Cleavage of Bacterial Flagellin with Cyanogen Bromide

CHEMICAL AND PHYSICAL PROPERTIES OF THE PROTEIN FRAGMENTS

By C. R. PARISH* AND G. L. ADA*

The Walter and Eliza Hall Institute of Medical Research, Royal Melbourne Hospital, Vic. 3050, Australia

(Received 28 October 1968)

1. Flagellin, isolated from the flagella of *Salmonella adelaide*, was shown by various criteria to be a pure protein. It had a molecular weight of about 40 000 and contained three methionine, six tyrosine, 11 arginine and 25 lysine residues/mol., of which 11 of the lysine residues were present as ϵ -N-methyl-lysine. 2. After treatment of flagellin with cyanogen bromide in formic acid, four main fragments (A, B, C and D) were obtained, with as many as six minor components that represented partial degradation products. The major fragments were estimated by amino acid analysis to have molecular weights of about 18 000 for fragment A, 12 000 for fragment B, 5500 for fragment C and 4500 for fragment D. Fragments A, B and D, but not fragment C, were recovered pure by gel chromatography as monitored by polyacrylamide-gel electrophoresis. 3. A complex between fragments C and D was also isolated (mol.wt. 10 000) after limited oxidation of flagellin by chloramine-T before digestion by cyanogen bromide. After oxidation essentially only two fragments were released from flagellin by cyanogen bromide: the 'C,D' complex and a presumed 'AB' fragment. 4. The sum of the amino acid analyses of fragments A and B and the 'C,D' complex gave residue values that agreed well with the amino acid composition of native flagellin. 5. Fragments A and D contained tyrosine, and ten of the 11 ϵ -N-methyl-lysine residues of the molecule were in fragment A. Reaction with [125 I]iodide at small extents of substitution showed that, in flagellin, the tyrosine residue of fragment D was more readily substituted than those of fragment A. By contrast, in polymerized flagellin, the tyrosine residues of fragment A were more readily substituted. 6. Treatment of flagellin with carboxypeptidases A and B revealed the C-terminal sequence -Leu-Leu-Leu-Arg. Arginine and leucine were released by carboxypeptidase from the 'C,D' complex but not from fragment D, indicating that fragment C was C-terminal. 7. On the basis of the results from amino acid analysis, carboxypeptidase digestion, N-terminal analysis, iodination studies and polyacrylamide-gel electrophoresis, the sequence of fragments in flagellin was considered to be B-A-D-C; in the polymer, fragment A was exposed. It is suggested that methylation of the lysine residues occurred in the organism after flagellin had polymerized.

There have been many investigations into the structure of naturally occurring antigens, but most of these studies have been concerned with the delineation of antigenic determinants and their position in the intact molecule (see review by Kabat, 1966). For example, with both tobacco-mosaic-virus protein (Young, Benjamini, Shimizu & Leung, 1966) and myoglobin (Crumpton, 1967*a,b*; Atassi & Saplin, 1968), the amino acid sequences of some of the antigenic determinants and their positions within the molecule are known. On the

other hand, comparatively little is known about some biological properties of these antigens, such as their immunogenicity, their ability to cause tolerance in animals and their behaviour *in vivo*.

The work from this laboratory in the past few years has in part been concerned with establishing the biological parameters of another antigen, flagellin, which is the basic unit of the flagellum, the organ of locomotion of certain bacteria. Flagellin is available for study in the form of flagella, as the monomeric unit flagellin of molecular weight about 40 000, or in a repolymerized form (called polymer) that can be shown to be similar in appearance and size to the flagellar particle (Ada, Nossal, Pye &

* Present address: Department of Microbiology, John Curtin School of Medical Research, Australian National University, Canberra, A.C.T. 2600, Australia.

Abbot, 1964a). In any of these forms the protein has been shown to be highly immunogenic, the injection of 1 μ g. or less in saline into adult rats causing a well-defined antibody response. On the other hand, injection of flagellin, but not the polymerized forms, into rats beginning at birth causes tolerance that is almost complete when tested by later injection of flagellin and partial when tested by later injection of the polymer. The fate of each form of the antigen has been studied in detail by trace labelling with radioactive iodine and by study of the distribution of radioactivity between and within rat tissues (for review see Ada, 1967). Finally, it has been shown (Diener & Armstrong, 1967) that, at appropriate dosages of the polymer, antibody production or tolerance can be achieved in cultures of spleen cells from CBA mice. These reasons were considered sufficient to justify a more detailed study of the chemical and physical properties of the flagellin molecule and its disposition within the flagellum or polymer.

In most previous studies, proteins have been degraded by enzymes into small peptides. We initially chose to degrade the protein in a more limited fashion and so obtain large fragments that might still retain some of the secondary and tertiary structure of the protein. Flagellin was treated with cyanogen bromide, which specifically and almost quantitatively splits the peptide chain on the carboxyl side of the methionine residues, converting the methionine into homoserine lactone. As flagellin has three methionine residues/mol., four fragments should be obtained. This paper describes the conditions of cleavage of the protein and some chemical and physical properties of the protein fragments. The following paper (Parish, Wistar & Ada, 1969) describes some of the biological properties of the fragments.

MATERIALS AND METHODS

Preparation of flagellin. *Salmonella adelaide* (strain, SW 1338; H antigen, fg; O antigen, 35) was used in this investigation. The procedure for preparing flagellin was as reported by Ada *et al.* (1964a). The flagella were removed from the bacteria by shearing in a blender and separated from the bacterial bodies by centrifugation. The flagella in the supernatant were then pelleted by high-speed centrifugation. When treated with acid, the flagella dissociated into flagellin (monomer) and acid-insoluble material. Flagellin was purified by polymerization in 15% saturated ammonium sulphate; this material was called 'polymer' and was also readily dissociated into flagellin when treated with acid. Flagellin prepared by acid dissociation of the polymer was used for all studies. The protein concentration of flagellin solutions was determined spectrophotometrically at 215 nm. $E_{1\%}^{1\text{cm}}$ at 215 nm. of flagellin was estimated to be 94.0 based on dry weight determinations.

