Supporting Material

Theoretical and Computational Investigation of Flagellin Translocation and Bacterial Flagellum Growth

David E. Tanner, Wen Ma, Zhongzhou Chen and Klaus Schulten

Movie S1: Flagellin Compression   (movie-S1-compression.mpg). This movie is generated from simulated compression of a flagellin segment (simulation 9C in Table 1). Shown are flagellin, confining cylinder and a dynamic plot of the flagellin properties. The flagellin segment atoms are colored from red (atoms pushed strongest by left and right walls) to green (not pushed by walls). The confining cylinder (9 Å radius) is shown in blue; atoms inside the cylinder appear darker, while atoms in contact with the cylinder wall appear lighter. The dynamic plot shows, in green, the axial pressure (y-axis), $p_{\text{axial}}(t)$, plotted against flagellin density (x-axis), $\rho(t)$; the x-axis spans 0.04 to 0.3 flagellin atoms/Å^3 while the y-axis spans 0.0 to 1.2 (kcal/mol)/Å^3. The plot also shows, in red, the radial pressure (y-axis), $p_{\text{radial}}(t)$, plotted against flagellin density (x-axis), $\rho(t)$; in this case the y-axis spans 0.0 to 0.45 (kcal/mol)/Å^3. The plot shows how stress, friction density and flagellin density increase as the protein is compressed.

