XML-IS

Our reference: YJTBI 5805 P-authorquery-v7

AUTHOR QUERY FORM

ELSEVIER	**Journal: YJTBI** **Article Number: 5805**	**Please e-mail or fax your responses and any corrections to:** **E-mail: corrections.essd@elsevier.macipd.com** **Fax: +44 1392 285878**

Dear Author,

Any queries or remarks that have arisen during the processing of your manuscript are listed below and highlighted by flags in the proof. Please check your proof carefully and mark all corrections at the appropriate place in the proof (e.g., by using on-screen annotation in the PDF file) or compile them in a separate list.

For correction or revision of any artwork, please consult http://www.elsevier.com/artworkinstructions.

Articles in Special Issues: Please ensure that the words 'this issue' are added (in the list and text) to any references to other articles in this Special Issue

Uncited references: References that occur in the reference list but not in the text – please position each reference in the text or delete it from the list.	
Missing references: References listed below were noted in the text but are missing from the reference list – please make the list complete or remove the references from the text.	
Location in article	**Query/remark** **Please insert your reply or correction at the corresponding line in the proof**
Q1	Please check e-mail id for correctness.

Electronic file usage
Sometimes we are unable to process the electronic file of your article and/or artwork. If this is the case, we have proceeded by:

☐ Scanning (parts of) your article ☐ Rekeying (parts of) your article ☐ Scanning the artwork

Thank you for your assistance.

Journal of Theoretical Biology ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

A molecular ruler mechanism for length control of extended protein structures in bacteria

J.P. Keener

Department of Mathematics, University of Utah, USA

ARTICLE INFO

Article history:
Received 11 June 2009
Received in revised form
8 December 2009
Accepted 11 December 2009

Keywords:
Flagellar assembly
Hook length control
Type III protein export

ABSTRACT

The lengths of the hook structure of flagellar motors and of the needle of the injectosome are both carefully controlled, by apparently similar mechanisms. In this paper we propose a novel mechanism for this length control and develop a mathematical model of this process which shows excellent agreement with published data on hook lengths.

The proposed mechanism for length control (described using biochemical nomenclature appropriate for hooks) is as follows: Hook growth is terminated when the C-terminus of the length control molecule FliK interacts with FlhB, the secretion gatekeeper. The probability of this interaction is an increasing function of the length of the hook for two reasons. First, FliK is secreted through the hook intermittently during hook growth. Second, the probability of interaction with FlhB is a function of the amount of time the C-terminus of a secreted FliK spends in the vicinity of FlhB. This time is short when the hook is short because the folding of FliK exiting the distal end of the hook acts to pull the FliK molecule through the hook rapidly. In contrast, this time is much longer when the hook is longer than the unfolded FliK polymer since movement through the tube is not enhanced by folding. Thus, it is much more likely that interaction will occur when the hook is long than when the hook is short.

1. Introduction

Q1 Every structure which could in principle be of any size but instead is found in a narrow distribution of sizes must be endowed with a size control mechanism. Two examples of such structures in the microbial world are the hook of the flagellum and the needle of the type III secretion injectosome. The bacterial flagellum is a molecular nanomachine that allows motility. The injectosome is a communication apparatus that allows bacteria docked on the surface of a eukaryotic cell membrane to inject effector proteins across the bacterial membranes and the eukaryotic cell membrane.

The injectosome and the flagellum have similar features (see Fig. 1 for a schematic diagram of flagellar motor structure). Both have a basal body embedded in the two bacterial membranes and contain a built-in type III secretion apparatus. In the injectosome, the structure directly above the basal body is the needle, while in the flagellum it is the hook. The hook functions like a universal joint to smoothly transmit torque from the motor to the filament, and the filament is the propellor. In both structures, the type III secretion apparatus exports the structural components to the distal end of the hollow growing structure where they polymerize. In the flagellum, it exports the hook subunits until the hook unit is the appropriate length (55 ± 6 nm), and then it exports filament subunits. For the injectosome, the needle is polymerized to its correct length (66 ± 8 nm), whereupon the type III secretion apparatus exports effector proteins.

The lengths of the hook and the injectosome are both genetically determined. Mutations of the *fliK* gene in *Salmonella typhimurium* commonly result in failure to switch secretion specificity from rod/hook-type to filament type substrate secretion. Consequently, many of the hook structures in mutants strains are much longer than in the wild type strain. Similarly, mutations in the *yscP* gene in *Yersinia* commonly lead to needles of indefinite length, and also to failure to secrete effectors.

Secretion specificity in flagellum is controlled by the protein FlhB, a membrane protein of the export apparatus which is localized at the base of the basal body. Substrate specificity switching is accompanied by proteolytic cleavage of the C-terminal cytoplasmic domain of FlhB, which occurs following binding of the C-terminus of the FliK with the C-terminus of FlhB (see Fig. 2). Similarly, the C-terminus of YscU, the homolog of FlhB in injectosomes, undergoes proteolytic cleavage following interaction with the C-terminus of YscP in *Yersinia*. Thus, the C-terminus region of both FliK and YscP acts as a substrate specificity switch. Indeed, when the type III secretion substrate specificity switch (T3S4) domain is overexpressed in *yscP* or *fliK* mutants, it restores the secretion switch, albeit with highly variable length.

E-mail address: keener@math.utah.edu

0022-5193/$ - see front matter © 2009 Published by Elsevier Ltd.
doi:10.1016/j.jtbi.2009.12.015

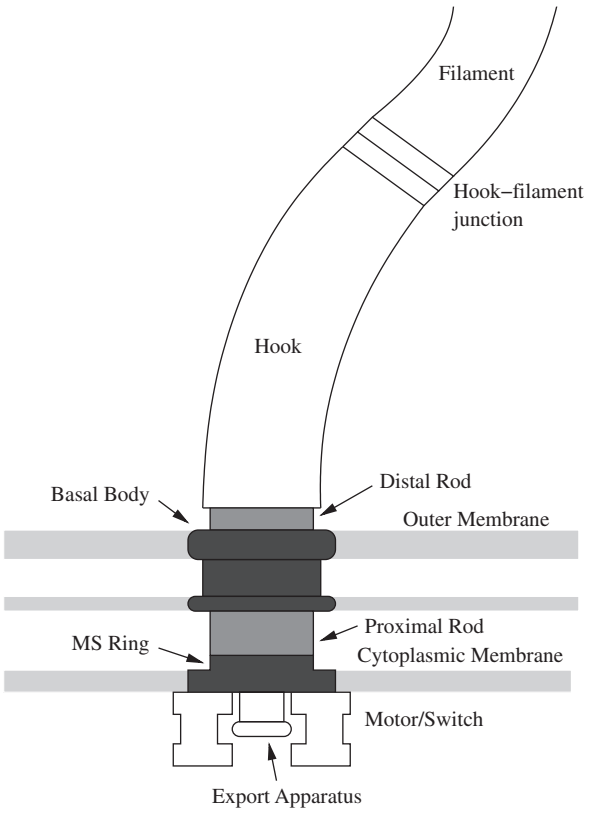

Fig. 1. Schematic diagram of flagellar motor construction.

In addition to providing a substrate specificity switch, FliK and YscP function as molecular rulers. This has been confirmed by experiments in which the lengths of these proteins was modified by insertions and deletions (Journet et al., 2003; Shibata et al., 2007). In both cases there was a strictly linear relationship between the length of the hook or needle structure and the number of residues added to or deleted from the protein.