Digestion of flagellin with cyanogen bromide. Concentrated solutions of polymer were obtained by allowing

the sedimented polymer to resuspend over several hours in small volumes of water. To 20–60 mg. of polymer in solution in water (5–50 mg./ml.) were added formic acid (100%) to give a final concentration of 70% (v/v) and 200–300 mg. of crystalline CNBr (Koch–Light Laboratories Ltd., Colnbrook, Bucks.)/100 mg. of protein. The mixture, in a tightly stoppered glass tube, was allowed to react at room temperature for 15 hr. The mixture was then dialysed for 2 hr. against water, which was changed every 30 min. The non-diffusible solution was called the CNBr digest. If the formic acid and CNBr were removed by freeze-drying instead of by dialysis, the dried digest was only slightly soluble in aqueous solvents. In practice, freeze-drying was only used to prepare samples for amino acid analysis. If the digest was to be fractionated on columns of Sephadex, the complete reaction mixture was applied directly to the column.

Chloramine-T oxidation of flagellin. Flagellin was oxidized by chloramine-T in the following manner. Approx. 30 mg. of polymer (15 mg./ml.) was dissociated into flagellin with 0.15 ml. of 1 M-HCl, and then neutralized with 0.15 ml. of 1 M-NaOH. To this solution was added 7.5 mg. of chloramine-T dissolved in 1.75 ml. of a 0.12 M-NaCl–0.02 M-sodium phosphate buffer, pH 7.3. Oxidation was carried out for 15 min. at room temperature, the reaction being halted by the addition of 0.75 ml. of potassium metabisulphite (20 mg./ml.). The reaction mixture was then dialysed overnight at 4° against water, freeze-dried, redissolved in 2 ml. of 70% formic acid and digested with CNBr in the standard manner. The CNBr digest was then fractionated by gel filtration as described below.

Fractionation of the cyanogen bromide digest by gel filtration. The CNBr digest (20–50 mg.) in 1–3 ml. of 70% formic acid was applied to a column of Sephadex (Pharmacia, Uppsala, Sweden), previously equilibrated with 50% (v/v) formic acid. The column contained an upper layer of Sephadex G-100 (4.4 cm. $\times$ 35 cm.) and a lower layer of Sephadex G-50 (4.4 cm. $\times$ 35 cm.). A fresh column of Sephadex was prepared and used every 3 weeks, as Sephadex stored in formic acid gave poor resolution. The column was eluted with 50% formic acid at room temperature. The flow rate was 10 ml./hr., and fractions (4 ml.) were collected. The extinction of the effluent solution was initially measured at 280 nm. However, since flagellin has a low content of aromatic amino acids, the protein was not readily detected at 280 nm. A more sensitive technique was to measure extinction at 215 nm. after removal of the formic acid from fractions by dialysis for 2 hr. Little loss of peptides occurred during this dialysis.

Electrophoresis. Electrophoresis on cellulose acetate strips (Gelman Instrument Co., Ann Arbor, Mich., U.S.A.) was performed in 0.045 M-veronal buffer, pH 8.6. A current of 1–2 mA was applied to each strip for 1½ hr. Strips were then stained with Amido Schwarz (G. T. Gurr, London S.W.6).

Polyacrylamide-gel disc electrophoresis was carried out at pH 8.9 (Davis, 1964) in the presence or absence of 8 M-urea (Reisfeld & Small, 1966), or at pH 2.7 in 8 M-urea (Neville, 1967) as a procedure to monitor the CNBr digests and subsequent purification of the fragments. The best resolution was obtained if a current of 1.5–2.0 mA was applied to each tube for 4–6 hr. Reagents were supplied by American Cyanamid Co., Pearl River, N.Y., U.S.A. Gels were stained for protein with a solution containing 0.25% (w/v) Amido

Schwarz, 0.5% (w/v) HgCl_2 and 7% (v/v) acetic acid. Gels were destained electrophoretically and stored in 7% (v/v) acetic acid (Davis, 1964).

Amino acid analysis. Samples (0.3–1 mg.) were hydrolysed *in vacuo* in 0.5 ml. of constant-boiling HCl (5.54 M) for 24, 48 and 72 hr. at 110° (Crestfield, Moore & Stein, 1963). Hydrolysates were redissolved in 0.3–1.0 ml. of 0.5 M-norleucine. Just before application to a Technicon amino acid analyser, hydrolysates were diluted with an equal volume of 60% (v/v) ethanediol in 0.01 M- HCl . This procedure increased the density of the sample and therefore ensured application in a narrow band to the column of the analyser. Information on the compositions of the buffers used in the analyser was obtained from Moore, Spackman & Stein (1958). ϵ -N-Methyl-lysine was separated from lysine on the 60 cm. column of the Technicon analyser, which was regenerated and developed with a Beckman pH 5.28 buffer (containing 177 mM-sodium citrate, 11 mM-phenol and 2 ml. of BRIJ 35 solution/l., titrated to pH 5.28 with 10 ml. of 6 M- HCl). The column temperature was 52.5° . Each analysis took approximately 2 hr. and the separation achieved is shown in Fig. 1; values for tyrosine, phenylalanine and ammonia were also obtained from these analyses. Authentic ϵ -N-methyl-lysine was received from Professor N. L. Benoiton (University of Ottawa), to whom we express our thanks. The colour values for lysine and ϵ -N-methyl-lysine were very similar, 0.825 and 0.823 respectively.

More accurate estimates of the methionine content were obtained from performic acid-oxidized flagellin. Approximately 0.1 ml. of flagellin (1 mg.) was mixed with 0.3 ml. of ice-cold freshly prepared performic acid (100% formic acid- H_2O_2 , 9:1, v/v) and left to react for 1 hr. at 4° . The reaction was stopped by the addition of 3.0 ml. of ice-cold water, and the solution was then freeze-dried and hydrolysed in the standard manner (see above). Separation of methionine sulphone was obtained by regenerating and developing the 60 cm. column of the Technicon amino acid analyser with a pH 2.25 buffer (containing 10 mM-sodium citrate, 5 mM- NaOH , 29 mM-thiodiglycol and 2 ml. of BRIJ 35 solution/l.,

titrated to pH 2.25 with 6 M- HCl) at 75° (see Technicon Instrument Co. 1966). Hydrolysates of performic acid-oxidized flagellin were applied to the analyser in the pH 2.25 buffer plus an equal volume of 60% (v/v) ethanediol. Under these conditions methionine sulphone was eluted within $1\frac{1}{2}$ hr., appearing just before aspartic acid.

