Amino Acid Sequence in Lysozyme

2. ELUTION CHROMATOGRAPHY OF PEPTIDES ON ION-EXCHANGE RESINS*

By ADRIENNE R. THOMPSON

Laboratory for the Study of Hereditary and Metabolic Disorders and the Departments of Biological Chemistry and Medicine, University of Utah College of Medicine, Salt Lake City, Utah, U.S.A., and the Biochemistry Unit, Wool Textile Research Laboratories, C.S.I.R.O., Melbourne, Australia

(Received 6 December 1954)

The previous paper in this series (Thompson, 1955) reported some amino acid sequences in lysozyme which were deduced from the structures of a number of small peptides isolated from partial acid hydrolysates by displacement chromatography on ion-exchange resins and further fractionation on paper chromatograms. Despite the success achieved by Partridge and co-workers (see Partridge & Brimley, 1952) in the separation of amino acids by displacement chromatography and the high capacity of such columns, it was apparent in these experiments that the relatively non-specific van der Waals interactions between the larger peptide molecules and the resin outweighed the effect of differences in the pK values of their ionizing groups, with the result that many peptides were similarly absorbed and thus not differentiated. Nevertheless, a large number of peptides were isolated and identified in these experiments.

Experiments using elution chromatography on ion-exchange resins for the preliminary separation of peptides from an acid hydrolysate of lysozyme are described in this paper. The high resolving power of this method for amino acids was elegantly demonstrated by Moore & Stein (1951). Dowmont & Fruton (1952) have shown that a number of simple peptides may be satisfactorily chromatographed in this way. Schroeder, Kay, Le Gette, Honnen & Green (1954) have chromatographed a partial acid hydrolysate of gelatin in Dowex-50 $\times$ 8 and examined the material in the eluate. After further resolution of the peptides as *N*-2,4-dinitrophenyl (DNP) derivatives on silicic acid-Celite columns a number of di- and tri-peptides were identified in gelatin hydrolysates. A similar examination was made of silk-fibroin hydrolysates by Kay & Schroeder (1954). This procedure has the advantage that no desalting step is necessary to remove buffer salts from the eluant fractions.

The present experiments have borne out the expectation that the elution procedure would possess the highest general resolving power of any method yet developed for the separation of

peptides. However, the complexity of a partial acid hydrolysate of a protein is such that it has not been possible to obtain pure peptides by this method alone. Nevertheless, the fractions representing peaks in the elution curve are fairly simple and can readily be further fractionated into pure peptides by two-dimensional paper chromatography. The identification of most of these peptides from an acid hydrolysate in this manner has allowed the compilation of further sequences in lysozyme.

Sulphonated polystyrene (Dowex-50) cross-linked with 4% divinylbenzene (cf. Dowmont & Fruton, 1952) has been used in the present experiments. Successful chromatography is achieved with this degree of crosslinkage, which allows penetration of peptides containing up to four or five residues, whereas the 8% crosslinked resin used for amino acid analysis is unsuitable (Moore & Stein, 1951).

METHODS

Preparation of hydrolysate of lysozyme. The egg-white lysozyme (Armour and Co.) used in these experiments was the same as that used previously (Thompson, 1955). Lysozyme (500 mg.) was hydrolysed at 37° for 4 days with 12*N*-HCl and treated as previously with performic acid to oxidize the disulphide bonds and convert the half-cystine residues into cysteic acid (Thompson, 1955). Methionine residues are converted into 'methionine sulphone' residues by this treatment.

Chromatography of lysozyme hydrolysate. The elution chromatogram described in this paper was run with gradient elution with ammonium formate and acetate buffers which were sublimed to give salt-free fractions (cf. Hirs, Moore & Stein, 1952). However, this procedure necessitated removal of ammonia from the samples of solution used to plot the elution peaks, since ammonia, like the amino acids and peptides, gives a ninhydrin colour. Alternatively, the chromatogram could have been run with sodium salts, when the elution curve could be plotted without preliminary treatment. However, each fraction would then have required separate desalting treatment (cf. Stein, 1953), and rather large columns would have been needed to handle the large volumes of effluent buffer in the less sharply defined peaks. The sodium cycle would also have necessitated individual neutralization of each eluate

* Part 1: Thompson (1955).

fraction in plotting the effluent peaks, or development of a ninhydrin reagent with sufficient buffering power to overcome this tedious operation (S. Moore & W. H. Stein, private communication).

A sample of sulphonated-polystyrene cation-exchange resin (Dowex-50, supplied by Dow Chemical Company) crosslinked with 4% divinylbenzene and graded 200-400 mesh/in. was first passed through a 40 mesh/in. sieve to remove coarse particles. In order to give a satisfactory flow rate in the column, it was found necessary to remove with a 270 mesh/in. sieve a large amount of fine material which otherwise reduced the flow rate seriously. The resin was prepared for chromatography according to the pro-

cedure were entirely filled with buffer, which was allowed to flow down the column at a rate of 8 ml./hr. Fractions of 8 ml. were collected with a Technicon fraction collector (Technicon Company, 215 E. 149th Street, New York 51) on a drop-counting basis. A preliminary experiment had indicated approximately the points at which changes of buffer should be made. The composition of the buffers used is given in Table 1.

The peptide and amino acid material in the tubes was located by taking a 0.2 ml. sample from every second tube and determining the colour yield with ninhydrin, according to the procedure of Moore & Stein (1948), 2 ml. ninhydrin solution and 10 ml. diluent being used.

Table 1. *Composition of buffers*

Based in part on the buffers used by Hirs, Moore & Stein (1952). 1.5% (v/v) of benzyl alcohol was added to all buffers.

Molarity of NH_4^+	pH	Anion	Buffer composition		Vol. of buffer used in elution run (ml.)
			Vol. of 2M acid	Vol. of M- NH_3	
			To final vol. 1 l. with water		
0.05	2.49	Formate	450	50	*
0.2	3.44	Formate	263	200	1500
0.2	3.67	Formate	200	200	1500
0.2	4.04	Formate	138	200	2000
0.2	4.28	Formate acetate	†	†	1500
0.2	4.50	Acetate	220	200	1500
0.4	4.93	Acetate	300	400	1000
0.4	5.11	Acetate	500	400	1500
0.6	5.62	Acetate	400	600	1500
0.6	6.27	Acetate	385	600	1000
0.6	6.81	Acetate	370	600	1000

* This buffer was used to equilibrate the column, add the hydrolysate to the column and fill the mixer reservoir and connecting tubing (approx. 600 ml.), and the next buffer (pH 3.44) was added to the separating funnel.

