PEPTIDES ISOLATED FROM THERMOLYSIN DIGEST OF HOG S-SULFO-PEPSIN

L. MORÁVEK and Č. KYSILKA*

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

Received February 1st, 1973

The thermolysin digest of S-sulfo-pepsin prepared from commercial pepsin, was fractionated by ion-exchange chromatography on Dowex 50. The peptides from the majority of fractions thus obtained were purified by paper chromatography and electrophoresis. Two more complex fractions were subjected to additional ion-exchange chromatography on Dowex 1. Of the total number of 90 peptides iso ated from this digest, selected peptides were used for detailed elucidation of amino–acid sequences of high molecular weight fragments, obtained by cyanogen bromide cleavage of S-sulfo-pepsin.

Limited cleavage giving rise to a few polypeptide fragments only is used to advantage in the sequential studies on proteins. The sequential information can be arranged in such a case within smaller regions of the chain of the protein studied; these regions can be investigated one by one and relatively independently. In this Laboratory, we prepared large fragments of hog pepsin by cyanogen bromide cleavage[1] of S-sulfo-pepsin and by tryptic digestion of aminoethylated pepsin[2]. Since the fractionation of high molecular weight fragments often represents a complicated and time-consuming task, we also used in addition to these two procedures thermolysin hydrolysis of S-sulfo-pepsin, described here, and chymotryptic digestion[3] of S-sulfo-pepsin. The small and medium-size peptides thus obtained can be isolated by routine procedures and in large yields because of the relative accessibility of large quantities of the substrate.

The aim of this study was to obtain thermolysin peptides in quantities necessary for the determination of amino acid sequences to be used as complementary information in the elucidation of complete amino–acid sequences of large cyanogen bromide fragments[1]. Of the number of peptides obtained from this hydrolysate of whole pepsin only those peptides were intended for additional sequential investigation which had been obtained from the thermolysin hydrolysates of the individual fragments in low yield. The enzymatically inactive S-sulfo-pepsin was used again as substrate for the thermolysin digestion, similarly to the cyanogen bromide cleavage.

* Present address: Laboratory of Physiology and Genetics of Animals, Czechoslovak Academy of Sciences, Liběchov.

TABLE I

Amino-Acid Analyses of Peptides Isolated

The values are not corrected; the half-cystine residue was determined as cysteic acid after oxidation. The analyses were made on 20 h-hydrolysates of peptides.

Designation of peptide	Number of amino acid residues																
	Lys	His	Arg	Cys	Asp	Thr	Ser	Glu	Pro	Gly	Ala	Val	Met	Ile	Leu	Tyr	Phe	Trp
Th1	—	—	—	1·8	—	—	1·1	1·1	—	2·0	1·9	—	—	1·0	—	—	—	—
Th2	—	—	—	1·0	—	—	2·6	—	—	—	—	0·5	—	0·5	—	—	—	—
Th2[a]	—	—	—	1·0	—	—	2·2	—	—	—	—	0·9	—	0·9	—	—	—	—
Th3	—	—	—	1·2	—	—	0·8	—	—	—	1·0	—	—	1·0	—	—	—	—
Th4	—	—	—	—	—	1·0	0·9	3·0	—	—	1·0	—	—	—	—	—	1·0	—
Th5	—	—	—	1·0	—	—	2·0	—	—	—	—	1·0	—	—	—	0·8	—	—
Th6	—	—	—	—	—	1·1	2·0	1·9	1·1	1·2	—	—	—	—	0·7	—	—	—
Th7	—	—	—	—	1·8	—	2·2	—	—	0·8	—	0·8	—	—	1·1	—	—	—
Th8	—	—	—	—	1·0	1·9	0·9	—	—	1·1	—	0·5	—	0·4	—	—	—	—
Th8[a]	—	—	—	—	1·0	1·7	0·9	—	—	1·1	—	0·8	—	0·8	—	—	—	—
Th9	—	—	—	—	2·3	—	2·0	2·0	—	1·1	0·8	—	0·8	—	—	—	—	—
Th10	—	—	—	—	1·1	0·9	1·9	1·0	1·1	1·0	—	0·9	0·6	—	—	—	—	—
Th11	—	—	—	—	2·0	—	1·2	—	1·2	—	—	—	—	0·9	0·9	—	—	—
Th12	—	—	—	—	2·0	0·9	1·0	1·0	—	—	—	—	—	1·0	—	—	—	—
Th13	—	—	—	—	—	1·0	1·0	—	—	1·9	—	—	—	—	—	—	—	—
Th14	—	—	—	—	1·0	1·0	—	1·0	1·0	2·1	1·0	—	—	2·2	—	—	—	—
Th15	—	—	—	—	—	1·7	1·0	—	1·0	1·0	1·1	—	—	—	0·8	—	—	—
Th16	—	—	—	—	—	—	2·0	—	1·3	—	—	—	—	—	0·9	—	—	—
Th17	—	—	—	—	1·1	—	—	—	—	1·0	—	—	—	0·7	0·9	—	—	—
Th18	—	—	—	—	—	—	—	—	1·0	—	1·2	—	—	—	0·8	—	—	—
Th19	—	—	—	—	0·9	—	0·9	1·0	—	—	—	—	—	1·1	—	—	—	—
Th20	—	—	—	—	1·1	—	1·0	1·1	—	—	—	1·0	—	—	0·9	—	—	—
Th21	—	—	—	—	1·0	—	0·9	1·0	—	—	—	0·8	—	—	—	—	—	—