The homoserine content of the CNBr fragments was determined by converting all homoserine lactone into homoserine (Ambler, 1965). Hydrolysates were treated with pyridine-acetate buffer, pH 6.5, at 105° for 1 hr., dried by rotary evaporation, redissolved in 30% (v/v) ethanediol in 0.01 M- HCl and applied to the column of the analyser. Best resolution of homoserine was achieved by regenerating and developing the column of the analyser (column temperature 52.5°) with the standard Beckman pH 3.28 buffer (containing 65 mM-sodium citrate, 128 mM- HCl , 48 mM-thiodiglycol, 11 mM-phenol and 2 ml. of BRIJ 35 solution/l., adjusted to pH 2.90 with 5.54 M- HCl). By using this system homoserine was well resolved from glutamic acid, but the threonine-serine separation was impaired.

Digestion of flagellin and fragments with carboxypeptidases A and B. Flagellin and its fragments were digested with carboxypeptidase A or B (or both) (Sigma Chemical Co., St Louis, Mo., U.S.A.) under the following conditions. Enzymes were added at $1\mu\text{g.}$ of carboxypeptidase A/5 nmol. of protein and $1\mu\text{g.}$ of carboxypeptidase B/nmol. of protein. Digestions were performed at 30° in 0.05 M-sodium phosphate buffer, pH 7.5, 25 nmol. protein samples being taken at times ranging from 15 min. to 4 hr. The reaction was stopped by adjusting the pH to 2 with dil. HCl , and the complete mixture was then applied to the amino acid analyser.

Optical rotatory dispersion. Measurements were made at 20° with a Perkin-Elmer spectropolarimeter, fitted with a 1 ml. quartz cell with a light-path of 1 dm. Protein concentrations of 2.5–3 mg./ml. were used and the optical rotation was measured at 365, 436, 546 and 578 nm. Helical content was calculated (Moffit & Yang, 1956) by assuming that λ_0 was 212 nm. and that b_0 (100% helix) was -630 . Three protein solutions were studied: flagellin, the CNBr digest of flagellin, and flagellin in 70% formic acid. In all cases blank readings were made on the appropriate solvents.

Polymerization of flagellin. Polymerization was followed in a spectrophotometer by measuring the increase in extinction at 250 nm. Extinctions were measured in 0.5 cm.-light-path quartz cuvettes against solvent blanks. Solutions containing approx. 2 mg. of flagellin/ml. were polymerized by adding saturated ammonium sulphate to give a final concentration of 15% saturation. The extinction was measured before and immediately after the addition of the ammonium sulphate, and then at intervals for the next 24 hr. A flagellin control was also included, to which water was added instead of ammonium sulphate. After 24 hr. the degree of polymerization was also checked by centrifugation of the solution at 70000g for 40 min. and the amount of non-polymerized protein in the supernatant was measured at 215 nm.

Iodination. Flagella, the polymer, flagellin and fragments from the CNBr digest were labelled with a mixture of [^{125}I]iodide (IMS3; The Radiochemical Centre, Amersham, Bucks.) and carrier iodide. The amount of the carrier iodide used was varied according to the required degree of substitution of tyrosine in the protein. The direct-oxidation technique (Greenwood, Hunter & Glover, 1963; Ada, Nossal

Fig. 1. Separation of a mixture of authentic lysine (30 nmol.) and ϵ -N-methyl-lysine (40 nmol.) at 52.5° on the 60 cm. column of a Technicon amino acid analyser, with a pH 5.28 buffer and the Technicon C-2 preparation of ion-exchange beads. Experimental details are given in the text.

& Pye, 1964b) was followed with the following modifications. For flagella and polymer, 80 μ g. (in 20 μ l. of 0.12M-NaCl-0.02M sodium phosphate buffer, pH 7.3) was added to 3 μ l. of mixed [125 I]iodide and carrier iodide and 2 μ l. of chloramine-T solution (14 μ mole). After 3 min., 2 μ l. of water containing an excess of potassium metabisulphite was added followed by 2 μ l. of 0.1M-carrier iodide, and the mixture was resolved by chromatography on a column (0.5 cm. $\times$ 8 cm.) of Sephadex G-100 eluted with 0.12M-NaCl-0.02M-sodium phosphate buffer, pH 7.3. The fragments were labelled under similar conditions but the mixtures were resolved on columns of similar size of Sephadex G-15 eluted with the same solvent. For flagellin one-hundredth of the above concentrations of oxidizing and reducing agents was used, but the time of oxidation was extended to 30 min. (McConahey & Dixon, 1966).

RESULTS

Characterization of flagellin

Electrophoretic behaviour. Electrophoresis of flagellin on cellulose acetate strips at pH 8.9 showed a single protein band, which migrated with a mobility between those of rat β - and γ -globulins. When it was examined by polyacrylamide-gel electrophoresis at pH 2.7 or 8.9 either in the presence or in the absence of 8M-urea, a single protein band was observed (Fig. 2a). Flagellin treated with 70% formic acid at room temperature

for 16 hr. also gave a single band on polyacrylamide-gel electrophoresis at pH 2.7 in 8M-urea.

Ultracentrifugal behaviour. Both flagellin and fragment A (2–5 mg./ml.) were examined by sedimentation analysis in the Spinco analytical ultracentrifuge, in a 0.1M-sodium chloride-0.01M-sodium phosphate buffer, pH 7.0. The solutions were spun at 59780 rev./min. for 240 min. at 20°. Both showed single symmetrical protein peaks with uncorrected $S_{20,w}$ values 2.24 and 2.07s respectively (Ada *et al.* 1964a).

Amino acid analysis. Flagellin contained approximately 380 amino acid residues/mol., was rich in glutamic acid, aspartic acid and alanine and poor in methionine, proline, tyrosine and phenylalanine, and lacked cysteine and histidine (Table 1, column 5). Others (e.g. McDonough, 1965) have shown tryptophan to be absent from flagellin. Of the 26 lysine residues/mol., 12 were present as ϵ -N-methyl-lysine. A low value for methionine (2.26 residues/mol.) was obtained after the standard acid hydrolysis of flagellin. After performic acid oxidation a value of 2.86 was obtained, and this is consistent with there being three methionine residues in a flagellin molecule. Based on the values for proline, methionine and tyrosine, a mol.wt. of 40000 for flagellin is reasonable and agrees with the value found by McDonough (1965).