† A mixture of ammonium formate buffer (pH 4.04) (2 vol.) and ammonium acetate buffer (pH 4.50) (1 vol.).

cedure of Hirs *et al.* (1952), the final wash being with 0.05M ammonium formate buffer (pH 2.49) (molarity refers to cation, see Table 1) until the pH remained constant.

A column of diameter 1.8 cm. and height 2 m. was packed in portions as described by Moore & Stein (1951), but without applying pressure. The column was again washed with 0.05M ammonium formate (pH 2.49) till the effluent pH was the same as that of the solution entering the column.

The column was set up for gradient elution by inserting a buffer mixer in the line between the reservoir funnel and the top of the column. This consisted of a 500 ml. flask with ground-glass joint carrying an inlet tube, connected by Tygon tubing (Arthur H. Thomas, Philadelphia 5, Pa., U.S.A.) to the reservoir, reaching half-way down the flask, and an outlet to the top of the column through Tygon tubing. The flask was stirred continuously by a magnetic stirrer. Thus discontinuous buffer changes resulted in smoothly continuous increase in the pH and ionic strength of the solution admitted to the top of the column.

The hydrolysate was taken up in the first buffer (10 ml.; pH 2.49), added to the top of the column, and washed in with a further 10 ml. of buffer. A certain amount of material remained as a sludge and was also transferred. The buffer mixer, connecting tubes and top of the column

As ammonia itself gives a colour with ninhydrin, the ammonium acetate from the buffer was first removed by evaporating the sample to dryness overnight at 37° in an evacuated desiccator over NaOH and H_2SO_4 , adding 0.5 ml. N-NaOH to the sample and re-evaporating overnight as before, and finally adding a drop of water to each tube and evaporating to dryness again over fresh H_2SO_4 . This last step was necessary as the sulphuric acid formed an ether- and acetone-soluble and water-insoluble scum, presumably a polymerized derivative of formic and acetic acids. This prevented the ammonia being completely absorbed and led to some variability in the results. The NaOH was neutralized with 1 ml. 0.5N-HCl before the addition of the ninhydrin solution. The tubes were heated in a boiling-water bath for 30 min. instead of 20 min. (since water boils at 93° at Salt Lake City and since the rate of reaction of the complex mixture of peptides with ninhydrin is not known).

An elution curve (Fig. 1) was obtained by plotting the colour density against the fraction number. Approximately fifty more-or-less sharply defined peaks were obtained and the eluate was divided into a corresponding number of fractions by combining the contents of the tubes under each peak.

Each fraction was evaporated to dryness and the material finally distributed as a film on the walls of a large

flask. The ammonium acetate was removed by sublimation on a cold-finger water condenser under high vacuum overnight, a trap cooled in solid CO_2 -ethanol being inserted before the pump. Solution and redistribution on the walls of the flask and resublimation gave efficient removal of virtually all the ammonium acetate.

The fractions were examined as previously described (Thompson, 1955). Two-dimensional chromatography was used to isolate the individual peptides, which were eluted after detection with dilute ninhydrin. A portion of each peptide fraction was hydrolysed to determine the constituent amino acids. The *N*-terminal end group was determined by the DNP method (Sanger & Thompson, 1953a). In suitable cases, when more material was avail-

able, carboxypeptidase was used to determine the *C*-terminal residue. As in the previous paper (Thompson, 1955), evidence for such a *C*-terminal amino acid is denoted by italicizing this residue. In a limited number of cases partial acid hydrolysis of the DNP-peptide was used to establish the sequence (cf. Sanger, 1949). The DNP-peptide was partially hydrolysed with 12*N*-HCl for 6 days at 37°. If the colour could not be extracted into ether or ethyl acetate, the DNP-peptide was freed from contaminating peptides and amino acids by adsorption on a column of talc (see Sanger, 1949). This step was necessary, otherwise the amino acids and peptides overlapped with the DNP derivatives. The latter were separated from each other on a paper chromatogram run with the *tert*.-pentanol

Fig. 1. Chromatogram of a partial acid hydrolysate of lysozyme on sulphonated-polystyrene resin (Dowex-50 $\times$ 4) eluted with ammonium formate and acetate buffers (Table 1). Gradient elution was used and the pH and molarity of the buffers are indicated below the curve. The vertical lines below the abscissa indicate the tubes which were combined to give fractions denoted by the letters above the curve. For other details see text.

(2-methylbutan-2-ol)-pH 6 buffer system used for DNP-amino acids. The DNP-peptide spots were eluted with acetone-*N*-HCl (1:1, by vol.), hydrolysed and the constituent amino acids examined on paper chromatograms. Identification of a DNP-dipeptide established the sequence of the first two residues in the original peptide and, in the case of a tripeptide, the complete sequence.

RESULTS

The results where the partial hydrolysis of DNP-peptides was studied are given in Table 2.

tion contained about eight components which the subsequent two-dimensional chromatogram was capable, in most instances, of resolving into pure peptides. The neutral amino acids emerged early (with leucine in fraction *U*) and in the same order as described by Moore & Stein (1951). The peptides in general were retarded somewhat in comparison (e.g. whereas threonine appears in fraction *J* and alanine in fraction *N*, Thr. Ala emerges in fraction *X*). The basic amino acids lysine and histidine occurred in fraction *AL* and arginine in fraction

Table 2. Sequences from partial acid hydrolysis of DNP-peptides

Peptide	Fraction (Fig. 1 and Table 4)	Amino acid identified in DNP-peptide	Sequence
Ser. (Asp, Arg)	<i>AP</i>	Asp	Ser. Asp. Arg
Thr. (Asp, Gly)	<i>R</i>	Gly	Thr. Gly. Asp
Thr. (Asp, Arg)	<i>AQ</i>	Asp	Thr. Asp. Arg
Gly. (CySO ₃ H, Asp)	<i>C</i>	CySO ₃ H	Gly. CySO ₃ H. Asp
Leu. (Gly, Ala, Val)	<i>AE</i>	Gly + + + + + Ala + +	Leu. Gly. Ala. Val*

* The large amount of glycine and the lesser amount of alanine are taken to indicate a mixture of DNP-Leu. Gly and DNP-Leu. Gly. Ala.

