ISOLATION AND CHARACTERIZATION OF CYANOGEN BROMIDE FRAGMENTS OF HOG PEPSIN

L. MORÁVEK and V. KOSTKA

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

Received December 10th, 1971

Hog pepsin was converted into its S-sulfo derivative and the latter was purified by gel filtration. The product obtained was subjected to cyanogen bromide cleavage. The hydrolysate was fractionated by gel filtration and by discontinuous recycling chromatography. The fractions obtained were characterized by their terminal sequences and by peptide maps. The location of some of the fragments in the pepsin chain was determined.

As we have reported[1], the hydrolysis of pepsin with cyanogen bromide at methionine residues[2] has been chosen in our Laboratory and also by other authors[3] for the preparation of large fragments of the protein. Commercial preparations of pepsin are contaminated by autolyzed products which would complicate considerably the process of isolation of the high molecular weight fragments. A very suitable starting material would be pepsin prepared by the activation of pepsinogen[4]. In view of the quantity of material required for the determination of the complete amino-acid sequence this approach is relatively costly. In this study enzymatically inactive S-sulfo-pepsin was used as starting material. It was prepared from commercial pepsin and a majority of the autolyzed products were removed by gel filtration. In spite of the considerable advance in fractionation techniques, the isolation of high molecular weight fragments remains a difficult task. A suitable procedure for this purpose is gel filtration in solutions containing 8M urea. A substantial improvement in the effectiveness of this fractionation procedure was achieved by the development of discontinuous recycling chromatography[5].

The analytical examination of the individual high molecular weight components constitutes an important part of the fractionation process. In this study we aimed at the fractionation of large fragments by a reproducible procedure. The technique of peptide maps has been employed in addition to end-group analysis as the main tool for analytical characterization. These maps provide a possibility to evaluate the fractionation process and also a basis for the localization of fragments obtained by other methods in well-defined regions of the pepsin molecule.

EXPERIMENTAL AND RESULTS

Material and Methods

Pepsin was a twice crystallized preparation of Worthington Biochemical Corporation, Freehold, New Jersey, U.S.A. Chymotrypsin was prepared by trypsin activation of chymotrypsinogen. The

latter was a five-times crystallized preparation, obtained from a commercial product of Léčiva, Prague. The α-chymotrypsin obtained was crystallized three times with ammonium sulfate, dialyzed against 1 mM-HCl, and lyophilized. Carboxypeptidase A, treated with diisopropyl phosphorofluoridate, was a commercial preparation of Calbiochem (Los Angeles, California 90054, U.S.A.). Soybean trypsin inhibitor was purchased from Koch-Light Laboratories Ltd. (London, England). Cyanogen bromide was from Fluka, A.G. Switzerland. Sephadex G-25 fine and Sephadex G-100 were from Pharmacia (Uppsala, Sweden). Dimethylallylamine was synthetized from dimethylamine hydrochloride and allyl bromide. Trifluoroacetic acid was refluxed with CrO_3, dried over $CaSO_4$, and redistilled. Organic solvents for paper chromatography and for thin-layer chromatography of 3-phenyl-2-thiohydantoin derivatives of amino acids, were prepared from G.R. grade reagents by distillation and drying by conventional procedures. Silufol, thin layers of silica gel with starch binder on aluminum sheets (15 × 15 cm) was obtained from Kavalier, Czechoslovakia. The remaining reagents used were of G.R. purity grade.