Fig. 2. Polyacrylamide-gel electrophoresis of flagellin before and after treatment with CNBr. The three buffers used were (from left to right in *a* and *b*) 0.05M-tris-0.038M-glycine, pH 8.9; 0.42M-tris-0.46M-glycine, pH 8.9, in 8M-urea; 7.5mm-KOH-0.58M-acetic acid, pH 2.7, in 8M-urea. O, Origin; M, dye marker. (*a*) Flagellin, showing in each case a single band. (*b*) CNBr-treated flagellin, showing the best resolution in the pH 2.7 system. In this gel, the five upper fine bands represent partially degraded flagellin, the next five bands are fragment A, fragment B, a presumed 'C,D' complex, fragment C and fragment D respectively. (*c*) Comparison of the fragments arising from CNBr digestion of flagellin and chloramine-T-oxidized flagellin. From left to right the gels are a CNBr digest of flagellin, a CNBr digest of flagellin after chloramine-T oxidation (the protein bands being undegraded flagellin, a presumed 'A,B' complex, and the 'C,D' complex, in order of increasing electrophoretic mobility), and the 'C,D' complex isolated from the digest of chloramine-T-oxidized flagellin by gel filtration. All gels were run in 7.5mm-KOH-0.58M-acetic acid, pH 2.7, in 8M-urea.

Table 1. *Amino acid analysis of flagellin and of fragments obtained from flagellin by cyanogen bromide digestion*

The values are given as residues/mol., assuming the molecular weight of flagellin to be 40000, that of fragment A to be 18000, that of fragment B to be 12000, that of the 'C,D' complex to be 10000 and that of fragment D to be 4500. The molecular weight of fragment C was calculated, by subtraction, to be 5500. Where not otherwise indicated, each value is the average of analyses of samples hydrolysed for 24, 48 and 72 hr. Values corrected to the nearest integer are given in parentheses.

Residue	Amino acid composition						
	Fragment A	Fragment B	'C,D' complex	Total A+B+'C,D'	Flagellin	Fragment D	Fragment C (C,D-D)
Cysteic acid	0.00	0.00	0.00	0.00	0.00	0.00	0
Aspartic acid	31.9 (32)	21.3 (21)	14.6 (15)	67.8 (68)	66.5 (67)	7.87 (8)	6.7 (7)
*Threonine	21.9 (22)	6.82 (7)	6.80 (7)	35.5 (36)	36.2 (36)	3.11 (3)	3.7 (4)
*Serine	8.10 (8)	14.8 (15)	12.1 (12)	35.0 (35)	34.7 (35)	5.88 (6)	6.2 (6)
Glutamic acid	11.1 (11)	16.6 (17)	9.09 (9)	36.8 (37)	36.8 (37)	2.17 (2)	6.9 (7)
Proline	1.99 (2)	0.00	1.98 (2)	3.97 (4)	3.79 (4)	*0.67 (1)	1.3 (1)
Glycine	*18.0 (18)	*5.10 (5)	4.13 (4)	27.2 (27)	26.9 (27)	1.83 (2)	2.3 (2)
*Alanine	22.0 (22)	8.78 (9)	13.9 (14)	44.7 (45)	44.0 (44)	6.21 (6)	7.7 (8)
†Valine	16.1 (16)	4.86 (5)	6.80 (7)	27.8 (28)	29.1 (29)	2.75 (3)	4.0 (4)
Cysteine	0.00	0.00	0.00	0.00	0.00	0.00	0
Isoleucine	†7.90 (8)	8.06 (8)	†5.70 (6)	21.7 (22)	†22.1 (22)	2.91 (3)	2.8 (3)
†Leucine	7.98 (8)	11.0 (11)	9.02 (9)	28.0 (28)	29.2 (29)	3.07 (3)	5.9 (6)
Tyrosine	4.84 (5)	0.00	0.95 (1)	5.79 (6)	6.14 (6)	*0.60 (1)	0.35 (0)
Phenylalanine	4.86 (5)	†1.82 (2)	†1.92 (2)	8.60 (9)	†8.54 (9)	0.97 (1)	0.95 (1)
Lysine	6.83 (7)	3.88 (4)	2.87 (3)	13.6 (14)	13.9 (14)	0.97 (1)	1.9 (2)
ε-N-Methyl-lysine	10.2 (10)	0.00	1.13 (1)	11.3 (11)	12.1 (12)	0.97 (1)	0.16 (0)
Histidine	0.00	0.00	0.00	0.00	0.00	0.00	0
Arginine	0.86 (1)	5.66 (6)	3.93 (4)	10.45 (11)	10.2 (10)	1.93 (2)	2.0 (2)
Tryptophan	0.00	0.00	0.00	0.00	0.00	0.00	0
Homoserine	0.81 (1)	0.86 (1)	0.00	1.67 (2)	0.00	0.53 (1)	0
Methionine	0.00	0.00	†0.81 (1)	0.81 (1)	†2.86 (3)	0.00	0
Total	176	111	97	384	384	44	53

* Extrapolated to zero time.

† Values plotted against the reciprocal of the time of hydrolysis and extrapolated to infinite time.

‡ Obtained by performic acid oxidation.

Cleavage of flagellin by cyanogen bromide

As methionine is a hydrophobic amino acid and flagellin contains only three methionine residues/mol., it was likely that these residues would be sterically masked so that the reaction of flagellin with cyanogen bromide would need to be carried out in a strong dissociating agent. Moreover, the cyanogen bromide reaction requires acidic conditions (Hofmann, 1964). The reaction of flagellin with cyanogen bromide in dilute hydrochloric acid (0.1–0.3 M) was incomplete, and at higher concentrations of hydrochloric acid the flagellin was irreversibly denatured and was precipitated from solution. Similarly, cleavage was incomplete in 0.1 M-hydrochloric acid–8 M-urea. Formic acid (70%) was found to be more suitable, probably because it is a strong dissociating agent and also maintains a low pH. Preliminary optical-rotatory-dispersion measurements suggested in fact that 70% formic acid did alter the conformation of the flagellin molecule. Values in Table 2 for flagellin, for flagellin in 70%

Table 2. *Optical-rotatory-dispersion measurements*

Values of a_0 , b_0 and helical content were calculated by the method of Moffit & Yang (1956), by assuming that λ_0 was 212 nm. and b_0 (100% helix) was –630.