Table 3. Peptides from fraction *C*

For the isolation of peptides, two-dimensional chromatograms were run with phenol in an atmosphere of ammonia (0.3%) in the first direction and with a butanol-acetic acid-water mixture (40:10:50, by vol.) (aged at least 3 days) in the second dimension. Whatman no. 3 paper, chromatographically washed with 2*N* acetic acid followed by water, was employed. Relative intensities of spots are indicated roughly by + signs.

Spot no.	Spot intensity	Distance moved in phenol-NH ₃ (cm.)	Distance moved in butanol-acetic acid (cm.)	Amino acids after hydrolysis	DNP treatment and hydrolysis		Structure
					Amino acids	DNP-amino acids	
1	++++	3.5	8	CySO ₃ H + + + Asp + + + Gly + + +	CySO ₃ H + + Asp + +	Gly (medium)	Gly. (CySO ₃ H, Asp)
2	+++	0.5	8	CySO ₃ H + + + Asp + + +	CySO ₃ H trace Asp + +	CySO ₃ H† (weak)	CySO ₃ H. Asp
3	+++	32.5	12	CySO ₃ H + + + Asp + + + Ileu + + + Pro *	CySO ₃ H + + + Ileu + + + Pro *	Asp (medium)	Asp. (CySO ₃ H, Pro, Ileu)
4	+	23	17	CySO ₃ H + + Asp + + Leu + +	CySO ₃ H + + Asp trace Leu + +	Asp (medium)	Asp. (CySO ₃ H, Leu)

* Yellow spot.

† Yellow aqueous layer.

The results of the examination of the peptides in a typical fraction (*C*) are given in Table 3.

Table 4 shows the peptides contained in each fraction, where they were obtained pure and the results of the determination of their structure were clear-cut.

Although fairly sharp, symmetrical peaks were obtained with a number of fractions, it was clear from subsequent analysis that each fraction, corresponding to one peak, did not represent a pure amino acid or peptide. On an average, each frac-

AR. The positions of the peptides in the elution sequence relative to each other correspond closely to what would be expected on the basis of their composition and the elution behaviour of the constituent amino acids. For instance, dipeptides containing *N*-terminal arginine linked with, as *C*-terminal residue, cysteic acid, aspartic acid, glycine or leucine, emerge in fractions *AF*, *AO* (and *AQ*), *AV* and *AX* respectively, in the order of emergence of these amino acids when free.

Table 4. *Peptides in elution fractions*

The major component of each fraction is indicated by an asterisk, and free amino acids are printed in capital letters. Where the sequence is unknown the amino acids are enclosed within brackets.

Fraction	Components	Fraction	Components
<i>A</i>	Artifact	<i>S</i>	—
<i>B</i>	CYSTEIC ACID*	<i>T</i>	ISOLEUCINE*
<i>C</i>	Gly.(CySO ₃ H, Asp)* CySO ₃ H. Asp Asp.(CySO ₃ H, Leu) Asp.(CySO ₃ H, Pro, Ileu)		Ala. MetO ₂ Gly. MetO ₂ * Asp.(Ala, MetO ₂) (Asp. Ser)
<i>D</i>	CySO ₃ H. Ala* CySO ₃ H. Glu CySO ₃ H. (Glu, Ala) CySO ₃ H. (Asp, Gly)	<i>U</i>	LEUCINE* Ala. MetO ₂ Thr.(Asp, Gly) Ser.(Asp, Gly) MetO ₂ . Asp (Asp, Leu) Asp. Ala (Asp, Glu, Ala, Val) (Asp, Glu, Val)
<i>E</i>	Ser.(CySO ₃ H, Asp, Val)* Asp.(CySO ₃ H, Pro, Ileu)* Val. CySO ₃ H (CySO ₃ H, Glu) (CySO ₃ H, Ala) Val.(CySO ₃ H, Ala)	<i>V</i>	Thr. Glu* (Glu, Thr, Ala) (Asp, Ser, Val) Ala. MetO ₂ Ser.(Asp, Ileu) Ser.(Asp, Gly) (Asp, Glu, Ala, Val) Asp. Leu
<i>F</i>	CySO ₃ H. (Glu, Leu)* Asp.(CySO ₃ H, Leu)* Leu. CySO ₃ H Asp.(CySO ₃ H, Ileu) (CySO ₃ H, Asp) Leu.(CySO ₃ H, Asp) CySO ₃ H. Ileu	<i>W</i>	(Asp, Glu, Ileu)* (Asp, Ileu)
<i>G</i>	Asp.(CySO ₃ H, Ileu)	<i>X</i>	Glu.(Asp, Ileu)* Asp. Ileu* Thr. Ala* Glu. Ala (Asp, Ala, Ileu) Ser. Ala Glu.(Asp, Ala)
<i>H</i>	Asp. Asp*	<i>Y</i>	Ala. Ala* (Asp, Ileu) Thr.(Asp, Glu, Ala, Val) Val.(Glu, Ala) Asp.(Gly, MetO ₂) Ser.(Asp, Val) (Asp, Ala, Val) (Ser, Ala) (Asp, Glu, Ala, Val)
<i>I</i>	ASPARTIC ACID*		(Pro, Gly) Glu.(Asp, Ileu) Val. Glu
<i>J</i>	SERINE* THREONINE METHIONINE SULPHONE	<i>AB</i>	(Pro, Gly)* Glu.(Asp, Ileu)* Val. Glu
<i>K</i>	GLUTAMIC ACID* PROLINE SERINE (trace) (Asp, Gly)	<i>AC</i>	Glu.(Ala, Leu)* CySO ₃ H. Lys* Ser. Val (Asp, Ileu)
<i>L</i>	GLUTAMIC ACID (small amount)* Asp. Ileu* Asp. Gly	<i>AD</i>	Thr.(Gly, Pro)* Val. Ala Val.(Thr, Gly, Pro, Ala) Thr.(Glu, Gly, Ala) CySO ₃ H.(Gly, Lys)
<i>M</i>	GLYCINE* ALANINE (Asp, Ileu)	<i>AE</i>	Glu.(Ala, Leu, Ileu)* (Ala, Leu)* Leu.(Gly, Ala, Val)* (CySO ₃ H, Gly, Lys)
<i>N</i>	ALANINE* GLYCINE Val.(CySO ₃ H, Asp, Ala) (Asp, Leu)		
<i>O</i>	GLYCINE* ALANINE Asp. Gly VALINE		
<i>P</i>	(Asp, Gly)* (Asp, Glu, Val)		
<i>Q</i>	Thr. Asp* (Asp, Ser, Gly) (Asp, Glu, Ala)		
<i>R</i>	Ser.(Asp, Gly)* Thr.(Asp, Gly) Asp. Ala Thr.(Asp, Gly, Val) Asp.(Val, Glu) Ser. Asp Val. Asp		

