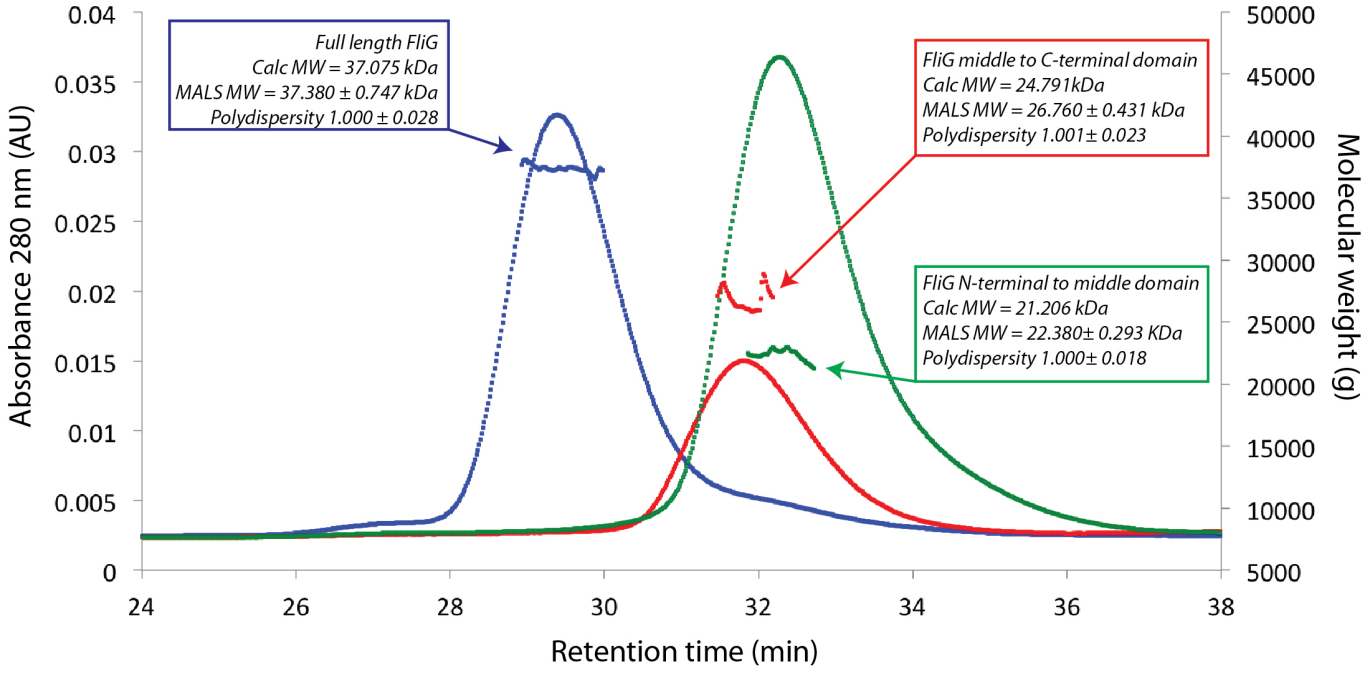
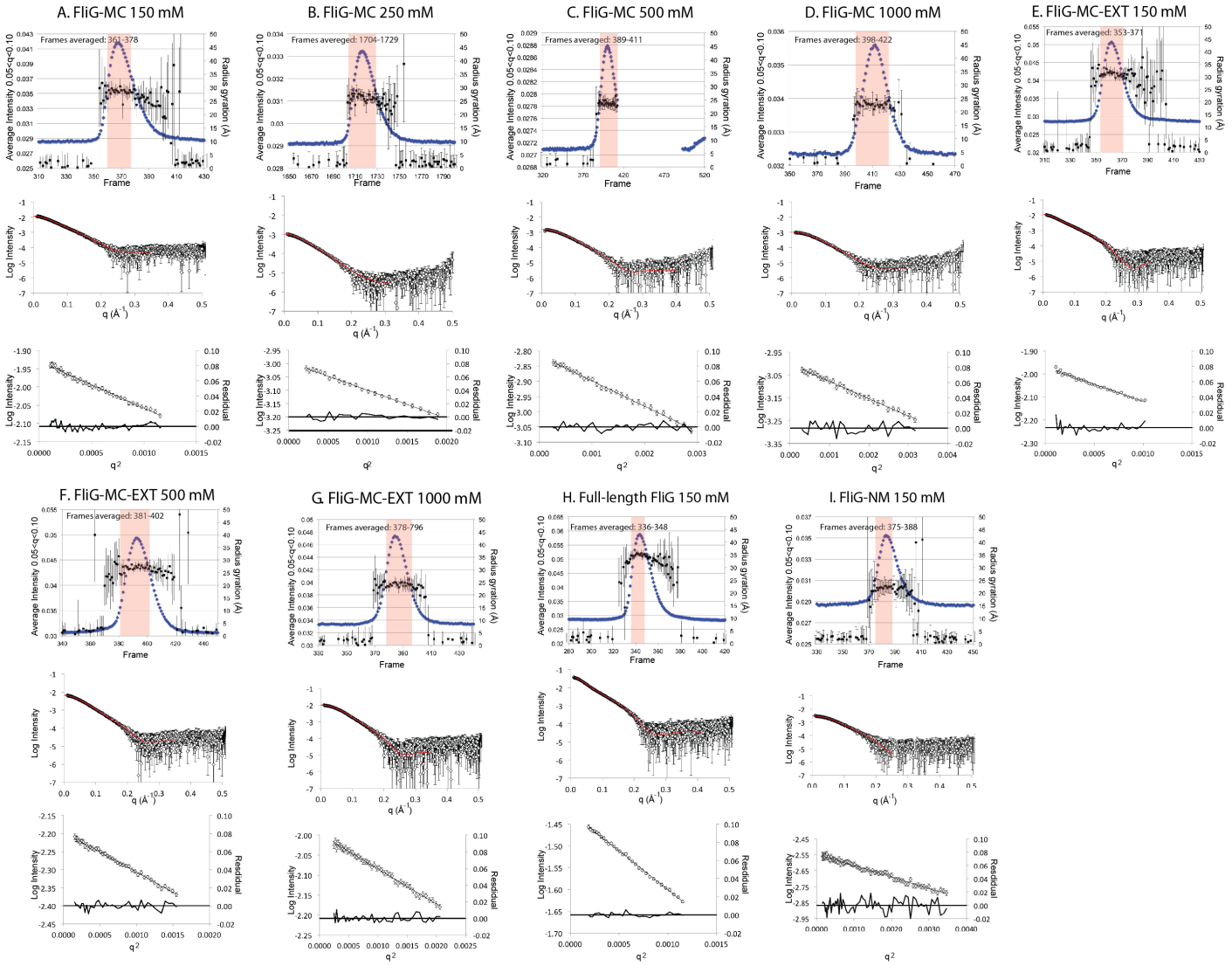


