J. Mol. Biol. (1990) **212**, 377–387

Stoichiometric Analysis of the Flagellar Hook–(Basal-Body) Complex of *Salmonella typhimurium*

Christopher J. Jones†, Robert M. Macnab‡

Department of Molecular Biophysics and Biochemistry
Yale University, New Haven, CT 06511, U.S.A.

Hiroshi Okino and Shin-Ichi Aizawa

ERATO, Research and Development Corporation of Japan
5-9-5 Tokodai, Tsukuba, Ibaraki 300-26, Japan

(Received 21 August 1989; accepted 17 October 1989)

The stoichiometries of components within the flagellar hook–(basal-body) complex of *Salmonella typhimurium* have been determined. The hook protein (FlgE), the most abundant protein in the complex, is present at approximately 130 subunits. Hook-associated protein 1 (FlgK) is present at approximately 12 subunits. The distal rod protein (FlgG) is present at approximately 26 subunits, while the proximal rod proteins (FlgB, FlgC and FlgF) are present at only approximately six subunits each. The stoichiometries of the proximal rod proteins and hook-associated protein 1 are, within experimental error, consistent with values of 5 or 6, and 11, respectively. Such values would correspond to either one or two turns of a helical structure with a basic helix of approximately 5·5 subunits per turn, which is the geometry of both the hook and the filament and, one supposes, the rod and hook-associated proteins. These stoichiometries may derive from rules for the heterologous interactions that occur when a helical structure consists of successive segments constructed from different proteins; the stoichiometries within the hook and the distal portion of the rod must, however, be set by different mechanisms. The stoichiometries for the ring proteins are approximately 26 subunits each for the M-ring protein (FliF), the P-ring protein (FlgI), and the L-ring protein (FlgH); the protein responsible for the S-ring feature is not known. The rings presumably have rotational rather than helical symmetry, in which case the stoichiometries would be directly constrained by the intersubunit bonding angle. The ring stoichiometries are discussed in light of other information concerning flagellar structure and function.

1. Introduction

Bacteria move by means of organelles called flagella (Macnab, 1987*a*,*b*). The known structure of the flagellum of *Salmonella typhimurium* consists of the basal body, the hook, a junction region between the hook and the filament, the filament, and a distal cap on the filament (Fig. 1). This structure has been characterized in terms of the proteins it contains, where they are located within the structure, and which genes encode them (Table 1; Horiguchi *et al.*, 1975; Suzuki *et al.*, 1978; Kutsukake *et al.*, 1979, 1980; Homma *et al.*, 1984*b*, 1985, 1987*a*,*b*,*c*, 1990*a*; Aizawa *et al.*, 1985; Homma & Iino, 1985; Ikeda *et al.*, 1985, 1987; Jones *et al.*, 1987, 1989; Ohnishi *et al.*, 1987; Okino *et al.*, 1989).

The components on the axis of the structure are (in order, from the cell membrane outward): (1) to (3) three proteins (FlgB, FlgC and FlgF) that are located in the cell-proximal part of the basal-body rod, but whose order in that structure is not known; (4) a distal rod protein (FlgG); (5) the hook protein (FlgE); (6) and (7) two filament–hook junction proteins or hook-associated proteins, HAP1 (FlgK) and HAP3 (FlgL); (8) the filament protein, flagellin (FliC); and (9) a protein, HAP2 (FliD), that caps the

† Present address: ERATO, Research and Development Corporation of Japan, 5-9-5 Tokodai, Tsukuba, Ibaraki 300-26, Japan.

‡ Author to whom all correspondence should be sent.

0022-2836/90/060377-11 $03.00/0

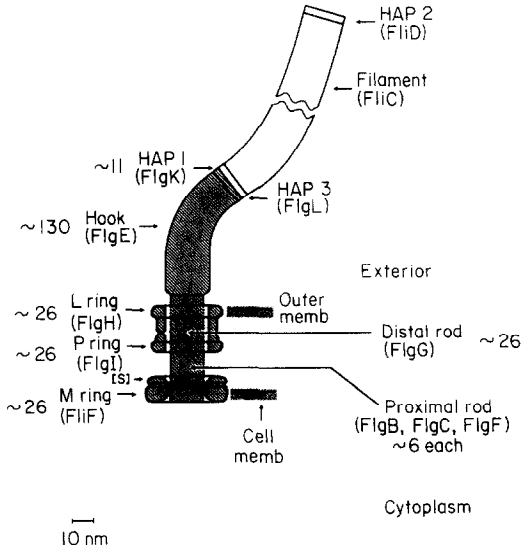

Figure 1. Cartoon of the flagellar filament-hook–(basal-body) complex of *S. typhimurium*, showing its various substructures and the proteins (indicated by the corresponding gene symbols) from which they are constructed. The basal body consists of an axial rod and 4 rings and is imbedded in the cell envelope, with the M and L rings associated with the cell membrane and the outer membrane, respectively. The hook–(basal-body) (HBB) complex (hatched areas) is the structure obtained after depolymerization of the filament, and consists of the basal body, the hook and HAP1. Stoichiometries of the various protein components of the HBB complex, measured in the present study, are indicated alongside the relevant structures.

filament. The conclusion that FlgB, FlgC and FlgF are rod components is based on several lines of evidence: their primary sequences are markedly similar to that of a known rod component (FlgG) (Okino *et al.*, 1989; Homma *et al.*, 1990*a*); they are basal-body components, and (along with FlgG) disappear when the rod structure is selectively dissociated (Aizawa *et al.*, 1985); the stoichiometry of FlgG alone (see Results) cannot account for the observed length of the rod.

Surrounding the basal-body rod are four rings (DePamphilis & Adler, 1971). These are (in order from the cell membrane outward): the M ring (FliF), the S ring (protein unknown), the P ring (FlgI), and the L ring (FlgH).

What is known of the symmetries of the various flagellar substructures? The hook and the filament have been characterized with respect to their symmetry and lattice dimensions, and three-dimensional reconstructions of both structures have been obtained that permit a description of the subunit shapes and the interactions among subunits (Shirakihara & Wakabayashi, 1979; Wagenknecht *et al.*, 1982; Trachtenberg & DeRosier, 1987). Knowledge of the structure of the basal body, and of the two filament–hook junction proteins (HAP1 and HAP3) and the filament cap (HAP2), is less detailed. The HAPs exist in quite short zones, making image analysis difficult. A cylindrically averaged model of the basal body has been obtained by image reconstruction (Stallmeyer *et al.*, 1989), but the averaging process prevents any statement from being made regarding rotational or helical symmetries of its components. Estimates of ring symmetries have been made (DePamphilis & Adler,

Table 1

Properties of flagellar hook–(basal-body) proteins of S. typhimurium

Protein	Substructure	Molecular mass (kDa)	Number of subunits/ HBB complex
Axial structures			
FlgB	Proximal rod	16	7·2 (1·2)
FlgC	Proximal rod	14	6·4 (1·1)
FlgF	Proximal rod	32	6·3 (1·4)
FlgG	Distal rod	30	25·8 (4·3)
FlgE	Hook	42	132 (21)
FlgK	HAP1	60	12·6 (2·5)
Ring structures			
FliF	M ring	65	27·2 (4·4)
FlgI	P ring	38	24·1 (4·3)
FlgH	L ring	27	28·3 (5·0)
			37·8 (6·7)

Identification of a basal-body protein of given molecular mass with the specific gene indicated had (with the exception of the FlgB protein) been established prior to this study on the basis of mutant analysis, cloning and product identification in minicells, or both (see the text for references). Direct evidence is provided in the present study by the correlation of N-terminal amino acid analysis and gene sequence analysis (see the text). The assignment of a given protein to a given morphological feature is based on references given in the text. The proteins are listed in order from the cell membrane outward, first for the axial components (rod, hook and HAP1) and then for the rings (cf. Fig. 1). The number of subunits per HBB complex was determined in the present study (see Table 2 and the text); the value for FlgH is ambiguous, but the smaller value is probably correct (see Discussion).