Th22	—	—	—	—	—	1·0	—	—	1·0	—	1·0	—	—	—	—	—	—	—
Th23	—	—	—	—	—	1·8	—	—	1·0	1·0	—	—	—	—	1·1	—	—	—
Th24	—	—	—	—	—	0·9	—	—	—	1·0	—	—	0·7	—	—	—	—	—
Th25	—	—	—	—	—	—	0·9	1·1	—	1·0	—	2·0	—	—	—	—	—	—
Th26	—	—	—	—	1·0	—	—	—	—	—	—	—	—	1·0	—	—	—	—
Th27	—	—	—	—	1·1	—	1·0	—	—	—	—	—	—	—	0·9	—	—	—
Th28	—	—	—	—	1·0	—	—	—	—	—	—	—	—	—	1·0	—	—	—
Th29	—	0·9	—	0·9	1·9	—	0·9	1·0	—	—	1·0	—	—	—	1·0	—	—	—
Th30	—	—	—	—	1·2	1·9	—	1·1	—	1·1	—	—	0·8	1·0	—	—	—	—
Th31	—	—	—	—	1·1	—	—	1·1	0·9	1·0	—	—	—	1·0	—	—	—	—
Th32	—	—	—	—	3·2	1·2	1·8	—	1·0	—	—	—	—	—	—	—	1·0	—
Th33	—	—	—	—	2·3	—	2·0	1·9	—	2·0	1·0	—	0·7	0·8	—	—	—	—
Th34	—	—	—	—	2·1	1·0	—	—	1·2	—	0·8	1·0	—	—	—	—	0·9	—
Th35	—	—	—	—	1·0	—	—	—	—	—	1·0	—	—	1·0	—	—	—	—
Th36	—	—	—	—	—	—	0·8	—	—	—	—	1·0	—	—	—	—	—	—
Th37	—	—	—	—	—	—	0·9	—	—	—	1·1	—	—	—	—	—	—	—
Th38	—	—	—	—	—	—	—	—	—	2·1	—	1·0	—	—	—	—	—	—
Th39	—	—	—	—	1·0	—	—	1·0	—	—	—	—	—	—	1·0	—	—	—
Th40	—	—	—	—	—	—	—	1·0	—	1·1	—	1·0	—	—	—	—	—	—
Th41	—	—	—	—	—	—	1·0	—	—	—	—	—	—	—	1·0	—	—	—
Th42	—	—	—	—	1·0	—	—	—	—	1·2	—	—	—	—	0·9	—	—	—
Th43	—	—	—	—	—	1·0	1·9	2·0	1·0	1·1	—	—	—	—	1·0	—	1·0	—
Th44	—	—	—	—	—	—	—	1·0	—	—	—	1·0	—	—	—	—	—	—
Th45	—	—	—	—	—	1·2	—	—	—	—	—	—	—	1·0	—	—	—	—
Th46	—	—	—	—	—	—	—	—	—	1·0	—	1·0	—	—	—	—	—	—
Th47	—	—	—	—	—	—	—	—	—	2·0	—	—	—	—	1·0	—	—	—
Th48	—	—	—	—	1·0	—	—	—	—	1·1	—	—	—	—	—	—	1·0	—
Th49	—	—	—	—	—	—	—	—	—	—	—	—	—	1·0	1·0	—	—	—
Th50	—	—	—	—	—	—	—	—	—	—	1·0	0·9	—	—	—	—	—	—
Th51	—	—	—	—	1·0	—	—	1·0	1·2	0·8	—	0·9	—	0·8	—	0·6	—	—
Th52	—	—	—	—	—	2·0	1·1	—	—	2·0	—	—	—	1·0	—	0·9	—	—
Th53	—	—	—	—	—	—	1·0	—	1·0	—	1·0	—	—	—	—	0·9	—	—