Peptide maps were prepared with chymotryptic digests of the analyzed material. The sample (1—3 mg) was dissolved in approximately 200 μl of distilled water. The pH of the solution was made alkaline (phenol red indicator) by the addition of 0·1M solution of ammonium carbonate. Twenty μl of a 0·1% solution of chymotrypsin containing soybean trypsin inhibitor (weight ratio inhibitor to chymotrypsin 1 : 50) was added to the substrate solution. The digestion was carried out for 2 h at 37°C. After 2 h another addition of 20 μl of the chymotrypsin solution was made. After another 2 h the digestion mixture was acidified by acetic acid (yellow color of the indicator) and dried in a desiccator. The separation of the digest on Whatman No. 3 paper in the first direction was carried out by high-voltage electrophoresis at 4000 V in a mixture of formic acid, acetic acid, and water (50 : 150 : 800, by volume) at pH 1·9 in an apparatus described by Prusík and Keil[6]. For the second direction descending chromatography in the system n-butanol–pyridine–acetic acid–water (30 : 20 : 60 : 4) was employed. Amino acid analysis was carried out on 20-h hydrolysates by the method of Spackmann, Moore, and Stein[7], using either a Beckmann Spinco Amino Acid Analyzer or an analyzer of Czechoslovak make. The tryptophan content was not determined quantitatively. The N-terminal end-group analysis by the dinitrophenylation technique[8] has been described elsewhere[9]. For carboxypeptidase A cleavage, the peptide (3 μmol) was dissolved in 0·1M piperidine (0·5 ml/1 mg). From this solution aliquots containing 0·5 μmol of the peptide were pipeted into conical test tubes. The aliquots were subsequently incubated for 2 h at 37°C and then lyophilized. (The C-terminal analysis of S-sulfopepsin was carried out with the omission of the incubation in piperidine). The dry aliquots were dissolved in 1 ml of 0·05M Tris-Cl buffer at pH 8·5 and 0·005 μmol of a freshly prepared carboxypeptidase A solution was added. The aliquots were then incubated for 0·5, 1, 2, 8, and 24 h at 25°C. At the end of the incubation 0·15 ml of glacial acetic acid was added to each aliquot; the precipitate formed was centrifuged off and the supernatant was set aside for further treatment. The precipitate was washed twice with 0·5 ml of buffer at pH 2·2 (ref.[7]). The washings and the supernatant were combined and the solution was lyophilized. The lyophilized samples were dissolved in 0·25 ml of buffer at pH 2·2 and 0·2 ml of the solution was analyzed using an amino acid analyzer equipped for the determination of homoserine. From values obtained the yield per 1 μmol of peptide (determined approximately) was calculated. Stepwise degradation was carried out by the technique of Niall and Edman[10]. The phenylthiohydantoins obtained in each degradation step were analyzed by thin layer chromatography on Silufol sheets. The entire procedure has been described in our previous paper[11].

The preparation of S-*sulfo-pepsin* was carried out by a modification of the procedure of Pechère and coworkers[12]. Several 5·5-g batches were prepared. Urea (480 g) was dissolved in 400 ml of 2·5 mM-HCl with heating. To this solution 100 ml of 1·5M-Na_2SO_3 was added and the pH was adjusted to 10·2 by the addition of conc. ammonia. Pepsin (5·5 g) was dissolved in this solution

and the pH was maintained at 10·2 by the addition of ammonia. A solution of 2M-$Cu(NO_3)_2$ was added afterwards and the pH was again kept constant at pH 10·2 with ammonia. The solution was made up to 1 000 ml with water and the reaction mixture was kept for 1 h at room temperature. After this period the reaction mixture was subjected to gel filtration on a column of Sephadex G-25 fine (9 × 80 cm) equilibrated with 0·2M ammonium carbonate. Fractions (250 ml) were collected and their absorbance at 280 nm was measured. Fractions containing the high molecular weight material were pooled and the solution was lyophilized. The dry material was freed of any remaining traces of ammonium carbonate by evacuation at 40°C for 4 h. The average yield of dry material was 4·8 g. Smaller peptides, if present in the starting material, were thus separated by gel filtration on Sephadex G-25. Larger products of pepsin autolysis were removed by gel filtration on Sephadex G-100. S-sulfo-pepsin (10 g) was dissolved in 500 ml of 0·3M ammonium acetate, pH 8·0, containing 8M urea. After the material had dissolved, the pH of the solution was adjusted to 5·0 with glacial acetic acid. The sample was then applied to a column of Sephadex G-100 (10 × 260 cm), equilibrated with 0·3M ammonium acetate, pH 6·0, containing 8M urea. The column was eluted with the same urea-containing eluant; the upward flow arrangement was used. The high molecular weight fraction was collected. The retarded fractions, containing the autolytic products, were discarded. The high molecular weight part of the material was desalted on a column of Sephadex G-25, equilibrated with dilute ammonia (pH 9), and lyophilized. The average yield of purified S-sulfo-pepsin was around 7 g. This material was hydrolyzed with cyanogen bromide.

Cyanogen bromide hydrolysis of S-sulfo-pepsin. S-sulfo-pepsin (10 g) was dissolved in 200 ml of 80% formic acid. To this solution 1·8 g of cyanogen bromide was added. The hydrolysis was allowed to proceed in a stoppered flask for 20 h at 37°C. After this period the solution was rotary evaporated to approximately 1/3 of its volume and then diluted with water. The white suspension which had formed was lyophilized.

