Paper chromatographic analyses of acid hydrolyzates of such fractions gave no evidence for peptide material other than the 11 peptides indicated in Table I.

With the methods used in this structure study, peptides might be overlooked that did not produce color with ninhydrin reagent. This possibility in itself is rather improbable. Furthermore, from the data on the rate of hydrolysis it has been possible to rather accurately predict the number of peptides formed in each incubation and the N-terminus of each peptide.

Discussion

Assuming that each peptide occurs only once in the chain, the distribution of the amino acids in the 11 peptides was identical to the amino acid composition of glucagon. Only peptide S-3, asp(tyr,leu), contained a measurable amount of impurity, having about 0.2 mole of threonine per mole of peptide. Dinitrophenylation and quantitative amino acid data, along with the behavior of the peptides on Dowex 50 columns and their generally good recovery, all pointed to the conclusion that these peptides were genuine fragments of glucagon arising from the hydrolytic action of subtilisin.

Integration of the 11 peptides from the subtilisin digestion into the glucagon chain is represented in Table II.

Acknowledgment.—The authors are indebted to H. L. Bird and C. T. Pugh for the paper chromatography, and to W. A. Tandy and R. G. Scheib for their technical assistance.

INDIANAPOLIS 6, INDIANA

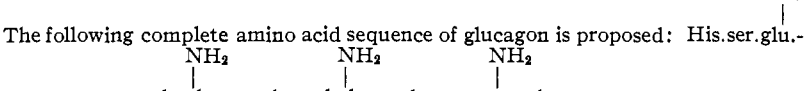
[CONTRIBUTION FROM THE LILLY RESEARCH LABORATORIES]

The Amino Acid Sequence of Glucagon. V. Location of Amide Groups, Acid Degradation Studies and Summary of Sequential Evidence

BY W. W. BROMER, L. G. SINN AND OTTO K. BEHRENS

RECEIVED NOVEMBER 19, 1956

Data is presented that elucidates the remaining unknown amino acid sequences in the glucagon chain, and that establishes the locations of the four amide groups. The following complete amino acid sequence of glucagon is proposed: His.ser.glu.-gly.thr.phe.thr.ser.asp.tyr.ser.lys.tyr.leu.asp.ser.arg.arg.ala.glu.asp.phe.val.glu.try.leu.met.asp.thr



Introduction

In preceding papers of this series[1-3] the degradation products of glucagon formed by the action of three different proteolytic enzymes have been described. These fragments of the glucagon chain can be arranged to fit together logically in only one manner; all of the data obtained are consistent with the proposed arrangement. However, the sequence of all the amino acids within the chain is not determined, and the location of the four amide groups can only be postulated from the behavior of the peptides on resin columns and from the carboxypeptidase data. This paper describes the elucidation of the unknown sequences, largely through partial acid degradation, and the final location of the amide linkages by means of chemical analysis.

Experimental

Partial Acid Degradation.—Peptides from enzymatic degradation of glucagon were dissolved in 11.7 *N* HCl at a concentration of about one μmole per ml., and the reaction mixture was incubated at 37° for 3 days. The excess acid was removed *in vacuo* over KOH pellets. The white residue was dissolved in water, and aliquots were subjected to paper chromatography in the solvent *n*-butanol:acetic acid:water, 4:1:1. The peptide degradation products were located by spraying an identification lane with ninhydrin reagent (*cf.* Fig. 1). Areas in the unreacted lanes corresponding to the ninhydrin-positive areas were excised, were eluted with 0.02 *N* HOAc, and were dinitrophenylated. The DNP-derivatives were hydrolyzed in acid, and the N-terminal residue and the amino acid composition were determined by paper chromatography.