Sample	a_0	b_0	% of right-handed α -helix
Flagellin in water	–680	–97	15
Flagellin in 70% formic acid	–427	–265	42
CNBr digest of flagellin in water	–488	–75	12

formic acid and for the digest showed that flagellin and the digest had similar helical contents. In 70% formic acid, however, the helical content of flagellin increased from about 15% to about 40%.

Cleavage by cyanogen bromide was monitored by polyacrylamide-gel electrophoresis (see below). Initially the reaction mixture was left for 18 hr., but

subsequent work showed that the reaction was complete within 30 min. Varying the formic acid concentration from 50% to 90% or increasing the cyanogen bromide concentration tenfold appeared to yield the same final products.

Products from cyanogen bromide digestion of flagellin

After cyanogen bromide digestion of flagellin in 70% formic acid for 18 hr., amino acid analysis revealed a complete disappearance of the methionine residues (Fig. 3) even when 2–3 times the usual amount of hydrolysate was analysed. Assuming flagellin had a mol.wt. 40000, 2.10 residues of homoserine/mol. were detected in the cyanogen bromide digest. In all other respects, the amino acid analyses of the digest and of flagellin were similar.

The cyanogen bromide digest was analysed most reliably by polyacrylamide-gel electrophoresis in a pH 2.7 buffer containing 8 M-urea. Under these conditions very little intact flagellin was detected. Ten protein bands appeared (Fig. 2b), four of which were subsequently found to be single cyanogen bromide fragments; the number of fragments expected if there were three methionine residues/flagellin molecule was four. These four bands were arbitrarily called fragments A, B, C and D in the order of their increasing mobility in the polyacrylamide gels. Five fainter bands appeared near the top of the gel. Two had a mobility similar to that of flagellin itself. Three occurred between these two bands and fragment A. These five bands were considered to be partially degraded flagellin molecules. A sixth band was present between fragments B and C and was believed to be a 'C,D' complex. Evidence for this is presented below. Gel chromatography of the digest yielded fractions that were rich in the five slow-moving components (Fig. 4). Amino acid analyses indicated that these fractions contained substantial amounts of methionine sulphone but no methionine. Methionine sulphone is not present in flagellin. The absence of methionine also indicated that methionine sulphoxide was not present, since acid hydrolysis would convert the sulphoxide back into methionine (Ray & Koshland, 1960). The results suggested that cyanogen bromide converted about 95% of the methionine residues into homoserine lactone, resulting in peptide-bond cleavage, whereas the remaining 5% of methionine residues were not susceptible to cleavage by cyanogen bromide, but were probably oxidized to methionine sulphone during acid hydrolysis. On this basis, the upper five minor bands are incompletely split peptides of flagellin, containing two, three or four of the fragments. Electrophoresis of the digest was also performed at pH 8.9, in the presence and absence of 8 M-urea.

Fig. 3. Portion of the amino acid-analysis profile of flagellin, showing the peaks for valine, methionine and isoleucine when 150 μ g. of an acid hydrolysate of flagellin was applied (—), and when 400 μ g. of an acid hydrolysate of a CNBr digest of flagellin was applied (---).

Under these conditions resolution of protein bands was very variable, the number of protein bands ranging from three to 12 (Fig. 2b). Subsequent tests revealed precipitating activity against anti-flagellin serum in all the protein bands resolved at pH 8.9 in the absence of 8 M-urea, suggesting that the 'bands' were aggregated material and not individual components.

Separation of the products of cyanogen bromide digestion

Gel filtration as described in the Materials and Methods section was used to separate the peptides resulting from cyanogen bromide digestion. A typical elution profile is shown in Fig. 4. Fractions were monitored by polyacrylamide-gel electrophoresis at pH 2.7 in 8 M-urea (Fig. 4). The partially degraded material was eluted first from the column, appearing in the leading edge of the first peak. Most of the protein in the first peak was fragment A, pure samples being obtained in the latter half of this peak. Fragment B appeared in the second peak and some was obtained free from all other fragments.

Fig. 4. Sephadex-gel chromatography of a CNBr digest of flagellin. The curve shows E_{215} of the different fractions after excess of CNBr and formic acid had been removed by dialysis. The lower part of the figure illustrates the components present in different fractions as monitored by polyacrylamide-gel electrophoresis at pH 2.7 in 8M-urea. Tubes 33-35 contained only fragment A, tubes 39-42 contained only fragment B and tubes 48-51 contained only fragment D. Fragment C and the 'C,D' complex were not obtained pure. Experimental details are given in the text.

The final peak contained pure fragment D. Fragment C and 'C,D' complex appeared between the second and third peaks but were never obtained free from fragments B and D. Some u.v.-absorbing material was detected in the portion of the column effluent in which low-molecular-weight substances would be expected to be present. This material was probably homoserine lactone. It reacted with ninhydrin, had the same R_F value and electrophoretic mobility as authentic lactone, and appeared in the same position as did the lactone during amino acid analysis. Since flagellin does not appear to contain adjacent methionine residues, the reason for the appearance of free homoserine lactone was uncertain. Possibly there was an acid-labile bond attaching homoserine lactone to one of the peptides.

Numerous attempts were made to purify fragment C and the postulated 'C,D' complex from a cyanogen bromide digest of flagellin by varying the conditions of fractionation of the digest. All were unsuccessful, and so advantage was taken of the observation that, when flagellin was exposed to relatively high concentrations of the oxidant chloramine-T and then treated with cyanogen bromide, a simpler pattern of breakdown products was observed on polyacrylamide-gel electrophoresis (Fig. 2c). Only three protein bands were seen, one

being a low-molecular-weight component corresponding in position to the postulated 'C,D' complex. It was isolated in a pure state (Fig. 2c) by chromatography on Sephadex columns, as described in the Materials and Methods section, in a yield of 20% of the theoretical maximum amount. Presumably chloramine-T oxidizes the more exposed methionine residues to methionine sulphoxide, which is not amenable to cleavage by cyanogen bromide. The methionine residue linking 'C,D' complex to the rest of the flagellin molecule is probably masked and therefore less readily oxidized.