* For note see p. 258.

Table 4 (cont.)

Fraction	Components	Fraction	Components
<i>AF</i>	Glu. Leu* Gly. Leu* Ileu. (Ser, Gly, Leu)* Arg. CySO ₃ H Ser. (Asp, Ileu)? Ileu. (Glu, Leu)	<i>AL</i>	LYSINE* HISTIDINE (His, Lys) Ser. (CySO ₃ H, Asp, Leu, Arg) Ser. (Asp, Thr, Ala, Ileu)? Asp. (Thr, Ala, Ileu)
<i>AG</i>	Ser. Leu* Ala. Leu* Ala. (Glu, Ileu, Leu) CySO ₃ H. Lys (Arg, CySO ₃ H, Glu) (Ser, Val)	<i>AM</i>	Asp. Arg*
<i>AH</i>	Phe. Asp* CySO ₃ H. Arg Asp. Arg (CySO ₃ H, Glu, Ala, Leu) (Tyr, Asp) Gly. (CySO ₃ H, Arg) Ileu. Pro	<i>AO</i>	Arg. Asp* Ala. Lys
<i>AI</i>	CySO ₃ H. (Asp, Gly, Arg) (Asp, Arg)	<i>AP</i>	Ala. Lys* Ala. (Gly, Lys)* Ser. (Asp, Arg) Ileu. (Thr, Ala)
<i>AJ</i>	Phe. Glu* Phe. Asp (Glu, Phe, Gly) Ser. (Asp, Phe) MetO ₂ . (CySO ₃ H, Lys) Asp. (Gly, Arg) (Asp, Arg)	<i>AQ</i>	Lys. Gly* Ala. (MetO ₂ , Lys) Arg. Asp Thr. (Asp, Arg)
<i>AK</i>	CySO ₃ H. (Arg, Leu)* Phe. (Asp, Glu)* (Asp, Ser, Phe)* Asp. (CySO ₃ H, Leu, Arg)? (Asp, Arg) Asp. (Gly, Arg) Ser. (CySO ₃ H, Asp, Leu, Arg) (CySO ₃ H, Asp, Leu, Arg)	<i>AR</i>	ARGININE* * (Asp, Arg)
		<i>AS</i>	MetO ₂ . Lys*
		<i>AT</i>	Val. (Ileu, Lys)* MetO ₂ . (CySO ₃ H, Lys, Arg)
		<i>AU</i>	Ser. Arg* (Asp, Thr, Arg)
		<i>AV</i>	Lys. (Phe, Glu)* (Arg, Gly) (CySO ₃ H, Lys, Arg)
		<i>AW</i>	Ileu. Arg* Arg. (CySO ₃ H, Gly, Lys)
		<i>AX</i>	Arg. Leu* (Arg, Ileu)

Note. R_F values of the peptides were not obtainable as the solvent was allowed to drip off the bottom of the paper; the chief aim of the experiment was the isolation of pure peptides rather than their characterization by R_F values. Thus only the distances moved by the peptides could be recorded. The peptides listed above were isolated from a number of different chromatograms and a comparison of distances moved by peptides on different chromatograms would be misleading. No common reference substance was included on chromatograms.

In general, any particular peptide did not spread over more than three consecutive fractions. For example, the peptide Ala. MetO₂ was identified in fractions *T*, *U* and *V*. This is to be expected on theoretical grounds if a peak does not return to the base-line on either side and the fraction is cut at the positions of minimum ninhydrin colour. This is true even when the substance is responsible for the shape of the elution peak; the position of the fraction cuts is even more arbitrary when a superimposed substance, present in higher concentration, determines the position of the cuts. In this particular instance free leucine, largely responsible for peak *U*, determined the cutting. In a few instances spreading over more than three fractions was observed. For example, peptides containing aspartic acid and isoleucine were found in fractions *L*, *M*, *W*, *X*, *Y* and *AC*. It was not always practicable to determine the *N*-terminal residue on

every such peptide. Hence some doubt exists as to whether all the peptides have the same sequence. Furthermore, with peptides containing aspartic and glutamic acids, uncertainty exists as to whether the presence of this amino acid in its amide form, owing to incomplete removal of the ammonia during partial acid hydrolysis, introduces a second species in the chromatogram. In this particular case it would certainly appear that the compound appearing in fractions *L* and *M* is different from that occurring in the later fractions which are widely separated from *L* and *M*.

In Table 5 are listed all the peptides that have been isolated from lysozyme in this work and

Note on Table 5 (opposite). In the preliminary elution experiment a few peptides were isolated which were not encountered in the main experiment. They are also listed as occurring in the corresponding fraction of the main experiment but are marked.†

Table 5. *Peptides from acid hydrolysis of lysozyme*

* Indicates that the peptide was isolated after displacement chromatography (Thompson, 1955).

