REFERENCES

Chaikoff, I. L. & Emanuel, C. F. (1957). *Biochim. biophys. Acta*, **24**, 254.
Deul, D., Slater, E. C. & Veldstra, L. (1958). *Biochim. biophys. Acta*, **27**, 133.
Lagnado, J. R. & Sourkes, T. L. (1957). *Canad. J. Biochem. Physiol.* **34**, 1095.
Parkes, A. S. (1946). *J. Hyg., Camb.*, **44**, 491.
Shelton, E. & Rice, M. E. (1957). *J. nat. Cancer Inst.* **18**, 117.
Slater, E. C. (1949). *Biochem. J.* **45**, 14.
Slater, T. F. (1959*a*). *Nature, Lond.*, **183**, 50.
Slater, T. F. (1959*b*). *Biochem. J.* **73**, 314.
Slater, T. F. & Planterose, D. N. (1960). *Biochem. J.* **74**, 584.

Amino Acid Sequence in Mycobacillin

By S. K. MAJUMDAR and S. K. BOSE
Department of Applied Chemistry, University College of Science and Technology, Calcutta, India

(*Received* 20 *July* 1959)

It has been observed from the study of physicochemical properties of mycobacillin that it is a polypeptide consisting of seven different amino acids (Majumdar & Bose, 1958). Studies on amino acid composition of mycobacillin show that the molecule is composed of five aspartic acid residues, two glutamic acid and two tyrosine residues and one residue each of serine, alanine, leucine and proline. The problem is how the amino acids are arranged in the mycobacillin molecule. The studies of Sanger and his collaborators (Sanger & Tuppy, 1951*a*, *b*; Sanger & Thompson, 1953) demonstrated the details of covalent structure of insulin by the application of fairly simple methods of separation and analysis, e.g. paper chromatography, paper electrophoresis and end-group analysis. This paper describes work done to obtain small peptides by a suitable method of hydrolysis followed by fractionation and identification of amino acid residues in each of the peptide fractions, and determination of *N*-terminal amino acids of the constituent peptides to discover the amino acid sequence of the mycobacillin molecule.

EXPERIMENTAL

Partial hydrolysis of mycobacillin. A sample of 30 mg. of mycobacillin was hydrolysed with 3 ml. of 11·4N-HCl for 7 days at 37° in sealed tubes. The hydrolysate was then freed from HCl by evaporating the sample to dryness *in vacuo* below 40° and the process was repeated thrice after addition of fresh water. Finally the hydrolysate was dissolved in 2 ml. of water.

Fractionation of peptide mixtures. After partial hydrolysis the next step was the separation of the complex mixture of peptides. By the application of descending two-dimensional paper chromatography with water-saturated phenol as the first solvent and butanol–acetic acid–water (4:1:1) as second, it was possible to have a satisfactory fractionation of the complex mixture of peptides. About 10 μl. of the sample was spotted at one corner (5 cm. away from both the edges) of the Whatman no. 1 filter paper (56 cm. × 46 cm.). The peptide mixture was allowed to run parallel in six pairs. One of each pair was developed with 0·2% of ninhydrin in acetone and used as a guide to locate peptide areas on the other undeveloped chromatogram. With the aid of the guide chromatogram, individual peptide areas were located, cut out and eluted with water (Consden, Gordon & Martin, 1947). The solution of each eluted peptide was evaporated *in vacuo* to give a residue.

Identification of constituent amino acids of a peptide. In the identification of constituent amino acids, a portion of the peptide residue eluted from the undeveloped chromatograms was completely hydrolysed in 0·5 ml. of 5·7N-HCl at 110° for 18 hr. The hydrolysate was then taken up in water and evaporated under reduced pressure to remove HCl, the operation being repeated twice. The final residue was dissolved in 0·05 ml. of 10% propan-2-ol. The solution so obtained was analysed chromatographically (Sanger & Tuppy, 1951*a*) and the constituent amino acids were identified.

Identification of N*-terminal amino acid of a peptide.* The *N*-terminal amino acid residue in each peptide from the partial acid hydrolysate of mycobacillin was determined in two ways: (*a*) by deamination with nitrosyl chloride followed by hydrolysis with hydrochloric acid; (*b*) by the dinitrophenyl (DNP) method of Sanger (1945, 1949).