Supplementary Figure 1

Multiangle light-scattering experiments.

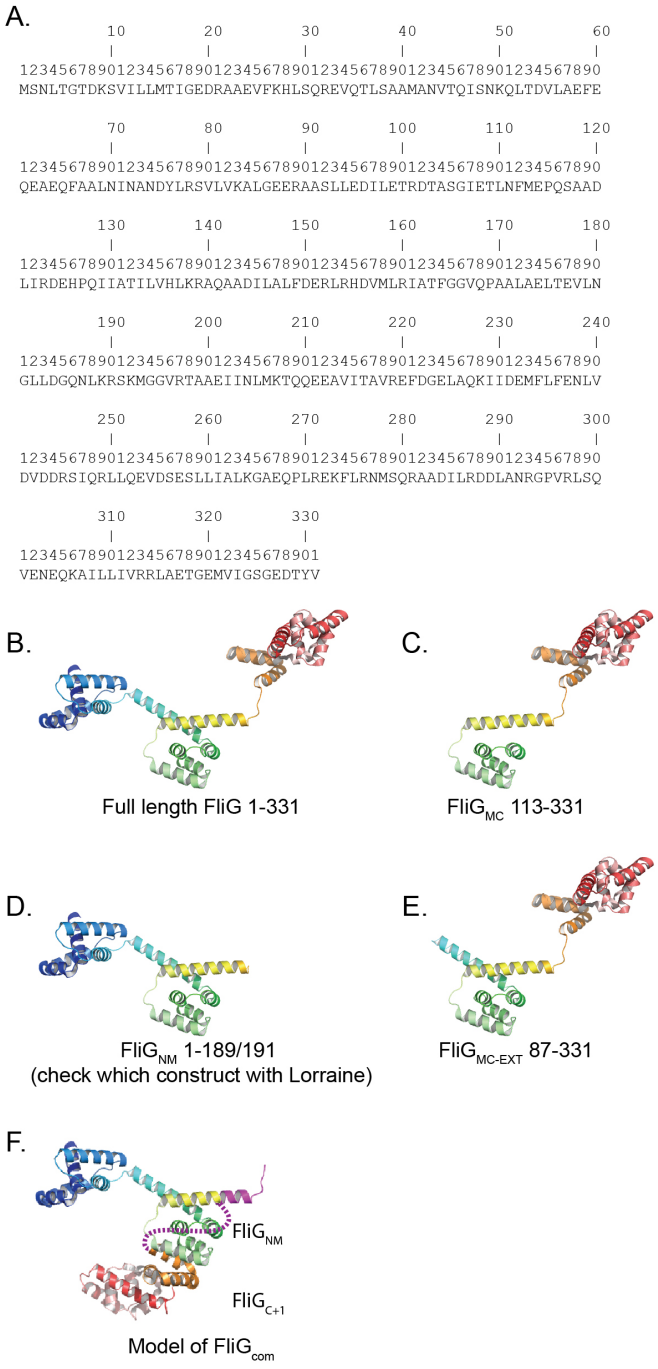
In line size exclusion chromatography and multi angle light scattering (SEC-MALS) experiments were performed on various FliG constructs. 0.5 ml of purified protein at a concentration of 5 mg /ml was injected onto a 23 ml Superdex S200 size exclusion chromatography column with a flow rate of 0.5 ml/min. For all constructs, the majority of protein eluted in a single dominant elution peak. The absorbance at 280 nm and the molecular weight deduced from light scattering data from these peaks are shown for full-length FliG (blue), $FliG_{MC}$ (red) and $FliG_{NM}$ (green). Average experimentally determined molecular weight and polydispersity are recorded in the boxes. The correspondence between experimental and theoretical molecular weights, and low measured polydispersity, demonstrate that FliG proteins in the major elution peak are monodisperse and monomeric.