Combined, this information suggests that length control consists of two steps, a length measurement and a substrate specificity switch. However, the question of how the length measurement is made is not resolved.

One proposal for how a molecular ruler might work is that one end of the ruler molecule is attached to the growing end of the structure and when the structure reaches the same length as the ruler molecule a signal is generated that stops growth. In fact, this is thought to be the mechanism by which the tail of phage lambda is controlled (Cornelis et al., 2006).

Two observations argue against this as the mechanism by which the length of hooks and needles are measured. First, multiple ruler molecules are needed to control the length of hooks and needles. It is estimated that between 4 and 8 molecules of FliK are needed to control the length of wild type hooks (Muramoto et al., 1998). It is a common feature of FliK and YscP that these are both exported by their respective secretion apparatuses, and released into the culture, before the substrate secretion switch occurs, but not after. This suggests that these molecules are internal, rather than external, rulers, that are secreted once in a while in order to test the length of the growing structure (Shibata et al., 2007).

Second, the length measurement is not precise. When FliK is overexpressed, hooks are shorter than wild type, while when they are underexpressed hooks are longer and much more variable ~~that~~ wild type. Thus, the secretion switch signal can occur at a wide variety of lengths and is dependent on the level of FliK expression,

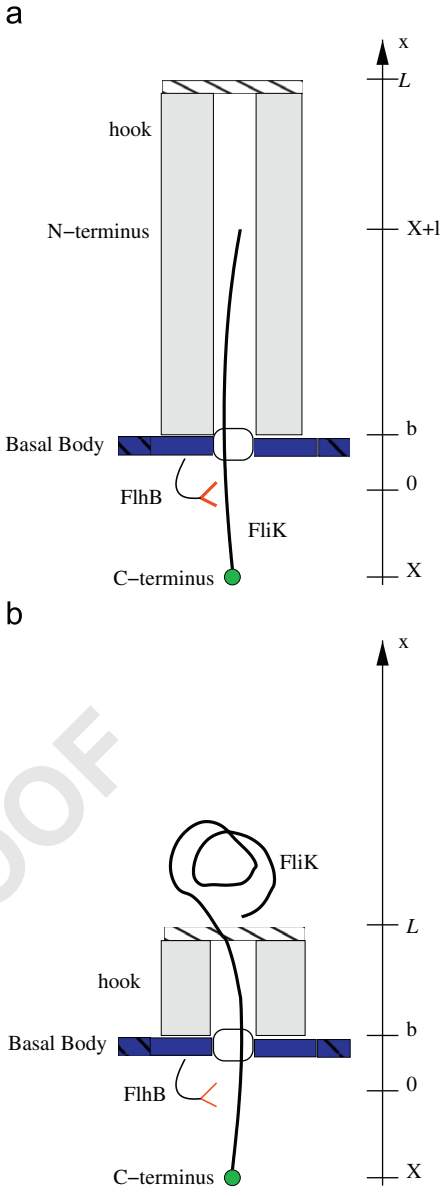

Fig. 2. Two configurations for a secreted FliK molecule. (Left) When the hook is long, movement of FliK is primarily via molecular diffusion, while (Right) when the hook is short, the motion is biased by the folding of FliK exiting the hook.

something that would not be expected for an external molecular ruler.

A more recent proposal for the length measurement is based on the observation that the N-terminus of FliK has a binding affinity for FliD, the hook cap molecule (Moriya et al., 2006). Thus, it is suggested that a secreted FliK molecule that is moving through a hook stalls when its N-terminus binds with FliD, and if at this length the C-terminus of FliK and FlhB are closely aligned, the secretion switch will occur, while if they are not aligned, the switch will not occur.

Two observations argue against this mechanism. First, this would provide for a secretion switch in some relatively narrow range of lengths, but not at much longer or shorter lengths. Second, while this might explain length control of hooks, it does not (according to current knowledge) apply to needles, as no cap molecule similar to FliD has been identified for needles.

In this paper we propose a different mechanism for length measurement. We propose that the length signal is via a

Please cite this article as: Keener, J.P., A molecular ruler mechanism for length control of extended protein structures in bacteria. J. Theor. Biol. (2009), doi:10.1016/j.jtbi.2009.12.015

stochastic process and the probability of interaction is an increasing function of hook length, so that for small lengths hook termination is unlikely, while for long lengths it is almost, but not quite, certain. A suggested mechanism for this comes from the observation that in order for molecules to be secreted, they must be partially or completely unfolded (Macnab, 2003). However, when the unfolded molecules exit the hook, they begin to fold (see Fig. 2) and the folding acts like a Brownian ratchet to pull the molecule through the hook. If a hook is short, this folding pulls the C-terminus past the FlhB cleavage site rapidly, preventing cleavage, while if the hook is long, the C-terminus has much more time to react with the cleavage site, because its motion is by diffusion only. Thus, cleavage is more likely to occur when hooks are long than when hooks are short.

In what follows, we use mathematical models to analyze different steps of this proposed hook length regulation process.

2. Hook length regulation without secretion

FliK must interact with FlhB in order for the secretion specificity switch to occur. It is thought that the secretion of FliK acts to carefully control this reaction by restricting access to the FlhB cleavage site. However, it is known that secretion is not necessary for cleavage to occur, probably because given enough time, FliK will find the FlhB cleavage site without additional guidance. Correspondingly, the hooks that result are quite long compared to those for which FliK is secreted.

Suppose length regulation occurs without secretion of FliK. We let $p(t)$ be the probability that the secretion switch has occurred by time t, and suppose that the reaction of cytoplasmic FliK with FlhB is a Poisson process. Thus,

$$\frac{dp}{dt} = k(1-p), \tag{1}$$

where k is the rate constant for the reaction. If the law of mass action holds, then k should be proportional to the number (or concentration) of FliK molecules in the cytoplasm. Now, the rate of growth of the hook is given simply by

$$\frac{dL}{dt} = \beta\Delta S(L), \tag{2}$$

where $S(L)$ is the rate of secretion of monomers, β is the percentage of secreted monomers that polymerize into hook, and Δ is the length of hook added per polymerizing monomer, about 0.5 nm.

It follows that

$$\frac{dp}{dL} = \frac{k(1-p)}{\beta\Delta S(L)}, \tag{3}$$

so that

$$-\ln(1-P(L)) = \frac{k}{\beta\Delta}\int_0^L \frac{1}{S(L)}\,dL, \tag{4}$$

where $P(L)$ is the probability that a hook has length less than L.

We use the published data of Hirano et al. (2005) (Fig. 8) to test this model. The data are reproduced in Fig. 3. Here are shown 6 histograms for the length of hooks for several different FliK mutants. The mutants were (a) FliKΔ13, (c) FliKΔ99, and (e) CFP-FliK. (FliKΔ13 and FliKΔ99 are FliK deletion variants for which the first 13 or 99, respectively, amino acids were truncated, and CFP-FliK is a FliK variant to which cyan fluorescent protein (CFP) was fused to the N-terminus of FliK.) None of these FliK monomers were secreted. Hook lengths were measured with normal FliK expression (a), (c), and (e), and with FliK levels overexpressed (b), (d), and (f), respectively.

