochymosin[42]) is known, 7 positions in this region are clearly conservative. These positions, 7, 8, 12, 15, 22, 24, and 27, are

Alignment of N-Terminal Sequences of Acid Protease Zymogens

		Sequence (positions 10 and 20 marked)
Pg	hog A	Leu - Val - Lys - Val - PRO - LEU - Val - Arg(10) - Lys - LYS - Ser - Leu - ARG - Gln - Asn - Leu - Ile - Lys(20) - Asp - GLY - Lys - LEU - Lys - Asp - PHE
	bovine A	Ser - Val - Val - Lys - Ile - PRO - LEU - Val - Lys - Lys - LYS - Ser - Leu - ARG - Gln - Asn - Leu - Ile - Glu - Asn - GLY - Lys - LEU - Lys - Glu - PHE
	bovine B	Leu - Val - Lys - Ile - PRO - LEU - Lys - Lys - Phe - LYS - Ser - Ile - ARG - Glu - Ile - Met - Lys - Glu - Lys - GLY - Leu - LEU - Glx - Asx - PHE
Pch	calf	Ala - Glu - Ile - Thr - Arg - Ile - PRO - LEU - Tyr - Lys - Gly - LYS - Ser - Leu - ARG - Lys - Ala - Leu - Lys - Glu - His - GLY - Leu - LEU - Glu - Asp - PHE
Pg	seal	Ala - Ile - Ile - Arg - Val - PRO - LEU - Val - Lys - Lys - LYS - Ala - Leu - ARG - Lys - Asx - Leu - Arg - Glx - His - GLY - Leu - LEU - Glx - Asx - PHE
	chicken	Ser - Ile - His - Arg - Val - PRO - LEU - Leu - Lys - Gly - LYS - Ser - Leu - ARG - Lys - Gln - Leu - Lys - Asp - His - GLY - Leu - LEU - Glu - Asp - PHE
	dogfish	Leu - Ile - Arg - Val - PRO - LEU - Ile - Lys

Pg = pepsinogen

Pch = prochymosin

FIG. 4

Alignment of N-Terminal Sequences of Acid Proteinase Zymogens

The system of numbering and alignment is that of Foltmann and Pedersen[42] summarizing data from papers[3,8,11,12,43]. Amino acids which occupy identical positions in all zymogens are set in capital letters.

in all zymogens occupied by the same amino acid. Moreover there are other positions which show a high degree of homology. A hypothesis has been voiced[46] that the conservative distribution of mainly basic amino acids especially in the N-terminal portions of the zymogen molecules is of functional importance and maintains inactive conformation of the zymogen through electrostatic interactions with negative charges in the enzyme part of the molecule. The same holds for C-terminal portions of the molecules. The presence of C-terminal sequence -Leu-Ser.COOH (obviously preceded by a proline residue as follows from the kinetics of the cleavage) in chicken pepsinogen suggests its homology with other acid proteases also in this region. By contrast, we have not been able to align the 7-residue N-terminal sequence of chicken pepsin determined here, Thr-Ala-(Thr)-Glu-Glu-Tyr-Asp-, with the corresponding sequences of other acid proteases.

We are indebted to Mr J. Neumann for the sedimentation analyses and to Dr F. Franěk for the interpretation of the obtained data. We thank Mrs E. Bulantová and Mrs J. Zaoralová for skillful technical assistance. We are indebted to Miss V. Himrová and Mrs E. Dršková for the amino acid analyses of the peptides and Mrs A. Čermáková for manual stepwise degradation of the activation peptides.