Fractionation of the cyanogen bromide hydrolysate by gel filtration. A 1-g sample of the cyanogen bromide hydrolysate was dissolved in 200 ml of 0·3M ammonium acetate, pH 6·0, containing 8M urea and applied to a column of Sephadex G-100 (11 × 70 cm), equilibrated with the same solution. The pooled fractions obtained, CB1 through CB5 (Fig. 1), were desalted on a Sephadex G-25 column equilibrated with dilute ammonia (pH 9) and lyophilized. A considerably more effective fractionation of the cyanogen bromide hydrolysate was achieved by discontinuous recycling chromatography[5]. For this fractionation, 6·2 g of the cyanogen bromide hydrolysate was dissolved in 200 ml of 0·3M ammonium acetate, pH 8·5, containing 8M urea. After the sample had dissolved completely, the pH of the solution was adjusted to 5·0 by the addition of glacial acetic acid. The entire volume of the solution was then pumped into the influent (bottom) part

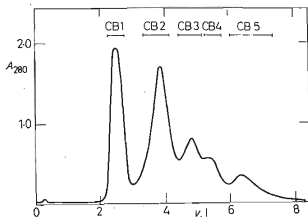


FIG. 1

Gel Filtration of Cyanogen Bromide Hydrolysate of S-Sulfo-pepsin on Sephadex G-100

The column (11 . 80 cm) was equilibrated with 0·3M ammonium acetate, pH 6·0, containing 8M urea. Elution by the same solution, flow rate 190 ml/h.

of the column (10 × 260 cm) equipped for recycling[5] and packed with Sephadex G-100; the column was equilibrated with the above described urea-containing eluant (pH 5·0). The column was eluted with the same solution; the upward-flow arrangement was used. One-hour fractions were collected, the flow rate was 100 ml/h. When fraction No 170 emerged from the column (as designated by an arrow in Fig. 2), the reapplication (as described elsewhere[5]) was begun. Fractions 52—165, which contained all the cyanogen fragments (CB1 through CB5) were re-applied. During this second cycle, in which the effective bed height achieved 520 cm, the separation of the components present had considerably improved. Fractions corresponding to individual peaks, CB1 through CB5, were pooled, desalted on Sephadex G-25 equilibrated with dilute ammonia (pH 9), lyophilized, and subjected to chemical characterization.

RESULTS

Characterization of S-Sulfo-pepsin and of its Cyanogen Bromide Hydrolysate

The material purified by gel filtration was subjected to chemical characterization. The N-terminal amino acid determined by the dinitrophenylation technique was isoleucine. The N-terminal amino-acid sequence, determined by Edman degradation, was Ile-Gly-Asp-Glu-Pro-Leu-Glu-Asn-Tyr. The C-terminal amino acid sequence, ...Val-Ala, was determined in experiments with carboxypeptidase A cleavage. The following values for alanine and valine (calculated per 1 μmol of S-sulfo-pepsin) were obtained.

Time, min	1	5	10	20	40	80	160	320	1 440
Ala μmol	0·42	0·49	0·65	0·64	0·62	0·66	0·65	0·64	0·66
Val	0·00	0·03	0·08	0·13	0·20	0·30	0·35	0·40	0·45

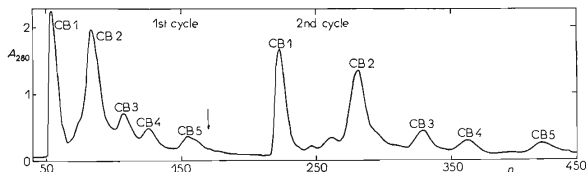


Fig. 2

Discontinuous Recycling Chromatography of Cyanogen Bromide Hydrolysate of S-Sulfo-pepsin

Fractionation of 6·2 g of sample dissolved in 200 ml of 0·3M ammonium acetate, pH 5·0, containing 8M urea, on a 10 . 260 cm column of Sephadex G-100, equilibrated and eluted with the same solution. Fractions 100 ml/h. n, fraction number. When fraction No 170 was emerging from the column (designated by an arrow in the graph), reapplication[5] of fractions No 52—165 was begun.

The values of the remaining amino acids were negligible. The amino acid-analysis of the 22-h hydrolysate gave the following values. Lys (0·9), His (0·8), Arg (1·7), Asp (42·0), Thr (25·0), Ser (41·2), Glu 25·1), Pro (15·2), Gly (34·2), Ala (15·9), Val (18·7), Met (3·9), Ile (21·9), Leu (24·9), Tyr (14·8), Phe (12·9), cysteic acid (5·9 in oxidized sample), Trp (not determined quantitatively).

The amino acid analysis of the lyophilized sample of the cyanogen bromide hydrolysate showed the content of 4·04 residues of homoserine based on 2·00 residues of arginine; only traces of methionine were detected. The presence of small peptides in the hydrolysate was not demonstrated using the technique of peptide maps described above.