According to our model, the only difference between these six datasets should be the parameter k. In other words, the curves $-\ln(1-P(L))$ for each of the datasets should differ only by a scale

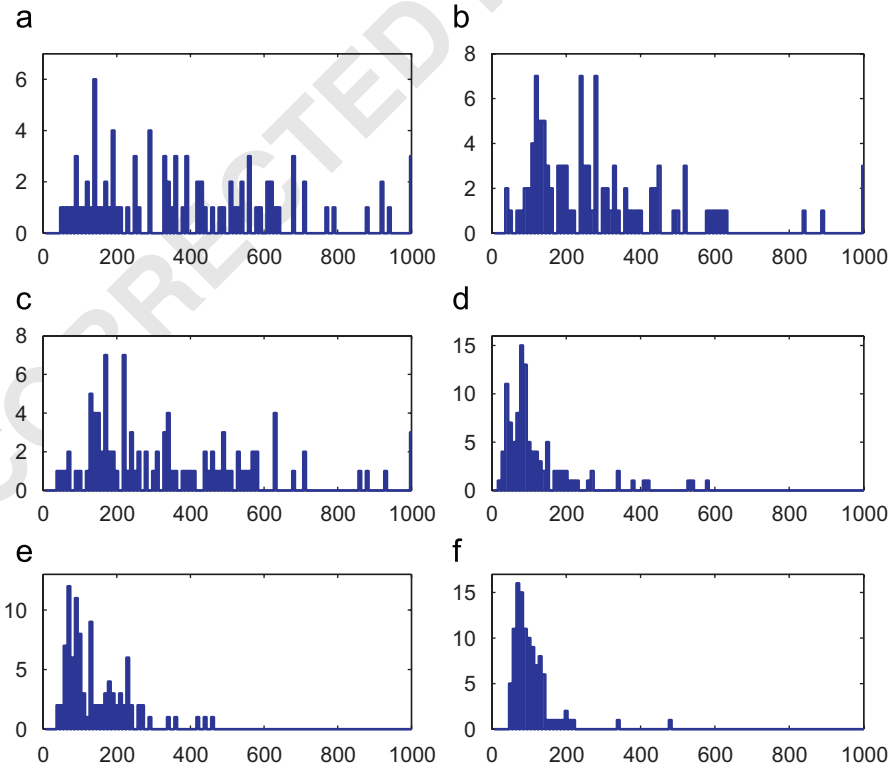

Fig. 3. Histogram of hook lengths for several non-secreted FliK mutants. (a) FliKΔ13, (b) FliKΔ13 overexpressed, (c) FliKΔ99, (d) FliKΔ99 overexpressed, (e) CFP-FliK, and (f) CFP-FliK overexpressed. Data taken from Fig. 8 of Hirano et al. (2005).

Please cite this article as: Keener, J.P., A molecular ruler mechanism for length control of extended protein structures in bacteria. J. Theor. Biol. (2009), doi:10.1016/j.jtbi.2009.12.015

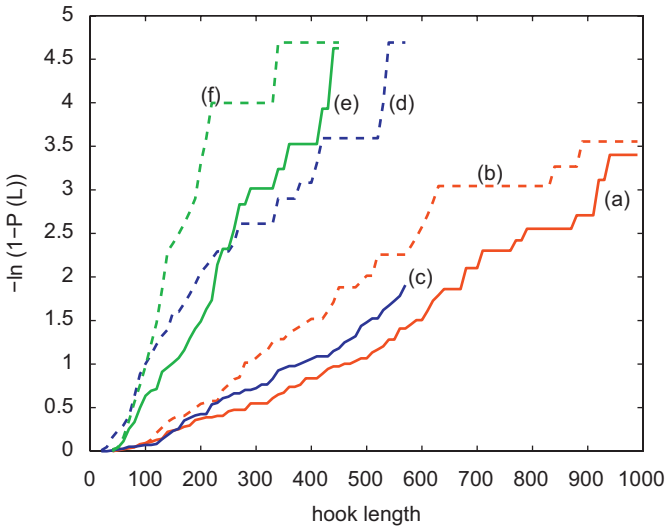

Fig. 4. Histogram data of Fig. 3 replotted as $-\ln(1-P(L))$ as a function of L.

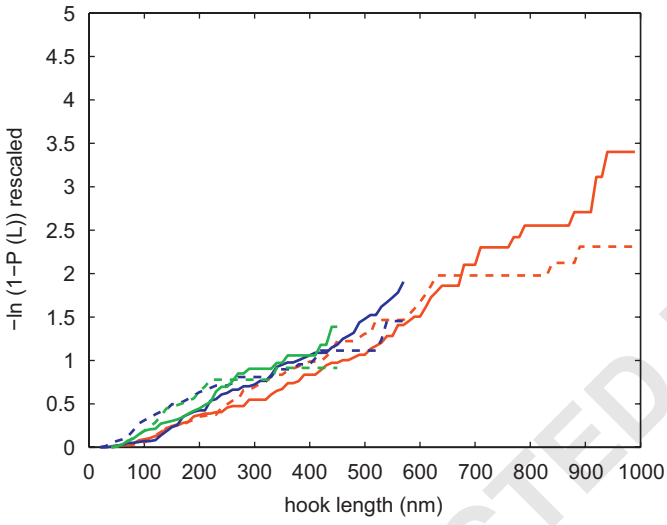

Fig. 5. Curves from Fig. 4 rescaled by 1 (a), 0.65 (b), 1 (c), 0.31 (d), 0.3 (e), and 0.195 (f).

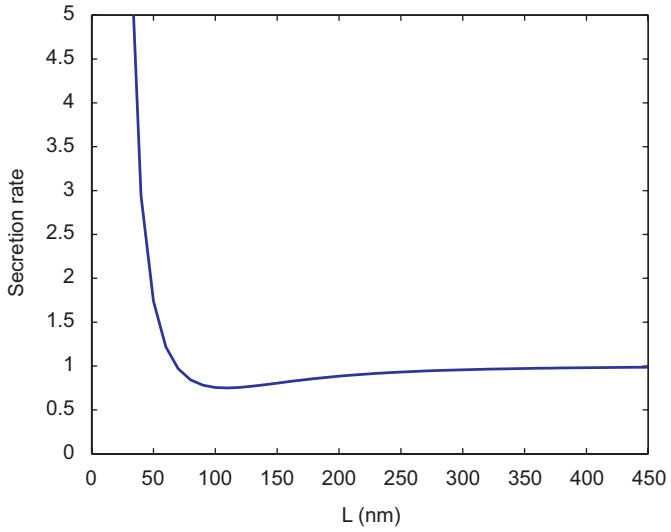

Fig. 6. Hook growth rate plotted as a function of L.

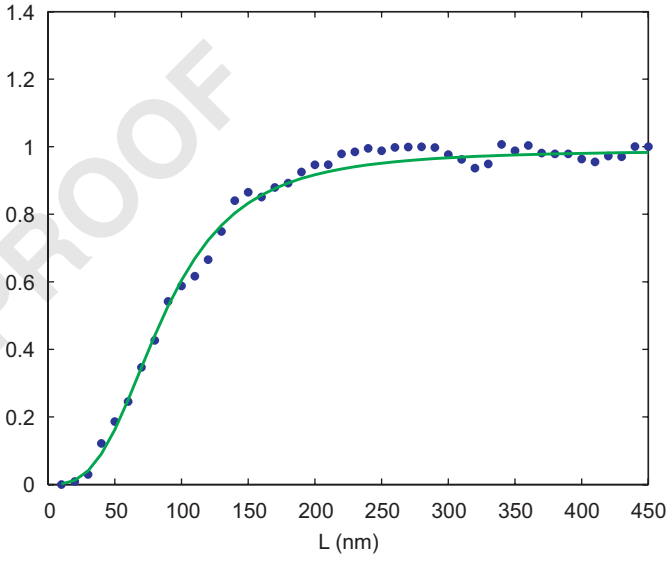

Fig. 7. Average of data from Fig. 5, divided by L, and scaled to have maximum value one (shown as points), compared to the Hill function $h(L)=0.99(L^3/K^3+L^3)$, with $K=86.1$ nm (solid curve).