Movie S2: Flagellin Translocation   (movie-S2-translocation.mpg). This movie is generated from a molecular dynamics simulation of flagellin pushed through the flagellum channel (simulation CD0 in Table 1). Shown are the CD0 helices (orange) which comprise the inner surface of the channel, the flagellin segment's backbone (green), basic residues ARG494 along the channel (blue) and acidic residues of the translocating flagellin (red). Acidic and basic residues of salt bridge formed transiently are highlighted in light red and light blue, respectively.
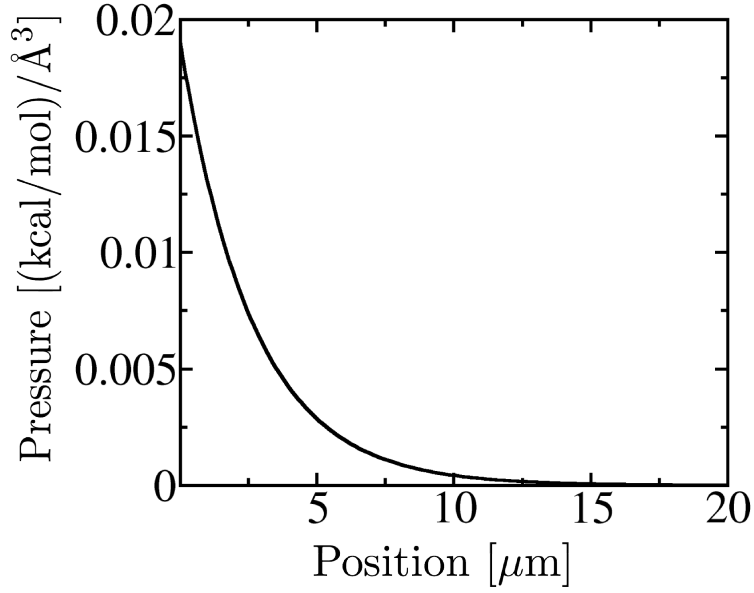
Figure S1: Flagellin Pressure. Using the system parameters listed in Table 2, including $P = 100$ (kcal/mol)/s, with Eq. 9 of the main text, flagellin pressure is shown as a function of position along the flagellum for the case $L = 20\mu$m. Pressure is highest at the flagellum base (position = 0) and lowest at the the flagellum tip (position = 20 μm). This plot reveals the flagellin pressure range relevant to the flagellum, namely $[1 \times 10^{-5}, 0.02]$ (kcal/mol)/Å^3.
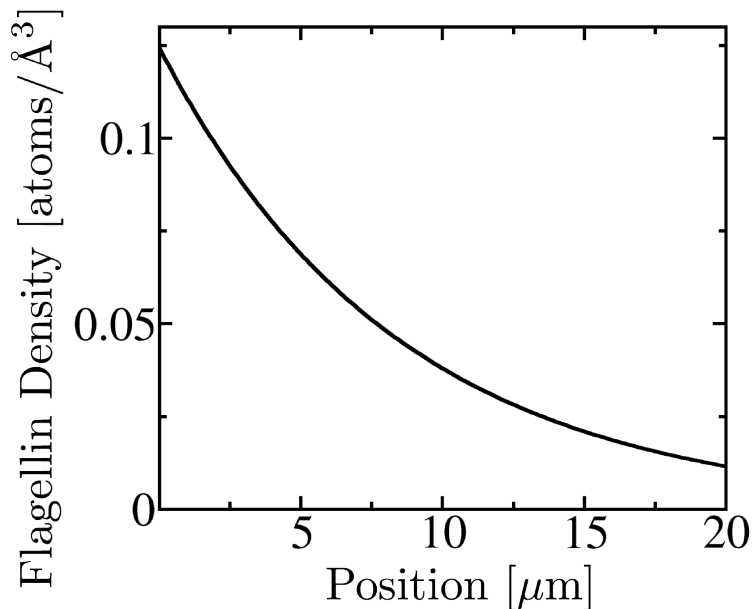
Figure S2: Flagellin Density. Using the system parameters listed in Table 2, including $P = 100$ (kcal/mol)/s, with Eq. 13 of the main text, flagellin density is shown as a function of position along the flagellum for the case $L = 20$ μm. Density is highest at the flagellum base (position = 0) and lowest at the flagellum tip (position = 20 μm). This plot reveals the flagellin density range relevant to the flagellum, namely density in the range [0.011, 0.12] atoms/Å^3. For comparison, the density of fully stretched flagellin in the channel is $\sim$0.01 atoms/Å^3, while the density of bulk water is $\sim$0.1 atoms/Å^3.
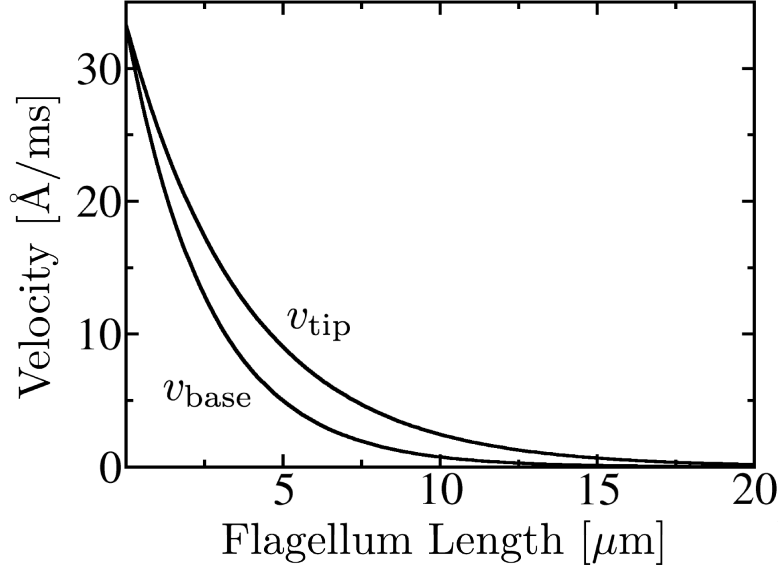
Figure S3: Flagellin Translocation Velocity. Using the system parameters listed in Table 2, including $P = 100$ (kcal/mol)/s, with Eq. 19 of the main text, flagellin translocation velocity is shown as a function of flagellum length. The upper curve plots $v_{\text{tip}} = v(x = 0, L)$, the translocation velocity at the flagellum tip, where velocity is always highest; the lower curve plots $v_{\text{base}} = v(x = L, L)$, the translocation velocity at the flagellum base, where velocity is always lowest. For $L = 0$ (nascent flagellum) tip and base translocation velocity are equal and highest since there is very little friction built up yet in the flagellum. As the flagellum grows to its maximal length, both v_{tip} and v_{base} decay to zero. The area between the two curves bounds the range of relevant translocation velocities as a function of flagellum length.
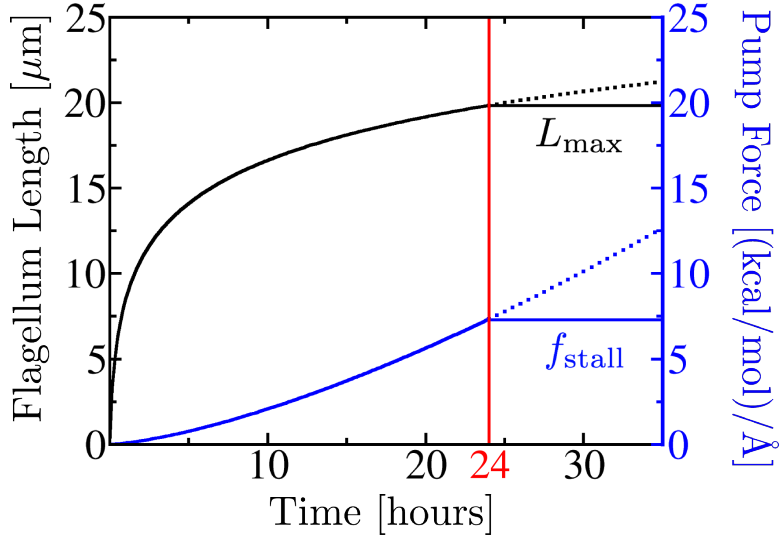
4