Supplementary Figure 2

SAXS data validation.

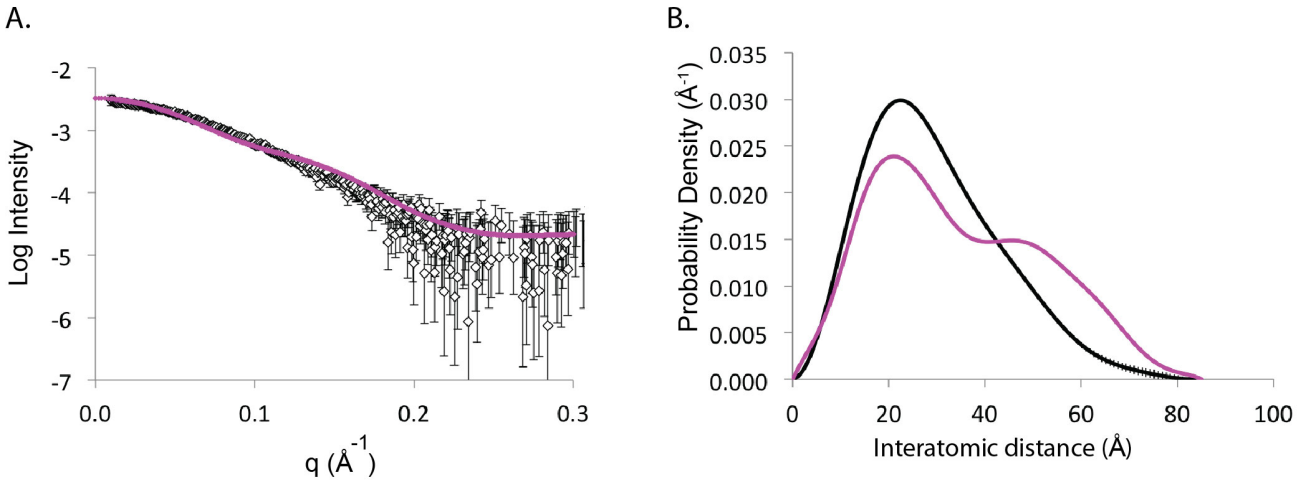
Plots of average low q intensity ($0.01 < q < 0.5$) and calculated radius gyration from background-subtracted data vs frame (top panel) where the frame indices that were scaled and averaged for further analysis are indicated at the top of the plot in text and with magenta overlay, subtracted scattering overlaid with fit used to generate P(r) vs r plots (middle panel) and linear fits to Guinier plots with residuals (bottom panel) are displayed for all datasets presented in this work: (A) $FliG_{MC}$ 150 mM NaCl, (B) $FliG_{MC}$ 250mM NaCl (C) $FliG_{MC}$ 500 mM NaCl, (D) $FliG_{MC}$ 1000 mM NaCl, (E) $FliG_{MC\text{-}EXT}$ 150 mM NaCl, (F) $FliG_{MC\text{-}EXT}$ 500 mM NaCl, (G) $FliG_{MC\text{-}EXT}$ 1000 mM NaCl, (H) full-length FliG 150 mM NaCl and (I) $FliG_{NM}$ in 150 mM NaCl.