1971; Stallmeyer *et al.*, 1989) but are of uncertain validity (see Discussion). Symmetry analysis of the rod, along the lines already achieved for the hook and filament, has not been attempted; it is likely to be difficult, given the fact that the rod is fairly short (approx. 30 nm, compared with 55 nm for the hook and up to 10 μm for the filament) and has at least four distinct protein components (FlgB, FlgC, FlgF and FlgG) at different positions along its length. Thus, for several reasons, determination of the symmetries of basal-body substructures by morphological examination is not readily achievable.

An alternative, albeit less detailed, approach to the question of symmetry is to determine biochemically how many subunits of a given protein are present within the structure. For components with well-defined stoichiometries, accurate estimates should not only define the protein composition of the structure but also give indications of possible symmetries. The assumption of well-defined stoichiometry is approximately valid for the hook, which has a fairly closely regulated length (Suzuki & Iino, 1981). It is probably valid for the HAPs, which exist in quite localized zones (Ikeda *et al.*, 1987), and almost certainly valid for the proteins within the basal body, whose various substructures are well-defined morphologically (DePamphilis & Adler, 1971; Aizawa *et al.*, 1985). Relative and absolute estimates of the number of subunits of the HAPs have been made by Homma *et al.* (1984*b*) and Ikeda *et al.* (1987), respectively, who determined the intensity of Coomassie stain in electrophoretograms of hook or filament–hook preparations. No estimates have been made for the stoichiometries of basal-body components.

A flagellar substructure consisting of the basal body, the hook and HAP1 (the hook–(basal-body) complex, or HBB† complex) can be isolated from *S. typhimurium* (Aizawa *et al.*, 1985). We describe here how we obtained stoichiometries for its components, by analyzing purified ^{35}S-labeled complexes, normalizing for the number of sulfur-containing residues each protein contains, and converting the relative values of stoichiometries into absolute values *via* the known helical lattice of the hook substructure of the complex.

(Flagellar gene designations in this report follow the nomenclature of Iino *et al.* (1988) and differ from those in the earlier literature.)

2. Materials and Methods

(a) *Bacteria*

S. typhimurium strain ST1, a motility-selected derivative of LT2, wild-type for flagellation (Aswad & Koshland, 1975), was used as the source of the HBB complexes. Derivatives of strain LT2 were used for the sequencing of all of the genes relevant to this paper except *fliF*, where SJW1673, a derivative of *S. typhimurium* strain TM2, also wild-type for flagellation, was used. SJW1103 (the parent of SJW1673) was used for the N-terminal amino acid sequence analyses.

(b) *Chemicals and enzymes*

Acrylamide and *N,N'*-methylenebisacrylamide were electrophoresis grade (Bio-Rad Laboratories, Richmond, CA). Chemicals for the N-terminal amino acid sequence analysis were from Applied Biosystems (Foster City, CA). $^{35}SO_4^{2-}$ was purchased as H_2SO_4 (carrier-free, in water) from ICN Radiochemicals, Irvine, CA. All other chemicals used were reagent grade or higher and were obtained from standard commercial sources.

(c) *Radiolabeling of cells*

Throughout the protocol, growth was at 37 °C with shaking. Cells were first grown overnight in a modified version of Vogel–Bonner citrate minimal medium (Vogel & Bonner, 1956) that was supplemented with glycerol to 1% (v/v) and contained 0·8 mM-$MgCl_2$ instead of 0·8 mM-$MgSO_4$. Sulfur was supplied as Na_2SO_4 at 200 μM. A total of 250 μl of this culture was used to inoculate 7·5 ml of fresh Vogel–Bonner citrate minimal medium lacking sulfate. This culture was incubated until cell growth began to slow down as a result of sulfur starvation (about 4 h). These sulfate-starved cells were used to inoculate, to a density of 10^6 cells ml^{-1}, 2 l of modified Vogel–Bonner citrate minimal medium supplemented with 50 μM-Na_2SO_4 and 5 mCi of $^{35}SO_4^{2-}$. Growth was then continued until sulfur limitation set in at a density of approx. 10^9 cells ml^{-1}.

(d) *Purification and electrophoretic analysis of HBB complexes*

HBB complexes were purified and analyzed by sodium dodecyl sulfate/polyacrylamide gel electrophoresis (SDS/PAGE), essentially as described (Aizawa *et al.*, 1985).

(e) *Autoradiography and densitometry*

Gels containing the separated radiolabeled HBB complexes were placed (without prior fixing) onto Whatman no. 1 filter paper, vacuum dried, and exposed to Kodak X-Omat film for 80 h at room temperature. Autoradiographs were scanned and analyzed as described (Wilson & Macnab, 1988).

(f) *Electron microscopy and distance estimation*

Electron microscopy was carried out essentially as described (Aizawa *et al.*, 1985). Measurements of the L to M-ring distance and hook length of HBB complexes were made from photographic enlargements using a Numonics 1224 Electronic Digitizer (Numonics Corp., Lansdale, PA).

(g) *N-Terminal amino acid analysis*

Purified HBB complexes were subjected to SDS/PAGE, using a Mini Protean II gel system (Bio-Rad Laboratories, Richmond, CA) with a gel size of 85 mm × 55 mm × 0·75 mm and a 10-well comb. A 2 l culture volume provided enough material to run 7 to 10 such gels. After staining the gel with Coomassie brilliant blue R,

† Abbreviations used: HBB, hook–(basal-body); SDS/PAGE, SDS/polyacrylamide gel electrophoresis.

bands of interest were excised and macerated by low-speed centrifugation through a fine mesh of 0·1 mm stainless steel wire. The gel pieces were washed repeatedly with 50% (v/v) ethanol until completely destained, soaked overnight with shaking in 70% (v/v) formic acid to elute the protein, centrifuged and the supernatant set aside. The gel pieces were subjected to several further cycles of elution using 80% formic acid. Supernatants were combined, concentrated by centrifugation under vacuum (Taiyo model VC36 centrifugal concentrator), and subjected to amino acid sequence analysis (model 477A, Applied Biosystems, Foster City, CA). The supernatants from 2 to 10 gels, depending on the abundance of the protein of interest, were sufficient for such analysis.