Figure S4: Flagellin Growth. Using the system parameters listed in Table 2, including $P = 100$ (kcal/mol)/s, flagellum length (black, c.f. Eq. 25) and pump force (blue, c.f. Eq. 9) is shown as a function of time. As the flagellum grows, the force required to pump additional flagellin into the channel increases (due to increasing $p(x = L)$). At $t = 24$ hrs (red) the pump force reaches the stall force, $f_{\text{stall}} = 7$ (kcal/mol)/Å, and the pump can not surmount the force required to pump additional flagellin into the channel (dotted blue). At this point, the flagellum can not grow longer (black dotted); instead it remains at its maximal length, $L_{\text{max}} = 20$ μm.
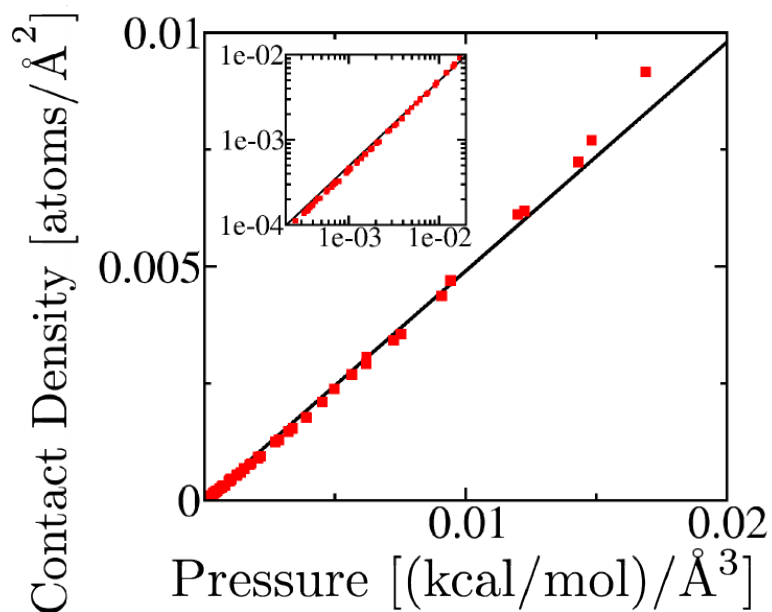
Figure S5: Surface Contact Density. Surface contact density, measured as number of atoms beyond the confining cylinder, versus flagellin pressure, measured as the total force of the confining cylinder on the flagellin segment, for each of the 60 equilibrium simulations. Samples at upper right are from simulations at higher density while samples at lower left are from simulations at lower density. Because atoms which contact the cylinder represent flagellin atoms which would interact with the flagellar channel (were it present), the density of contacts between unfolded flagellin and the confining channel are proportional to flagellin pressure, confirming Condition 2. Inset: The same data is presented in a log-log plot to highlight the proportional relationship across a wide range of flagellin densities.