Peptide	Fraction	Peptide	Fraction
Ala. Ala	Y	*Leu. Ala	AG
*Ala. Leu		*Leu or Ileu. Glu	
Ala. Lys	AO, AP	Leu. CySO ₃ H	F
*Ala. MetO ₂	U, V	Leu. Leu	†AJ
Ala. (Gly, Lys)	AP	Leu. (CySO ₃ H, Asp)	F
Ala. (MetO ₂ , Lys)	AQ	*Leu. (Thr, Ala)	
Ala. (Glu, Ileu, Leu)	AG	Lys. Gly	AQ
*Ala. (Asp, Glu, Ala?)		Lys. (Phe, Glu)	AV
*Arg. Asp	AO, AQ	MetO ₂ . Asp	U
*Arg. CySO ₃ H	AF	MetO ₂ . Lys	AS
Arg. Gly	AV	MetO ₂ . (CySO ₃ H, Lys)	AJ
Arg. Leu	AX	MetO ₂ . (CySO ₃ H, Lys, Arg)	AT
*Arg. (CySO ₃ H, Asp)		Phe. Asp	AH
*Arg. (CySO ₃ H, Glu)	AG	Phe. Glu	AJ
*Arg. (CySO ₃ H, Lys)	AV	*Phe. (Asp, Glu)	AK
Arg. (CySO ₃ H, Gly, Lys)	AW		
*Asp. Ala	R, U	*Ser. Ala	Y
*Asp. Arg	AH, AK, AM	Ser. Arg	AU
Asp. Asp	H	*Ser. Asp	R
*Asp. Gly	L, O	*Ser. Leu	AG
*Asp. Ileu	L, X	Ser. Val	AC
Asp. Leu	V	Ser. (Asp, Arg)	AP
Asp. (Ala, MetO ₂)	T	*Ser. (Asp, Gly)	R, U, V
Asp. (CySO ₃ H, Ileu)	F, G	Ser. (Asp, Ileu)	V, AF
Asp. (CySO ₃ H, Leu)	C, F	*Ser. (Asp, Leu)	
Asp. (CySO ₃ H, Arg)	N	Ser. (Asp, Phe)	AJ
Asp. (Gly, Arg)	AJ, AK	*Ser. (Asp, Val)	Y
*Asp. (Gly, MetO ₂)	Y	Ser. (Gly, Ala)	X
*Asp. (Glu, Ala)		*Ser. (Gly, Leu)	†AG
*Asp. (Glu, Ileu)		*Ser. (Leu, Arg)	
*Asp. (Glu, Val)	R	*Ser. (Asp, Gly, Ala)	
Asp. (CySO ₃ H, Leu, Arg)	AK	*Ser. (Asp, Gly, MetO ₂)	
Asp. (CySO ₃ H, Pro, Ileu)	C, E	*Ser. (CySO ₃ H, Asp, Val)	E
Asp. (Gly, Leu, Arg)	AJ	Ser. (CySO ₃ H, Ala, Val, Lys)	AF
*Asp. (Glu, Ala, Val)		Ser. (CySO ₃ H, Asp, Leu, Arg)	AK, AL
Asp. (Thr, Ala, Ileu)	AL	*Ser. (Gly, Ala, Leu?, Leu)	
CySO ₃ H. Ala	D	Ser. (Asp, Glu, Ala, Ileu, Leu)	AG
CySO ₃ H. Arg	†AH	Leu. (Gly, Ala, Val)	AE
CySO ₃ H. Asp	C	*Leu or Ileu. (Glu, Ala, Leu, Leu or Ileu)	
CySO ₃ H. Glu	D		
CySO ₃ H. Ileu	F	*Thr. Ala	X
CySO ₃ H. Lys	AC	*Thr. Asp	Q
*CySO ₃ H. (Ala, Lys)		Thr. Gly	†AD
CySO ₃ H. (Arg, Leu)	AK	*Thr. Glu	V
CySO ₃ H. (Asp, Gly)	D	Thr. Pro	†AA
*CySO ₃ H. (Gly, Lys)	AD, AE	Thr. (Asp, Arg)	AQ
CySO ₃ H. (Glu, Ala)	D	*Thr. (Asp, Gly)	R, U
CySO ₃ H. (Asp, Gly, Arg)	AI	Thr. (Asp, Val)	†R
Gly. Leu	AF	*Thr. (Glu, Ala)	†Q
Gly. MetO ₂	†T	*Thr. (Asp, Arg, Arg)	AU
Gly. (CySO ₃ H, Arg)	AH	*Thr. (Asp, Glu, Val)	†P
Gly. (CySO ₃ H, Asp)	C	Thr. (Asp, Gly, Val)	R
*Glu. Ala	X	Thr. (Glu, Gly, Ala)	AD
*Glu. Leu	AF	Thr. (Asp, Glu, Ala, Val)	Y
*Glu. (Ala, Leu)	AC	Thr. (Gly, Pro)	AD
Glu. (Asp, Ala)	X	Val. Ala	AD
*Glu. (Asp, Ileu)	X	Val. Asp	R
Ileu. Arg	AW	*Val. CySO ₃ H	E
*Ileu. Asp		Val. Glu	AA, AB
Ileu. Pro	AH	*Val. (Asp, Glu)	
Ileu. Val	†AH	*Val. (CySO ₃ H, Ala)	E
Ileu. (Asp, Ala)	†Y	Val. (Glu, Ala)	Y
Ileu. (Glu, Leu)	AF	Val. (Ileu, Lys)	AT
Ileu. (Thr, Ala)	AP	Val. (CySO ₃ H, Asp, Ala)	N
Ileu. (Ser, Gly, Leu)	AF	Val. (Gly, Thr, Ala, Pro)	AD

also in earlier experiments, using displacement chromatography on ion-exchange resins (Thompson, 1955).

Fractions in the preliminary experiment corresponding exactly to peaks *A* and *B* of the main experiment were examined. Peak *B* consisted practically entirely of free cysteic acid. Peak *A* also contains cysteic acid but in very much smaller quantity than peak *B* and it seems probable that peak *A* is an artifact occurring at the point of emergence of the first hold-up volume of the column.