Amide Analysis.—Two to 5 μmoles of a selected peptide was dissolved in dilute alkali at *p*H 8.5 to 9, and the solution was evaporated to dryness *in vacuo* over H_2SO_4. A solution of approximately 2 μmoles of peptide per ml. of 11.7 *N* HCl was prepared and was incubated at 37° for as long as 144 hr. At appropriate intervals, aliquots corresponding to 0.2 to 1 μmole of peptide were removed for ammonia analysis conducted either according to the method of Rees[4] or the procedure of Russell.[5] Each of the 7 peptides which were studied contained only one potential amide linkage.

In all cases a mixture of amino acids devoid of amide groups but otherwise simulating the peptide was treated similarly. Such control hydrolyses provided evidence for the amount of interfering substances or non-amide ammonia liberated.

Hydrolysis of Peptide C-3 with Carboxypeptidase.—A solution containing 0.02% peptide C-3, val(glu,try), and 0.024% carboxypeptidase was incubated under conditions identical to those employed in a previous report[6] enzyme: substrate mole ratio, about 1:65. Aliquots of the incubate were removed at various time intervals for dinitrophenylation and subsequent paper chromatography.[7]

Results

Partial Acid Degradation.—Three of the four remaining sequences were determined by characterizing the partial degradation products of the two enzymatically obtained peptides C-4, his(ser,glu)-

(1) Reference to Paper II of this series, THIS JOURNAL, **79**, 2798 (1957).

(2) Reference to Paper III of this series, *ibid.*, **79**, 2801 (1957).

(3) Reference to Paper IV of this series, *ibid.*, **79**, 2805 (1957).

(4) M. W. Rees, *Biochem. J.*, **40**, 632 (1946).

(5) J. A. Russell, *J. Biol. Chem.*, **156**, 457 (1944).

(6) Reference to Paper I of this series, THIS JOURNAL, **79**, 2794 (1957).

(7) A. L. Levy, *Nature*, **174**, 126 (1954).