TABLE I
(*Continued*)

Designation of peptide	Number of amino-acid residues																	
	Lys	His	Arg	Cys	Asp	Thr	Ser	Glu	Pro	Gly	Ala	Val	Met	Ile	Leu	Tyr	Phe	Trp
Th54	—	—	—	—	—	—	—	—	—	—	1·0	—	—	—	1·0	—	—	—
Th55	—	—	—	—	—	—	0·9	—	1·0	—	1·0	—	—	—	1·0	0·9	—	—
Th56	—	—	—	—	—	—	—	—	—	1·0	—	—	—	0·8	1·1	—	—	—
Th57	—	—	—	—	—	1·0	1·0	—	—	1·1	—	—	—	—	—	0·7	—	—
Th58	—	—	—	—	—	1·0	—	—	—	1·1	—	—	—	—	—	—	0·9	—
Th59	—	—	—	—	1·0	1·0	—	—	—	—	—	—	—	—	—	0·8	—	—
Th60	—	—	—	—	1·1	—	1·9	—	—	—	—	—	—	1·0	—	0·8	—	—
Th61	—	—	—	—	1·0	0·9	—	0·9	—	—	—	—	—	—	1·0	0·9	—	—
Th62	—	—	—	—	1·1	—	—	—	—	—	—	—	—	—	—	—	0·9	—
Th63	—	—	—	—	—	—	—	1·0	1·0	—	—	1·1	—	—	—	0·8	—	—
Th64	—	—	—	—	—	—	1·0	—	—	—	—	—	—	—	—	—	1·0	—
Th65	—	—	—	—	—	1·1	—	—	—	—	—	—	—	—	—	—	0·9	—
Th66	—	—	—	—	—	0·9	—	—	—	—	—	0·5	—	0·4	—	—	1·0	—
Th66[a]	—	—	—	—	—	1·0	—	—	—	—	—	1·0	—	0·8	—	—	1·0	—
Th67	—	—	—	—	—	—	—	1·1	—	1·0	—	—	—	—	—	—	0·6	—
Th68	—	—	—	—	1·3	—	1·9	—	—	2·1	—	—	—	0·9	0·8	1·1	—	—
Th69	—	—	—	—	—	—	—	—	1·0	—	1·0	—	—	—	—	—	—	—
Th70	—	—	—	—	—	—	—	1·2	—	1·1	—	1·0	—	—	—	0·6	—	—
Th71	—	—	—	—	—	—	—	—	1·0	—	1·0	—	—	—	—	0·7	—	—
Th72	—	—	—	—	1·1	1·0	—	1·3	—	1·0	—	—	—	1·0	1·2	0·9	—	—
Th73	—	—	—	—	0·7	—	—	—	—	1·1	—	1·0	—	0·9	1·2	—	0·9	—
Th74	—	—	—	—	1·0	—	—	1·1	—	—	—	—	—	—	1·2	0·8	—	—
Th75	—	—	—	—	—	—	—	1·1	—	—	—	—	—	—	—	—	1·0	—
Th76	—	—	—	—	—	1·0	—	—	—	1·1	—	—	—	1·0	—	0·9	—	—
Th77	—	—	—	—	—	1·0	—	—	—	—	—	—	—	—	—	0·7	—	—
Th78	—	—	—	—	—	1·0	—	—	—	—	—	—	—	1·1	—	0·6	—	—