(*a*) Deamination with nitrosyl chloride. In the deamination of each peptide, the method as followed by Consden *et al.* (1947) was used. A portion of the eluted peptide was dissolved in 0·5 ml. of 5·7N-HCl and deaminated with nitrosyl chloride at 30° for 30 min. The deaminated product was dried over solid KOH *in vacuo*, dissolved in water and evaporated *in vacuo* before complete hydrolysis. Finally, identification of amino acids remaining after hydrolysis of deaminated peptides with 5·7N-HCl at 110° for 18 hr. indicated the amino acid destroyed at the *N*-terminal end of the original peptide.

(*b*) Dinitrophenyl method. The DNP method of Sanger was next tried for amino end-group analysis of peptides. In practice, the peptides eluted from four paper chromatograms were dissolved in 1 ml. of water. To the solution was added 100 mg. of $NaHCO_3$ and then a solution of 0·1 ml. of dinitrofluorobenzene in 2 ml. of absolute ethanol (Schroeder & Legette, 1953). The DNP-peptide residue thus formed

was hydrolysed with 0·2 ml. of 5·7 N-HCl at 105° for 16 hr. to give the DNP-amino acid, which was then extracted with peroxide-free ether. DNP-amino acids were next identified by ascending one-dimensional paper chromatography, with (1) the toluene–pyridine–ethylene chlorohydrin–aq. 0·8 N-NH_3 soln. (5:1:3:3) system of Biserte & Osteux (1951) and (2) the *tert.*-amyl alcohol–phthalate system of Blackburn & Lowther (1951). In each paper chromatogram, control samples of DNP-amino acids, synthesized previously by the method of Sanger (1945), were also allowed to run simultaneously.

Identification of N-terminal amino acids in mycobacillin. A separate experiment was carried out to identify *N*-terminal amino acids in mycobacillin by the DNP method of Sanger as described previously, and also by the phenyl-*iso*thiocyanate method of Edman (1950).

Amino acids. For definition of the abbreviations used for amino acids in this paper see *Biochem. J.* (1953), **55**, 5.

Table 1. *Peptides of a partial hydrolysate of mycobacillin*

Peptide spot no. (Fig. 1)	R_F of peptides: In phenol–water	R_F of peptides: In butanol–acetic acid–water (4:1:1)	Amino acids present	Concn.* of amino acids after hydrolysis of peptides: Before deamination	Concn.* of amino acids after hydrolysis of peptides: After deamination	DNP-amino acids identified	Suggested structure†
1	0·18	0·11	Asp	× × × ×	× ×	Asp	Asp.Glu
			Glu	× ×	× ×		
2	0·26	0·12	Asp	×	×	Ser	Ser.Asp
			Ser	×	—		
3	0·26	0·21	Asp	× ×	× ×		
			Glu	× × × ×	×	Glu	Glu.(Asp, Ala)
			Ala	× ×	× ×		
4	0·30	0·15	Asp	×	×	Glu	Glu.Asp
			Glu	×	?		
5	0·35	0·28	Asp	× × ×	×	Asp	Asp.Ala
			Ala	× × ×	× ×		
6	0·36	0·34	Asp	× ×	?	Asp	Asp.Tyr
			Tyr	× ×	×		
7	0·40	0·40	Asp	×	×	Ala	Ala.Asp
			Ala	×	—		
8	0·48	0·30	Asp	× ×	× ×		
			Glu	× ×	× ×	Asp	Asp.(Glu, Tyr)
			Tyr	× ×	× ×		
9	0·50	0·41	Glu	×	?	Glu	Glu.Tyr
			Tyr	×	×		
10	0·52	0·36	Asp	× × ×	× × ×		
			Glu	× × ×	× × ×	Leu	Leu.(Asp, Glu)
			Leu	× × ×	×		
11	0·60	0·42	Asp	× × ×	× ×		
			Glu	× × ×	× ×	Leu	Leu.(Asp, Glu, Ala)
			Ala	× × ×	× ×		
			Leu	× × ×	?		
12	0·74	0·47	Asp	× × ×	×		
			Ser	× × ×	× ×	Asp	Asp.(Ser, Tyr)
			Tyr	× × ×	× ×		
13	0·76	0·56	Asp	× × ×	× ×		
			Glu	× × ×	× ×	Ser	Ser.(Asp, Glu, Leu)
			Ser	× × ×	—		
			Leu	× × ×	× ×		
14	0·77	0·67	Asp	× × ×	× ×		
			Ser	× × ×	—	Ser	Ser.(Asp, Leu)
			Leu	× × ×	× ×		
15	0·80	0·34	Asp	× ×	× ×		
			Ala	× ×	—	Ala	Ala.(Asp, Pro)
			Pro	× ×	× ×		

* The number of crosses indicates the relative intensities of colour given by each amino acid spot with ninhydrin; ?, very weak spot.