Supplementary Figure 3

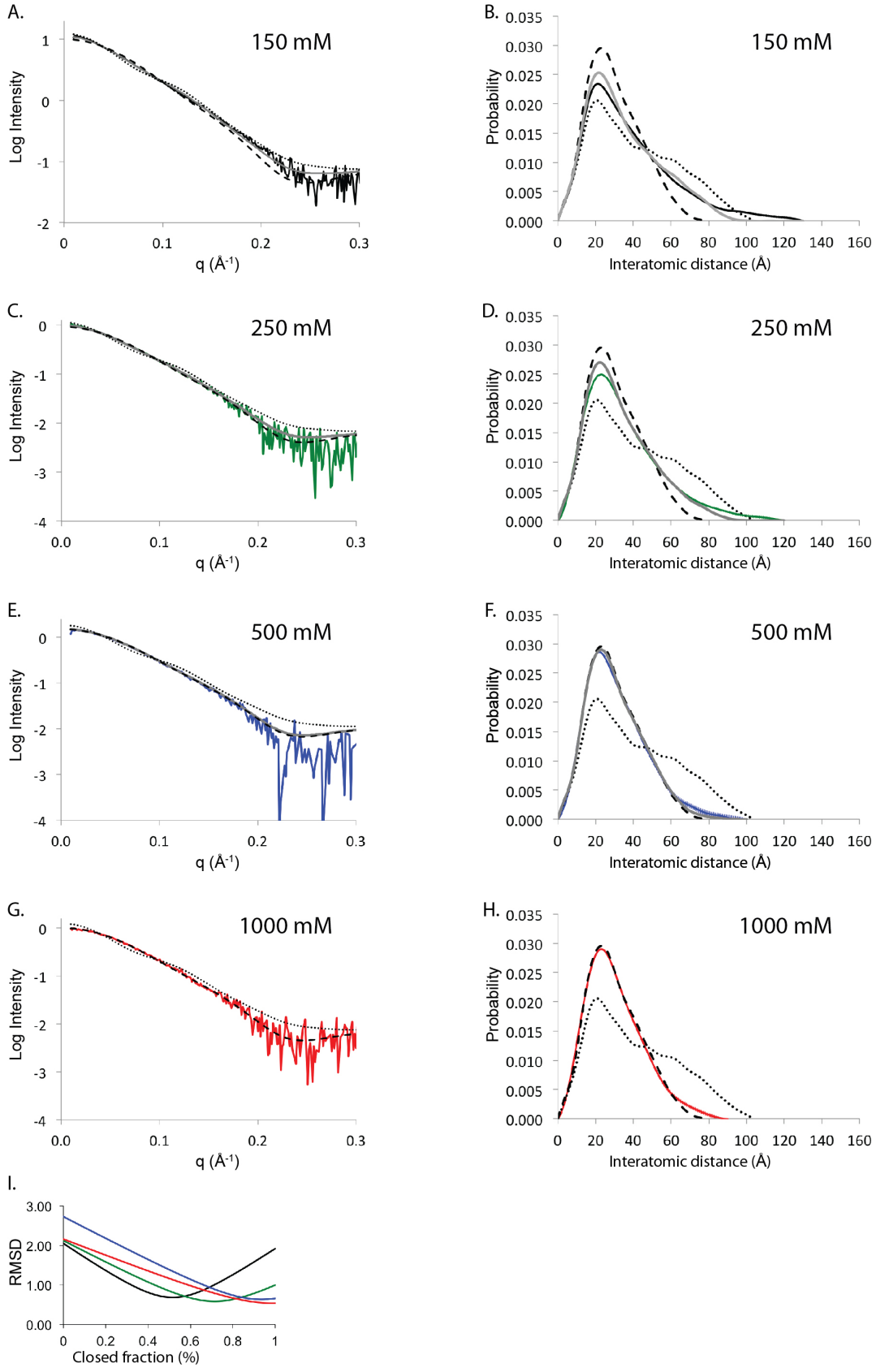
Details of protein constructs used in SAXS experiments.

The sequence of FliG from *E. coli* is shown in (A). Panels (B-E) show the amino acids in the $FliG_{FL}$, $FliG_{MC}$, $FliG_{NM}$, and $FliG_{MC-EXT}$ constructs and the equivalent truncated structure from the *A. aeolicus* crystal structure (PDBID:3HJL) respectively. The model of the compact conformation of FliG used in the SAXS analysis is shown in (F). This was generated from $FliG_{NM}$ and $FliG_{C}$ domains from neighbouring molecules that form the ARM_{M}-ARM_{C} interface in the crystal. For the protein to adopt the compact conformation, the amino acids shown in magenta must unravel and were not included in structural analysis.



Supplementary Figure 4

Comparison of solution and crystal structures of $FliG_{NM}$.
Scattering data (A) and atom pair distribution profile (B) of $FliG_{NM}$ from SAXS measurements (black) and calculated from the $FliG_{NM}$ portion of the FliG crystal structure from *A. aeolicus* (magenta) (PDBID: 3HJL). The bimodal distribution in (B) from crystal structure but not SAXS data, indicates that in solution the two distinct $FliG_N$ and $FliG_M$ domains seen in the crystal structure form a single domain in solution.