Th79	—	—	—	—	—	—	—	1·0	—	—	0·9	—	—	1·1	—	—	—	—
Th80	—	—	—	—	—	—	—	—	1·0	—	—	—	0·7	—	—	—	—	—
Th81	—	—	—	—	—	0·9	—	0·9	—	1·0	—	1·0	—	—	—	—	—	+ [b]
Th82	1·0	—	—	2·0	—	—	—	—	0·9	—	—	—	—	—	—	—	—	—
Th83	—	—	—	1·1	—	—	1·0	—	1·1	—	—	—	1·0	—	—	—	—	+
Th84	—	—	—	1·0	—	—	—	1·1	—	—	1·0	—	0·9	—	—	—	—	+
Th85	—	—	1·2	1·0	—	—	—	—	—	—	0·8	—	—	—	—	0·8	—	—
Th86	—	—	1·0	—	—	—	1·1	—	—	—	—	—	0·7	—	—	—	—	—
Th87	1·0	—	1·0	2·9	0·9	—	—	—	—	0·9	0·9	—	—	—	—	0·9	—	—
Th88	1·0	—	1·0	2·8	—	—	—	—	—	1·4	—	—	—	—	—	0·8	—	—
Th89	1·0	—	1·0	3·0	—	—	—	—	—	1·0	0·9	—	—	—	—	0·9	—	—
Th90	—	—	—	—	—	—	—	—	—	—	—	—	1·0	—	—	—	—	+

[a] Analyzed after 70 h hydrolysis. [b] The tryptophan content was determined qualitatively only.

TABLE II

Summary of Peptides Isolated and their Location in Cyanogen Bromide Fragments of Pepsin

For the sake of easy alignment with the cyanogen bromide fragments, the amino acids are ordered according to our present knowledge of the amino acid sequence of pepsin. The methods of N-terminal end-group analysis are marked by the following symbols: DNS — dansyl technique, SD — sequential degradation[11]. Unless specified in detail, the N-terminal amino acids were derived from the known specificity of thermolysin.