Characterization of the products of cyanogen bromide digestion

Molecular-weight estimates. Estimates of the molecular weight of each fraction, based on the area under the peaks of the elution profile (Fig. 4), were made relative to the molecular weight of flagellin by assuming a value of 40000 as found by McDonough (1965). These estimates were valid if the following conditions obtained; (1) complete digestion of flagellin by cyanogen bromide; (2) no adsorption of fragments on the Sephadex; (3) good resolution of peaks; (4) E_{215} being a quantitative measure of protein, irrespective of the amino acid composition of each fragment.

These conditions were approximately fulfilled for fragments A and B and yielded values of about 20000 and 10000 respectively. The presence of partially degraded molecules may have made the estimate for A a little too high. Then the third Sephadex peak, fragment D, was estimated to have mol.wt. 4000, which left a residual value of 6000, presumably the molecular weight of fragment C.

Amino acid analysis. Amino acid analyses were performed on fragments A, B and D and on the 'C,D' complex (Table 1). The sum of the analysis of fragments A, B and the 'C,D' complex agrees well with the composition of the complete flagellin molecule (see Table 1, columns 4 and 5). There are only minor discrepancies, the most important being 11 ϵ -N-methyl-lysine and 11 arginine residues/mol. in the cyanogen bromide fragments compared with 12 and 10 residues/mol. respectively in the flagellin analysis. In these cases we have assumed that the values obtained from the fragments are more reliable than the flagellin analysis.

Fragment A contained five of the six tyrosine residues and ten of the 11 ϵ -N-methyl-lysine residues of the molecule, but was low in glutamic acid, isoleucine and leucine, and contained only one of the 11 arginine residues. Its molecular weight was estimated to be 18000, based on the arginine, proline, tyrosine and phenylalanine contents. Fragment B contained no ϵ -N-methyl-lysine, tyrosine or proline, but possessed six of the 11

arginine residues and 17 of the 37 glutamic acid residues of the molecule. The threonine and serine contents of fragments A and B showed large differences. The molecular weight of fragment B was 12000, based on the glycine, isoleucine, leucine, valine, phenylalanine and lysine contents. The 'C,D' complex had a molecular weight of about 10000 as judged by the proline, glycine, tyrosine, phenylalanine, lysine and arginine contents. One residue of methionine/mol. was recovered from the 'C,D' complex, which was presumably methionine sulphoxide during digestion with cyanogen bromide, but was reconverted into methionine by acid hydrolysis. Fragment D had a molecular weight of 4500, this estimate being based on the phenylalanine, lysine, ϵ -N-methyl-lysine and arginine contents. This leaves a molecular weight of 5500 for fragment C, which accords well with its position of banding in polyacrylamide-gel electrophoresis, and its contents of arginine, lysine and phenylalanine estimated from the difference between the compositions of fragment D and the 'C,D' complex. Thus the estimates of the molecular weights obtained from the gel chromatography are in reasonable agreement with the values from amino acid analysis.

Homoserine was found to be present in fragments A, B and D, but not in the 'C,D' complex (Table 1).

Action of carboxypeptidases A and B on flagellin and flagellin fragments. The main results of these investigations are presented in Table 3, from which the following points are made. (1) No amino acids were released from flagellin during the 2 hr. treatment with carboxypeptidase A. (2) After 4 hr. treatment with carboxypeptidase B, only leucine and arginine were released. (3) In the presence of both enzymes, a 15 min. treatment also only released leucine and arginine but a more prolonged treatment released a large number of different amino acids. Because of the known

specificities of these enzymes, it seems most likely that arginine is the C-terminal amino acid of flagellin. Variation in the ratio of leucine to arginine suggests a sequence of -Leu-Leu-Leu-Arg. Because of this finding, the flagellin fragments were only treated with carboxypeptidase B. Fragments A and B yielded no amino acids after 2 hr. digestion, possibly due to the presence of C-terminal homoserine lactone. Leucine and arginine were released from the 'C,D' complex but not from fragment D, which is consistent with the interpretation that in the flagellin molecule fragment C is in the C-terminal position. However, other amino acids were released from fragment D and the 'C,D' complex in small but very similar amounts by this enzymic treatment. This suggests that carboxypeptidase B had an endopeptidase activity (fragment D also contained homoserine lactone) which was displayed when reacting with fragment D either by itself or in the 'C,D' complex, but not with native flagellin.

Iodination of flagellin fragments. Both for the interpretation of the amino acid analyses and for subsequent work with labelled fragments, it was desirable to see which of these fragments could be labelled with radioactive iodine. Fragment A and a mixture of fragments B, C and D were trace-labelled with [125 I]iodide, mixed with unlabelled complete digest, and subjected to gel electrophoresis at pH 2.7 in 8M-urea. The gels were stained as usual, destained and sliced into about 55 pieces. The radioactivity of each was determined, and in Fig. 5 this is compared with the protein pattern. Allowing for slight differences in the protein patterns, the fragment A preparation showed a major peak of radioactivity at the position of fragment A with very minor peaks corresponding in position to fragment D and the 'C,D' complex. The mixture of fragments B, C and D showed major peaks of radioactivity corresponding to fragment D and the 'C,D' complex. To be sure that the

Table 3. Release of amino acids from flagellin and fragments by carboxypeptidase A or B or both

Carboxypeptidase	Protein	Time (min.)	Amino acid released (moles/mole of protein)		Leucine/arginine ratio
			Leucine	Arginine	
A	Flagellin	120	< 0.02	< 0.02	
B	Flagellin	15	0.41	0.34	1.21
		60	0.98	0.47	2.10
		240	1.73	0.86	2.01
A + B	Flagellin	15	0.79	0.28	2.85
		120	1.71	0.58	2.95
B	Fragment A	120	< 0.02	< 0.02	
	Fragment B	120	< 0.02	< 0.02	
	Fragment D	120	< 0.02	< 0.02	
	'C,D' complex	120	0.54	0.48	1.13

radioactivity on 'C,D' complex was not adsorbed, purified fragment D was labelled and added to unlabelled cyanogen bromide digest, and electrophoresis was carried out and the distribution of radioactivity determined. A single peak of radioactivity corresponding to the position of fraction D in the gel was found (Fig. 5). This suggested that fragment C, which was not available in purified form to test chemically, and fragment B did not contain a tyrosine residue, whereas fragment D

(and the 'C,D' complex) contained one or more tyrosine residues that could be substituted with iodide.