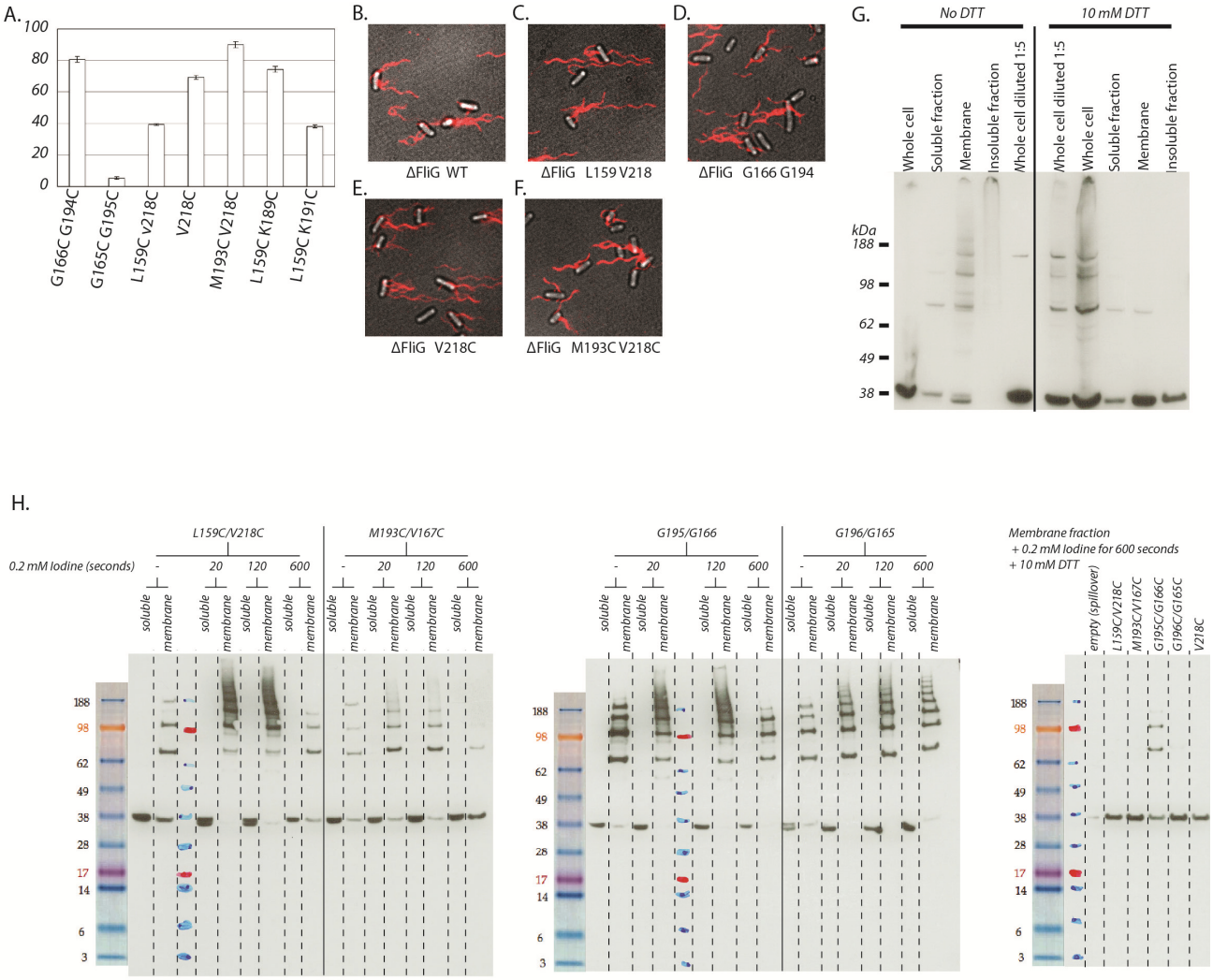


Nature Structural & Molecular Biology: doi:10.1038/nsmb.3172

Supplementary Figure 5

Estimation of fraction of FliG in an extended conformation in different NaCl concentrations.