Characterization of Individual Pooled Fractions after Recycling

Fraction CB1. The results of amino-acid analysis of this fraction (as well as of the other fractions characterized), are given in Table I. The N-terminal amino acid,

TABLE I

Quantitative Amino-Acid Analysis of Fractions CB1—CB5

The values for individual amino acids, obtained with 20-h hydrolysates of the samples, are given in μmol. The tryptophan content was not determined. None of the fragments contained methionine.

Amino acid	CB1	CB2	CB3	CB4	CB5
Lysine	0·188	—	—	—	—
Histidine	—	0·048	—	0·055	—
Arginine	0·423	—	—	—	—
Cysteic acid[a]	0·016	0·078	0·014	0·111	0·145
Aspartic acid	0·417	1·046	0·482	0·802	0·522
Threonine[b]	0·180	0·684	0·287	0·545	0·302
Serine[b]	0·222	1·069	0·534	0·775	0·539
Glutamic acid	0·196	0·663	0·307	0·543	0·361
Proline	0·176	0·344	0·172	0·268	0·172
Glycine	0·268	0·910	0·491	0·562	0·442
Alanine	0·224	0·313	0·147	0·267	0·275
Valine	0·369	0·527	0·290	0·261	0·130
Isoleucine	0·180	0·514	0·255	0·372	0·374
Leucine	0·250	0·754	0·399	0·442	0·224
Tyrosine	0·170	0·453	0·246	0·278	0·101
Phenylalanine	0·163	0·395	0·182	0·303	0·086
Homoserine[c]	0·008	0·076	0·037	0·067	0·093

[a] Oxidized sample, [b] values not corrected, [c] Homoserine lactone was converted into homoserine by incubation of the sample hydrolysate with 0·1M piperidine, 1 h at 37°C (ref.[13]).

aspartic acid or asparagine, was determined by the dinitrophenylation technique. The sequence of 17 amino acids[11] at the N-terminus was determined by Edman degradation as Asp-Val-Pro-Thr-Ser-Ser-Gly-Glu-Leu-Trp-Ile-Leu-Gly-Asp-Val-Phe-Ile. Its peptide map (Fig. 3) includes several typical basic peptides.

Fraction CB2. The N-terminal amino acid determined by the dinitrophenylation technique was isoleucine. The N-terminal amino-acid sequence, established by Edman degradation, Ile-Gly-Asp-Glu-Pro, is identical with the N-terminal amino-acid sequence of the whole pepsin molecule. The C-terminal amino-acid sequence, Ile-Thr-Hse (homoserine) was determined in experiments with carboxypeptidase A digestion. The following values were found for individual amino acids liberated by carboxypeptidase A:

Time, h	0·5	1	2	8	24
Hse, μmol	0·50	0·52	0·56	0·57	0·58
Thr	0·19	0·33	0·46	+	0·60
Ile	0·15	0·28	0·39	0·42	0·55

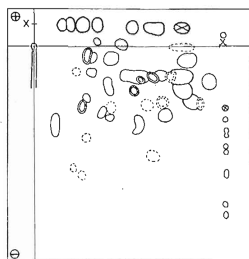
+ not determined

The liberation of small quantities of serine was also observed; for technical reasons, however, the value for this amino acid was not determined quantitatively.

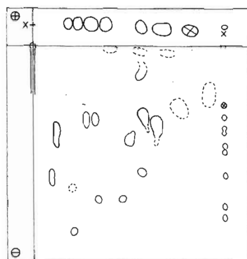
Fractions CB3. The peptide map (Fig. 3) shows some of the typical peptides, characteristic of fraction CB2, yet certain major peptides are missing. The results of N-terminal end-group analysis by the dinitrophenylation technique were ambiguous whereas the amino acid sequence
$\genfrac{}{}{0pt}{}{\text{Thr}}{\text{(Ile)}}\text{-Gly-}\genfrac{}{}{0pt}{}{\text{Ile}}{\text{(Asp)}}\text{-Leu-Gly-}\genfrac{}{}{0pt}{}{\text{Tyr}}{\text{(Gly)}}$ was established by Edman degradation. The symbols in brackets stand for weak side spots observed on chromatography of the thiohydantoins. The C-terminal amino-acid sequence established from carboxypeptidase A digestion was Ser(Ile,Thr) Hse. The values for aspartic acid or asparagine are relatively low and thus its possible assignment to position No. 5 with respect to the C-terminus requires additional evidence. The following values for individual amino acids were found.