† In referring to peptides of known structure, the abbreviations for the residues are joined by a point and where the arrangements of residues are unknown the residues are included in parentheses.

RESULTS

Peptides of a partial hydrolysate of mycobacillin. Results showing the *N*-terminal amino acids of different peptides (Fig. 1) of a partial hydrolysate of mycobacillin are given in Table 1.

It appears from Table 1 and Fig. 1 that of the 15 peptides isolated, there are seven dipeptides, six tripeptides and two tetrapeptides. Identification of *N*-terminal amino acids by the deamination method has been possible for 10 out of the 15 peptides studied, and in peptides 1, 3, 5, 8 and 12 no conclusion can be drawn about the identity of *N*-terminal amino acid because of incomplete deamination. The DNP method, on the other hand, completely establishes the identity of *N*-terminal residues of all the 15 peptides. Both of these methods give identical results in all cases where it has been possible to establish the identity of *N*-terminal amino acids. These findings lead to a possible sequence of amino acids in these peptides as indicated in the last column of Table 1. Aspartic acid and glutamic acid are present in a higher concentration respectively in peptides 1 and 3, although they are not repeated in the sequence, as indicated by their absence in the aqueous layer after ethereal extraction of DNP-peptide hydrolysate. This may be taken to mean that peptides 1 and 3 are not pure and contain free aspartic acid and glutamic acid respectively.

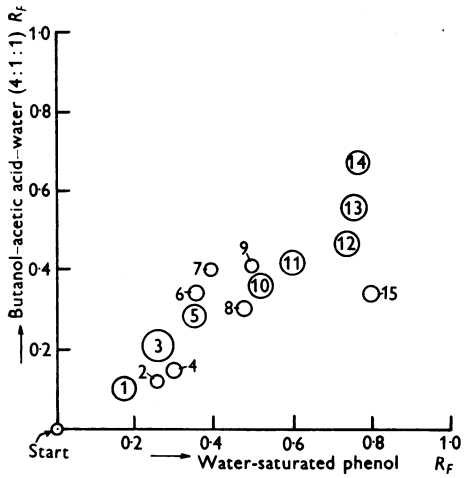


Fig. 1. Two-dimensional chromatograms of peptides in a partial acid hydrolysate of mycobacillin. Intensities of spots with ninhydrin are represented approximately by the areas of the circles.

The knowledge of the amino acid composition and molecular weight (about 1775) of mycobacillin, coupled with the results of determination of *N*-terminal amino acids in each peptide, indicates a possible sequence of amino acids in the molecule as shown in Table 2. It may be noted that mycobacillin contains no free amino group, as shown by the absence of its reaction with ninhydrin and the failure to identify *N*-terminal amino acids in the molecule by the DNP method of Sanger and the phenylthiocarbamyl method of Edman (1950).

DISCUSSION

It will be evident from Table 1 that *N*-terminal amino acids of 15 peptides of a partial acid hydrolysate of mycobacillin are definitely known by the DNP method whereas the method of deamination of peptides followed by hydrolysis has confirmed the *N*-terminal position in 10 out of the 15 peptides.