Designation of peptide	Isolated from fraction No	Net charge at pH 5·6	N-terminal end-group analysis	Composition of peptide	Fragment
Th1	I		DNS	Ile(Ala,Cys,Ser,Gly,Gly,Cys,Glx,Ala)	CB6
Th2	III		DNS	Val(Ile,Ser,Cys,Ser,Ser)	CB5
Th3	III		DNS	Ile(Ala,Cys,Ser)	CB6
Th4	IV		SD	Phe(Glx,Ala,Thr,Ser,Glx,Glx)	CB4
Th5	VIII		DNS	Val(Tyr,Cys,Ser,Ser)	CB4
Th6	X			Leu(Ser,Glx,Thr,Glx,Pro,Gly,Ser)	CB3
Th7	X	—		(Val,Ser,Ser,Asx,Asx,Gly,Leu)	CB3
Th8	X	—	DNS	Ile(Val,Asp,Thr,Gly,Thr,Ser)	CB6
Th9	XI		DNS	Ala(Ser,Glx,Asx,Ser,Asx,Gly,Glx,Met)	CB6
Th10	XI	—		(Met,Asx,Val,Pro,Thr,Ser,Ser,Gly,Glx)	CB5/CB1
Th11	XI		DNS	Ile(Asx,Ser,Leu,Pro,Asx)	CB5
Th12	XII		DNS	Ile(Ser,Asx,Thr,Asx,Glx)	CB3
Th13	XII		SD	Gly(Thr,Gly,Ser)	CB4
Th14	XII	—		Ile(Gly,Ile,Gly,Thr,Pro,Ala,Glx,Asx)	CB4
Th15	XII		SD	Leu(Thr,Gly,Pro,Thr,Ser,Ala)	CB6
Th16	XIII			Leu(Ser,Pro,Ser)	CB5
Th17	XIII	—	DNS	Ile(Leu,Gly,Asp)	CB1
Th18	XIII		DNS	Leu(Ala,Pro)	CB1
Th19	XIV	—	DNS	Ile(Glx,Ser,Asx)	CB6
Th20	XIV	—	DNS	Leu(Val,Ser,Glx,Asx)	CB3
Th21	XIV			Val(Ser,Glx,Asx)	CB3
Th22	XIV		SD	Ala(Thr,Pro)	CB3
Th23	XIV		DNS	Leu(Thr,Gly,Pro,Thr)	CB6
Th24	XV		DNS	Met(Thr,Gly)	CB4/CB3
Th25	XV	—		Val(Ser,Glu,Gly,Val)	CB3
Th26	XV	—		Ile.Asp	—
Th27	XV	—	SD	Leu(Asp,Ser)	CB3
Th28	XV	—		Leu.Asp	CB4
Th29	XV		SD	Leu(Ala,Cys,Ser,Asx,His,Asx,Glx)	CB4
Th30	XV		DNS	Ile(Thr,Met,Asx,Gly,Glx,Thr)	CB3/CB6
Th31	XV	—		Ile(Gly,Asx,Glx,Pro)	CB4
Th32	XV		SD	Phe(Asx,Pro,Asx,Asx,Ser,Ser,Thr)	CB4
Th33	XV		DNS	Ile(Gly,Ala,Ser,Glx,Asx,Ser,Asx,Gly,Glx,Met)	CB6

TABLE II
(*continued*)

Designation of peptide	Isolated from fraction No	Net charge at pH 5·6	N-Terminal end-group analysis	Composition of peptide	Fragment
Th34	XVII	—	DNS	Ala(Thr,Pro,Val,Phe,Asx,Asx)	CB3
Th25	XVII	—		Ile(Ala,Asn)	CB6
Th36	XVII			Val.Ser	CB3
Th37	XVIII			Ala.Ser	—
Th38	XVIII			Val.Gly.Gly	CB3
Th39	XIX	—	DNS	Leu(Asx,Glx)	—
Th40	XIX	—	DNS	Val(Glu,Gly)	—
Th41	XIX			Leu.Ser	CB4
Th42	XIX	—		Leu(Gly,Asp)	CB1
Th43	XX	—	DNS	Leu(Ser,Glx,Thr,Glx,Pro,Gly,Ser,Phe)	CB3
Th44	XX	0		Val.Gln	CB5
Th45	XX			Ile.Thr	CB3
Th46	XX			Val.Gly	CB1
Th47	XXI			Leu.Gly.Gly	CB3
Th48	XXI	—	DNS	Phe(Asp,Gly)	CB3
Th49	XXII			(Ile,Leu)	—
Th50	XXII		DNS	Val.Ala	CB1
Th51	XXIII	—	DNS	Ile(Asx,Gly,Val,Glx,Tyr,Pro)	CB5
Th52	XXV		DNS	Ile(Thr,Tyr,Gly,Thr,Gly,Ser)	CB4
Th53	XXV		DNS	Ala(Tyr,Pro,Ser)	CB3
Th54	XXV			(Leu,Ala)	—
Th55	XXVI		DNS	Leu(Ala,Tyr,Pro,Ser)	CB3
Th56	XXVI		SD	Ile(Leu,Gly)	CB3
Th57	XXVII		SD	Tyr(Thr,Gly,Ser)	CB3
Th58	XXVII		DNS	Phe(Gly,Thr)	CB4
Th59	XXIX	—	SD	Tyr(Asp,Thr)	CB3
Th60	XXIX	—	DNS	Ile(Asp,Ser,Ser,Tyr)	CB3
Th61	XXIX	—	DNS	Leu(Asx,Thr,Glx,Tyr)	CB4
Th62	XXIX	—		Phe.Asp	CB4
Th63	XXX	0	DNS	Val(Gln,Tyr,Pro)	CB5
Th64	XXX			Phe.Ser	CB3
Th65	XXX			Phe.Thr	CB4
Th66	XXX		SD	Ile(Val,Phe,Thr)	CB5
Th67	XXXI	—		Phe(Glu,Gly)	CB5
Th68	XXXI		DNS	Leu(Gly,Gly,Ile,Asx,Ser,Ser,Tyr,Tyr)	CB3
Th69	XXXII		DNS	Ala.Pro	CB1
Th70	XXXIII	—		Val(Glu,Gly,Tyr)	CB3
Th71	XXXIII		SD	Tyr(Ala,Pro)	CB3