Time, h	0·5	2	8	24
Hse, μmol	0·23	0·34	0·33	0·40
Thr	0·03	0·27	0·31	0·47
Ile	0·03	0·28	0·36	0·46
Ser	0·00	0·03	0·10	0·26
Asp	0·00	+	0·07	0·18

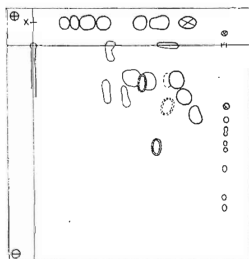
+ not determined



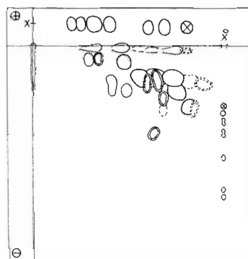
a



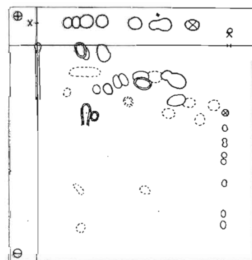
b



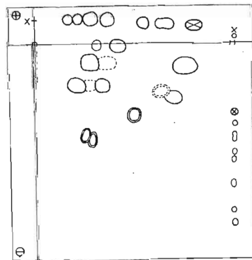
c



d



e



f

Fraction CB4. The peptide map of this fraction (Fig. 3) as well as the results of N-terminal end-group analysis by dinitrophenylation, are relatively complex. By Edman degradation two amino acids, alanine and glycine, were determined in addition to the N-terminal isoleucine.

Fraction CB5. This fraction yields a relatively simple peptide map with some typical yellow spots of peptides. The N-terminal valine was determined by dinitrophenylation. The N-terminal amino-acid sequence, determined by Edman degradation, was $\substack{\text{Val}\\(\text{Asp})}$-$\substack{\text{Ile}\\(\text{Gly})}$-$\substack{\text{Ser}\\(\text{Glu})}$-(Thr). The symbols in brackets denote weak side spots observed on chromatography, in addition to the major spots corresponding to the amino-acid sequence Val-Ile-Ser-. No major spot was detected in the fourth step. Digestion of fraction CB5 by carboxypeptidase A liberated only the C-terminal homoserine, the values for the remaining amino acids being too low to determine. The following values were obtained.

Time, h	0·5	1	2	8	24
Hse, μmol	0·30	0·29	0·31	0·38	0·42

DISCUSSION

The first problem to be solved before the sequential studies on pepsin were begun was the preparation of a suitable starting material. Rajagopalan, Moore, and Stein[4] examined commercially available samples of hog pepsin and found that these samples were contaminated with products of pepsin autolysis. These authors described the preparation of pepsin by the activation of pepsinogen. The product obtained was found to be sufficiently homogeneous for sequential work.

The aim of the work described here was to convert commercial pepsin into an appropriate, enzymatically inactive derivative, which could be used in subsequent studies. The majority of contaminants were removed by gel filtration in 8M urea. The interruption of the disulfide bridges by the modification of the active enzyme contri-

FIG. 3

Peptide Maps of Recycled Fractions and of Original Cyanogen Bromide Hydrolysate of S-Sulfopepsin

Chymotryptic digests of the analyzed samples were applied at the origin on Whatman No 3 paper. A reference mixture of amino acids was applied at the margin. Neutral red and ε-DNP-lysine, which served as markers for paper chromatography and high-voltage electrophoresis, respectively, are marked ⊗. First direction, high-voltage electrophoresis at pH 1·9, second direction, descending chromatography in n-butanol–pyridine–acetic acid–water (30 : 20 : 60 : 4, by vol.). The spots stained yellow with ninhydrin are marked by double contours.

butes to the separation of those autolytical products which are linked together by a disulfide bond in the starting material. The choice of the method of cleavage, which would lead to fragments suitable for sequential studies, depends on the amino-acid composition of the given protein. Pepsin contains only two arginine residues and one lysine residue[4]. As follows from the data of several authors[14,15], these amino acids are located within the C-terminal 20-residue region. Trypsin cleavage thus leaves the predominant part of the pepsin molecule, containing 300 amino-acid residues, intact. Cleavage of pepsin with other, less specific enzymes, can be expected to yield a great number of peptides. These peptides are difficult to arrange and combine within the whole molecule. The cleavage of the chain at a small number of sites is an alternative method. This approach may cause difficulties during the fractionation of the fragments. On the other hand, this approach offers considerable advantages since it is much easier to arrange and link together peptides within smaller regions of the peptide chain.