factor. In Fig. 4 are shown the curves $-\ln(1-P(L))$ for each of the six datasets, where $P(L)$ represents the percentage of hooks in the dataset with length less than L. In Fig. 5 are shown these same six curves, rescaled by 1 (a), 0.65 (b), 1 (c), 0.31 (d), 0.3 (e), and 0.195 (f). (Fig. 6)

The striking agreement of these six curves gives credence to the hypothesis that the reaction of cytosolic FliK with FlhB is by a Poisson process, and that overexpression of a particular FliK mutant simply increases the reaction rate. This plot also suggests that the reaction rates of FliKΔ13 and FliKΔ99 with FlhB are quite similar, while the reaction of CFP-FliK with FlhB is about 3.3 times faster. This increased reaction rate may be explained by the fact that the N-terminus of CFP-FliK is not modified, as is the N-terminus of FliKΔ13 and FliKΔ99, and this could lead to a higher affinity of CFP-FliK for the secretion apparatus, leading to enhanced localization of CFP-FliK in the vicinity of FlhB.

There is a second piece of information that can be gleaned from this data. Notice that according to this model $-\ln(1-P(L))$ is proportional to $\int_0^L 1/S(L)\,dL$, where $S(L)$ is the secretion rate, proportional to the hook growth rate. In Fig. 7 is shown the average of the data from Fig. 5, divided by hook length L, and scaled to have maximum value 1, compared to a Hill function fit of this data

$$h(L)=0.99\frac{L^3}{K^3+L^3}, \tag{5}$$

where $K=86.1$ nm.

Using this Hill function we can estimate the (dimensionless) secretion rate $S(L)$ to be

$$S(L)=\frac{(K^3+L^3)^2}{L^3(4K^3+L^3)}. \tag{6}$$

A plot of this function is shown in Fig. 6.

In Koroyasu et al. (2003) the hook growth rate was estimated using a population balance model. The hook growth velocity curve found here has two features that are similar to the hook growth velocity curve that was found in Koroyasu et al. (2003). In

Please cite this article as: Keener, J.P., A molecular ruler mechanism for length control of extended protein structures in bacteria. J. Theor. Biol. (2009), doi:10.1016/j.jtbi.2009.12.015

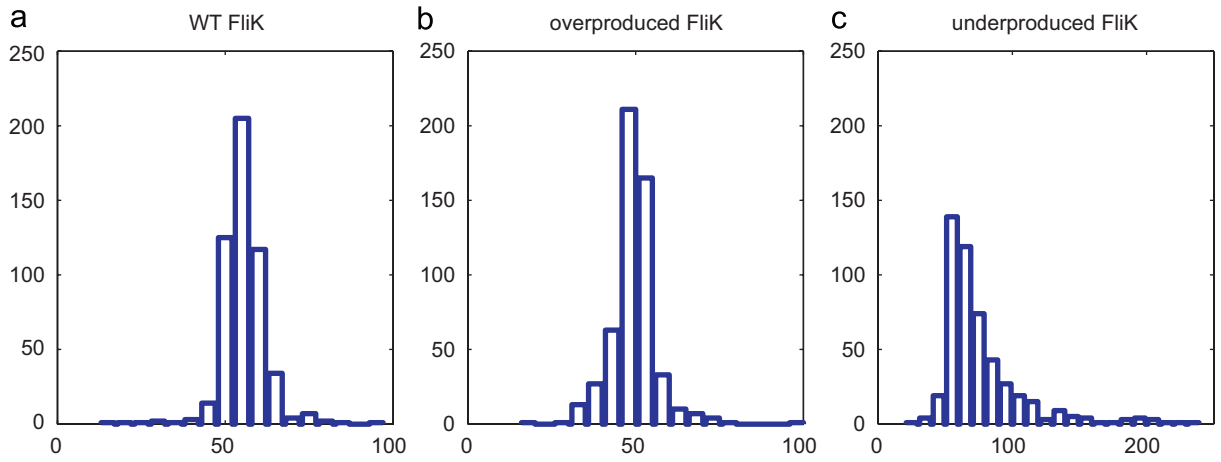

Fig. 8. Histograms of hooks lengths of wild type (a), with FliK overproduced (b), and with FliK underproduced (c), from Muramoto et al. (1998).

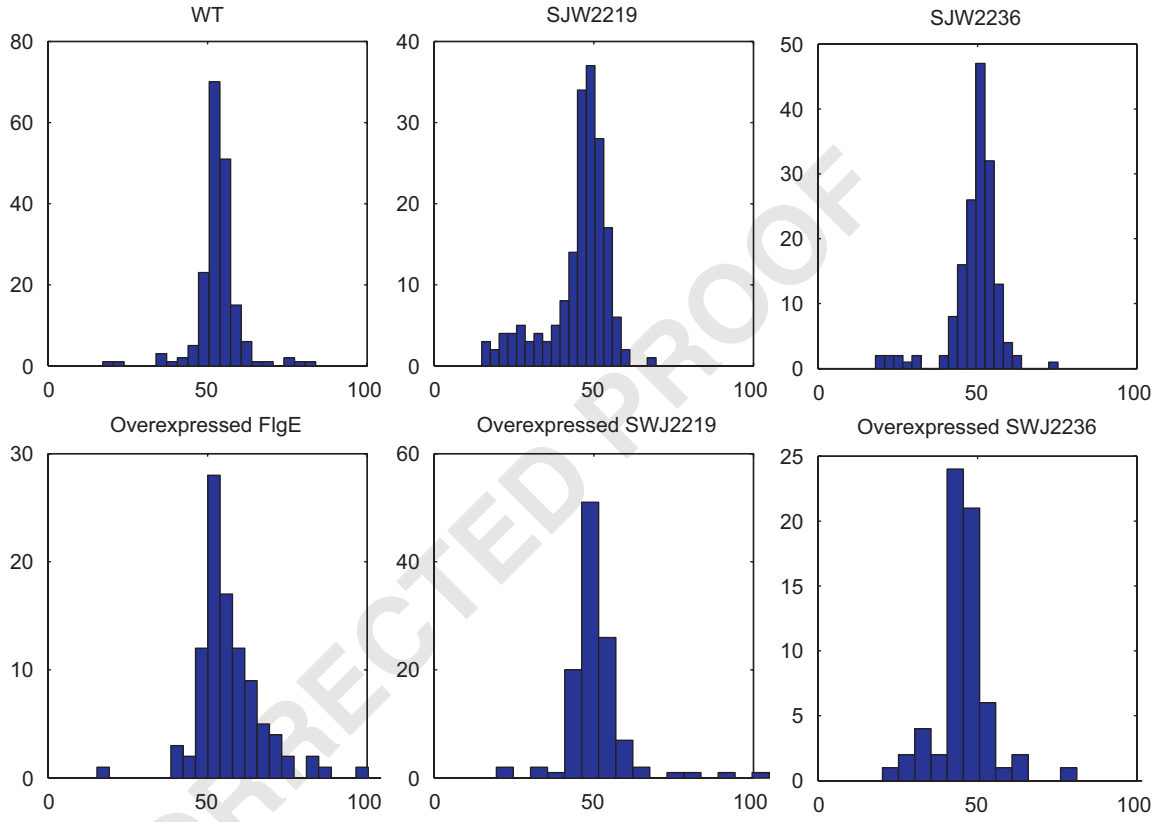

Fig. 9. Histograms of hooks lengths of FlgE polymerization mutants, from Moriya et al. (2006).