TABLE II
(*continued*)

Designation of peptide	Isolated from fraction No	Net charge at pH 5·6	N-terminal end-group analysis	Composition of peptide	Fragment
Th72	XXXIV		DNS	Ile(Leu,Gly,Tyr,Asx,Thr,Glx)	CB3
Th73	XXXIV	—	DNS	Ile(Leu,Gly,Asp,Val,Phe)	CB1
Th74	XXXV	—	DNS	Leu(Glx,Asx,Tyr)	CB4
Th75	XXXV	—		Phe.Glu	CB4
Th76	XXXV		SD	Ile(Thr,Tyr,Gly)	CB4
Th77	XXXVI		SD	Tyr.Thr	CB1
Th78	XXXVII		DNS	Ile(Thr,Tyr)	CB4
Th29	XXXVIII		SD	Ile(Phe,Gly)	CB3
Th80	XL		DNS	Ala.Tyr	CB5
Th81	XLIII		DNS	Leu(Trp,Val,Pro,Ser)	CB4
Th82	XLIII	+		Ala(Asn,Asn,Lys)	CB1
Th83	XLIV		DNS	Leu(Trp,Asx,Glx,Gly)	CB3
Th84	XLVI	0	SD	Leu(Asn,Trp,Val,Pro)	CB3
Th85	L	0	DNS	Val(Phe,Asp,Arg)	CB1
Th86	LI	+		Ile(Arg,Gln)	CB1
Th87	LIII		DNS	Thr(Val,Phe,Asx,Arg,Ala,Asx,Asx,Lys)	CB1
Th88	LV	+	DNS	Phe(Asx,Arg,Ala,Asx,Asx,Lys)	CB1
Th89	LV	+	DNS	Val(Phe,Asx,Arg,Ala,Asx,Asx,Lys)	CB1
Th90	LV			Leu.Trp	CB4

EXPERIMENTAL

Material. S-Sulfo-pepsin was prepared from a twice crystallized product of Worthington Biochemical Corporation, Freehold, New Jersey, U.S.A. by a modification of the procedure of Pechère and coworkers[4], as described in an earlier study[1]. The preparation obtained was freed of high molecular weight products of pepsin autolysis by gel filtration on Sephadex G-100 in 8M urea[1]. Thermolysin was a B-grade product of Calbiochem (Los Angeles, California, U.S.A.). Dowex 50 W-X2 (200–400 mesh) and Dowex 1-X2 (200–400 mesh) were from Fluka (Switzerland), Sephadex G-25 fine and G-100 were products of 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. The organic solvents for paper chromatography and for thin-layer chromatography of 3-phenyl-2-thiohydantoin derivatives of amino acids were prepared from analytical grade preparations by distillation and drying by conventional techniques. Silufol, a silica gel thin layer with starch binder on aluminum sheets (15 . 15 cm), was purchased from Kavalier, Czechoslovakia. The remaining chemicals were of analytical purity grade.