In this Laboratory, we have chosen two such lines of approach which follow from the amino acid composition of pepsin. The hydrolysis by cyanogen bromide[2] at the four methionine residues present in pepsin[4], which is the subject of this paper, gives rise to five specific fragments. The other line of approach represented by tryptic cleavage of reduced and aminoethylated pepsin (RAE-pepsin), will be reported elsewhere[16]. By reduction and aminoethylation[17] of the half-cystine residues of pepsin, six additional sites susceptible to tryptic cleavage[18] are introduced into the molecule. Fragments of two types then provide links overlapping half-cystine and methionine residues and offer alternative possibilities of prepairing individual parts of the pepsin chain. It has been observed[19] that under certain conditions of aminoethylation the methionine residues are partly modified. This could interfere with the subsequent cyanogen bromide hydrolysis of such derivatives. For this reason, the S-sulfo derivative which can be subsequently reduced and aminoethylated was chosen. The conversion of commercial pepsin into the S-sulfo derivative was effected by a modification of the procedure of Pechère and coworkers[12]. Pepsin was dissolved at pH 10·2 so that autolysis was eliminated. The dialysis of the reaction mixture was replaced by gel filtration on Sephadex G-25 equilibrated with 0·2M-$(NH_4)_2CO_3$. At lower concentrations of ammonium carbonate a part of the copper compounds remained associated with S-sulfo-pepsin. Gel filtration of S-sulfo-pepsin on Sephadex G-100 was not expected to completely remove all peptide contaminants. The amino-acid analysis of the 22-h hydrolysate of S-sulfo-pepsin, which was of orienting character, gave values relatively close to those characterizing pepsin obtained by the activation of pepsinogen[4]. Purified S-sulfo-pepsin was found to be homogeneous when judged by the sequence of the N- and C-terminals. Whereas only one N-terminal end group, isoleucine or leucine was determined by the dinitrophenylation technique, Edman degradation permitted us to determine the N-terminal amino-acid sequence Ile-Gly-Asp-Glu-Pro-Leu-Glu-Asn-Tyr, which is identical with the 9-residue N-terminal

sequence of pepsin reported by Tang[20]. Carboxypeptidase A cleavage liberated only alanine and valine. Rajagopalan, Moore, and Stein[4] have ascribed the liberation of valine from commercial pepsin to the presence of autolytic products. These authors observed that only alanine was liberated from pepsin prepared from pepsinogen. From our results, we conclude that valine is liberated from the C-terminal amino-acid sequence...Pro-Val-Ala. This interpretation is supported by the kinetics of carboxypeptidase A cleavage showing that the liberation of valine is considerably slower than that of alanine. We explain the difference in these observations by different conditions of carboxypeptidase A cleavage since valine is liberated in addition to alanine also from the pure C-terminal heptapeptide Val-Gly-Leu-Ala-Pro-Val-Ala(CB-1-T4) (ref.[11]).

The cyanogen bromide hydrolysis was carried out in 80% formic acid in which S-sulfo-pepsin was soluble. The amino acid analysis of the hydrolysate showed the presence of 4·04 residues of homoserine. Traces of methionine were detected. No solubility problems were encountered when this material was fractionated in solutions containing 8M urea.

For the fractionation by gel filtration Sephadex G-100 equilibrated with 0·3M ammonium acetate (pH 6·0), containing 8M urea was used. This eluant increases the solubility of the material and also to a certain degree supresses the association of the individual components. Five fractions, designated CB1 – CB5, were obtained. Some of them were relatively well separated (CB1, CB2, CB5) whereas the separation of fractions CB3 and CB4 was incomplete. The application of discontinuous recycling chromatography[5] resulted in a good separation of all peaks, detectable by absorbance measurement at 280 nm.

The information on the N-terminal amino acids characterizes each individual component present in the fraction. On the other hand the values obtained by amino acid analysis represent a sum of the amino-acid composition of all components present in the given fraction. Hence, even analyses of fractions, containing only one main component, could be affected by a relatively small amount of contaminants. Therefore the amino-acid composition of the fragments is presented in μmol as determined by the analysis. A more detailed analytical characterization will be a part of a subsequent examination of individual fragments. The results of amino-acid analyses expressed in this manner provide information on the distribution of some low proportion amino acids, as will be discussed later.

The peptide mapping of the chymotryptic digests of individual fragments was found to be the most appropriate technique for checking the course of the fractionation. The fragments show characteristic patterns of peptides so that the fragments can be detected even when they are present in heterogeneous mixtures. The knowledge of the peptide maps of the cyanogen bromide fragments will make possible a correlation of fragments of different type.