particular, this shows a growth rate that is very large for small hook lengths dropping rapidly to a nearly constant rate of growth for larger hook lengths. It is noteworthy that the transition between the two behaviors occurs in the vicinity of $L = 50$ nm. A possible explanation for this is as follows. Secreted molecules must be partially unfolded in order to fit into the 2 nm diameter hook structure. As they exit the hook or polymerize onto the forming end of the hook, folding will cause the C-terminus of the molecule to be pulled more rapidly than if the molecule were moving by diffusion alone. Thus, when the hook is short, the time it takes for a single monomer to be secreted is much smaller than when the forming tube is long. Since the secretion rate (number of molecules secreted per unit time) is certain to be affected by the amount of time it takes for a single molecule to be secreted, it is reasonable that the rate of secretion should be higher when the hook is short compared to when it is long.

The effect of folding on the movement of a monomer is also likely to be significant for hook length regulation, as is described below.

3. The infrequent molecular ruler mechanism

The above model examined hook length in the case that cytoplasmic FliK molecules are not secreted but react with FlhB in a purely random fashion, independent of the length of the hook.

Please cite this article as: Keener, J.P., A molecular ruler mechanism for length control of extended protein structures in bacteria. J. Theor. Biol. (2009), doi:10.1016/j.jtbi.2009.12.015

The suggestion of the infrequent molecular ruler mechanism is that while hook is being constructed by FlgE secretion, every once in a while a FliK molecule is also secreted. Secretion of FliK has the effect of carefully controlling the interaction of FliK with FlhB, and we propose that the probability of interaction of FliK with FlhB is some function of hook length. A possible mechanism for this length dependence is described in the next section.

Suppose that the rate of growth of the hook is given by (2). As before, the secretion switch occurs when a FliK molecule interacts with FlhB, however, now we assume that only a secreted FliK molecule can interact with FlhB. Thus, the probability that the secretion switch takes place by time t, $p(t)$, is determined by

$$\frac{dp}{dt} = \alpha S(L) P_c(L)(1-p), \tag{7}$$

where α is the fraction of secreted molecules that are FliK, and $P_c(L)$ is the probability that the secreted FliK molecule interacts with FlhB when the hook length is L. Combining this with (2), we find that

$$\frac{dp}{dL} = \frac{\alpha S(L) P_c(L)(1-P)}{\beta \Delta S(L)} \tag{8}$$

so that

$$-\ln(1-P(L)) = \frac{\alpha}{\beta\Delta}\int_0^L P_c(L)\,dL, \tag{9}$$

where $P(L)$ is the probability of hook length less than L.

We test this model using two different sets of data of hook lengths. The first, shown in Fig. 8 taken directly from Muramoto et al. (1998), shows histograms of hook lengths of wild type (a), with FliK overproduced (b), and with FliK underproduced (c). The second data set, from Moriya et al. (2006) (with raw data kindly provided by Profs. K. Namba and T. Minamino) show histograms of hook lengths for certain FlgE mutants. In particular, SJW2219 and SJW2236 are FlgE mutants for which polymerization of FlgE into hook is less efficient than wild type.

The only difference between the hooks in Fig. 8 is presumably in the percentage of secreted molecules that are FliK, represented by the parameter α, while the only difference between the hooks in Fig. 9 is the percentage of secreted molecules that polymerize into hook represented by β. Thus, according to the infrequent molecular ruler hypothesis, the only difference between these different data sets should be in the scale factor $\alpha/\beta\Delta$.

In Fig. 10 (top panel) are shown the curves $-\ln(1-P(L))$ for each of the datasets in Fig. 5, with $-\kappa\ln(1-P(L))$ shown in the bottom panel for $\kappa = 1, 0.9, 4$ for the data (a), (b), and (c), respectively.

Similarly, in Fig. 11 are shown the curves $-\ln(1-P(L))$ for each of the datasets in Fig. 8, while in Fig. 12 are shown the rescaled curves with $\kappa = 1, 0.5, 0.7, 2.3, 1$, and 0.7, for the curves (a) through (e).

We take the agreement between the rescaled curves in Fig. 10 as strong evidence in favor of the infrequent molecular ruler mechanism. The agreement in Fig. 12 is not as good, and there are

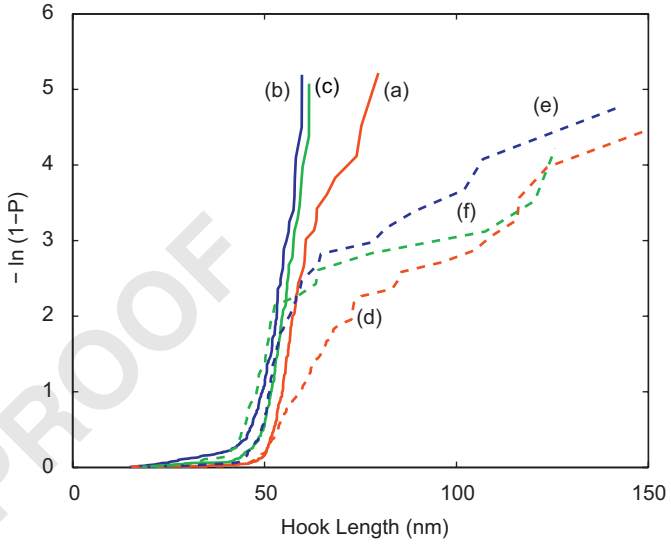

Fig. 11. Data from Fig. 8 plotted as $-\ln(1-P(L))$.

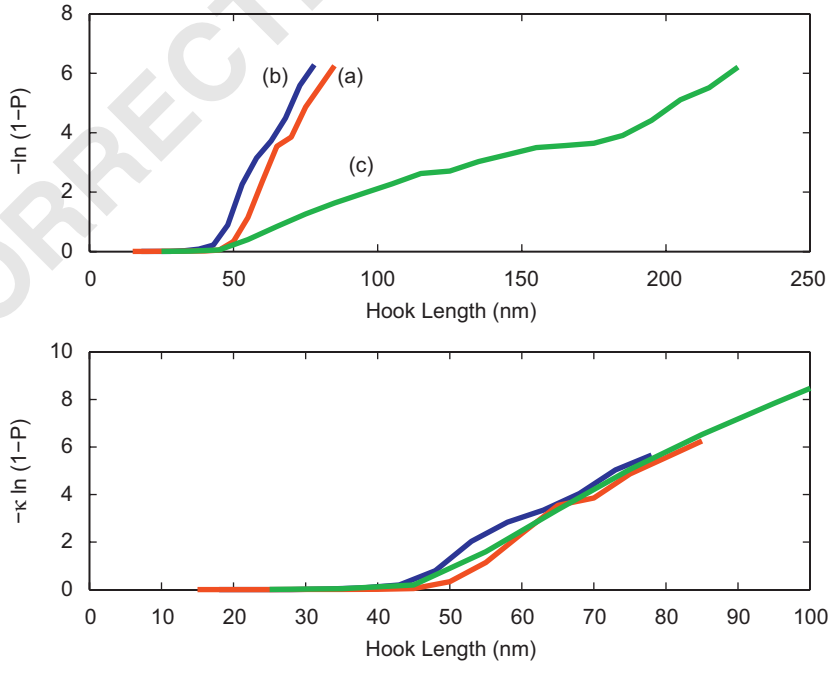

Fig. 10. Data from Fig. 5 plotted as $-\ln(1-P(L))$ (top) and rescaled by 1 (a), 0.9 (b), 4 (c).