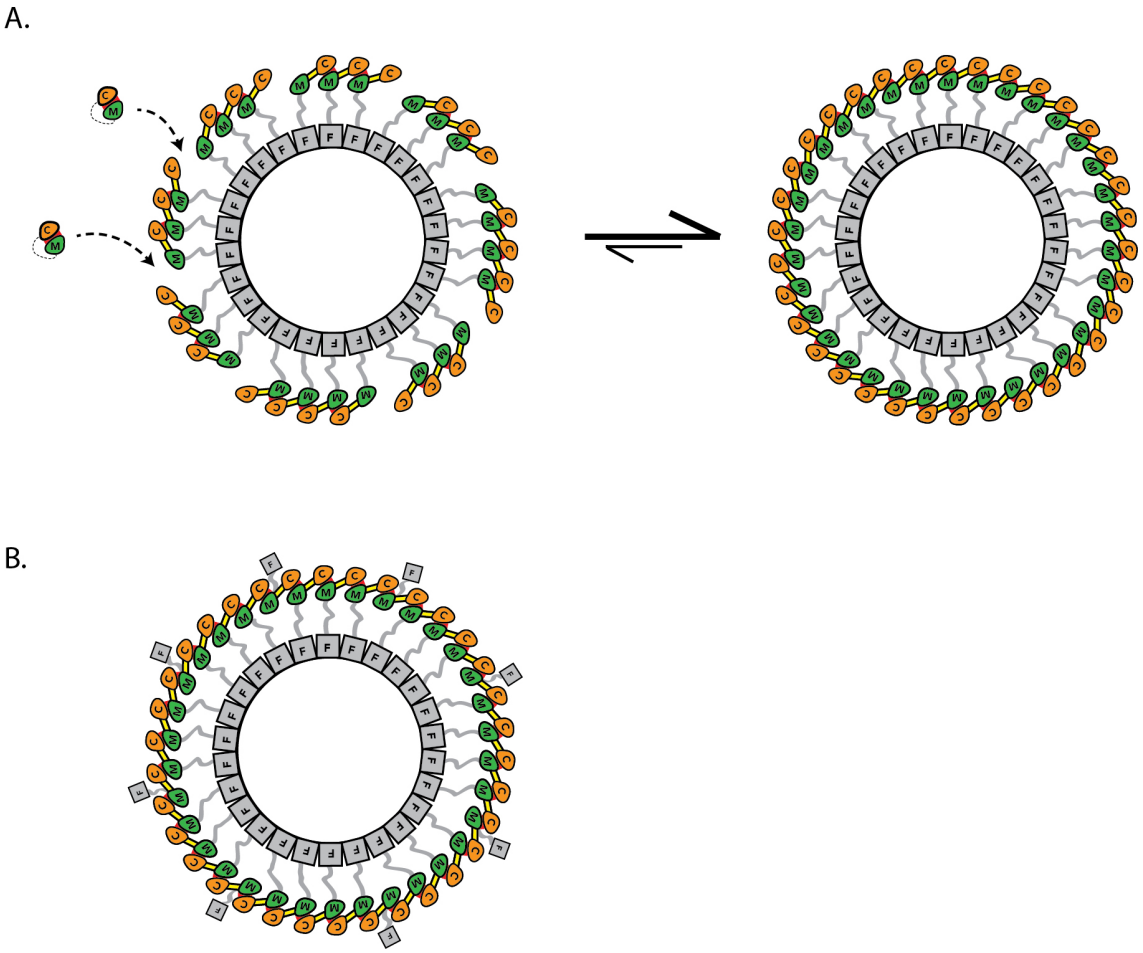
Background-subtracted SAXS scattering (left panels) and corresponding interatomic distance distributions (right panels) from $FliG_{MC}$ collected in 150 mM (A and B), 250 mM (C and D), 500 mM (E and F) and 1000 mM NaCl (G and H) with experimental data in black, green, blue and red for each NaCl concentration respectively. Theoretical scattering for the compact and extended conformation are shown as dashed and dotted lines respectively. Grey curves are generated from a linear combination of these theoretical scattering curves: $I_{fit}(q) = f \times I_{compact}(q) + (1-f) \times I_{extended}(q)$ where f represents the fractional contribution from the compact monomer. The fit was optimised by minimising the root mean squared deviation from the experimental scattering curve. (I) shows a plot of normalized RMSD vs f in each NaCl concentration The best fits correspond to a fractional population of the extended conformation of 0.48, 0.38, 0.07 and 0.01 for NaCl concentrations of 150 mM, 250 mM, 500 mM and 1000 mM NaCl respectively.



Supplementary Figure 6

Functional characterization of FliG cysteine mutants.

A) Swarm diameter of ΔFliG cells with mutant FliG expression plasmids displayed as a percentage of swarm diameter for wild-type FliG. B-F) Fluorescence images of fluorophore-stained flagellar filaments overlaid with bright-field image depicting cell body of ΔFliG cells with WT and mutant FliG expression plasmids in Salmonella. G) Immunoblots of L159C V218C mutants performed in conditions previously published (ref 27 main text) with FliG protein expressed in T7 expression vectors in cells grown in rich LB media overnight are shown in the presence and absence of DTT in (C). In keeping with previous results, in these conditions, no higher order molecular weight species are observable in whole cell samples. However upon fractionation, FliG ladders are seen in the membrane fraction but not the soluble fraction. Interestingly, there is a substantial amount of insoluble crosslinked aggregate that is too large to enter the polyacrylamide gel but which can be observed upon the addition of DTT. Finally, in these conditions, protein expression levels are substantially higher than the conditions used in the rest of this study. A 1 in 5 dilution shows that the quantity of protein is well beyond the dynamic range of the immunoblot. H. Shows uncropped SDS PAGE gels for all immunoblots presented in Figure 3.



Supplementary Figure 7

Symmetry mismatch.

The FliF template exists as a 26-fold symmetry ring, while the C-ring, consisting of FliG, FliM and FliN, exists as a 45 nm 34-fold symmetric ring. Thus there exists a mismatch in the symmetry between the two rings. This could manifest itself as gaps in the FliG ring (A, left), but our model predicts that there is a strong thermodynamic drive to fill these gaps (A, right). If these gaps are indeed plugged by FliG, then in the functional FliFG fusion construct there could be 8 dangling FliF subunits that are not incorporated into the MS-ring template, but are fused to the FliG molecules that are plugging gaps in the C-ring (B).