Characterization of Fractions Obtained by Discontinuous Recycling Chromatography

Fraction CB1. The results of dinitrophenylation analysis, Edman degradation, the peptide map pattern, and the results of amino-acid analysis permit us to conclude that fraction CB1 contains one main component, designated fragment CB1. The absence of homoserine indicates that the fragment represents the C-terminal portion of the whole pepsin molecule. Arginine and lysine were found in this fragment only. This is in agreement with previous data[14,15] indicating that all residues of these amino acids are present in the C-terminal region of the pepsin molecule. Fragment CB1 was studied in detail and the complete sequence[11] of its 37 amino-acid residues was determined. The small elution volume of the fragment on gel filtration is in disagreement with the relatively low molecular weight of the fragment. It may be assumed that fragment CB1 is in aggregated form under the conditions of the separation. Amino-acid analysis of this fragments has shown that the content of certain amino acids in the preparation used is greater than that present in the known sequence. These discrepancies have been ascribed[11] to the presence of contaminants in the material analyzed. Additional experiments[22] to remove these contaminants were unsuccessful.

Fraction CB2. The results of N-terminal end group analysis and of Edman degradation indicate that fraction CB2 contains one main component only, fragment CB2. The N-terminal sequence determined by Edman degradation, i.e. Ile-Gly-Asp-Glu-Pro, is identical with the N-terminal amino-acid sequence of the whole pepsin molecule. Fragment CB2 can be therefore assigned to the N-terminus. Since fragment CB2 contains the only histidine of pepsin, the two half-cystine residues derived from the adjacent short disulfide loop[23,24] can be expected in fragment CB2. The cysteic acid content found is in agreement with this deduction. The value for homoserine, however, is higher than that corresponding to the expected single residue. The approximate ratio of histidine to cysteic acid to homoserine in fragment CB2, determined by amino-acid analysis, is 1 : 2 : 2.

Fraction CB3. The peptide map of fraction CB3 shows a part of the peptide pattern characteristic of fragment CB2, yet some of the main peptides are missing. From a comparison of the peptide maps, the main component of fraction CB3 appears to be a part of fragment CB2. The carboxyl-terminal amino-acid sequence determined for CB3, Ser(Ile,Thr)Hse, seems to be identical with the C-terminal end of fragment CB2,...Ile-Thr-Hse. The assignment of aspartic acid to position No 5 with respect to the C-terminal and of fraction CB3 was confirmed by the isolation and sequential analysis of a peptide, Asp-Ser-Ile-Thr-Hse, obtained from a chymotryptic digest[22] of fragment CB2. Since the C-terminal amino-acid sequences of fragments CB2 and CB3 are identical, fragment CB3 obviously represents the C-terminal region of the larger fragment CB2.

The results of Edman degradation $\frac{\text{Thr}}{(\text{Ile})}$-Gly-$\frac{\text{Ile}}{(\text{Asp})}$-Leu-Gly-$\frac{\text{Tyr}}{(\text{Gly})}$ can be interpreted as indicating the main sequence Thr-Gly-Ile-Leu-Gly-Tyr and a secondary sequence Ile-Gly-Asp..., showing the presence of some material from the N-terminal region of the pepsin molecule.

Fraction CB4. The presence of histidine and the ratio of histidine to cysteic acid to homoserine, 1 : 2 : 1, determined by amino-acid analysis, indicates that fragment CB4 should be a part of fragment CB2. The N-terminal end group of fragment CB4, isoleucine, is identical with the N-terminal isoleucine of fragment CB2 and thus of the whole pepsin molecule.

Fragment CB5. Edman degradation revealed the presence of an N-terminal sequence Asp-Gly-Glu-Thr in addition to the N-terminal amino-acid sequence Val-Ile-Ser... of the main component (referred to further as CB5). This sequence (Asp-Gly-Glu-Thr) belongs to another component, designated fragment CB6. Its N-terminal sequence is in agreement with the pentapeptide sequence Asp-Gly-Glu-Thr-Ile, determined by Ostoslavskaja and coworkers[3] at the N-terminal end of their peptide B-5. In our Laboratory, the peptide Ile(Thr, Met, Asx, Gly, Glx, Thr) has been isolated from the thermolysin digest[25] of S-sulfo-pepsin. The amino acid composition of our peptide accounts for the presence of the C-terminal methionine sequence Ile-Thr-Met and of the above mentioned N-terminal sequence Asp-Gly-Glu-Thr. This peptide would result from the cleavage at the amino side of two isoleucine residues, in agreement with the known specificity of thermolysin. Together with the determined sequence Asp-Ser-Ile-Thr-Met (*cf.* fraction CB3), these results add up to a longer methionine sequence, Asp-Ser-Ile-Thr-Met-Asp-Gly-Glu-Thr. This finding is reconcilable with the methionine-containing sequence Asp-Ser-Ile-Thr-Met(Asx,Glx, Gly) (Ala,Tyr), reported by Tang and Hartley[24], with the exception of the last two amino-acid residues.