Pressure - Density Relationship of a Gaussian Chain Polymer Confined to a Cylinder

In the following we seek to show that an exponent $\beta \approx 3$ results in the pressure - density relationship $p = \gamma \rho^\beta$ for a polymer confined to a cylinder of radius a and length L. The confinement applies to the unfolded flagellin in the flagellar channel as studied in the present paper. In our description we consider the protein to be a linear polymer, i.e., we neglect the detailed structure of the amino acid side chains.

The polymer property is accounted for through a so-called Gaussian chain model (1, 2). In this model a polymer consists of N segments represented through vectors $\mathbf{r}_j$, $j = 1, 2, \ldots N$, such that the vector connecting the beginning with the end of the polymer is given by $\sum_{j=1}^{N} \mathbf{r}_j$. The $\mathbf{r}_j$ are distributed isotropically according to a Gaussian distribution

$$\phi(\mathbf{r}_j) = \left[3/2\pi b^2\right]^{\frac{3}{2}} \exp\left[-3r_j^2/2b^2\right] . \tag{1}$$

Since the distributions $\phi(\mathbf{r}_j)$ are independent of each other, the polymer can intersect with itself. The Gaussian model, therefore, does not apply well to a description where each segment described through $\mathbf{r}_j$ is literally a polymer segment, but rather applies best, though still only very approximately, when the $\mathbf{r}_j$ represent large polymer units with a length $|\mathbf{r}_j| = b$; b is suggested to be two times the so-called persistence length of the polymer (2) (see Sect. 2.6.2). The persistence length is the polymer correlation length characteristic of the directional order; beyond the persistence length, directional correlation of the chain is lost.

The Gaussian chain polymer behaves like an entropic spring governed by the Hamiltonian (2) (see Sect. 2.3)

$$H_0 = \frac{3k_{\mathrm{B}}T}{2b^2} \sum_{i=1}^{N} |\mathbf{R}_{i+1} - \mathbf{R}_i|^2 . \tag{2}$$

This behavior follows from the fact that for a polymer with independent segments distributed according to Eq. 1 any polymer piece involving segments $j, j+1, \ldots i-1$ has an end-end distribution (1) (see Chap. II)

$$\phi(\mathbf{R}_i - \mathbf{R}_j, i - j) = \left[3/2\pi b^2|i - j|\right]^{3/2} \exp\left[-3\left(\mathbf{R}_i - \mathbf{R}_j\right)^2/2|i - j|b^2\right] . \tag{3}$$

Here $\mathbf{R}_i$ denotes the starting point of segment i.