Please cite this article as: Keener, J.P., A molecular ruler mechanism for length control of extended protein structures in bacteria. J. Theor. Biol. (2009), doi:10.1016/j.jtbi.2009.12.015

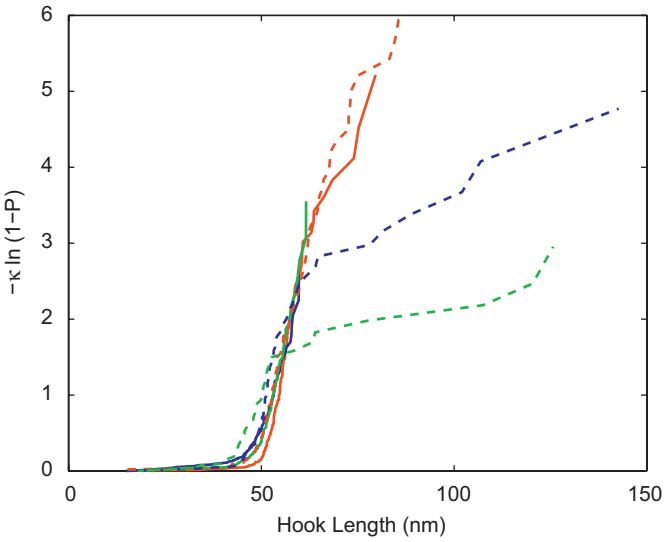

Fig. 12. Data from Fig. 8 plotted as $-\kappa \ln(1-P(L))$, with $\kappa = 1, 0.5, 0.7, 2.3, 1$, and 0.7.

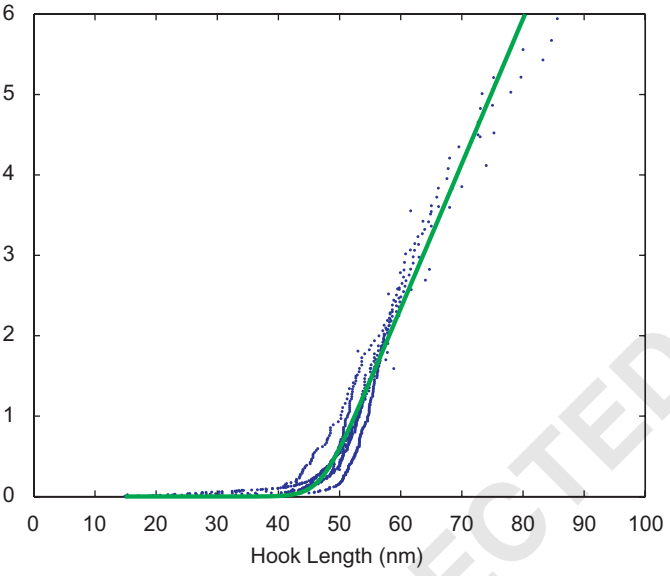

Fig. 13. Composite of Figs. 11 and 12 compared with the curve $\int_0^L P_c(L)\,dL$.

several possible explanations for this. First, the total number of hooks in these datasets is not particularly large, so that one must be careful not to overfit, or overinterpret the data. In particular, the tails of the distributions do not have large sample numbers and as a result, information about the tails of the distributions is lacking. Second, with these mutants, growth of hooks was quite slow compared to wild type. As a result there was much more opportunity for the FlhB cleavage to occur by means other than via a secreted FliK molecule. For example, the Poisson process described in the previous section, although slow when growth of hook is normal, could give rise to an increased number of short hooks when growth is slowed. This might explain the larger than expected number of short hooks seen with mutant SJW2219.

From this data we can estimate the function $P_c(L)$. In fact, a good fit of the data can be obtained with

$$P_c(L) = \frac{1}{1+\exp(a(L-L_0))}, \tag{10}$$

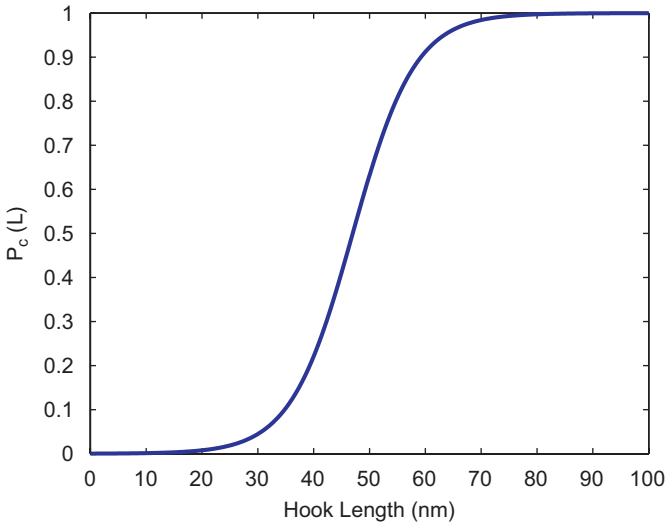

Fig. 14. Plot of $P_c(L)$ given by (10).

with $a = 0.72\,\mathrm{nm}^{-1}$ and $L_0 = 47\,\mathrm{nm}$. In Fig. 13, the curve $\int_0^L P_c(L)\,dL$ is shown in comparison with the data from Figs. 11 and 12. The curve $P_c(L)$ is shown in Fig. 14. The scale of $P_c(L)$ cannot be determined from these data, so we have arbitrarily chosen the scale so that $P_c(\infty) = 1$.

4. Length dependent FlhB cleavage

To understand how the probability of FlhB cleavage might be length dependent, we must know something about how a secreted FliK molecule moves and how its C-terminus interacts with FlhB to cause cleavage. Unfortunately, it is not known how secretion works, so we make some (hopefully reasonable) assumptions. Because the inner diameter of the forming hook-filament complex is approximately 2 nm (Yonekura et al., 2003), the secreted molecules must be unfolded. The action of the ATPase FliI is apparently significant here, but the details of how this ATPase works are also not known. Secretion does not occur without the action of an ATPase, so presumably the energy of ATP hydrolysis is necessary to unfold the secretant molecule or to push it over some high energy barrier or both. Thus, we suppose that there is something that prevents the monomer from exiting the tube at its proximal end once secretion begins. One possibility is that for a molecule to be secreted, its N-terminus must cross a large energy barrier in the vicinity of the entrance, and it is the hydrolysis of ATP that enables this energy barrier to be crossed. However, to reexit the tube at the proximal end, this same energy barrier must be recrossed (in the retrograde direction), but this is highly unlikely to occur.

Recent evidence (Minamino and Namba, 2008; Paul et al., 2008) suggests that secretion is accomplished by a protonmotive force (PMF), perhaps related to the force that drives the rotation of the motor. But the details of how this might work or what it accomplishes are not known.

The second driving force is the result of the tendency of monomers to fold. For example, when FlgE polymerizes it does so by folding into the forming tube, facilitated by FliD. Similarly, when FliK exits the distal end of the tube, the free energy gained in folding prevents diffusion back into the hook. Thus, the overall motion of diffusing molecules is biased by the exit restriction at the proximal end and by the folding energy at the distal end. Consequently, when the hook is short, the secretion of FliK is fast, because its distal end is folding even before the C-terminus has entered the tube. However, for long hooks, the motion of FliK is

much slower, and this allows much more time for the C-terminus to interact with FlhB.