Summarizing all these results we obtain preliminary information on the order of individual cyanogen bromide fragments.

The N-terminal region of the pepsin molecule is present in the largest fragment CB2, containing two homoserine residues per one histidine residue. The results of amino-acid analyses, peptide maps, and the N-terminal amino acid sequences indicate that fragments CB3 and CB4 are linked together in fragment CB2 and that their order should be CB4–CB3. Fragment CB2 can be accounted for by assuming incomplete cleavage at the first methionine residue in spite of the fact that conversion of methionine to homoserine was complete. The second methionine residue of pepsin is most likely involved in the methionine containing sequence Asp-Ser-Ile-Thr-Met-Asp-Gly-Glu-Thr. The amino-acid sequence at the carboxyl side of this methionine, characteristic of fragment CB6, would place this fragment as the third one in the molecule from the N-terminal end. The assignment of fragment CB1 to the C-terminal

end of the pepsin molecule leaves only one possible position for fragment CB5, namely that between the third and the fourth methionine residues. In view of this preliminary data we expect the order of cyanogen bromide fragments in the molecule to be: CB4 – CB3 – CB6 – CB5 – CB1. These results have been supported by a more detailed study[26] of amino-acid sequences around the methionine residues.

With the exception of fragment CB1, which has been sequenced completely[11], all the remaining fractions obtained from the cyanogen bromide hydrolysate are being purified and studied.

We wish to express our thanks to Prof. F. Šorm for stimulating discussions of the problem studied and for the interest with which he followed our experiments. We thank Mrs J. Viková and Mrs E. Bulantová for skillful technical assistance. We are indebted to Mr K. Grüner for the Edman degradation experiments as well as to Mr J. Zbrožek, Miss V. Himrová, and Mrs E. Dršková for the amino-acid analyses.

REFERENCES

1. Kostka V., Morávek L., Kluh I., Keil B.: Biochim. Biophys. Acta *175*, 459 (1969).
2. Gross E., Witkop B.: J. Biol. Chem. *237*, 1856 (1962).
3. Ostoslavskaja V. I., Pugačeva I. B., Vachitova E. A., Krivcov V. F., Muratova G. L., Levin E. D., Stepanov V. M.: Biochimija *33*, 331 (1968).
4. Rajagopalan T. G., Moore S., Stein W. H.: J. Biol. Chem. *241*, 1190 (1966).
5. Morávek L.: J. Chromatog. *59*, 343 (1971).
6. Prusík Z., Keil B.: This Journal *25*, 2049 (1960).
7. Spackman D. H., Stein W. H., Moore S.: Anal. Chem. *30*, 1190 (1958).
8. Sanger F. in the book: *Advances in Protein Chemistry* Vol. VII, p. 1., (M. L. Anson, H. Bailey, J. T. Edsall, Eds.). Academic Press, New York 1952.
9. Trufanov V. A., Kostka V., Keil B., Šorm F.: European J. Biochem. *7*, 544 (1969).
10. Niall H., Edman P.: J. Gen. Physiol. *45*, 185 (1962).
11. Kostka V., Morávek L., Šorm F.: European J. Biochem. *13*, 447 (1970).
12. Pechère J. F., Dixon G. H., Maybury R. H., Neurath H.: J. Biol. Chem. *233*, 1364 (1958).
13. Gross E.: Personal communication.
14. Dopheide T. A. A., Moore S., Stein W. H.: J. Biol. Chem. *242*, 1833 (1967).
15. Perham R. N., Jones G. M. T.: European J. Biochem. *2*, 84 (1967).
16. Kostka V.: Unpublished results.
17. Raftery M. A., Cole R. D.: Biochem. Biophys. Res. Commun. *10*, 467 (1963).
18. Lindley H.: Nature *178*, 647 (1956).
19. Schroeder W. A., Shelton J. R., Robberson B.: Biochim. Biophys. Acta *147*, 590 (1967).
20. Tang J.: Biochem. Biophys. Res. Commun. *41*, 697 (1970).
21. Li C. H., Liu W. K., Dixon J. S.: J. Am. Chem. Soc. *88*, 2050 (1966).
22. Morávek L.: Unpublished results.
23. Keil B., Morávek L., Šorm F.: This Journal *32*, 1968 (1967).
24. Tang J., Hartley B. S.: Biochem. J. *118*, 611 (1970).
25. Morávek L., Kysilka Č.: Unpublished results.
26. Morávek L., Kostka V.: Unpublished results.

Translated by the author (V. K.).