One of the major components in fraction *AD* gave only glycine on hydrolysis. However, glycine was identified in fraction *M*, which clearly corresponds to its correct position in the elution sequence (cf. Moore & Stein, 1951). It was at first thought that the compound found in fraction *AD* might be Gly.Gly. This was disproved by comparison of the *R* values of the DNP-derivatives on a Celite column buffered at pH 4 and with chloroform as solvent (Perrone, 1951). (*R* of DNP derivative of the glycine compound in fraction *AD*, 0.18; *R* of authentic DNP-glycine, 0.16; *R* of authentic DNP-Gly.Gly, 0.03.) It thus appears that the compound is probably free glycine and the most plausible explanation for its anomalous occurrence in fraction *AD* would seem to be its formation by cleavage of some labile compound in the fraction after the chromatography on ion-exchange resin.

The dipeptide Gly.Lys, detected in experiments using displacement chromatography, was listed in previous communications (Thompson, 1954, 1955). There is now, however, reason to believe that it was, in fact, Lys.Gly. On end-group examination of a peptide containing glycine and lysine in fraction *AQ*, DNP-glycine was again detected in small quantity and it was at first thought that glycine was in the *N*-terminal position, the poor yield of DNP-glycine being attributed to the known lability of this compound to acid hydrolysis (Sanger, 1945). However, on exposing the chromatogram to the vapour of boiling 20 % hydrochloric acid (Sanger & Thompson, 1953*a*), a rather elongated spot was revealed under the position occupied by dinitrophenol and another artifact, always present on such chromatograms (run with *tert*.-pentanol-pH 6 buffer) and corresponding to a bis-DNP-lysine control. The dipeptide is therefore Lys.Gly, and the large amount of glycine and virtual absence of lysine or ϵ -DNP-lysine in the aqueous layer confirm this conclusion.

Acher *et al.* (1954) have reported the presence of the peptide Leu.Pro in lysozyme. Since Ileu.Pro was detected in fraction *AH* and there are only two proline residues in this protein, the other being preceded by threonine (cf. Thr.Pro in Table 5), it appears that these workers have failed to distinguish between the two leucine isomers. In the

present work clear proof that isoleucine and not leucine was present in this peptide was provided by paper chromatography of a hydrolysate with *tert*.-pentanol-water as solvent system. This conclusion is substantiated by the isolation of the peptide Asp.(CySO₃H,Pro,Ileu), in which the isoleucine residue was determined in the same way, and by the establishment of the sequence Asp.Ileu.Pro.-CySO₃H (see Table 6).

DISCUSSION

The results of the experiments described in this paper indicate that elution chromatography on ion-exchange resins has the highest resolving power for the separation of peptides from a partial acid hydrolysate of any method hitherto evaluated. That this procedure alone did not yield pure peptides is a reflexion of the complexity of such a hydrolysate. It may be noted that some 130 peptides have been identified in the lysozyme hydrolysate and the actual number present is probably considerably higher.

These results emphasize that the usual conditions for partial acid hydrolysis produce fairly random splitting of a protein molecule and a peptide mixture of far too great a complexity, so that it is desirable that other methods be worked out for the determination of amino acid sequence of proteins. In the absence of efficient methods for determining sequence by stepwise degradation along a protein chain, more specific methods of cleavage of the chains must be sought to provide a lesser number of fragments for separation and identification. Digestion by proteolytic enzymes to some extent provides such a method, but the cleavage of every susceptible bond is by no means complete (Sanger & Tuppy, 1951; Sanger & Thompson, 1953*b*), with consequent increase in the complexity of the hydrolysate.

Two problems arise in using methods which give rise to longer-chain fragments: The first is the separation of such peptides. It seems likely that elution chromatography on a less highly crosslinked resin (such as '2 % crosslinked' Dowex-50) will be successful for this purpose. A preliminary elution chromatogram of a trypsin digest of lysozyme on this resin showed a simpler curve, with a smaller number of well-defined peaks, when compared with the partial acid hydrolysate. It also indicated that successful chromatography of such a peptide mixture can probably be achieved. The second problem is that of determining the structure of these larger peptides, which is far more difficult than with the small peptides from an acid hydrolysate. A preliminary necessity would be a quantitative analysis of each peptide and the isolation of a reasonable quantity of it for the amount of work

Table 6. *Sequences of residues in lysozyme*

1. Thr. Asp. Val. Glu. Ala*
2. Ileu. Glu. Leu. Ala. Leu*
3. Asp. Glu. Ala*
4. Leu. Thr. Ala*
5. Glu. Asp. Ileu*
6. **Thr. Glu. Ala. Gly**
Thr. (Glu, Ala, Gly)
Thr. Glu. Ala*
7. **Ser. Asp. Gly. MetO₂. Asp**
Ser. Asp. Gly. MetO₂*
MetO₂. Asp
8. **Asp. Ala. MetO₂. Lys. CySO₃H. Arg**
MetO₂. (Lys, CySO₃H, Arg)
Asp. (Ala, MetO₂)
Ala. (MetO₂, Lys)
MetO₂. (Lys, CySO₃H)
Asp. Ala. MetO₂. Lys
Ala. MetO₂. CySO₃H. Arg
9. **Val. Thr. Pro. Gly. Ala**
Val. (Thr, Pro, Gly, Ala)
Thr. (Pro, Gly)
Thr. Pro
(Pro, Gly)
10. **Ser. Asp. Arg**
Ser. (Asp, Arg)†
Ser. Asp
Asp. Arg
11. **Lys. Phe. Glu. Gly**
Lys. (Phe, Glu)
(Phe, Glu, Gly)
Phe. Glu
12. **Arg. CySO₃H. Glu. Ala**
Arg. (CySO₃H, Glu)
CySO₃H. (Glu, Ala)
Arg. CySO₃H. Glu. Ala
CySO₃H. Glu
13. **Ser. Phe. Asp. Glu**
Ser. (Phe, Asp)
Phe. (Asp, Glu)
Phe. Asp
14. **Thr. Asp. Arg. Arg**
Thr. (Asp, Arg, Arg)
Thr. (Asp, Arg)†
Thr. Asp
Asp. Arg
15. **Thr. Gly. Asp. Val**
Thr. (Gly, Asp, Val)
Thr. (Gly, Asp)†
Thr. Gly
16. **Ser. Val. CySO₃H. Ala. Lys. Gly**
Ser. (Val. CySO₃H, Ala, Lys)?
Val. (CySO₃H, Ala)†
CySO₃H. (Ala, Lys)
Ala. (Lys, Gly)
Ser. Val. CySO₃H. Ala. Lys. Gly
Val. CySO₃H. Ala. Lys
17. **Gly. CySO₃H. Asp**
Gly. (CySO₃H, Asp)†
CySO₃H. Asp
18. **Leu. Gly. Ala. Val**
Leu. (Gly, Ala, Val)†
19. **Asp. Ileu. Pro. CySO₃H**
Asp. (Ileu, Pro, CySO₃H)
Leu or Ileu } (Pro, CySO₃H)§
Ileu. Pro
Asp. Ileu
20. **Arg. CySO₃H. Lys. Gly**
Arg. (CySO₃H, Lys, Gly)
Arg. (CySO₃H, Lys)
CySO₃H. (Lys, Gly)
Arg. CySO₃H. Lys. Gly
CySO₃H. Lys
21. **Ser. Val. Asp. CySO₃H. Ala**
Ser. (Val, Asp, CySO₃H)
Val. (Asp, CySO₃H, Ala)
Ser. (Val, Asp)
Ser. Val. CySO₃H. Ala
Val. Asp
22. **Asp. Leu. CySO₃H. Asp**
Asp. (Leu, CySO₃H)
Leu. (CySO₃H, Asp)
Leu. CySO₃H
Asp. Leu. CySO₃H. Asp
23. **Arg. Asp. CySO₃H. Ileu**
Arg. (Asp, CySO₃H)
Asp. (CySO₃H, Ileu)
Arg. Asp. CySO₃H. Ileu
24. **Ser. Arg. Leu**
Ser. (Arg, Leu)
Ser. Arg
Arg. Leu
25. **Ser. Asp. CySO₃H. Arg. Leu**
Ser. (Asp, CySO₃H, Arg, Leu)
Asp. (CySO₃H, Arg, Leu)
Asp. (CySO₃H, Arg)
CySO₃H. (Arg, Leu)
Ser. Asp. Arg. Leu
CySO₃H. Arg