3. Results

(a) *N-Terminal amino acid sequence and sulfur content of HBB proteins*

The genes, gene sequences and gene products for many (probably most) of the HBB components are known (see Introduction, Fig. 1, and Tables 1 and 2). We wished to obtain the N-terminal amino acid sequences of these components, for several reasons: (1) to confirm predicted gene starts; (2) to determine the nature and extent of any post-translational processing at the N termini of the products; and (3) to determine the cysteine + methionine content of the mature proteins for use in the present stoichiometric analysis.

Approximately the first 15 N-terminal residues of each basal-body protein were determined. For all of the proteins except the L-ring protein (see below) unambiguous identification of most residues was possible. The results confirmed the starts that had been predicted by DNA sequencing for the M and P-ring genes (Jones *et al.*, 1989), the rod genes (Homma *et al.*, 1990*a*), and the hook and HAP1 genes (Homma *et al.*, 1990*b*). They also enabled us to determine the N termini of the mature proteins.

The N-terminal sequences of the various proteins and the positions of their cleavage sites (if any) are given in Table 2. The axial components are either not cleaved at all (the FlgC, FlgF and FlgG rod proteins) or have only their N-terminal methionine removed (the FlgB rod protein, hook protein and HAP1). The M-ring protein has its N-terminal methionine removed.

It was known from previous work that both the P and L-ring proteins are processed by removal of a signal peptide during their export across the cell

Table 2

Analysis of N-terminal amino acid sequence and determination of stoichiometry of flagellar hook–(basal-body) proteins

Protein	N-terminal sequence	Processing event	N_S	Stoichiometry	
				Relative	Absolute
FlgB	M ↑ LDRL . .	Terminal methionine	6	0·0549 (0·004)	7·2 (1·2)
FlgC	MALLN . .	None	5	0·0486 (0·003)	6·4 (1·1)
FlgF	MDHAI . .	None	9	0·0480 (0·007)	6·3 (1·4)
FlgG	MISSL . .	None	6	0·1952 (0·008)	25·8 (4·3)
FlgE	M ↑ SFSQ . .	Terminal methionine	8	1·0000 (0·020)	132 (21)
FlgK	M ↑ SSLI . .	Terminal methionine	6	0·0954 (0·011)	12·6 (2·5)
FliF	M ↑ SATA . .	Terminal methionine	9	0·2063 (0·005)	27·2 (4·4)
FlgI	MX_{16}HA ↑ ER . .	Signal peptide	13	0·1827 (0.015)	24·1 (4·3)
FlgH	MX_{18}TG ↑ CA WI . .	Signal peptide	4	0·2146 (0·016)	28·3 (5·0)
	or ↑		3	0·2861 (0·022)	37·8 (6·7)

The molecular masses and locations of these proteins within the HBB complex are given in Table 1. Complete amino acid sequences are based on the translated gene sequences from the following sources: FlgB, FlgC, FlgF and FlgG (Homma *et al.*, 1990*a*); FlgE and FlgK (Homma *et al.*, 1990*b*); FliF, FlgI, and FlgH (Jones *et al.*, 1989). N-terminal cleavage sites (↑) are based on amino acid analysis of mature proteins (this study), except in the case of FlgH, where the N terminus is blocked; in the latter case, FlgH is known to be processed by cleavage of a signal peptide (Homma *et al.*, 1987*c*) and 2 potential cleavage sites based on the translated gene sequence are shown. Signal peptide sequences are abbreviated, with X_n representing a stretch of n residues. N_S is the number of sulfur-containing residues (Cys + Met) in the mature protein, calculated from the translated gene sequence and knowledge of the N terminus of the mature protein. Relative stoichiometries are calculated from autoradiographic intensities and N_S (see the text), and are normalized to a value of 1 for the hook protein. The absolute stoichiometry of the hook protein, FlgE, is based on measurement of hook length and knowledge of the rise per subunit (see the text); absolute stoichiometries of other components are derived from their stoichiometries relative to the hook protein, and its absolute stoichiometry. Standard errors of the mean are given in parentheses. In the case of relative stoichiometries, error estimates are based on uncertainties in densitometry; in the case of absolute stoichiometries, they are based additionally on uncertainties in the measurement of hook length.

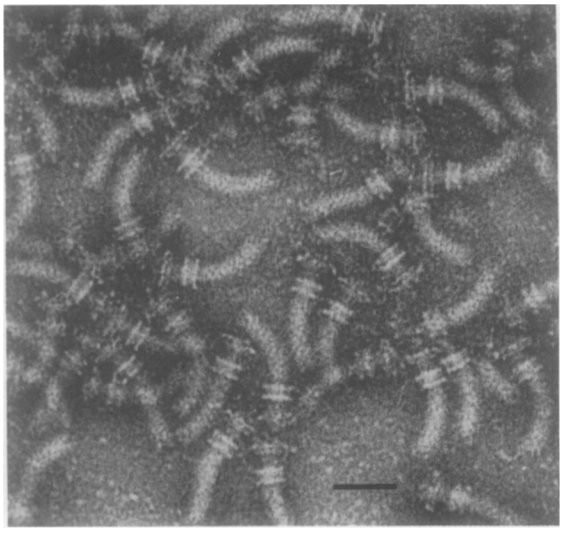

Figure 2. Electron micrograph of the preparation of HBB complexes used for the stoichiometric analysis. The bar represents 50 nm.

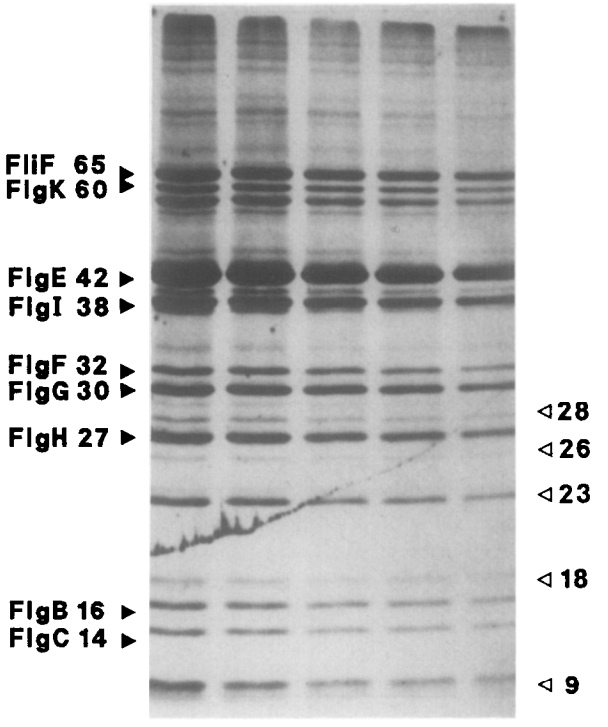

Figure 3. Autoradiograph of an electrophoretic gel of ^{35}S-radiolabeled HBB complexes. Known HBB proteins (filled arrowheads) are described by their gene symbols and their apparent molecular masses (in kDa); see Table 1. Other proteins that are believed to be HBB components (open arrowheads) are indicated by their apparent molecular masses only, since the genes encoding these proteins are not known. (The major band just below FlgK is flagellin, present in residual amounts after depolymerization of the filament.) Successive lanes differ in load by a factor of $\sqrt{2}$, and are representative of the data that were used to construct the intensity *versus* load graph (Fig. 4).

membrane (Homma *et al.*, 1987*b*,*c*; Jones *et al.*, 1989). N-Terminal analysis confirmed the cleavage site for the P-ring protein but failed to do so for the L-ring protein, whose N terminus was blocked. For the latter protein, two potential cleavage sites are indicated by comparison of the translated sequence to consensus sites (Jones *et al.*, 1989). Depending on which is actually used, the N-terminal sequence of the mature L-ring protein will be either Cys-Ala-Trp-Ile . . . or Trp-Ile . . ., with a difference of one sulfur-containing residue.