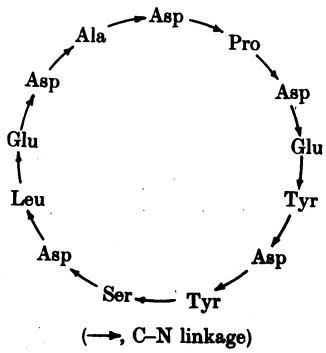
Table 2. *Sequence of amino acids in mycobacillin*

	*1.	Asp.Glu
	8.	Asp.(Glu, Tyr)
	9.	Glu.Tyr
Sequence (I)		Asp.Glu.Tyr
	6.	Asp.Tyr
	12.	Asp.(Tyr, Ser)
	2.	Ser.Asp
	14.	Ser.(Asp, Leu)
	13.	Ser.(Asp, Leu, Glu)
	10.	Leu.(Glu, Asp)
	11.	Leu.(Glu, Asp, Ala)
	3.	Glu.(Asp, Ala)
	4.	Glu.Asp
	5.	Asp.Ala
	7.	Ala.Asp
	15.	Ala.(Asp, Pro)
Sequence (II)		Asp.Tyr.Ser.Asp.Leu.Glu.Asp.Ala.Asp.Pro
		Asp.Glu.Tyr.Asp.Tyr.Ser.Asp.Leu.Glu.Asp.Ala.Asp.Pro†

* Adjacent numbers are peptide spot numbers (Table 1 and Fig. 1).

† The proof of the structure suggested depends on there being two sequences as shown in (I) and (II) above; the relationship between these is not directly proved but, since the peptide is cyclic, they can be joined only in the way shown.

Determination of the chemical structure of tri- and tetra-peptides without the application of the stepwise-degradation method or any method for determining *C*-terminal amino acids has been possible because some of the amino acids such as serine, alanine, leucine and proline are present in the proportion of one molecule each per molecule of mycobacillin. Since these amino acids occur only once in the molecule, the known structure of different dipeptides containing any one of these amino acids has been helpful in determining the possible arrangements of amino acids in tripeptides and tetrapeptides. For example, it will be evident from Table 2 that the dipeptide Ser.Asp gives the possible arrangements of amino acids in the tripeptide Ser.Asp.Leu, and so in the tetrapeptide Ser.Asp.Leu.Glu. Again, the dipeptides Glu.Asp and Asp.Ala show the arrangement of amino acids in the tripeptide Glu.Asp.Ala. Similarly, the structures of other tripeptides and of the tetrapeptide Leu.Glu.Asp.Ala have been established. This knowledge of chemical structure of di-, tri- and tetra-peptides has made it possible to fit together 15 peptide fragments from 13 amino acid residues of mycobacillin molecule into a unique sequence as depicted below. In the absence of a free α-amino group in the molecule, a cyclic structure has been deduced.



The amino acid sequence of antifungal polypeptide antibiotics from *Bacillus subtilis* has not been reported yet, except for bacillomycin, for which only a partial sequence of amino acids has been worked out (Turner, 1956).

SUMMARY

1. Mycobacillin was partially hydrolysed and the hydrolysate fractionated by two-dimensional paper chromatography to give a separation of complex mixtures of peptides.

2. Altogether 15 peptides, including seven dipeptides, six tripeptides and two tetrapeptides, were isolated.

3. The *N*-terminal residue of each peptide was determined by both the deamination method and the dinitrophenylation method.

4. As the mycobacillin molecule does not contain an *N*-terminal residue, a cyclic structure has been suggested.

The authors wish to express their thanks to Professor B. C. Guha for his kind advice and interest in this work and to the Council of Scientific and Industrial Research for the grant of a Fellowship to one of the authors (S.K.M.).

REFERENCES

Biserte, G. & Osteux, R. (1951). *Bull. Soc. Chim. biol., Paris*, **33**, 50.

Blackburn, S. & Lowther, A. G. (1951). *Biochem. J.* **48**, 126.

Consden, R., Gordon, A. H. & Martin, A. J. P. (1947). *Biochem. J.* **41**, 590.

Edman, P. (1950). *Acta chem. scand.* **4**, 283.

Majumdar, S. K. & Bose, S. K. (1958). *Nature, Lond.*, **181**, 134.

Sanger, F. (1945). *Biochem. J.* **39**, 507.

Sanger, F. (1949). *Biochem. J.* **45**, 563.

Sanger, F. & Thompson, E. O. P. (1953). *Biochem. J.* **53**, 353.

Sanger, F. & Tuppy, H. (1951*a*). *Biochem. J.* **49**, 463.

Sanger, F. & Tuppy, H. (1951*b*). *Biochem. J.* **49**, 481.

Schroeder, W. A. & Legette, J. (1953). *J. Amer. chem. Soc.* **75**, 4612.

Turner, R. A. (1956). *Arch. Biochem. Biophys.* **60**, 364.