Additional forces on the secreted molecule could come from binding or unbinding to the wall of the forming tube or the FliD cap (Moriya et al., 2006). It is known that FliK associates with FliD *in vitro*, however, as this association probably requires more folding of FliK than can occur inside the forming hook, we do not include this effect in our model.

Our second hypothesis is that the cleavage of FlhB through interaction with the C-terminus of FliK is a Poisson process. That is, there is some rate of reaction that is a function of the distance between the C-terminus and the cleavage site of FlhB. When the two are in close proximity, their reaction probability is highest, but when they are separated spatially, the probability of reaction is reduced.

To explore the consequences of these assumptions, we develop and analyze a mathematical model of this process. Since movement of the secretant molecule is essentially one-dimensional, we let x measure the distance along the axis of the hook, and suppose that the FlhB cleavage site is located at $x=0$, the proximal end. The polymerizing end of the hook is at $x=L$, while the high energy barrier at the entrance of the basal body is located at $x=b>0$. (In this notation, L is the total length of the basal body and the hook combined, the length of the basal body being about 40 nm.) The FliK molecule is assumed to be of unfolded length l. Its C-terminus is located at $x=X$ and its N-terminus is located at $x=X+l$ and whenever the N-terminus is inside the forming hook, $X+l<L$. If $X+l>L$, then the N-terminus has exited the distal end of the hook. We also assume that once secretion has begun, the N-terminus cannot escape from the proximal end, so that X is always greater than $b-l=a$, and that if the C-terminus passes $x=b$, i.e., if $X>b$, it cannot reexit through the proximal end, i.e., secretion is complete and interaction with FlhB is no longer possible. (See Fig. 2). Note that we have idealized the location of the C-terminus to be a point when in fact it is a region (residues 265-405 in Flik and 403-515 in YscP). Furthermore, l is taken to be the unfolded length of the non-reacting region of the molecule (residues 1-264 in FliK and 1-402 in YscP).

We follow the position of the C-terminus using the stochastic differential equation

$$v\,dX = F(X)\,dt + \sqrt{2k_bTv}\,dW, \tag{11}$$

where $F(X)$ represents the force acting on the unfolded (rod-like) FliK molecule, $W(t)$ is Brownian white noise, a Wiener process, k_b is Boltzmann's constant, and T is absolute temperature. Next we assume that the probability of cleavage is a function of x, say $g(x)$. Now we let $P(x,t)$ be the probability density that the C-terminus is at position x and FlhB has not been cleaved at time t, and we let $Q(t)$ be the probability that the FlhB molecule has been cleaved at time t. It follows that

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x}\left(\frac{1}{v}F(x)P\right) + D\frac{\partial^2 P}{\partial x^2} - g(x)P, \tag{12}$$

and

$$\frac{dQ}{dt} = \int_a^b g(x)P\,dx, \tag{13}$$

where $D=k_bT/v$. We also suppose that once secretion has begun, the molecule is prevented from exiting through the proximal end of the hook, and that once the C-terminus of the molecule enters the hook, it cannot reemerge (an absorbing boundary condition), so that the possibility of cleavage is ended. This places boundary conditions on $P(x,t)$, namely

$$DP_x(a,t) - \frac{1}{v}F(a)P(a,t) = 0, \quad P(b,t) = 0 \quad \text{for } t>0. \tag{14}$$

We are interested in knowing the probability of FlhB cleavage. We define $\pi_c(x)$ to be the probability that cleavage occurs and let $\pi_b(x)$ be the probability that the molecule is secreted without cleavage by exiting at $x=b$, starting the process from position x. Of course, we ultimately want to know the probability that cleavage occurs starting the process at position a, i.e. $\pi_c(a)$, but we carry along the variable x in order to do some analysis. Clearly, $\pi_c(x)+\pi_b(x)=1$.

Define $g_b(x,t)$ to be the probability that the molecule exits at $x=b$ without cleaving after time t,

$$g_b(x,t) = \int_t^\infty J(b,t'|x,0)\,dt', \tag{15}$$

where $J(b,t'|x,0)$ is the probability flux through $x=b$ at time t'. Of course,

$$J(b,t|x,0) = \frac{1}{v}F(b)P(b,t|x,0) - D\frac{\partial P(y,t|x,0)}{\partial y}\bigg|_{y=b}. \tag{16}$$

Because this is a time independent process, $P(b,t|x,0) = P(b,0|x,-t)$, so that $P_b(x,t)=P(b,t|x,0)$ satisfies the backward Kolmogorov equation

$$\frac{\partial P_b}{\partial t} = -g(x)P_b + \frac{1}{v}F(x)\frac{\partial P_b}{\partial x} + D\frac{\partial^2 P_b}{\partial x^2}. \tag{17}$$

Boundary conditions on P_b are that $\partial P_b(a,t)/\partial x = 0$ and $P_b(b,t)=0$. It follows that

$$\frac{\partial}{\partial t}J(b,t|x,0) = \left(\frac{1}{v}F(b) - D\frac{\partial}{\partial b}\right)\frac{\partial P_b}{\partial t}, \tag{18}$$

$$= -g(x)J + \frac{1}{v}F(x)\frac{\partial}{\partial x}J(b,t|x,0) + D\frac{\partial^2}{\partial x^2}J(b,t|x,0). \tag{19}$$

Thus,

$$\frac{\partial g_b(x,t)}{\partial t} = -J(b,t|x,0) = \int_t^\infty J_{t'}(b,t'|x,0)\,dt' \tag{20}$$

$$= -g(x)g_b + \frac{1}{v}F(x)\frac{\partial g_b}{\partial x} + D\frac{\partial^2 g_b}{\partial x^2}. \tag{21}$$

Now, notice that $J(b,0|x,0)=0$ if $x\neq b$, so taking the limit $t\to 0$ and defining $g_b(x,0)=\pi_b(x)$ gives

$$D\frac{d^2\pi_b}{dx^2} + \frac{1}{v}F(x)\frac{d\pi_b}{dx} - g(x)\pi_b = 0. \tag{22}$$

Boundary conditions are $\pi_b'(a)=0$ and $\pi_b(b)=1$.

This second order linear boundary value problem is easily solved numerically. In fact, with

$$\frac{1}{D}g(x) = \frac{\alpha}{\sigma}\exp\left(-\frac{x^2}{2\sigma^2}\right), \tag{23}$$

(a gaussian curve) and

$$\frac{1}{kT}F(x) = f_0H(x-L+l), \tag{24}$$

where H is the Heaviside function, and with parameters $\sigma=14$ nm, $\alpha=0.5\,\text{nm}^{-1}$, $f_0=50-100\,\text{nm}^{-1}$, $a=-95$ nm, $b=10$ nm, the curve $\pi_c(L)$ and the curve $P_c(L)$ shown in Fig. 14 are nearly indistinguishable. Thus, this model of length dependent cleavage gives a reasonable fit to the data shown in Figs. 8 and 9.

It is also worth noting that changing the length of the unfolded FliK molecule merely shifts the curve $\pi_c(L)$ by the same amount, showing that the length of the hook is linearly related to the unfolded length of FliK.