The sulfur contents (N_S) of the various proteins, derived from the translated gene sequences and the N-terminal amino acid sequence analyses, are given in Table 2. In the case of the M-ring protein, the gene sequence was, for technical reasons, obtained from a derivative of a different wild-type strain (TM2) than that used for isolation of the HBB complexes (LT2); however, since even in the case of two different genera (*Salmonella* and *Escherichia*) the sequences exhibit a high degree of identity (86%; Jones *et al.*, 1989), it is unlikely that the sequences from the two different wild-type strains of *S. typhimurium* differ appreciably.

(b) *Analysis of radiolabeled HBB complexes*

Cells were grown for approximately ten generations in minimal medium containing $^{35}SO_4^{2-}$, harvested and their HBB complexes purified. The homogeneity of the material was verified by electron microscopy, with essentially all particles consisting of morphologically intact complexes (Fig. 2). The HBB complexes were then separated by SDS/PAGE and the gels autoradiographed (Fig. 3). Because there is considerable variation both in the abundance of individual components within the complex and in the number of sulfur-containing residues per molecule, a wide range of dilutions was used to ensure that data within the linear range of the film would be available for each component. The data that were finally used spanned a total load factor of 128, with successive lanes differing by a factor of $\sqrt{2}$. The autoradiographs were digitized and the integrated intensities of the bands corresponding to known and suspected HBB components were measured.

(c) *Relative stoichiometries of HBB components*

Within the linear range of the film, the integrated intensity of a band for any given protein should be proportional to the amount of protein in that band. A plot of log [intensity] *versus* log [load] should therefore give a line with a slope of unity for each protein. The data fitted this expectation within experimental error, with the slopes in all cases lying within the range 0·9 to 1·1. The deviations from linearity consisted of random scatter rather than curvature. We therefore constrained the data to a slope of unity (Fig. 4), and found that the ordinate

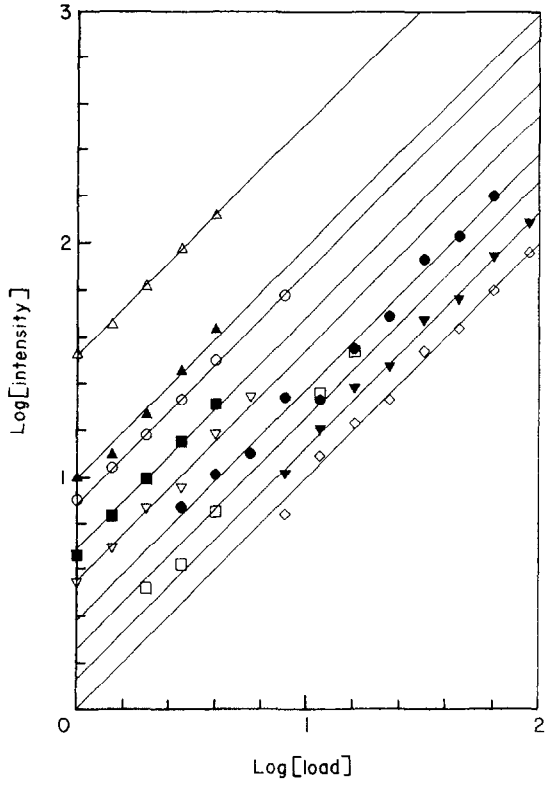

Figure 4. Plot of data from autoradiography of electrophoretic gels of HBB complexes (cf. Fig. 3). Protein components are indicated by their gene symbols. The ordinate represents the integrated intensity for a given protein and the abscissa represents the sample load. Data are presented on a logarithmic scale; lines are regressions constrained to have slopes of unity (see the text). (▼) FlgB; (◇) FlgC; (□) FlgF; (■) FlgG; (△) FlgE; (●) FlgK; (○) FliF; (▲) FlgI; (▽) FlgH.

intercept had a standard error of the mean that was typically about 0·04, corresponding to an uncertainty of about ±10% in the intensity; minimum and maximum values of this uncertainty were about ±2% and ±16%. The relative intensities of the various protein bands were then extracted from the ordinate values at an arbitrary sample load; since all slopes were constrained to unity, the choice of load was immaterial.

The data at this stage corresponded to relative stoichiometries of sulfur-containing residues in the various components of the HBB complex. The relative stoichiometry of each protein was then determined by taking into account its cysteine + methionine content (Table 2).

In addition to the well-characterized HBB proteins listed in the Tables, there are several other proteins present in HBB preparations (Fig. 3, open arrowheads), whose synthesis is dependent on the master operon of the flagellar regulon, and which are therefore presumed to be authentic flagellar components (Jones & Macnab, 1990); however, their physical location and encoding genes are currently unknown. The intensities of the bands corresponding to these proteins were also measured (data not shown); once the encoding genes have been identified and sequenced, this intensity information can be converted into molecular stoichiometries.

(d) *Absolute stoichiometry of the hook protein*

Diffraction analysis of electron micrographs of the flagellar hook of *S. typhimurium* has revealed that one turn of the basic helix contains 5·56 monomers and has a pitch of 2·3 nm; i.e. the rise per subunit is 0·41 nm (Wagenknecht *et al.*, 1982). A measurement of hook length can therefore be converted into a measurement of number of subunits.

Using a field of HBB complexes (of which Fig. 2 is a subset), we measured the contour distance D from the hook–rod junction to the tip of the particle, and also the center-to-center distance between the L and the M rings. Taking the value for the latter to be 27·5 nm (Stallmeyer *et al.*, 1989), we obtained a mean value of D of 60·6 nm (standard error of the mean = 8·8 nm, sample size = 25). However, D represents the contour length of the hook plus the HAP1 structure that is retained at the distal end of the hook in HBB preparations (Aizawa *et al.*, 1985). Given the relative stoichiometries of hook and HAP1 (Table 2), and the assumption that the lattice geometries of the two proteins are similar in terms of rise per subunit, we estimate that only 90% of D corresponds to the length of the hook, with the remaining 10% corresponding to the length of the HAP1 structure. This then gives a mean estimate of 54·5(8·8) nm for the length of the hook, and 132(21) for the number of hook subunits per HBB complex.