Our strategy to determine the pressure - density relationship for the present polymer model will be to exploit the relationship

$$p = -\frac{1}{\pi a^2} \frac{dF}{dL} \ . \tag{4}$$

Here F is the free energy of the system related to the partition function Z through $F = -k_{\mathrm{B}}T \ln Z$. Z can be expressed through the end-end distribution function $\rho(\mathbf{R}, \mathbf{R}', N)$ where $\mathbf{R}'$ is the starting point of the polymer, $\mathbf{R}$ is the end point of the polymer, and N denotes the number of polymer segments

$$Z = \int\int d\mathbf{R}\,d\mathbf{R}'\,\rho(\mathbf{R}, \mathbf{R}', N)\ . \tag{5}$$

The polymer described in Eq. 4 and Eq. 5 has to be confined to the cylinder of radius a and length L. However, the polymer described through Eq. 1, Eq. 2 and Eq. 3 is not confined. In order to obtain $\rho(\mathbf{R}, \mathbf{R}', N)$ for a confined polymer we apply, following (2), to the unconfined polymer a confining potential $V(\mathbf{R}_j)$, to be specified further below. For this purpose we modify the Hamiltonian H_0 to

$$H = \frac{3k_{\mathrm{B}}T}{2b^2} \sum_{i=1}^{N} |\mathbf{R}_{i+1} - \mathbf{R}_i|^2 + \sum_{j=0}^{N+1} V(\mathbf{R}_j)\ . \tag{6}$$

The sum over segments is extended from N to $N+1$ to include in the potential the term $V(\mathbf{R}_{N+1})$ for the end point of the polymer. The polymer subjected to the potential $V(\mathbf{R}_j)$ exhibits a distribution of its conformations $\{\mathbf{R}_1, \mathbf{R}_2, \ldots \mathbf{R}_{N+1}\}$

$$Q(\mathbf{R}_1, \mathbf{R}_2, \ldots \mathbf{R}_{N+1}) \sim \exp\left[-\frac{3}{2b^2} \sum_{i=1}^{N} |\mathbf{R}_{i+1} - \mathbf{R}_i|^2 - \sum_{j=0}^{N+1} \frac{1}{k_{\mathrm{B}}T} V(\mathbf{R}_j)\right] \tag{7}$$

The corresponding end-end distribution function $\rho(\mathbf{R}, \mathbf{R}', N)$, where $\mathbf{R}_1 = \mathbf{R}$ and $\mathbf{R}_{N+1} = \mathbf{R}'$, is obtained by averaging over the intermediate polymer bead positions $\{\mathbf{R}_2, \mathbf{R}_3, \ldots \mathbf{R}_N\}$, i.e.,

$$\rho(\mathbf{R}, \mathbf{R}', N) = \int d\mathbf{R}_2 \int d\mathbf{R}_3 \cdots \int d\mathbf{R}_N \, Q(\mathbf{R}, \mathbf{R}_2, \ldots \mathbf{R}_N, \mathbf{R}')\ . \tag{8}$$

Using Eq. 7 this can be written symbolically

$$\rho(\mathbf{R} = \mathbf{R}_1, \mathbf{R}' = \mathbf{R}_{N+1}, N) \sim \int d\mathbf{R}_2 \int d\mathbf{R}_3 \cdots \int d\mathbf{R}_N \times$$

$$\times \exp\left[-\frac{3}{2b^2} \sum_{i=1}^{N} |\mathbf{R}_{i+1} - \mathbf{R}_i|^2 - \sum_{j=0}^{N+1} \frac{1}{k_{\mathrm{B}}T} V(\mathbf{R}_j)\right] \qquad (9)$$

Carrying out this calculation for $V \equiv 0$ yields

$$\rho_0(\mathbf{R}, \mathbf{R}', N) = \left[3/2\pi b^2 N\right]^{3/2} \exp\left[-3\left(\mathbf{R} - \mathbf{R}'\right)^2/2Nb^2\right] \qquad (10)$$

which agrees, as expected, with the unconfined polymer distribution function in Eq. 3.