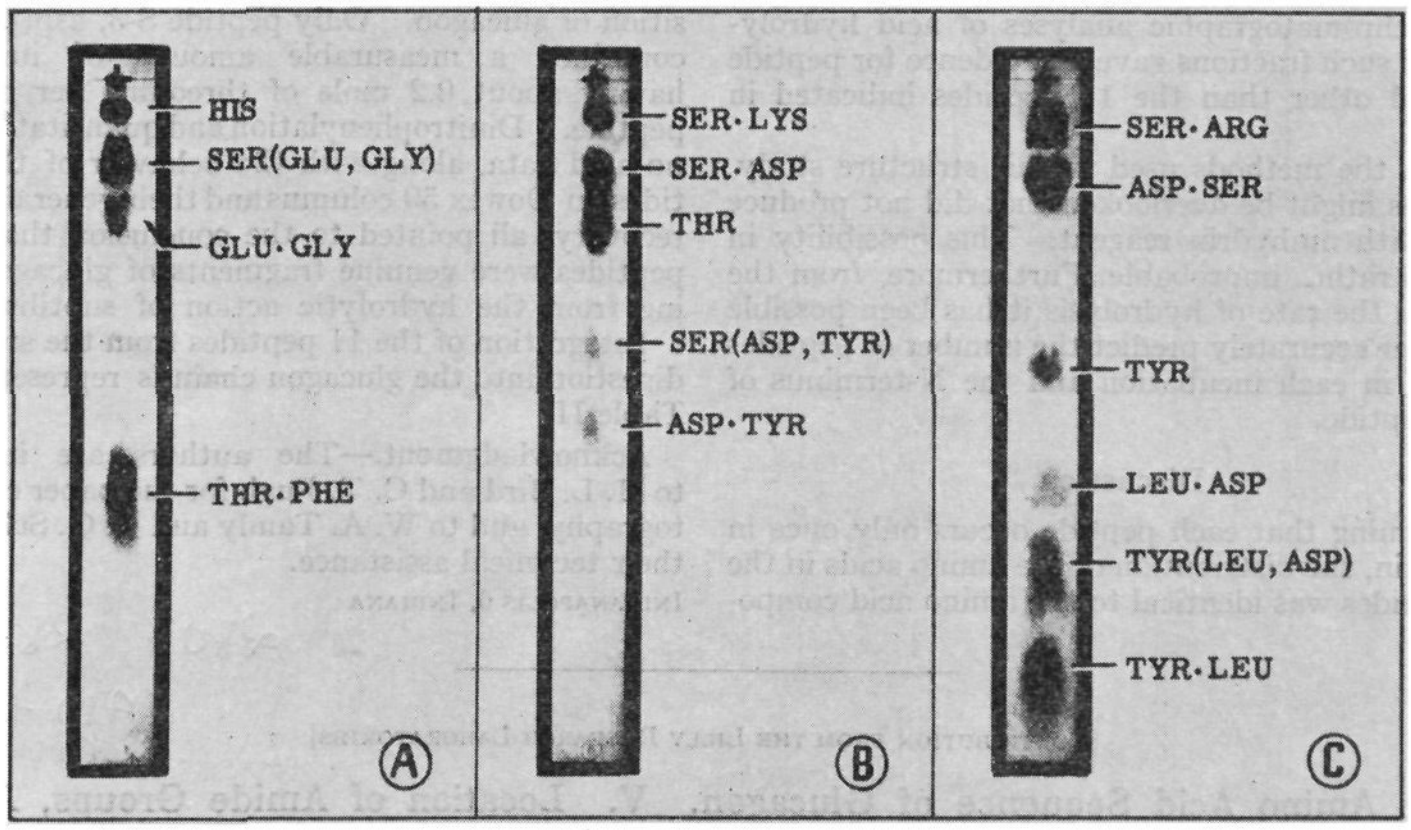


Fig. 1.—Photographs of the paper chromatographic separation of the acid hydrolysis products from peptides C-4, LT-3 and LT-4. The chromatograms were developed with *n*-butanol:acetic acid:water (4:1:1), and were treated with ninhydrin reagent: A, hydrolysis products of peptide C-4, His(ser,glu,gly,thr,phe); B, hydrolysis products of peptide LT-3, Thr(ser,asp,tyr,ser,lys); C, hydrolysis products of peptide LT-4, Tyr(leu,asp,ser,arg).

gly(thr,phe), and LT-4, tyr.leu(asp,ser,arg). In addition, peptide LT-3, thr.ser.asp.tyr,ser.lys, was similarly degraded. The quantitative amino acid composition of these peptides had been determined in the earlier described work.[1,2] Since acid degradation is not considered to promote rearrangements, such data elucidated the unknown sequences and also provided additional evidence for certain sequences which were proposed largely on the basis of data obtained from enzyme degradation.

Presented in Fig. 1 are photographs of the paper chromatographic separation of the acid hydrolysis products of the above three peptides.

The chromatograms were treated with ninhydrin reagent. In cases where overlapping of the spots occurred, the chromatograms were carefully excised to exclude the overlapping area. The analysis and recovery of the degradation products are presented in Table I. The figures on amino acid analysis are based on single or duplicate paper chromatographic analyses. Precise analytical data were not required to ascertain the amino acid composition of the degradation products since the parent peptides had been carefully analyzed and only one, LT-3, contained more than one residue of a given amino acid.

Small quantities of impurities were present in some of the hydrolytic products; this observation had been expected because of the non-specific degradative approach. Estimation of the purity of the fragments was based on end group analysis using the dinitrophenylation method. In the majority of cases at least 0.4 mole of a given amino acid residue was recovered per mole of starting peptide.