To determine the site in the flagellin molecule where iodination preferentially occurred in trace-labelling procedures, flagellin and polymer were iodinated and then mixed with carrier flagellin or polymer. Each preparation was treated with cyanogen bromide and subjected to electrophoresis in polyacrylamide gels as usual. The iodination patterns of flagellin and polymer differed (Table 4). In flagellin, at low degrees of substitution, fragments A and D and 'C,D' complex were labelled, but fragment D was labelled to a greater extent than fragment A. At higher degrees of substitution, fragment A carried a greater proportion of the label. On the other hand, fragment A was preferentially labelled in polymer. The five faint bands with lower electrophoretic mobility than fragment A became progressively more heavily labelled in polymer and in flagellin as the proportion of radioactivity in fragment A increased.

Polymerization studies. Flagellin can be readily polymerized to form a structure similar to native flagella. Polymerization of flagellin occurs rapidly in 15% saturated ammonium sulphate. Flagellin, previously exposed to 70% formic acid or to 8M-urea for 16hr. at 20°, and the cyanogen bromide digest were brought to 15% saturation with ammonium sulphate. The increased opacity of each solution was followed for 24hr. by reading extinction values at 250nm. After 24hr., the polymerized protein was removed by centrifugation and the soluble protein in the supernatant was estimated. Flagellin polymerized more rapidly than either formic acid- or urea-treated flagellin. By 50min. the degree of polymerization that had occurred relative to the 24hr. value was as follows: flagellin, 84%; formic acid-exposed flagellin, 68%; urea-treated flagellin, 47%.

In each case 90% or more of the protein had polymerized by 24hr. Neither flagellin in water

Fig. 5. Polyacrylamide-gel electrophoresis of labelled CNBr fragments of flagellin. Fragment A (—), a mixture of fragments B, C and D (---), and fragment D (.....) were labelled with ¹²⁵I iodide, mixed with a complete digest of unlabelled flagellin and subjected to polyacrylamide-gel electrophoresis at pH 2.7 in 8M-urea. Each gel was stained, excess of stain was then removed to allow location of the protein bands, and the gel was sliced into about 55 pieces. The radioactivity (curves) of each was estimated and compared with the positions of the protein bands (lower part). There was almost complete identity in the protein patterns, but the position of the dye marker varied slightly.

Table 4. Distribution of radioactivity between fragments A and D of flagellin and polymerized after labelling with radioactive iodide at different extents of substitution

The value for the percentage distribution of iodide neglects the radioactivity associated with the five minor components with electrophoretic mobilities less than fragment A.

Substance	Extent of substitution (g.atom of iodide/40 000 g. of protein)	% distribution of iodide between fragments	
		Fragment A	'C,D' complex + fragment D
Flagellin	0.19	19	81
	0.70	46	54
Polymerized flagellin	0.16	72	28
	1.46	90	10
	6.0	93	7

Fig. 6. Proposed sequence of the CNBr fragments in the polypeptide chain of flagellin. The *N*-terminal and the four *C*-terminal amino acid residues of flagellin are marked. The molecular weight of each fragment is also recorded. The distribution of ϵ -*N*-methyl-lysine residues is designated by ●.

nor the digest in ammonium sulphate polymerized to a significant extent. The digest also did not prevent polymerization of flagellin.

DISCUSSION

The evidence presented in this paper is consistent with the belief that flagellin, as isolated from purified flagella preparations from *S. adelaide*, is a homogeneous protein. Ultracentrifugation studies showed a single sedimenting component with an estimated molecular weight of about 40000. This is in good agreement with molecular-weight estimates of flagellins from other *Salmonella* strains. McDonough (1965) found a value of 40000 for flagellin from a *Salmonella typhimurium* mutant and a value of 40000 has been reported for flagellin from *Bacillus subtilis* and *Spirillum serpens* (Martinez, Brown & Glazer, 1967). The amino acid composition of *S. adelaide* flagellin is similar to analyses of other flagellins (McDonough, 1965). All show the absence of tryptophan and cysteine and the absence or low content of histidine. As tryptophan and cysteine are commonly found in proteins, this can also be taken as evidence for the lack of contamination of the flagellin by extraneous proteins. The preparation may contain carbohydrate, but if so the content is less than 0.03% (Ada *et al.* 1964a). Finally, polyacrylamide-gel electrophoresis of flagellin under stringent conditions yielded a single staining band (see also Martinez *et al.* 1967).

Flagellin from *S. adelwidi* contains ϵ -*N*-methyl-lysine, a derivative of lysine that has so far been found only in some bacterial flagellins and in histones. The reason for its presence is uncertain. Flagellins can be obtained which differ apparently only in the presence or absence of ϵ -*N*-methyl-lysine, and *N*-methylation of lysine can be effected by transduction (Stocker, McDonough & Ambler, 1961). The lack of cysteine in the molecule suggests not only the presence of a single peptide chain but also the absence of any cross-linking, and this makes degradation studies simpler and increases the likelihood of recovery of intact antigenic determinants in the protein fragments.

Cyanogen bromide is the only known chemical

that specifically and quantitatively cleaves a protein at a particular amino acid residue. As a result it has been used increasingly to study the primary structure of proteins. In many cases a low pH and a strong dissociating agent are necessary to allow all methionine residues to be available for reaction. Formic acid has been found to be effective in this and other work (Steers, Craven, Anfinsen & Bethune, 1965). However, formic acid has a disadvantage in that after its removal aggregation of protein fragments becomes troublesome. With flagellin, the reaction appeared to be quantitative but the cleavage was incomplete. Some of the fragments in the partial degradation products were linked by a derivative of methionine, which was not, apparently, susceptible to the action of cyanogen bromide (suggesting that the sulphur was not bivalent), but which was recovered as methionine sulphone after both cyanogen bromide treatment and acid hydrolysis. At present the nature of this derivative is unknown, but preliminary results suggest that it is probably not methionine sulphoxide. The four major fragments were resolved by polyacrylamide-gel electrophoresis but only three (fragments A, B and D) were resolved by gel filtration. A complete amino acid analysis for the fragments of flagellin was only made possible by the isolation of one of the partial breakdown products ('C,D' complex), which was obtained after treatment of flagellin with chloramine-T followed by cyanogen bromide. The results obtained from the polyacrylamide - gel - electrophoretic analysis, amino acid analyses and carboxypeptidase treatment of flagellin and fragments allow an interpretation of the possible sequence of the fragments in the flagellin molecule. Fragments C and D are joined in the flagellin molecule as, in electrophoresis experiments, the band between fragments B and C was a 'C,D' complex. Two results supported the interpretation that fragment C was C-terminal. (i) Fragments A, B and D but not the 'C,D' complex contained homoserine. (ii) The treatment of the 'C,D' complex and other fragments with carboxypeptidase B showed that fragment C contained the C-terminal amino acids present in flagellin. The two possible sequences are then B-A-D-C or A-B-D-C. Two results supported the interpretation that fragment B was N-terminal.