In order to determine $\rho(\mathbf{R}, \mathbf{R}', N)$ for an external potential $V(\mathbf{R}_j)$ we do not carry out the integration in Eq. 9, but rather follow a procedure suggested in (2) (see Sect. 3.2) that determines $\rho(\mathbf{R}, \mathbf{R}', N)$ as a solution of the partial differential equation

$$\left[\frac{\partial}{\partial N} - \frac{b^2}{6}\nabla^2 + \frac{1}{k_{\mathrm{B}}T} V(\mathbf{R})\right] \rho(\mathbf{R}, \mathbf{R}', N) = 0 \qquad (11)$$

where ∇^2 is the Laplacian with respect to $\mathbf{R}$. This solution must obey the condition

$$\rho(\mathbf{R}, \mathbf{R}', 0) = \delta(\mathbf{R} - \mathbf{R}') . \qquad (12)$$

In the present case, the confining potential for a cylinder is defined through

$$V(\mathbf{R}) = \begin{cases} 0 & \mathbf{R} \text{ inside cylinder} \\ \infty & \mathbf{R} \text{ outside cylinder} \end{cases} . \qquad (13)$$

One can conclude readily from this definition that $\rho(\mathbf{R}, \mathbf{R}', N)$ should obey then

$$\left[\frac{\partial}{\partial N} - \frac{b^2}{6}\nabla^2\right] \rho(\mathbf{R}, \mathbf{R}', N) = 0 \qquad (14)$$

and vanish for $\mathbf{R}$ and $\mathbf{R}'$ on the cylinder surface. It should also obey condition Eq. 12.

Equation 14 with $N = t$ (time) and $b^2/6 = D$ (diffusion coefficient), is equivalent to the free diffusion equation, solutions of which are well known. In fact, the solution of Eq. 14 for a cylindrical boundary condition can be

found in (3) (see Sect. 8.6). According to Eq. 5 we need to determine actually the quantity

$$\sigma(\mathbf{R}, N) = \int d\mathbf{R}' \, \rho(\mathbf{R}, \mathbf{R}', N) \, . \tag{15}$$

In the case of axial symmetry, that applies in the present case, one obtains the solution as stated in (3) after some simple algebra

$$\sigma(\mathbf{R}, N) = \sum_{\mu} \sum_{m=1}^{\infty} A_{\mu m} \, J_0(\mu r) \, \sin[m\pi z/L] \, \exp[-D\{\mu^2 + (m^2\pi^2/L^2)\} N] \tag{16}$$

where we used the cylindrical coordinates $\mathbf{R} = (r, \phi, z)$ (ϕ drops out due to axial symmetry). The expansion coefficients $A_{\mu m}$ are

$$A_{\mu m} = \frac{8}{a^2 L [J_1(\mu a)]^2} \int_0^a r \, J_0(\mu r) \, dr \int_0^L \sin[m\pi z/L] dz \, . \tag{17}$$

In Eq. 16 and Eq. 17, J_0 and J_1 are the cylindrical Bessel functions of zeroth and first order; the summation index μ denotes the zeroes of $J_0(\mu a)$. Using

$$\int_0^a r \, J_0(\mu r) \, dr = \frac{a}{\mu} \, J_1(\mu a) \tag{18}$$

and

$$\int_0^L \sin[m\pi z/L] dz = \frac{L}{m\pi}[1 - (-1)^m] \tag{19}$$

one obtains finally

$$A_{\mu m} = \frac{8\,[1 - (-1)^m]}{a\,\mu\,m\,\pi\,[J_1(\mu a)]} \, . \tag{20}$$

We need to determine now (c.f. Eq. 15 and Eq. 5)

$$Z = \int d\mathbf{R} \, \sigma(\mathbf{R}, N) \, . \tag{21}$$

where $\sigma(\mathbf{R}, N)$ is given by Eq. 16. Using again the integral values from Eq. 18 and Eq. 19 one obtains

$$Z = \sum_{\mu} \sum_{m=1,3,5...} \frac{64L}{m^2\mu^2\pi} \, \exp[-D\{\mu^2 + (m^2\pi^2/L^2)\} N] \, . \tag{22}$$