This recovery was considered acceptable since large losses had been expected due to the lengthy manipulation consisting of partial acid hydrolysis, paper chromatography, elution from the paper, dinitrophenylation, complete acid hydrolysis and

TABLE I

PAPER CHROMATOGRAPHIC ANALYSIS OF THE ACID DEGRADATION PRODUCTS

No. of parent peptide	Compn. and analysis of acid degrad. products	R_f	Estimated purity %	Major impurity
C-4	his	0.04	90	his(ser,glu)?
C-4	ser(glu,gly)			
	moles 1.2 1.3 0.6	.12	80	gly
C-4	glu.gly			
	moles 1.2 0.8	.20	100	...
C-4	thr.phe			
	moles 1.1 0.9	.60	100	...
LT-3	ser.lys			
	moles 1.0 1.0	.05	80	lys
LT-3	ser.asp			
	moles 1.1 0.9	.14	80	asp
LT-3	thr	.21	100	...
LT-3	ser(asp,tyr)			
	qual.	.37	65	asp(tyr.?)
LT-3	asp.tyr			
	moles 1.2 0.8	.50	100	...
LT-4	ser.arg			
	moles 0.9 1.1	.07	90	arg
LT-4	asp.ser			
	qual.	.15	100	...
LT-4	tyr	.41	100	...
LT-4	leu.asp			
	moles 0.8 1.2	.60	100	...
LT-4	tyr(leu,asp)			
	moles 1.0 0.6 1.4	.72	100	...
LT-4	tyr.leu			
	moles 1.0 1.0	.86	100	...

paper chromatographic analysis of both the free and the DNP-amino acids. The degradation of the parent peptides into relatively few fragments may be attributed at least in part to the high content of serine and threonine in the peptides. Bonds involving the amino groups of the hydroxy amino acids are known[8] to be particularly labile in concentrated acids at rather low temperatures.

Sequence of Peptide C-3.—Paper chromatography[7] of the dinitrophenylated aliquots from the carboxypeptidase time study of peptide C-3 revealed the presence of two major DNP-derivatives in approximately equal amount. DNP-tryptophan was recognized readily; however, the second derivative did not correspond to any known DNP-amino acid. Quantitative analysis[9] demonstrated the complete absence of tryptophan in the unknown DNP-derivative. After hydrolysis, followed by paper chromatographic separation, the unknown spot was identified as DNP-val.glu. The course of the hydrolysis is represented by the curves in Fig. 2, with the DNP-peptide calculated as if it

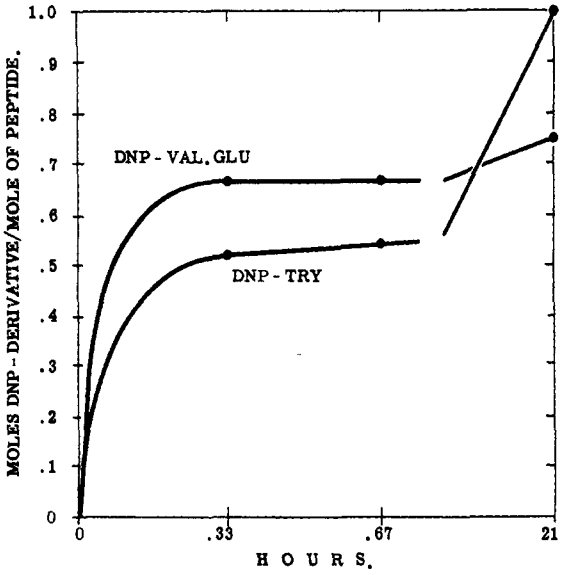


Fig. 2.—The rate of hydrolysis of peptide C-3, val(glu, try), by carboxypeptidase as determined by the dinitrophenylation method.

were DNP-valine. While the quantitative aspects of this experiment leave something to be desired, there can be little doubt that the sequence in peptide C-3 is val.glu.try. Both products of hydrolysis point to this conclusion; furthermore, this conclusion is in agreement with the usual specificity of chymotrypsin.