Preparation and basic fractionation of thermolysin digest of S-sulfo-pepsin. S-sulfo-pepsin (5·7 g) was dissolved in 570 ml of 0·1M sodium bicarbonate, 0·001M in $CaCl_2$. Thermolysin (114 mg) was added and the solution was incubated at 37°C. Another addition of 114 mg of thermolysin was made after 2 h and the incubation was continued for a total period of 18 h. The solution was then acidified by glacial acetic acid, using phenol red as indicator, and evaporated to dryness. Ammonium acetate was removed afterwards by sublimation *in vacuo* at 60°C. The digest was dissolved in water adjusted to pH 8·0 by ammonium hydroxide and lyophilized; the lyophilized product was readily soluble. The basic fractionation of the digest was effected by ion-exchange chromatography. The lyophilized material (5·1 g) was dissolved in 150 ml of 0·05M pyridine formate buffer at pH 2·8; the pH of the solution was adjusted to 2·6 by the addition of formic acid. The sample was applied to a column of Dowex 50 W-X2 equilibrated with the same buffer (pH 2·8). The column was eluted by volatile pyridine formate and pyridine acetate buffers at room temperature. The buffers were mixed to develop two linear concentration gradients. The first gradient was prepared by mixing 10 1-volumes of 0·05M pyridine formate (pH 2·8) and 0·20M pyridine acetate (pH 3·5); the next gradient was developed by mixing 5 1-volumes of 0·20M pyridine acetate (pH 3·5) and 2·0M pyridine acetate (pH 5·0). Fractions of 50 ml/h were collected. An 0·3 ml aliquot was withdrawn from each fraction and taken to dryness in a desiccator over NaOH and H_2SO_4. The dry residues were applied to Whatman No 3 paper and chromatographed in the solvent system n-butanol–pyridine–acetic acid–water (30 : 20 : 60 : 4) (system S1). Individual fractions from the column were pooled according to the chromatographic evaluation to give 56 pooled fractions, designated by Roman numerals I through LVI. The pooled fractions were subsequently analyzed by a two-dimensional combination of electrophoresis and paper chromatography, and subjected to additional fractionation if necessary.

Fine resolution of pooled fractions and purification of peptides. Pooled fractions XIII and XV, which were relatively more complex, were subjected to additional ion-exchange chromatography on a column of Dowex 1-X2 (60 . 0·9 cm) following the procedure described by Guest and coworkers[5]. The fractions were evaluated by paper chromatography of 1 ml aliquots as described above. The final purification of peptides was carried out by paper chromatography in system S1 and by two procedures of paper electrophoresis. Procedure E1 was carried out in the vertical arrangement[6] at a potential of 1 500 V in pyridine acetate buffer at pH 5·6. Procedure E2 was horizontal high voltage electrophoresis with cooling[7], performed at a potential of 4—5000 V in acetate–formate solution at pH 1·9. Whatman No 3 paper served as a support in both cases. For the evaluation of fractions and individual peptides by the two-dimensional technique, a combination of electrophoresis E2 and paper chromatography in system S1 was used. The net charge of peptides, from which the presence or absence of amides of acidic amino acids could be determined in certain cases, was judged by the mobility of the peptides on electrophoresis E1.

Amino-acid analysis of peptides was carried out on 20- or 70-h hydrolysates by the method of Spackman, Stein, and Moore[8] in Beckmann Spinco Amino Acid Analyzer or in an analyzer of Czechoslovak make. The tryptophan content was not determined quantitatively. The N-terminal amino acids were determined by the dansyl technique[9] (as 1-dimethylamino-naphthalene-5--sulfonyl derivatives), using the procedure of Novotný and Franěk[10], or from the first step of stepwise degradation, carried out by the technique of Niall and Edman[11]. The phenylthiohydantoins from the individual degradation steps were analyzed by thin-layer chromatography on Silufol plates. The entire procedure was described before[12]. The N-terminal amino acid of certain peptides could be derived from the composition of the peptide and the known specificity of thermolysin, which predominantly cleaved bonds involving the amino group of leucine, isoleucine, and valine, and to a certain degree also the amino group of phenylalanine and alanine. The tryptophan content of the peptides was assayed by the qualitative test by dipping the chromato-

gram in a 1% solution of *p*-dimethylaminobenzaldehyde in 9 parts of acetone and 1 part of concentrated hydrochloric acid[13]; tryptophan-containing peptides were stained violet.