Supplementary Table 1. Summary of derived SAXS structural parameters

	$FliG_{FL}$	**$FliG_{MC}$**				**$FliG_{MC-EXT}$**			**$FliG_{NM}$**
[NaCl] (mM)	150	150	250	500	1000	150	500	1000	150
Data collection									
Wavelength (Å)	1.127	1.127	1.127	1.127	1.127	1.127	1.127	1.127	1.127
Flux (photons/sec)	4×10^{12}	4×10^{12}	4×10^{12}	4×10^{12}	4×10^{12}	4×10^{12}	4×10^{12}	4×10^{12}	4×10^{12}
Temperature (K)	293	293	293	293	293	293	293	293	293
Exposure time (sec)	5	5	5	5	5	5	5	5	5
Beam dimensions (FWHM μm)	250 × 150	250 × 150	250 × 150	250 × 150	250 × 150	250 × 150	250 × 150	250 × 150	250 × 150
Data Processing									
Q×Rg range for Guiner	0.48 – 1.19	0.52 – 1.16	0.40 – 1.17	0.39 – 1.29	0.40 – 1.29	0.43 – 1.19	0.34 – 1.09	0.24 – 1.31	0.78 – 1.34
Q range for P(r)	0.013 – 0.417	0.011 – 0.384	0.017 – 0.307	0.016 – 0.396	0.017 – 0.332	0.012 – 0.345	0.012 – 0.345	0.017 – 0.326	0.012 – 0.252
I(0)(cm^{-1})[from P(r)]	$3.8\times10^{-2}\pm1.7\times10^{-5}$	$1.2\times10^{-2}\pm5.2\times10^{-5}$	$1.2\times10^{-2}\pm8.5\times10^{-6}$	$1.5\times10^{-3}\pm1.2\times10^{-5}$	$1.0\times10^{-3}\pm7.3\times10^{-6}$	$1.1\times10^{-2}\pm3.7\times10^{-5}$	$6.4\times10^{-3}\pm2.8\times10^{-5}$	$1.0\times10^{-2}\pm3.6\times10^{-5}$	$2.9\times10^{-3}\pm2.5\times10^{-5}$
Rg (Å) [from P(r)]	37.2 ± 0.3	31.4 ± 0.3	28.8 ± 0.5	24.8 ± 0.5	24.1 ± 0.3	33.5 ± 0.2	29.4 ± 0.4	24.9 ± 0.2	23.1 ± 0.3
Rg (Å) [from Gunier]	35.1± 0.2	*not interpretable*	27.0 ± 0.4	24.0 ± 0.3	22.9 ± 0.3	*not interpretable*	27.7 ± 0.3	24.6 ± 0.3	22.3 ± 0.3
Dmax (Å)	150	130	120	100	90	130	130	90	83.5
Porod volume ($Å^3$)	74744	47905.4	47393.3	53744.4	46047.1	56058.2	59749.7	44713.1	37237.8
Molecular mass estimates									
MW (kDa) [theoretical]	37.1	24.8	24.8	24.8	24.8	27.4	27.4	27.4	21.2
MW (kDa) [MALS]	37.4 ± 0.7	26.8 ± 0.4	*Not measured*	*Not measured*	*Not measured*	*Not measured*	*Not measured*	*Not measured*	22.4 ± 0.3
MW (kDa) [Porod]	44.0	28.2	27.9	31.6	27.1	33.0	35.1	26.3	21.9

Supplementary Notes

Supplementary Note 1: Derivation of equations describing the formation of an intramolecular ARM_M-ARM_C interaction as a pseudo bimolecular interaction.

We denote a linear oligomer comprising *n* FliG subunits, with two exposed ARM motifs and (*n*-1) intermolecular interfaces, as $FliG_n$ and define ΔG_p^0 as the standard state free-energy change for the chain-extension step:

$$FliG_n + FliG_{ext} \xrightarrow{\Delta G_p^0} FliG_{n+1}. \quad (1)$$

ΔG_p^0 is assumed to be independent of *n* and is defined by:

$$\Delta G_p^0 \stackrel{\text{def}}{=} -k_B T \ln\left(\frac{[FliG_{n+1}]^*}{[FliG_n]^*[FliG_{ext}]^*}\right), \quad (2)$$

where [*X*]* is the equilibrium concentration of *X* relative to the standard state concentration of 1 M, k_B is Boltzmann's constant and *T* the absolute temperature. Similarly, the standard free energy change associated with the monomer transition from an extended to a closed state ($\Delta G_{e\to c}^0$):

$$FliG_{ext} \xrightarrow{\Delta G_{e\to c}^0} FliG_{com}. \quad (3)$$

where $\Delta G_{e\to c}^0$ is defined by:

$$\Delta G_{e\to c}^0 \stackrel{\text{def}}{=} -k_B T \ln\left(\frac{[FliG_{com}]^*}{[FliG_{ext}]^*}\right). \quad (4)$$

Equations 2 and 4 respectively are equivalent to the following equations in the main text:

$$\frac{[FliG_{n+1}]^*}{[FliG_n]^*[FliG_{ext}]^*} = \exp\left(-\frac{\Delta G_p^0}{k_B T}\right), \quad (5)$$

and

$$\frac{[FliG_{com}]^*}{[FliG_{ext}]^*} = \exp\left(-\frac{\Delta G_{e\to c}^0}{k_B T}\right). \quad (6)$$