Location of the Amide Linkages.—The curves presented in Fig. 3 are representative of the results obtained from the amide analysis of peptides C-3, C-4, S-2, S-3, S-4, S-7 and S-9. In some cases the zero-time values for ammonia were abnormally high despite efforts to remove extraneous ammonia by concentration to dryness *in vacuo* from a weakly basic solution. Because of this difficulty it is possible that the quantitative recovery of the ammo-

(8) F. Sanger, *Advances in Protein Chem.*, **7**, 1 (1952).
(9) J. R. Spies and D. C. Chambers, *Anal. Chem.*, **20**, 30 (1948).

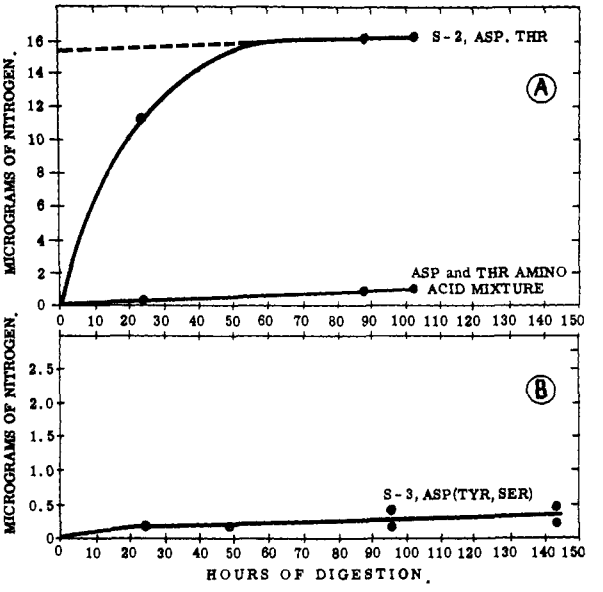


Fig. 3.—Amide analysis of peptides S-2 and S-3: A, ammonia nitrogen liberated from 1.2-μM aliquots of peptide S-2, and from equal quantities of a mixture of aspartic acid and threonine. Extrapolation to zero time provides a value of approximately one μM of NH_3 formed per μM of peptide S-2. B, Ammonia nitrogen liberated from 0.2 μM-aliquots of peptide S-3. Less than 0.1 μM of NH_3 was formed per μM of peptide S-3.

nia was masked by the initial high value, particularly in the cases of peptides C-4 and S-9

Despite this factor, the gradual release of ammonia during the digestion clearly showed which peptides contained the amide linkages. The quantitative data are presented in Table II. Of the 7

TABLE II

AMIDE ANALYSIS OF PEPTIDES FROM GLUCAGON DEGRADATION

Peptide no.	Peptide compn.	Moles of amide per mole of peptide
C-3	NH_2 | Val(glu,try)	1.1
C-4	NH_2 | His(ser,glu,gly,thr,phe)	0.6
S-2	NH_2 | Asp.thr	.9
S-3	Asp(tyr,ser)	.03
S-4	Asp,phe	.03
S-7	Leu(asp,ser,arg)	.0
S-9	NH_2 | Arg(ala,glu)	.5

peptides tested, only C-3, C-4, S-9 and S-2 yielded significant quantities of ammonia when incubated with strong acid. This is in agreement with the number of amide linkages found by chemical analysis of glucagon, and with the behavior of the peptides on resin columns. The data presented earlier on the release of amino acids from glucagon by the action of carboxypeptidase are also consistent with the conclusions that have been presented. Three amide-containing residues were liberated.