RESULTS AND DISCUSSION

The chief line of approach to sequential studies on hog pepsin in this Laboratory has been directed to the isolation of large fragments which make easier the arrangement of the sequential data obtained. In addition to that, we decided to prepare small and medium-size peptides from a relatively large quantity of S-sulfo-pepsin. These peptides were used wherever required for the determination of detailed sequences within the individual large fragments. The content of only two arginine residues and one lysine residue in the pepsin molecule[14] does not permit to use for the preparation of these peptides tryptic hydrolysis, which is the most valuable tool of sequential analysis because of its high specificity. The fact that the use of other enzymes with relatively high specificity (*i.e.* chymotrypsin and thermolysin) does not guarantee the obtaining of sufficient sequential overlaps, *e.g.* around the phenylalanine and leucine residues, makes the achievement of the final goal considerably difficult. Since pepsin has a relatively high content[14] of amino acids with hydrophobic side chains, which are potential sites of thermolysin attack, a great number or relatively small peptides were expected.

We isolated from this digest a total of 90 peptides whose amino-acid compositions are shown in Table I. The characteristics of these peptides are summarized in Table II.

The peptides isolated bear symbols composed of letters "Th" for thermolysin digestion and numbers increasing in the order of the emergence of the peptides from the column during the basic fractionation. In case that the peptide was isolated from more than one of the pooled fractions, it is listed here only once. Free amino acids, isolated from the digest, are not listed either. For these reasons the summary does not contain all of the basic pooled fractions since some of them contained additional portions of repeating peptides. In this study, the fundamental characteristics of the peptides, *i.e.* their amino-acid composition, N-terminal amino acid, and, in certain cases, their net charge were determined. The elucidation of complete amino acid sequences of the cyanogen bromide fragments, also including the sequential analysis of selected thermolysin peptides from this digest, will be described in forthcoming papers.

The result of these complex studies is a tentative amino acid sequence of pepsin, including at present the complete structures of four of the five specific caynogen bromide fragments[1] of pepsin; these structures, with the exception of the C-terminal fragment[12], will be reported in near future. The composition of the peptides isolated is given in Table II in a form permitting their easy alignment with the tentative amino-acid sequence. The location of the peptides in the individual cyanogen bromide fragments is given in the last column of Table II. In case of alternate possibilities

of alignment (especially of small peptides), the peptides were allocated to individual fragments with regard to their origin as expected from the known specificity of thermolysin.

We thank Mrs J. Viková for skillful technical help, Mr K. Grüner and Mrs A. Čermáková for the Edman degradation experiments, and Mr J. Zbrožek, Miss V. Himrová and Mrs E. Dršková for the amino-acid analyses. We are indebted to Dr J. Gašperík for the end-group analyses of some peptides by the dansyl technique.

REFERENCES

1. Morávek L., Kostka V.: This Journal *38*, 304 (1973).
2. Kostka V.: Unpublished results.
3. Kostka V., Salvetová A.: This Journal, in press.
4. Pechère J. F., Dixon G. H., Maybury R. H., Neurath H.: J. Biol. Chem. *233*, 1364 (1958).
5. Guest J. R., Carlton B. C., Yanofsky C.: J. Biol. Chem. *242*, 5397 (1967).
6. Mikeš O.: This Journal *22*, 831 (1957).
7. Prusík Z., Keil B.: This Journal *25*, 2049 (1960).
8. Spackman D. H., Stein W. H., Moore S.: Anal. Chem. *30*, 1190 (1958).
9. Gray W. R., Hartley B. S.: Biochem. J. *89*, 59P (1963).
10. Novotný J., Franěk F.: Chem. listy *62*, 995 (1968).
11. Niall H., Edman P.: J. Gen. Physiol. *45*, 185 (1962).
12. Kostka V., Morávek L., Šorm F.: Eur. J. Biochem. *13*, 447 (1970).
13. Smith I.: Nature *171*, 43 (1953).
14. Rajagopalan T. G., Moore S., Stein W. H.: J. Biol. Chem. *241*, 4940 (1966).

Translated by V. Kostka.