(1) The *N*-terminal amino acid of flagellin (Ada *et al.* 1964a) and of fragment B, but not of fragment A (B. E. Davidson, personal communication) is alanine. (2) As shown best in the case of polymer, when fragment D was only lightly labelled, the five minor components with a lower mobility than fragment A all contained fragment A because all were radioactive in proportion to the radioactivity in fragment A itself. It can be easily demonstrated that five components containing fragment A are only possible if fragment A is not *N*-terminal. The sequence therefore proposed is B-A-D-C (see Fig. 6), but amino acid-sequence studies will be necessary to establish this.

The differences in the patterns of iodination of flagellin and of polymer suggest strongly that, in the polymer, fragment A is exposed but fragment D is masked. The amino acid compositions of the fragments are also of interest. Although fragment A contained only half of the total lysine it contained ten of the 11 ϵ -*N*-methyl-lysine residues of the molecule (see Fig. 6). As fragment A appears to be exposed in polymer and methylation of lysine is enzymically mediated (Kim & Paik, 1965), the possibility must be seriously considered that methylation of the flagellin lysine residues may occur only after polymerization of the flagellin. It has previously been suggested on genetic evidence (Stocker *et al.* 1961) that methylation of the lysine residues occurs after flagellin synthesis. The other marked difference is that fragment A contains only one of the 11 arginine residues of the molecule. It has also been shown that methylation of the lysine residues in peptides decreases the susceptibility of the peptide to tryptic action (Benoiton & Deneault, 1966). Therefore fragment A may be relatively resistant to a trypsin-like enzyme because it contains only one arginine residue and many methylated lysine residues. Preliminary studies with trypsin have shown that this is so. Flagella resistant to tryptic digestion would be advantageous for *Salmonella* organisms since flagella are an extracellular protein and these organisms thrive in the mammalian gut, an environment rich in tryptic enzymes.

This is Publication no. 1235 from The Walter and Eliza Hall Institute. This work was supported by grants from

the National Health and Medical Research Council and the Australian Research Grants Committee.

REFERENCES

- Ada, G. L. (1967). In *Immunity, Cancer and Chemotherapy*, p. 17. Ed. by Mihich, E. New York: Academic Press Inc.
- Ada, G. L., Nossal, G. J. V. & Pye, J. (1964b). *Aust. J. exp. Biol. med. Sci.* **42**, 295.
- Ada, G. L., Nossal, G. J. V., Pye, J. & Abbot, A. (1964a). *Aust. J. exp. Biol. med. Sci.* **42**, 267.
- Ambler, R. P. (1965). *Biochem. J.* **96**, 32 p.
- Atassi, M. Z. & Saplin, B. J. (1968). *Biochemistry*, **7**, 688.
- Benoiton, L. & Deneault, J. (1966). *Biochim. biophys. Acta*, **113**, 613.
- Crestfield, A. M., Moore, S. & Stein, W. H. (1963). *J. biol. Chem.* **238**, 622.
- Crompton, M. J. (1967a). In *Antibodies to Biologically Active Molecules*, p. 61. Ed. by Cinader, B. Oxford and New York: Pergamon Press.
- Crompton, M. J. (1967b). *Nature, Lond.*, **215**, 17.
- Davis, B. J. (1964). *Ann. N.Y. Acad. Sci.* **121**, 404.
- Diener, E. & Armstrong, W. D. (1967). *Lancet*, ii, 1281.
- Greenwood, F. C., Hunter, W. M. & Glover, J. S. (1963). *Biochem. J.* **89**, 114.
- Hofmann, T. (1964). *Biochemistry*, **3**, 356.
- Kabat, E. A. (1966). *J. Immunol.* **97**, 1.
- Kim, S. & Paik, W. K. (1965). *J. biol. Chem.* **240**, 4629.
- McConahey, P. J. & Dixon, F. J. (1966). *Int. Arch. Allergy*, **29**, 185.
- McDonough, M. W. (1965). *J. molec. Biol.* **12**, 342.
- Martinez, R. G., Brown, D. M. & Glazer, A. N. (1967). *J. molec. Biol.* **28**, 45.
- Moffit, W. & Yang, J. T. (1956). *Proc. nat. Acad. Sci., Wash.*, **42**, 596.
- Moore, S., Spackman, D. H. & Stein, W. H. (1958). *Analyt. Chem.* **30**, 1185.
- Neville, D. M. (1967). *Biochim. biophys. Acta*, **133**, 168.
- Parish, C. R., Wistar, R., jun. & Ada, G. L. (1969). *Biochem. J.* **113**, 501.
- Ray, W. J. & Koshland, D. E. (1960). *Brookhaven Symp. Biol.* **13**, 135.
- Reisfeld, R. A. & Small, P. A. (1966). *Science*, **152**, 1253.
- Steers, E., Craven, G. R., Anfinsen, C. B. & Bethune, J. L. (1965). *J. biol. Chem.* **240**, 2478.
- Stocker, B. A., McDonough, M. W. & Ambler, R. P. (1961). *Nature, Lond.*, **189**, 556.
- Technicon Instrument Co. (1966). *Proc. 4th Amino-Acid Colloquium*, p. 18. Chauncey, N.Y.: Technicon Instrument Co.
- Young, D. J., Benjamini, E., Shimizu, M. & Leung, C. Y. (1966). *Biochemistry*, **5**, 1481.