* Evidence has been presented for these sequences in a previous paper (Thompson, 1955).

† Evidence for these sequences was obtained by partial hydrolysis of the DNP-peptides (see Table 2).

‡ On hydrolysis of the DNP-peptide during end-group examination, a second yellow spot was observed in addition to DNP-valine running a little faster than DNP-cysteic acid in *tert*-pentanol-pH 6 buffer but readily separated in *n*-butanol-acetic acid. On hydrolysis it yielded DNP-valine and cysteic acid and was thus DNP-Val. CySO₃H. This establishes the sequence Val. CySO₃H. Ala in the tripeptide.

§ Evidence for this peptide was obtained in the experiments using displacement chromatography (Thompson, 1955) but, as insufficient material was available for confirmation, it was not listed as a definite peptide.

|| The sequence Ser. Phe. Asp has also been established by Acher, Laurila, Thauereaux & Fromageot (1954).

which structural examination would require. The previous examination of a partial acid hydrolysate does have the advantage of aiding in the interpretation of the results on longer-chain peptides.

The sequences of amino acids in lysozyme deduced from peptides from partial acid hydrolysis identified in this and earlier work (Thompson, 1955) are listed in Table 6, and beneath them the structure of the peptides from which evidence for the sequence is derived. Peptides not used in compiling these sequences suggested a number of additional sequences. They cannot be logically deduced with all other possibilities of combination eliminated but are statistically probable. Further confirmation will be required. They are listed below with the peptides providing evidence for them:

Ileu.Ser.Leu.Gly
 Ileu.(Ser, Leu, Gly)
 Ser.(Leu, Gly)
 Ser.Leu

This sequence perhaps adjoins sequence 18 of Table 6 to give

Ileu.Ser.Leu.Gly.Ala.Val
Asp.Arg.Gly.Leu
 Asp.(Arg, Gly, Leu)
 Asp.(Arg, Gly)
 Arg.Gly
 Asp.Arg Gly.Leu
CySO₃H.Asp.Gly.Arg
 CySO₃H.(Asp, Gly, Arg)
 CySO₃H.(Asp, Gly)
 Asp.(Gly, Arg)
 CySO₃H.Asp
 Asp.Gly

This sequence probably adjoins either sequence 17 or 22 of Table 6.

Ileu.Asp.Ala
 Ileu.(Asp, Ala)
 Ileu.Asp
 Asp.Ala
Asp.Ileu.Thr.Ala
 Asp.(Ileu, Thr, Ala)
 Ileu.(Thr, Ala)
 Asp.Ileu Thr, Ala
Ser.Ala.Gly.Asp
 Ser.(Ala, Gly, Asp)
 Ser.(Ala, Gly)
 Ser.Ala

Two other sequences in lysozyme, the *N*-terminal one, Lys.Val.Phe.Gly.Arg (Schroeder, 1952; Acher *et al.* 1954) and Arg.His.Lys (Acher, Thureaux, Crocker & Fromageot, 1952) have been established.

It is of interest to assess at this stage the extent to which these sequences (1-25, Table 6) contribute to the knowledge of the lysozyme molecule as a whole. It may be observed that the sequences

around the two proline and two methionine residues have been established; also three of the three or four phenylalanine residues, the one histidine residue (Acher *et al.* 1952), the ten cysteic acid residues, five of the six lysine residues, and the seven threonine residues. Sequences containing glutamic acid appear to involve seven glutamic acid residues, though the analytical figure for glutamic acid in lysozyme is uncertain (Thompson, 1955).

In addition to the tripeptide and higher sequences deduced above, a number of dipeptides have been isolated but not used in compiling these sequences (1-25, Table 6). Of course they also represent bonds in the lysozyme molecule and are listed below:

Ala.Ala	Gly.Leu	Leu.Leu
Arg.Asp	Ileu.Arg	Ser.Ala
Arg.Gly	Ileu.Asp	Ser.Leu
Asp.Asp	Ileu.Val	Val.Ala

In addition, the dipeptides Tyr.Gly and Gly.Tyr have been reported by Acher *et al.* (1952).