We can get a rough estimate of f_0 as follows: In Thompson et al. (2002) it was estimated that the entropic folding energy of a polyprotein is in the range of 2.1–2.5 kT/amino acid residue. This

Please cite this article as: Keener, J.P., A molecular ruler mechanism for length control of extended protein structures in bacteria. J. Theor. Biol. (2009), doi:10.1016/j.jtbi.2009.12.015

suggests that the folding energy for a 264 residue N-terminal domain of FliK is about 550–660 kT. A completely unfolded amino acid contributes about 0.4 nm of length to a protein so in its elongated state the 264 residue chain is about 105 nm long. Thus, for such a molecule, the entropic folding force per unit length is in the range of 5.3–6.3 kT/nm. However, since the FliK in the forming hook is probably not in its completely extended state, the entropy gain upon exiting the hook is likely to be somewhat less than this.

This is estimate is low compared to the value of f_0 needed to get a good fit with the data. In fact, with f_0 in the range of 5.3–6.3/nm, the model produces a broader distribution of hook lengths with more short hooks than are seen in the data.

At present there is not enough known about the FlhB cleavage process to estimate the other parameters of the model.

5. Discussion

In this paper we develop a mathematical model of the hook termination process. We propose that hook termination results from cleavage of FlhB by FliK, that FliK is secreted intermittently, and that the probability that cleavage occurs is an increasing function of the length of the hook. The mechanism for this length dependent cleavage rate is proposed to be related to the rate at which a secreted FliK molecule moves through the forming hook or more specifically, the amount of time the C-terminus of FliK spends in the vicinity of the FlhB binding site. This time is short when the hook is short because the folding of FliK exiting the distal end of the hook acts to pull the FliK molecule through the hook rapidly. In contrast, this time is much longer when the hook is longer than the unfolded FliK polymer since movement through the tube is not enhanced by folding. The transition between these two modes occurs when the length of the unfolded FliK molecule is about the same length as the hook plus basal body, approximately 95 nm. Furthermore, a consequence of this hypothesis is that hook length should be linearly correlated to the length of the FliK molecule, which it is Shibata et al. (2007).

A strength of this proposal is that it gives excellent qualitative agreement with known data, but at the same time, reasonable estimates of the entropic folding force are smaller than needed to give close quantitative agreement. Indeed, with the estimated entropic folding force of 5–6/nm, this mechanism would produce more short hooks than are actually observed, because the C-terminus is not sufficiently prevented from reacting with FlhB. Of course, there might be other forces acting on FliK that are not taken into account here, including a force from the FliI ATPase-chaperone, a protonmotive force driving secretion, or other folding forces in addition to the entropic folding force. Or, it could be that the length measurement is by a different mechanism.

A different, but related hypothesis was explored in Keener (2005). There the idea was, following the suggestion of Muramoto et al. (1998), that the export of FlgE occurs in three distinct stages. In stage 1, secretion by the ATP-ase is rate limiting so that there is no more than one monomeric subunit in the forming channel at any given time. In stage 2, the increased hook length has caused diffusion to slow so that it is rate limiting and there are several hook subunits in the channel causing interference in the diffusion process. In stage 3, the channel is completely congested and secretion export is coupled to assembly at the distal end. This coupling causes a delay in the ATP-ase transport cycle, allowing the length control protein FliK to bind FlhB, leading to a change of the specificity of the transporter.

We are now of the opinion that this earlier proposal it is likely not to be correct, primarily because the length of the completed hook is only long enough to contain one unfolded molecule at a time. A mechanism, such as the one proposed in this paper, that involves movement of a single molecule, seems more reasonable.

Another recent proposal for the length measurement is based on the observation that the N-terminus of FliK has a binding affinity for FliD, the hook cap molecule (Moriya et al., 2006). It is suggested that a secreted FliK molecule that is moving through a hook stalls when its N-terminus binds with FliD, and if at this length the C-terminus of FliK and FlhB are closely aligned, the secretion switch will occur, while if they are not aligned, the switch will not occur.

A drawback to this proposal is that if FliK is not secreted when the hook is of the right length, the growth will not be terminated. In fact, one can easily construct a stochastic model for this mechanism, similar to the model constructed here, and determine that $\pi_c(L)$ is similar to a gaussian distribution, rather than the monotone increasing function found here. Thus, another mechanism for termination must be invoked for example, when FliK is in short supply. In Moriya et al. (2006), the suggestion was made that this termination occurs because of an auto-cleavage reaction of FlhB that is unrelated to the length of the hook, so that if FliK termination does not occur during the appropriate window of length opportunity, auto-cleavage comes to the rescue.

This proposal also has the disadvantage that to date, no cap molecule similar to FliD has been found for injectosomes, so that a different mechanism is needed to explain injectosome growth termination. We note that the model proposed in this paper works equally well to explain the termination of injectosome growth in *Yersinia*.

Thus, while we obviously favor the hypothesis presented in this paper, it remains to be determined which, if any, of these hypotheses is correct.

Acknowledgment

This research was supported in part by NSF Grant DMS-0718036.

References

Cornelis, G.R., Agrain, C., Sorg, I., 2006. Length control of extended protein structures in bacteria and bacteriophages. Current Opinion in Microbiology 9, 201–206.

Hirano, T., Shibata, S., Ohnishi, K., Tani, T., Aizawa, S.-I., 2005. N-terminal signal region of FliK is dispensable for length control of the flagellar hook. Molecular Microbiology 56, 346–360.

Journet, L., Agrain, C., Broz, P., Cornelis, G.R., 2003. The needle length of bacterial injectisomes is determined by a molecular ruler. Science 302, 1757–1760.

Keener, J.P., 2005. A model for length control of flagellar hooks of *Salmonella typhimurium*. Journal of Theoretical Biology 234, 263–275.

Koroyasu, S., Yamazoto, M., Hirano, T., Aizawa, S.-I., 2003. Kinetic analysis of the growth rate of the flagellar hook in *Salmonella typhimurium* by the population balance method. Biophysical Journal 74, 436–443.

Macnab, R.M., 2003. How bacteria assemble flagella. Annual Review of Microbiology 57, 77–100.

Minamino, T., Namba, K., 2008. Distinct roles of the FliI ATPase and proton motive force in bacterial flagellar protein export. Nature 451, 485–488.

Moriya, N., Minamino, T., Hughes, K.T., Macnab, R.M., 2006. The type III flagellar export specificity switch is dependent on FliK and a molecular clock. Journal of Molecular Biology 359, 466–477.

Muramoto, K., Makishima, S., Aizawa, S.-I., Macnab, R.M., 1998. Effect of cellular level of FliK on flagellar hook and filament assembly in *Salmonella typhimurium*. Journal of Molecular Biology 277, 871–882.

Paul, K., Erhardt, M., Hirano, T., Blair, D.F., Hughes, K.T., 2008. Energy source of flagellar type III secretion. Nature 451, 489–492.

Shibata, S., Takahashi, N., Chevance, F.F.V., Karlinsey, J.E., Hughes, K.T., Aizawa, S.-I., 2007. FliK regulates flagellar hook length as an internal ruler. Molecular Microbiology 64, 1404–1415.

Thompson, J.B., Hansma, H.G., Hansma, P.K., Plaxco, K.W., 2002. The backbone conformational entropy of protein folding: experimental measures from atomic force microscopy. Journal of Molecular Biology 322, 645–652.

Yonekura, K., Maki-Yonekura, S., Namba, K., 2003. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. Nature 424, 643–650.

Please cite this article as: Keener, J.P., A molecular ruler mechanism for length control of extended protein structures in bacteria. J. Theor. Biol. (2009), doi:10.1016/j.jtbi.2009.12.015