TABLE III

Proposed partial sequence	His(ser,glu)gly(thr,phe)thr.ser.asp.tyr.ser.lys.tyr.leu(asp,ser,arg)arg.ala.glu.asp.phe.val(glu,try)leu.met.asp.thr
Carboxypeptidase	val.glu.try
	NH_2 NH_2 NH_2 NH_2
Amide groups	His(ser,glu,gly,thr,phe) asp(tyr,ser) leu(asp,ser,arg)arg.(ala,glu)asp.phe val (glu,try) asp.thr
Acid degradation of certain peptides	His thr tyr.
	ser(glu.gly) ser(asp,tyr) tyr.leu
	glu.gly ser.asp tyr(leu,asp)
	thr.phe asp,tyr leu.asp
	ser.lys asp.ser
	ser.arg
	NH_2 NH_2 NH_2 NH_2
Summary	His.ser.glu.gly.thr.phe.thr.ser.asp.tyr.ser.lys.tyr.leu.asp.ser.arg.arg.ala.glu.asp.phe.val.glu.try.leu.met.asp.thr

TABLE IV

SUMMARY OF EVIDENCE LEADING TO THE AMINO ACID SEQUENCE OF GLUCAGON

Peptide no.[a]	Degradation products
ST-3	His(ser,glu,gly,thr,phe,thr,ser,asp,tyr,ser,lys)
S-8A	His(ser,glu)
C-4	NH_2
LT-5A	His.ser.glu.gly.thr.phe
S-6	gly(thr,phe)
LT-3	thr.ser.asp.tyr.ser.lys
S-1	thr.ser
C-2	thr(ser,asp,tyr)
S-3	asp(tyr,ser)
C-6	ser(lys,tyr)
S-10	lys,tyr
ST-1	tyr.leu.asp.ser.arg
LT-4	
S-7	leu(asp,ser,arg)
C-5	leu(asp,ser,arg,arg,ala,glu,asp,phe)
ST-2	arg
LT-5B	NH_2
S-9	arg(ala,glu)
ST-PA	arg(ala,glu,asp,phe,val,glu,try,leu,met,asp)thr
ST-PB	ala(glu,asp,phe,val,glu,try,leu,met,asp)thr
LT-PB	
LT-2	ala(glu,asp,phe,val,glu,try)
S-4	asp.phe
	NH_2
C-3	val.glu.try
S-8B	
S-5	leu.met
C-1	leu(met,asp)thr
LT-1	NH_2
S-2	asp.thr
Amino acids liberated by carboxypeptidase	NH_2 NH_2 NH_2
	ala(glu,asp,phe,val,glu,try,leu,met,asp)thr
	NH_2 NH_2 NH_2 NH_2
Proposed sequence	His.ser.glu.gly.thr.phe.thr.ser.asp.tyr.ser.lys.tyr.leu.asp.ser.arg.arg.ala.glu.asp.phe.val.glu.try.leu.met.asp.thr

[a] Prefix of the peptide number indicates the hydrolytic agent: ST = trypsin (2.25 hr.), S = subtilisin, C = chymotrypsin, LT = trypsin (50 hr.).

Two, as glutamine, occur in peptides C-3 and S-9. One, as asparagine, has been found in the terminal peptide, S-2. Finally, a residue of aspartic acid was released slowly by carboxypeptidase; the identification of this residue with the aspartic acid occurring in peptide S-4 is a logical interpretation.

Discussion

The results obtained by degradation of certain peptides with acid and with carboxypeptidase have been integrated into existing knowledge of the glucagon structure as shown in Table III. At least one line of evidence is available for each sequence in the chain, and in many cases more than one proof has been presented. In no case was it necessary to rely on the specificity of a proteolytic enzyme for a structure proof; however, with the exception of 2 or 3 minor splits caused by trypsin, the accepted specificity patterns of trypsin and chymotrypsin were confirmed. Of critical importance is the fact that all data, with no exceptions, are consistent with the final structure as proposed in Table III. A summary of all the evidence described in this series of reports is presented in Table IV.

Acknowledgment.—The authors gratefully acknowledge the assistance of E. R. Diller, E. E. Logsdon, H. L. Bird and C. T. Pugh. The authors also wish to thank C. W. Pettinga for the many helpful discussions and encouragement throughout this work.

INDIANAPOLIS 6, INDIANA