(e) *Absolute stoichiometries of HBB components*

Using the absolute stoichiometry estimate for the hook protein, the relative stoichiometries of the other components were converted into absolute values (Table 2). For convenience, these are repeated in Table 1, which contains other pertinent information regarding the various proteins.

4. Discussion

Determination of subunit stoichiometries is an essential part of the characterization of any macromolecular assembly. By analyzing the purified flagellar hook–(basal-body) complex, we have been able to estimate the relative numbers of protein subunits present in its major substructures (Table 2 and Fig. 1). Knowledge of the absolute stoichiometry of the hook protein has enabled these relative stoichiometries to be converted into absolute ones.

(a) *Reliability of estimated stoichiometries*

Because the deduced amino acid sequences of many of the HBB components were known from their gene sequences, and the extent of any N-terminal processing was known from N-terminal amino acid analysis (Table 2). ^{35}S-radiolabeling and

autoradiography could be used to assay the amount of each protein in the HBB complex. This is a more accurate method than staining with an agent such as Coomassie blue, where affinities of different proteins for the stain can vary considerably (Bio-Rad Technical Bulletin 1069, Bio-Rad Laboratories).

A critical issue is the quality of the HBB samples that were analyzed. Had they contained many damaged or partial structures, the measured stoichiometries would not be a valid indication of the stoichiometries of the intact complex. Electron microscopic examination (Fig. 2) established that the great majority of particles in the samples were intact HBB complexes, a contention that is further supported by the reproducibility in SDS/PAGE of the relative amounts of various components of the complex, in numerous preparations we have carried out in the course of this and other studies. As noted in Results, the HBB preparations contained a number of proteins (open arrowheads in Fig. 3) that are of unknown genetic origin, but whose synthesis is under control of the flagellar master operon and therefore presumably are flagellum-related. We think it unlikely that they are degradation products of known HBB components, since only large macromolecular complexes should carry through the purification protocol, after which the sample was immediately subjected to SDS/PAGE and examined by electron microscopy. A few unidentified bands were present in the gels and might represent contaminants; porins deriving from the outer membrane are known examples of this (Aizawa *et al.*, 1985). However, their presence did not affect estimates of stoichiometries from bands corresponding to known flagellar proteins.

For the calculation of relative stoichiometries, the principal source of error was in the estimates of the amount of radioactivity associated with each component. These estimates had a precision that ranged from about $\pm 2\%$ for the most abundant component present (the hook protein) to about $\pm 16\%$ for the FlgF protein. However, an additional uncertainty of about $\pm 16\%$ in the calibration standard (the length of the hook) reduced the precision of the absolute stoichiometries. In considering these absolute stoichiometries, it is therefore important to distinguish between the absolute value itself for a given component (where the precision was about ± 4 to 5 subunits for the more abundant basal-body components), and a comparison of the absolute values of various components (where the precision was about ± 1 to 2 subunits). In other words, stoichiometric comparisons between components are more precise than the stoichiometries themselves.

Ikeda *et al.* (1987) estimated the stoichiometries of all three HAPs, based on Coomassie stain intensity of gels of purified filament–hook complexes. The estimates cover rather large ranges, in part because of sample variability, but also because the authors combined two different measures of hook length for the calibration of absolute stoichiometries. Our estimate of hook length (55 nm) agrees better with their lower estimate (50 nm) than their higher one (70 nm); we consider the latter estimate to be unreliable, because it was based on inspection of published electron micrographs from another study (Hilmen & Simon, 1976), and the complexes were from another species, *Escherichia coli*. After adjustment of the data of Ikeda *et al.* to employ only the 50 nm value, their estimate for the stoichiometry of HAP1 (11 to 16 subunits) is consistent with ours ($12{\cdot}6 \pm 2{\cdot}5$).

There is an ambiguity in the stoichiometry of the L-ring protein, which arises because the site of cleavage of its signal peptide has not been established; although the two possible sites are only two residues apart, one of these is a sulfur-containing residue, cysteine. Since the P and L rings together form the outer cylinder of the basal body and appear to be in direct physical contact with each other, the simplest assumption is that they possess the same rotational symmetry. The lower estimate for the stoichiometry of the L-ring protein satisfies this assumption, whereas the higher estimate does not. We consider that the lower estimate is likely to be correct, and that the outer cylinder is constructed from L-ring/P-ring dimers present with a rotational symmetry of approximately 26.

(b) *Stoichiometries of axial components*

The rod and HAP proteins are present in the flagellum in fairly small numbers, and it is instructive to consider possible implications of these numbers in terms of the known lattice geometry of the major components, hook protein and the filament protein flagellin, which are present in quite large numbers (approx. 130 and 20,000, subunits, respectively).

The filament and hook are both constructed from subunits occupying a helical lattice with a basic (1-start) helix of close to 5·5 subunits per turn (O'Brien & Bennett, 1972; Wagenknecht *et al.*, 1982). Nearest-neighbor interactions are on 5 and 6-start helices (both at an angle of roughly 60° to the particle axis) and on an 11-start helix (almost parallel to the particle axis). The overall packing can thus be described as a cylindrical pseudohexagonal lattice (Fig. 5(a)). Within the hook and the filament, subunit interactions are homologous (hook protein to hook protein, or flagellin to flagellin). If the hook and filament structures were to abut each other directly this would introduce, at the junction, heterologous interactions (hook protein to flagellin) that would probably be unfavorable, given that the two proteins have to fulfil different structural and mechanical requirements. These include quite different macroscopic helical parameters (pitch and amplitude), and the known resistance to conformational phase cycling of the filament *versus* the presumed ease of such cycling in the hook during its function as a universal joint (Macnab & Aizawa, 1984). The mismatch of hook and flagellin subunits could be alleviated by the use of subunits of a