It is of interest to estimate the proportion of the lysozyme molecule which these sequences represent. A very rough assessment may be made by counting the number of bonds allocated in the sequences. The figure is of course high, as condensation of two sequences together may reduce the number of bonds. However, as no alternative method can be readily applied, an approximate estimate of the present position is given in this way. Of the number of bonds in lysozyme, which is of the order of 125-130, 79 have been deduced from the above experiments (see Table 6), a further 6 in the sequences established by other workers, another 14 in the above dipeptides, making a total of 99 bonds allocated. As tryptophan is destroyed during partial acid hydrolysis, one would not expect to locate the 16 bonds involving these residues. The same applies to the 14 bonds involving the amino groups of serine and threonine (9 Ser+7 Thr-2 intact Thr bonds), which are preferentially destroyed by acid. Thus about 30 bonds cannot be determined in this hydrolysate. These factors considered, it may be seen that the extent of our knowledge of the lysozyme molecule is approaching completeness in so far as examination of a partial acid hydrolysate can provide such information.

A noticeable blank in the results of the elution experiments is the absence of peptides containing tyrosine [except (Asp, Tyr) in fraction AH]. The reason for failure to detect such peptides is not known.

It was previously reported (Thompson, 1954) that the sequence Gly.Phe.Glu.Asp.Ileu occurs in lysozyme. The portion Gly.Phe.Glu was deduced

on the basis of peptides Phe.Glu and (Gly, Phe, Glu), and the peptide Phe.(Asp, Glu) which was thought to belong to the same sequence, leading to the tetrapeptide sequence Gly.Phe.Glu.Asp. The peptide Glu.(Asp, Ileu) extended this to Gly.Phe.-Glu.Asp.Ileu. However, on subsequent isolation of the peptide Lys.(Phe, Glu), and since lysozyme contains probably three phenylalanine residues, it appears that the glycine of (Phe, Glu, Gly) follows the glutamic acid residue, yielding Lys.Phe.Glu.-Gly and the Phe.(Asp, Glu) forms part of the sequence Ser.Phe.Asp.Glu. Evidence is insufficient to indicate whether the next residues are -Asp.Ileu (from Glu.Asp.Ileu, sequence 5 of Table 6) or -Ala (from Asp.Glu.Ala, sequence 3 of Table 6).

Five of the six sequences containing lysine have been established (Table 6). The sixth involves the peptide Val.(Ileu, Lys), but the order of the isoleucine and lysine residues cannot be fixed on available evidence. If the order is Val.Ileu.Lys then the sequence Lys.Phe.Glu.Gly (sequence 11 of Table 6) may follow either this sequence or the sequence Arg.His.Lys established by Acher *et al.* (1952). If the order is Val.Lys.Ileu then it must follow the Arg.His.Lys sequence, giving the extended sequence Arg.His.Lys.Phe.Glu.Gly.

On inspection of the sequences thus established, no regularity in the structure of this protein has become apparent. The arrangement of amino acids follows very much a random distribution (except that the recurrence of the dipeptide sequence Glu.Ala on four occasions would appear to be more than chance distribution would yield).

Although the full structure of the molecule is not known, it seems unlikely that amino acid sequence determination alone will provide the clue as to the requirements of the enzyme for its activity. Such an understanding will only be reached when further studies, testing the effect of modification of the molecule on its biological activity, have been carried out. As a basis for this, knowledge of the sequence of amino acids in the molecule is of prime importance.

SUMMARY

1. A partial acid hydrolysate of egg-white lysozyme has been fractionated by gradient elution

chromatography on sulphonated-polystyrene resin with ammonium formate and acetate buffers.

2. A large number of peptides have been isolated by chromatography of these fractions on paper, and their structures determined.

3. A number of sequences in the lysozyme molecule have been deduced from these results.

The author wishes to thank Dr E. L. Smith for his hospitality and interest in this work during the portion of these investigations carried out at the College of Medicine, University of Utah, which was aided by grants from the United States Public Health Service and the Rockefeller Foundation. It is a pleasure to acknowledge several helpful discussions with Dr S. Moore and Dr W. H. Stein of the Rockefeller Institute for Medical Research, New York, on aspects of elution chromatography.

REFERENCES

- Acher, R., Chauvet, J., Crocker, C., Laurila, U. R., Thureaux, J. & Fromageot, C. (1954). *Bull. Soc. chim. Biol., Paris*, **36**, 167.
- Acher, R., Laurila, U. R., Thureaux, J. & Fromageot, C. (1954). *Biochim. biophys. Acta*, **14**, 151.
- Acher, R., Thureaux, J., Crocker, C. & Fromageot, C. (1952). *Biochim. biophys. Acta*, **9**, 339.
- Dowmont, Y. P. & Fruton, J. S. (1952). *J. biol. Chem.* **197**, 271.
- Hirs, C. W. H., Moore, S. & Stein, W. H. (1952). *J. biol. Chem.* **195**, 669.
- Kay, L. M. & Schroeder, W. A. (1954). *J. Amer. chem. Soc.* **76**, 3564.
- Moore, S. & Stein, W. H. (1948). *J. biol. Chem.* **176**, 367.
- Moore, S. & Stein, W. H. (1951). *J. biol. Chem.* **192**, 663.
- Partridge, S. M. & Brimley, R. C. (1952). *Biochem. J.* **51**, 628.
- Perrone, J. C. (1951). *Nature, Lond.*, **167**, 513.
- Sanger, F. (1945). *Biochem. J.* **39**, 507.
- Sanger, F. (1949). *Biochem. J.* **45**, 563.
- Sanger, F. & Thompson, E. O. P. (1953a). *Biochem. J.* **53**, 353.
- Sanger, F. & Thompson, E. O. P. (1953b). *Biochem. J.* **53**, 366.
- Sanger, F. & Tuppy, H. (1951). *Biochem. J.* **49**, 481.
- Schroeder, W. A. (1952). *J. Amer. chem. Soc.* **74**, 5118.
- Schroeder, W. A., Kay, L. M., Le Gette, J., Honnen, L. & Green, L. C. (1954). *J. Amer. chem. Soc.* **76**, 3556.
- Stein, W. H. (1953). *J. biol. Chem.* **201**, 45.
- Thompson, A. R. (1954). *Biochim. biophys. Acta*, **14**, 581.
- Thompson, A. R. (1955). *Biochem. J.* **60**, 507.