In general, concentrations are not 1 M, and the free energy change for reaction 1 is:

$$\Delta G_p = \Delta G_p^0 + k_B T \ln\left(\frac{[FliG_{n+1}]}{[FliG_n][FliG_{ext}]}\right), \quad (7)$$

where [*X*] is the molar concentration of *X* which is not necessarily an equilibrium concentration. Similarly for an extended monomer undergoing a transition to a closed state:

$$\Delta G_{e\to c} = \Delta G_{e\to c}^0 + k_B T \ln\left(\frac{[FliG_{com}]}{[FliG_{ext}]}\right). \quad (8)$$

The intramolecular transition from extended to closed conformation is similar to the chain-extension reaction 1 in that the intramolecular ARM_M-ARM_C interaction that drives it is structurally very similar to the inter-molecular interaction that links together the polymer. We can consider the intramolecular conformational transition (Eq. 3) as a pseudo-bimolecular reaction between the two domains, $FliG_M$ and $FliG_C$,

and define an effective concentration c_{eff} of one domain of the extended monomer relative to the other. c_{eff} accounts for the difference between the free energy changes for formation of inter- and intramolecular ARM_M-ARM_C interactions in chain extension and monomer closing transitions respectively:

$$\Delta G_p^0 - \Delta G_{e\to c}^0 \stackrel{\text{def}}{=} k_B T \ln(c_{eff}). \quad (9)$$

We substitute equation 9 into equation 8 to give:

$$\Delta G_{e\to c} = \Delta G_p^0 + k_B T \ln\left(\frac{[FliG_{com}]}{[FliG_{ext}]c_{eff}}\right). \quad (10)$$

Thus we can consider the formation of an intramolecular ARM_M-ARM_C as a pseudo bimolecular reaction equivalent to chain extension, with an associated free energy change (cf. equation 7). In equilibrium:

$$\frac{[FliG_{com}]^*}{c_{eff}[FliG_{ext}]^*} = \exp\left(-\frac{\Delta G_p^0}{k_B T}\right), \quad (11)$$

which is analogous to equation 5.

Supplementary Note 2: Defining threshold concentrations for polymerization

Reaction 1 is the addition of an extended monomer to a polymer. The free energy change upon adding a closed monomer to a polymer ($\Delta G_{p,C}$) is:

$$\Delta G_{p,C} = \Delta G_p - \Delta G_{e\to c}. \quad (12)$$

Subtraction of equation 10 from equation 7 gives the free energy for adding a closed monomer to a polymer:

$$\Delta G_{p,C} = k_B T \ln\left(\frac{[FliG_{n+1}]}{[FliG_n]}\frac{C_{eff}}{[FliG_{com}]}\right). \quad (13)$$

Setting $\Delta G_{p,C} = 0$ in equation 13 gives:

$$\frac{[FliG_{n+1}]^*}{[FliG_n]^*} = \frac{[FliG_{com}]^*}{C_{eff}}, \quad (14)$$

for the system in equilibrium. Thus, c_{eff} sets a threshold concentration monomer concentration of $FliG_{com}$ such that if $[FliG_{com}] < c_{eff}$ polymers of different lengths n co-exist in a stable equilibrium at concentrations that decrease exponentially with length n. Otherwise, the equilibrium concentration is unbounded at large n and polymerization will continue until the closed monomer concentration falls below c_{eff}.

From equation (1), in the limit of excess monomer we can fix $[FliG_{ext}]$ and consider a steady-state system with an equilibrium between $FliG_n$ and $FliG_{n+1}$, defined by $\Delta G_p = 0$ in Eqn. 5:

$$k_B T \ln\left(\frac{[FliG_{n+1}]^*}{[FliG_n]^*}\right) = -\Delta G_p^0 + k_B T \ln([FliG_{ext}]). \quad (15)$$

Equation 11 allows us to express the threshold monomer concentration of $FliG_{ext}$ above which spontaneous polymer extension ($FliG_{n+1}$) is favored:

$$[FliG_{ext}]_{thresh} = exp\left(\frac{\Delta G_p^0}{k_B T}\right). \quad (16)$$

By adding the energy required to open a compact monomer, allows us to define the threshold concentration of $FliG_{com}$ for $[FliG_{com}]_{thresh}$ for polymerization, which is equal to c_{eff} as written in the main text:

$$[FliG_{com}]_{thresh} = \exp\left(\left(\Delta G_p^0 - \Delta G_{e\to c}^0\right)/k_B T\right) = c_{eff}. \quad (17)$$

Supplementary Note 3: Calculation of effective concentration of one domain of FliG at the location of its linked partner domain in an extended FliG molecule

We approximate the 12 amino acid peptide linking the Middle and C-terminal domains of FliG as as a freely-jointed chain with contour length L = 12×0.38 nm and Kuhn length K = 0.8 nm[52]. The effective concentration of one end of the chain at the other is:

$$c_{eff} \approx \frac{1}{N_A}\left(\frac{3}{2\pi E[\mathrm{R}^2]}\right)^{\frac{3}{2}} = 0.08\mathrm{M}$$

where $E[\mathrm{R}^2] = KL$ is the mean squared distance between the two ends.

Supplementary references

52. Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J. M. & Gaub, H. E. Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science* **276,** 1109–1112 (1997).