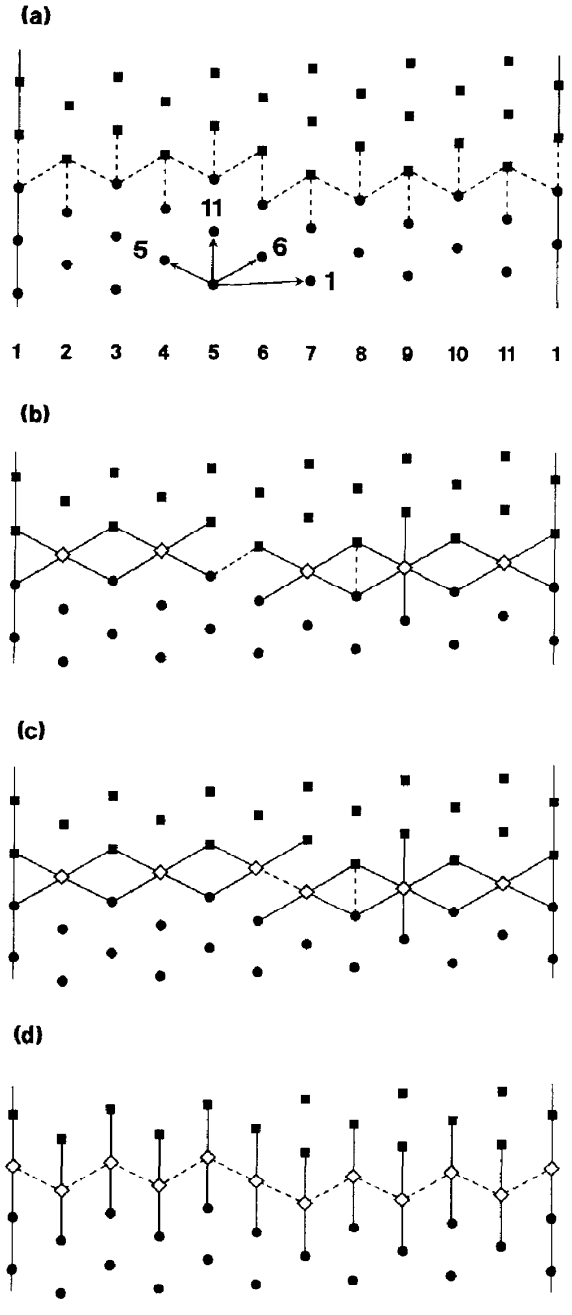

Figure 5. Insertion of subunits of a junction protein J (◇) between 2 extensive cylindrical structures such as the bacterial flagellar hook and filament, composed of distinct proteins A (●) and B (■). All 3 types of proteins are assumed to lie on similar helical lattices with close to 5·5 subunits per turn of the basic (1-start) helix, generating a cylindrical pseudohexagonal lattice with nearest-neighbor interactions occurring on the 5, 6 and 11-start helical directions (arrows). (The lattice is represented here by the planar lattice that would result if the cylinder were slit parallel to the cylinder axis and opened out.). Homologous A–A and B–B interactions are considered to be favorable, as are heterologous A–J and J–B interactions, and are represented by continuous lines. Conversely, heterologous A–B interactions and homologous J–J interactions are considered to be unfavorable, and are represented by broken lines. (a) No subunits of junction protein J: a total of 22 unfavorable heterologous interactions (5 in the 5-start direction, 6 in the 6-start direction, and 11 in the 11-start direction) exist at the junction; (b) 5 subunits of J: all unfavorable A–B interactions in the 5-start direction are eliminated, as are all but one in the 6-start direction; 6 out of 11 remain in the 11-start direction (only one example each of an unfavorable and a favorable interaction are shown explicitly); (c) 6 subunits of J: all unfavorable A–B interactions in the 5 and 6-start directions are eliminated, but one unfavorable J–J interaction in the 5-start direction is introduced; 5 out of 11 unfavorable A–B interactions remain in the 11-start direction; (d) 11 subunits of J: all unfavourable A–B interactions in the 5, 6 and 11-start directions are eliminated, but 11 unfavorable J–J interactions (5 in the 6-start direction, 6 in the 5-start direction) are introduced.

junction protein (or proteins), whose lattice arrangement was the same as those of hook protein and flagellin, and whose structure was compatible with hook protein in the proximal direction and with flagellin in the distal direction. Thus, the design constraints for such heterologous interactions would be satisfied by adaptor molecules that did not have the additional constraints described above for either hook or filament. Such considerations probably explain the existence of HAPs 1 and 3 at the hook-filament junction (Ikeda *et al.*, 1987).

What would be the consequences of the use of a given number of subunits of a junction protein J (say, HAP1) between structures composed of proteins A and B (say, hook protein and flagellin)? We assume that the interactions of importance occur in the 5, 6 and 11-start directions. We also assume that homologous interactions are favorable in the case of A and B but unfavorable in the case of J, whereas heterologous interactions are unfavorable between A and B (Fig. 5(a)), but favorable between A and J and between J and B.

The use of either five subunits of J (Fig. 5(b)) or six subunits of J (Fig. 5(c)) largely eliminates potentially unfavorable interactions in the 5 and the 6-start directions but only eliminates about half of those in the 11-start direction. The use of 11 subunits eliminates all potentially unfavorable interactions in the 11-start direction but introduces ones in the 5 and 6-start directions. The optimal choice for the number of subunits of J will thus depend on the detailed energetics of the combined homologous and heterologous interactions in the various directions.

As discussed above, HAPs 1 and 3 probably occupy the same general lattice arrangement as hook protein and flagellin (cf. Ikeda *et al.*, 1987). Circumstantial evidence suggests that the rod, too, is constructed with the same lattice arrangement, since its component proteins show strong primary sequence similarities to hook protein and to HAP1 (Homma *et al.*, 1990*a*,*b*). The HAPs and rod proteins are present in the flagellum in relatively small numbers, which can be interpreted in terms of numbers of turns of a basic helix with 5·5 subunits per turn.

The stoichiometry of HAP1 is statistically consistent with two turns of such a helix (11 subunits), suggesting that interactions in the 11-start direction dominate over 5 or 6-start interactions at the hook-filament junction (Fig. 5(d)). The importance of 11-

start interactions in both the hook and the filament structures is underscored by the fact that their macroscopic helical waveforms, which are essential for motility, result from a quasi-equivalence of quaternary interactions among subunits and a co-operativity in the 11-start direction (Calladine, 1982; Kamiya *et al.*, 1982). In the study that first revealed the existence of the HAPs, Homma *et al.* (1984*b*) concluded that HAP1 and HAP3 were present in essentially equal amounts. We postulate that the hook–filament junction zone, which is constructed from HAPs 1 and 3 (Ikeda *et al.*, 1987), consists of two successive two-turn segments of these proteins. The existence of two distinct proteins at the junction probably indicates that the hook and filament structures are too different for a single species of protein to provide a suitable interface between them.

Homma *et al.* (1984*b*) also estimated that the number of subunits of the capping protein, HAP2, was about half that of either HAP1 or HAP3. This would place its stoichiometry at about 5 or 6. If it occupies a similar lattice as flagellin, this would correspond to one turn of basic helix. However, given its location at the filament tip (Homma & Iino, 1985), HAP2 cannot be viewed as an adaptor molecule like HAPs 1 or 3, and so the assumptions made above for the lattice geometry do not necessarily apply to it. In view of the remarkably flat appearance of the cap (Ikeda *et al.*, 1985), and its role in permitting insertion of newly exported flagellin monomers into growing filament (Homma *et al.*, 1984*a*,*b*), it might have rotational rather than helical symmetry.

We next consider the rod structure. The stoichiometry of the distal rod protein is consistent with about four turns, although beyond two turns there is no reason to assume that the number must be integral, or for that matter well-defined; this issue is discussed further below. The stoichiometries of the three proximal rod proteins are all consistent with either just less than one turn (5 subunits) or just greater than one turn (6 subunits) indicating that, within the rod, 5 and 6-start interactions are more important than 11-start interactions. It is interesting in this regard to note that, since the rod is straight, subunit interactions are equivalent rather than just quasi-equivalent and co-operativity in the 11-start direction is therefore not an obligatory feature of the structure as it is in the hook or filament.