10

According to Eq. 4 the pressure is

$$p = \frac{k_{\mathrm{B}}T}{\pi a^2} \frac{\partial \ln Z}{\partial L} \; . \tag{23}$$

Factoring Z into an L-independent and an L-dependent term

$$Z = Z_1 \, Z_2(L) \tag{24}$$

where

$$Z_1 = \sum_{\mu} \frac{64}{\mu^2 \pi} \exp(-D \, \mu^2 \, N) \tag{25}$$

and

$$Z_2(L) = L \sum_{m=1,3,5\ldots} \frac{1}{m^2} \exp(-D \, N m^2 \pi^2 / L^2) \tag{26}$$

one can write

$$p = \frac{k_{\mathrm{B}}T}{\pi a^2} \frac{1}{Z_2(L)} \frac{\partial Z_2(L)}{\partial L} \; . \tag{27}$$

From this results

$$p = \frac{k_{\mathrm{B}}T}{\pi a^2} \left\{ \frac{1}{L} + \frac{N\pi^2 b^2}{3L^3} \frac{\sum_{m=1,3,5\ldots} \exp(-b^2\pi^2 N m^2/6L^2)}{\sum_{n=1,3,5\ldots} (1/n^2) \, \exp(-b^2\pi^2 N n^2/6L^2)} \right\} \; . \tag{28}$$

One can recognize that the two series arising in this expression converge quickly for $\alpha = 3N\pi^2 b^2/2L^2 \gg 1$. In the case of the system simulated in the present case typical α values obey $\alpha > 15$ and, hence, one may limit the two series after the first term, yielding

$$p = \frac{k_{\mathrm{B}}T}{\pi a^2} \left[\frac{1}{L} + \frac{N\pi^2 b^2}{3L^3} \right] \; . \tag{29}$$

In order to compare Eq. 29 to the simulation data in Fig. 5 we employ the expression for the atomic density

$$\rho_{\mathrm{atom}} = N_{\mathrm{atom}}/\pi a^2 L \tag{30}$$

which permits one to express

$$L = N_{\mathrm{atom}}/\pi a^2 \rho_{\mathrm{atom}} \; . \tag{31}$$

One can also express the number of polymer segments N by L_0, the length of the totally stretched polymer main chain,

$$N = L_0/b .$$

(32)

Employing Eq. 31 and Eq. 32 in Eq. 29 results in

$$p = k_{\mathrm{B}}T \left[\frac{\rho_{\mathrm{atom}}}{N_{\mathrm{atom}}} + \frac{\pi^4 a^4 b\, L_0}{3} \left(\frac{\rho_{\mathrm{atom}}}{N_{\mathrm{atom}}} \right)^3 \right] .$$

(33)

Introducing

$$\tilde{\rho} = \rho_{\mathrm{atom}}/N_{\mathrm{atom}} ,$$

(34)

which stands for the density of one polymer chain, one obtains finally the pressure - density relationship for the Gaussian chain model polymer confined to a cylinder

$$p = k_{\mathrm{B}}T \left[\tilde{\rho} + \frac{1}{3}\pi^4 a^4 b\, L_0\, \tilde{\rho}^3 \right] .$$

(35)

For physical parameters typical in the flagellum, the cubic term of Eq. 35 is one to two orders of magnitude larger than the linear term; pressure, therefore, is approximately proportional to the cube of the density, i.e., $p \propto \tilde{\rho}^3$.

References

1. Yamakawa, H. 1971. *Modern Theory of Polymer Solutions.* Harper and Row, New York.

2. Kawakatsu, T. 2004. *Statistical Physics of Polymers : An Introduction.* Springer, Berlin.

3. Carslaw, H. and J. Jaeger. 1959. *Conduction of Heat in Solids.* Oxford University Press, Oxford, England, 2nd edition.