REFERENCES

1. Morávek L., Kostka V.: FEBS (Fed. Eur. Biochem. Soc.) Lett. *43*, 207 (1974).
2. Tang J., Sepulveda P., Marciniszyn K. jr, Chem K. C. S., Huang W.-Y., Tao N., Liu D., Lanier J. P.: Proc. Nat. Acad. Sci. U.S.A. *70*, 3437 (1973).
3. Harboe M., Andersen P. M., Foltmann B., Kay J., Kassel B.: J. Biol. Chem. *249*, 4487 (1974).
4. Harboe M. K., Foltmann B.: FEBS (Fed. Eur. Biochem. Soc.) Lett. *60*, 133 (1975).
5. Meitner P. A.: Biochem. J. *124*, 673 (1971).
6. Rassmussen K. T., Foltmann B.: Acta Chem. Scand. *25*, 3873 (1971).
7. Harboe M. K.: Unpublished results.
8. Klemm P., Poulsen F., Harboe M. K., Foltmann B.: Acta Chem. Scand., Ser. B *30*, 979 (1976).
9. Sepulveda P., Jackson K. W., Tang J.: Biochem. Biophys. Res. Commun. *63*, 1106 (1975).
10. Huang W.-T., Tang J.: J. Biol. Chem. *245*, 2189 (1970).
11. Pedersen V. B., Foltmann B.: FEBS (Fed. Eur. Biochem. Soc.) Lett. *35*, 250 (1973).
12. Pedersen V. B., Foltmann B.: Eur. J. Biochem. *55*, 95 (1975).
13. Foltmann B., Hartley B. S.: Biochem. J. *104*, 1064 (1967).
14. Foltmann B., Pedersen V. B., Jacobsen H., Kauffman D., Wybrandt G.: Proc. Nat. Acad. Sci. U.S.A. *74*, 2321 (1977).
15. Cunningham A., Wang H.-M., Jones S. R., Kurosky A., Rao L., Harris C. I., Rhee S. H., Hofmann T.: Can. J. Biochem. *54*, 902 (1976).
16. Bohak Z.: J. Biol. Chem. *244*, 4638 (1969).
17. Keilová H., Kostka V.: This Journal *40*, 574 (1975).
18. Keilová H., Salvetová A., Kostka V.: This Journal *40*, 580 (1975).
19. Spackmann D. H., Stein W. H., Moore S.: Anal. Chem. *30*, 1190 (1958).
20. Benson Jr J. V., Patterson J. A.: Anal. Chem. *37*, 1108 (1965).
21. Moore S.: J. Biol. Chem. *238*, 235 (1963).
22. Edelhoch H.: Biochemistry *6*, 1948 (1967).

23. Chervenka C. H.: Anal. Biochem. *34*, 24 (1970).
24. Niall H., Edman P.: J. Gen. Physiol. *45*, 185 (1962).
25. Kostka V., Morávek L., Šorm F.: Eur. J. Biochem. *13*, 447 (1970).
26. Edman P., Henschen A. in the book: *Protein Sequence Determination* (S. B. Needleman, Ed.) *Molecular Biology, Biochemistry and Biophysics* 8, p. 232. Springer, Berlin, Heidelberg, New York 1975.
27. Grüner K.: Chem. Listy *64*, 1160 (1970).
28. Klebe J. F., Finkbeiner H., White D. M.: J. Amer. Chem. Soc. *88*, 3390 (1966).
29. Kostka V., Salvetová A.: This Journal *39*, 2291 (1974).
30. Raftery M. A., Cole R. D.: J. Biol. Chem. *241*, 3457 (1966).
31. Dopheide T. A. A., Moore S., Stein W. H.: J. Biol. Chem. *242*, 1833 (1967).
32. Robyt J. F., Ackerman R. J., Chittenden C. G.: Arch. Biochem. Biophys. *147*, 262 (1971).
33. Pechére J. F., Dixon G. H., Maybury R. H., Neurath H.: J. Biol. Chem. *233*, 1364 (1958).
34. Hirs C. H. W.: J. Biol. Chem. *219*, 611 (1956).
35. Rajagopalan T. G., Moore S., Stein W. H.: J. Biol. Chem. *241*, 4940 (1965).
36. Green M. L., Llewelin J. M.: Biochem. J. *133*, 105 (1973).
37. Kay J., Dykes C. W. in the book: *Acid Proteases, Structure, Function and Biology* (J. Tang, Ed.), p. 103. Plenum Press, New York 1977.
38. Plummer T. H. jr: Anal. Biochem. *73*, 532 (1976).
39. Donta S. T., Van Vunakis H.: Biochemistry *9*, 2791 (1970).
40. Bohak Z.: Methods Enzymol. *19*, 347 (1970).
41. Dykes C. W., Kay J.: Biochem. J. *153*, 141 (1976).
42. Foltmann B., Pedersen V. B. in the book: *Acid Proteases, Structure, Function and Biology* (J. Tang, Ed.), p. 3. Plenum Press, New York 1977.
43. Ong E. B., Perlmann G. E.: J. Biol. Chem. *243*, 6104 (1968).
44. Keilová H., Kostka V., Kay J.: Biochem. J., in press.
45. Hofmann T.: Advan. Chem. Ser. *136*, 146 (1974).
46. Foltmann B.: Ccmpt. Rend. Trav. Lab. Carlsberg *35*, 143 (1966).

Translated by the author (V. K.).