If we assume that the rise per subunit for the distal rod protein is about 0·4 nm (the value for the hook protein), its observed stoichiometry would result in a length of about 11 nm for the distal portion of the rod, consistent with the estimate of <20 nm for this feature by Okino *et al.* (1989). By combining the stoichiometries of all four rod proteins and again assuming a rise per subunit of about 0·4 nm, a length estimate of about 18 nm is obtained. This is approximately the distance from the hook–rod junction to the proximal face of the P ring, and is considerably smaller than the estimate of 27·5 nm for the distance between the M and L rings (Stallmeyer *et al.*, 1989). Other rod components may remain to be identified.

Structures that have helical symmetry can, in principle, assemble indefinitely. This in fact happens in the flagellar filament, where the only constraint on length seems to be a logistical one (Iino, 1974). However, all of the other known or presumed helical components are present with fairly closely defined stoichiometries. In the case of the hook, the controlling mechanism seems to be in the assembly process, since mutants with defects in the *fliK* gene (which is not the hook structural gene) grow hooks of indefinite length (Patterson-Delafield *et al.*, 1973; Suzuki & Iino, 1981). Thus, it seems that the stoichiometry of the hook protein (like that of flagellin) is not constrained by structural considerations.

The stoichiometry of the distal rod protein is sufficient to generate roughly four turns of the basic helix (assuming the lattice geometry is similar to that of the hook and filament). Once more than two turns of the basic helix have been generated, it is unclear whether structural constraints could result in termination at a well-defined length. An accumulated-strain mechanism (such that addition of successive subunits becomes progressively unfavorable) does not seem plausible for a structure like the rod, which has to bear a heavy torsional load as a result of motor rotation. It seems likely therefore that the length of the FlgG portion of the rod is determined by considerations (for example, interaction with the L ring or the outer membrane) other than its own internal structure.

Thus, among the axial components of the HBB complex, it is likely that only the HAPs and minor rod proteins have stoichiometries that are structurally constrained, the constraints arising from the bonding rules discussed above.

(c) *Stoichiometries of ring components*

On the basis of their appearance and their inferred functions, the basal-body rings almost certainly have rotational rather than helical symmetry, and consist of single annuli of subunits. We make this assumption in interpreting the stoichiometries of the ring proteins, which are then absolutely constrained by geometrical considerations, i.e. by the bonding angles between subunits. This is in contrast to the geometrically unconstrained stoichiometries of structures with helical symmetry, discussed in the previous section.

The stoichiometries of the M, P and L-ring proteins are (within experimental error) the same. Identical stoichiometries for the P and L-ring proteins can be readily interpreted in terms of their joint formation of the P–L ring outer cylinder, with the same rotational symmetry for its two structural components. It is not obvious, however, why the M ring and the outer cylinder should possess the same symmetry, since the only known interactions between them are indirect (*via* the rod). It may

simply reflect the fact that they are annular structures of about the same size. Or it might reflect interactions *via* structures that have not yet been detected; for example, the nature of the anchoring of the flagellar motor to the cell surface, necessary to permit the generation of torque externally, is not yet known.

How do the observed ring-protein stoichiometries relate to the geometry of the basal body? The P and L rings have their centers of mass located at around 12 nm radius (Stallmeyer *et al.*, 1989). A spherical 30 kDa protein could then exist in a ring of about 18 subunits. The observed stoichiometry of about 26 is readily reconciled with this value by assuming that the detailed subunit shape permits closer packing. Especially in the case of the P and L rings, the subunits may also have some prolate character with the long axis roughly parallel to the flagellar axis (Stallmeyer *et al.*, 1989).

How the M and S rings interact with the rod is unclear (Stallmeyer *et al.*, 1989), but it seems reasonable to assume that at least one of them rotates rigidly with it, in which case there must be strong quaternary interactions between the two types of structure and their respective symmetries must interface in a rather specific manner.

(d) *Ring protein stoichiometries* versus *other flagellar periodicities*

We consider briefly other observations of structural or dynamic properties of the flagellum that exhibit periodicities, and compare these with the molecular stoichiometry of approximately 26 that we have measured for the M, P and L-ring proteins.

Inspection of electron microscopic images of isolated rings led DePamphilis & Adler (1971) to estimate a rotational symmetry of approximately 16; however, the images showed considerable elliptical distortion, and the method of analysis (simple rotational reinforcement) is not a reliable one. Stallmeyer *et al.* (1989) made a rough estimate of 12 for the rotational symmetry of axial views of the P–L ring outer cylinder, but not all of their images exhibited rotational symmetry, and those that did only provided it for a portion of the circumference. The inferences from both of these studies must therefore be regarded as tentative but, if approximately correct, could be reconciled with our observed molecular stoichiometries by assuming that the asymmetrical subunit is a molecular dimer.

Rings of particles, with a rotational symmetry of approximately 11, have been observed in freeze-fracture images of the cell membrane of *E. coli* (Khan *et al.*, 1988) and are thought to be constructed from the MotA and MotB proteins. The nature of the interaction between these particles and the rest of the flagellum is not well understood. Specifically, it is not known whether they interact with the M ring. If they do, the differences in stoichiometry between the particles and M-ring proteins indicate that only about every second or third M-ring protein could be in equivalent interaction with a particle.

The experiments of Block & Berg (1984) and, more recently, Blair & Berg (1988) indicate that a maximum of approximately eight force-generating units can operate within the motor, but that a lesser number will still work (with a corresponding reduction in torque). Thus, rotation does not require a fixed stoichiometry of force-generating units. It is easy therefore to imagine that it is the physical size of the force-generating units, *versus* the space available for their interaction with the basal body, that determines the maximum number that can be accommodated.

(e) *Concluding comments*

The HBB complex consists of at least nine proteins organized into subassemblies such as hook, rod and rings, which have different geometries and subunit stoichiometries. We have suggested constraints that may be responsible for some of these stoichiometries, but a number of puzzles remain. How are the lengths of the hook and the distal portion of the rod determined? If the rod has helical symmetry, how is it joined to the rings, which have rotational symmetry? (Or if the rod has rotational symmetry, how is it joined to the hook, which has helical symmetry?) Stoichiometric analysis can raise questions such as these but ultimately structural analysis will be needed to answer them.

It must also be realized that the present study does not completely define the stoichiometry of the basal body, which contains other proteins whose genes remain to be identified (Fig. 3). Furthermore, the flagellar motor undoubtedly contains components in addition to the basal body. A major missing piece of information concerns the flagellar switch, which is believed to be a complex consisting of the FliG, FliM and FliN proteins (Yamaguchi *et al.*, 1986) mounted on the M ring, but which has not yet been visualized, or isolated and characterized.

We are grateful to M. Homma, D. DeRosier and K. Kutsukake for permitting us to use translated DNA sequences prior to publication; to S. Kanto for assistance with the amino acid sequence analysis; to H. Hotani for encouragement; to E. Egelman for advice with the densitometry; and to D. DeRosier for helpful discussions and critical comment on the manuscript. This work was supported by USPHS grant AI12202 to R.M.M.

References

Aizawa, S.-I., Dean, G. E., Jones, C. J., Macnab, R. M. & Yamaguchi, S. (1985). *J. Bacteriol.* **161**, 836–849.

Aswad, D. & Koshland, D. E., Jr (1975). *J. Mol. Biol.* **97**, 225–235.

Blair, D. F. & Berg, H. C. (1988). *Science*, **242**, 1678–1681.

Block, S. M. & Berg, H. C. (1984). *Nature (London)*, **309**, 470–472.

Calladine, C. R. (1982). In *Prokaryotic and Eukaryotic Flagella* (Amos, W. B. & Duckett, J. G., eds), pp. 33–51. Cambridge University Press, Cambridge, U.K.

DePamphilis, M. L. & Adler, J. (1971). *J. Bacteriol.* **105**, 384–395.

Hilmen, M. & Simon, M. (1976). In *Cell Motility* (Goldman, R., Pollard, T. & Rosenbaum, J., eds), pp. 35–45. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Homma, M. & Iino, T. (1985). *J. Bacteriol.* **162**, 183–189.

Homma, M., Fujita, H., Yamaguchi, S. & Iino, T. (1984*a*). *J. Bacteriol.* **159**, 1056–1059.

Homma, M., Kutsukake, K., Iino, T. & Yamaguchi, S. (1984*b*). *J. Bacteriol.* **157**, 100–108.

Homma, M., Kutsukake, K. & Iino, T. (1985). *J. Bacteriol.* **163**, 464–471.

Homma, M., Aizawa, S.-I., Dean, G. E. & Macnab, R. M. (1987*a*). *Proc. Nat. Acad. Sci., U.S.A.* **84**, 7483–7487.

Homma, M., Komeda, Y., Iino, T. & Macnab, R. M. (1987*b*). *J. Bacteriol.* **169**, 1493–1498.

Homma, M., Ohnishi, K., Iino, T. & Macnab, R. M. (1987*c*). *J. Bacteriol.* **169**, 3617–3624.

Homma, M., Kutsukake, K., Hasebe, M., Iino, T. & Macnab, R. M. (1990*a*). *J. Mol. Biol.* **211**, 465–478.

Homma, M., DeRosier, D. J. & Macnab, R. M. (1990*b*). *J. Mol. Biol.* In the press.

Horiguchi, T., Yamaguchi, S., Yao, K., Taira, T. & Iino, T. (1975). *J. Gen. Microbiol.* **91**, 139–149.

Iino, T. (1974). *J. Supramol. Struct.* **2**, 372–384.

Iino, T., Komeda, Y., Kutsukake, K., Macnab, R. M., Matsumura, P., Parkinson, J. S., Simon, M. I. & Yamaguchi, S. (1988). *Microbiol. Rev.* **52**, 533–535.

Ikeda, T., Asakura, S. & Kamiya, R. (1985). *J. Mol. Biol.* **184**, 735–737.

Ikeda, T., Homma, M., Iino, T., Asakura, S. & Kamiya, R. (1987). *J. Bacteriol.* **169**, 1168–1173.

Jones, C. J. & Macnab, R. M. (1990). *J. Bacteriol.* In the press.

Jones, C. J., Homma, M. & Macnab, R. M. (1987). *J. Bacteriol.* **169**, 1489–1492.

Jones, C. J., Homma, M. & Macnab, R. M. (1989). *J. Bacteriol.* **171**, 3890–3900.

Kamiya, R., Hotani, H. & Asakura, S. (1982). In *Prokaryotic and Eukaryotic Flagella* (Amos, W. B. & Duckett, J. G., eds), pp. 53–76, Cambridge University Press, Cambridge, U.K.

Khan, S., Dapice, M. & Reese, T. S. (1988). *J. Mol. Biol.* **202**, 575–584.

Kutsukake, K., Suzuki, T., Yamaguchi, S. & Iino, T. (1979). *J. Bacteriol.* **140**, 267–275.

Kutsukake, K., Iino, T., Komeda, Y. & Yamaguchi, S. (1980). *Mol. Gen. Genet.* **178**, 59–67.

Macnab, R. M. (1987*a*). In Escherichia coli *and* Salmonella typhimurium: *Cellular and Molecular Biology*, (Neidhardt, F. C., Ingraham, J., Low, K. B., Magasanik, B., Schaechter, M. & Umbarger, H. E., eds), vol. 1, pp. 70–83, American Society for Microbiology Publications, Washington, DC.

Macnab, R. M. (1987*b*). In Escherichia coli *and* Salmonella typhimurium: *Cellular and Molecular Biology*, (Neidhardt, F. C., Ingraham, J., Low, K. B., Magasanik, B., Schaechter, M. & Umbarger, H. E., eds), vol. 1, pp. 732–759, American Society for Microbiology Publications, Washington, DC.

Macnab, R. M. & Aizawa, S.-I. (1984). *Annu. Rev. Biophys. Bioeng.* **13**, 51–83.

O'Brien, E. J. & Bennett, P. M. (1972). *J. Mol. Biol.* **70**, 133–152.

Ohnishi, K., Homma, M., Kutsukake, K. & Iino, T. (1987). *J. Bacteriol.* **169**, 1485–1488.

Okino, H., Isomura, M., Yamaguchi, S., Magariyama, Y., Kudo, S. & Aizawa, S.-I. (1989). *J. Bacteriol.* **171**, 2075–2082.

Patterson-Delafield, J., Martinez, R. J., Stocker, B. A. D. & Yamaguchi, S. (1973). *Arch. Mikrobiol.* **90**, 107–120.

Shirakihara, Y. & Wakabayashi, T. (1979). *J. Mol. Biol.* **131**, 485–507.

Stallmeyer, M. J. B., Aizawa, S.-I., Macnab, R. M. & DeRosier, D. J. (1989). *J. Mol. Biol.* **205**, 519–528.

Suzuki, T. & Iino, T. (1981). *J. Bacteriol.* **148**, 973–979.

Suzuki, T., Iino, T., Horiguchi, T. & Yamaguchi, S. (1978). *J. Bacteriol.* **133**, 904–915.

Trachtenberg, S. & DeRosier, D. J. (1987). *J. Mol. Biol.* **195**, 581–601.

Vogel, H. J. & Bonner, D. M. (1956). *J. Biol. Chem.* **218**, 97–106.

Wagenknecht, T., DeRosier, D. J., Aizawa, S.-I. & Macnab, R. M. (1982). *J. Mol. Biol.* **162**, 69–87.

Wilson, M. L. & Macnab, R. M. (1988). *J. Bacteriol.* **170**, 588–597.

Yamaguchi, S., Aizawa, S.-I., Kihara, M., Isomura, M., Jones, C. J. & Macnab, R. M. (1986). *J. Bacteriol.* **168**, 1172–1179.

Edited by D. DeRosier