Supplemental Material for:

The range of interaction between DNA-bending proteins is controlled by the second-longest correlation length for bending fluctuations

Houyin Zhang and John F. Marko

S1. TRANSFER MATRIX IN SPHERICAL HARMONIC BASIS

By using symmetries, the definition of the Wigner 3J symbols, and partial-wave decomposition of $e^{\beta b f \hat{t}\cdot\hat{z}}$, we re-expressed the two matrices $\mathbf{A}$ and $\mathbf{B}$ in a space with the spherical harmonics $\{|l,m\rangle\}$ as the basis,

$$\begin{aligned}\langle lm|\mathbf{A}|l'm'\rangle = (-1)^m\delta_{mm'}\sqrt{(2l+1)(2l'+1)}\sqrt{\frac{8\pi}{a'}} \\ \times\int_0^{+\infty} dq[j_{l'}(q)e^{-\frac{q^2}{2a'}}(e^{i(\frac{\pi}{2}l'-q\gamma)}+e^{-i(\frac{\pi}{2}l'-q\gamma)})] \\ \times\sum_{l_1} i_{l_1}(\beta bf)(2l_1+1)\begin{pmatrix} l & l_1 & l' \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l & l_1 & l' \\ -m & 0 & m\end{pmatrix}\ ,\end{aligned} \tag{1}$$

and

$$\begin{aligned}\langle lm|\mathbf{B}|l'm'\rangle = (-1)^m\delta_{mm'}\sqrt{(2l+1)(2l'+1)}[4\pi e^{-a}i_{l'}(a)] \\ \times\sum_{l_1} i_{l_1}(\beta bf)(2l_1+1)\begin{pmatrix} l & l_1 & l' \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l & l_1 & l' \\ -m & 0 & m\end{pmatrix}\ ,\end{aligned} \tag{2}$$

where the $j_l(x)$ are spherical Bessel functions of the first kind, and the $i_l(x)$ are modified spherical Bessel functions of the first kind. Both of these two matrices are block diagonal.

S2. TANGENT VECTOR OPERATORS IN SPHERICAL HARMONIC BASIS

Choosing spherical harmonics $\{|l,m\rangle\}$ as the basis, by making use of Wigner 3J symbols, we obtain the matrix elements of the three tangent vector component operators,

$$\begin{aligned}\langle l,m,n|\hat{\mathbf{t}}^+|l',m',n'\rangle &= \delta_{nn'}(-1)^{m+1}\sqrt{2(2l+1)(2l'+1)}\begin{pmatrix} l & 1 & l' \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l & 1 & l' \\ -m & 1 & m'\end{pmatrix}\ , \\ \langle l,m,n|\hat{\mathbf{t}}^-|l',m',n'\rangle &= \delta_{nn'}(-1)^{m}\sqrt{2(2l+1)(2l'+1)}\begin{pmatrix} l & 1 & l' \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l & 1 & l' \\ -m & -1 & m'\end{pmatrix}\ , \\ \langle l,m,n|\hat{\mathbf{t}}^z|l',m',n'\rangle &= \delta_{nn'}(-1)^{m+1}\sqrt{(2l+1)(2l'+1)}\begin{pmatrix} l & 1 & l' \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l & 1 & l' \\ -m & 0 & m'\end{pmatrix}\ .\end{aligned} \tag{3}$$

S3. TANGENT VECTOR CORRELATION FUNCTIONS FOR NAKED DNA

Figure 1 shows the transverse and longitudinal tangent vector correlation functions and corresponding correlation lengths for naked DNA (DNA in the absence of protein).

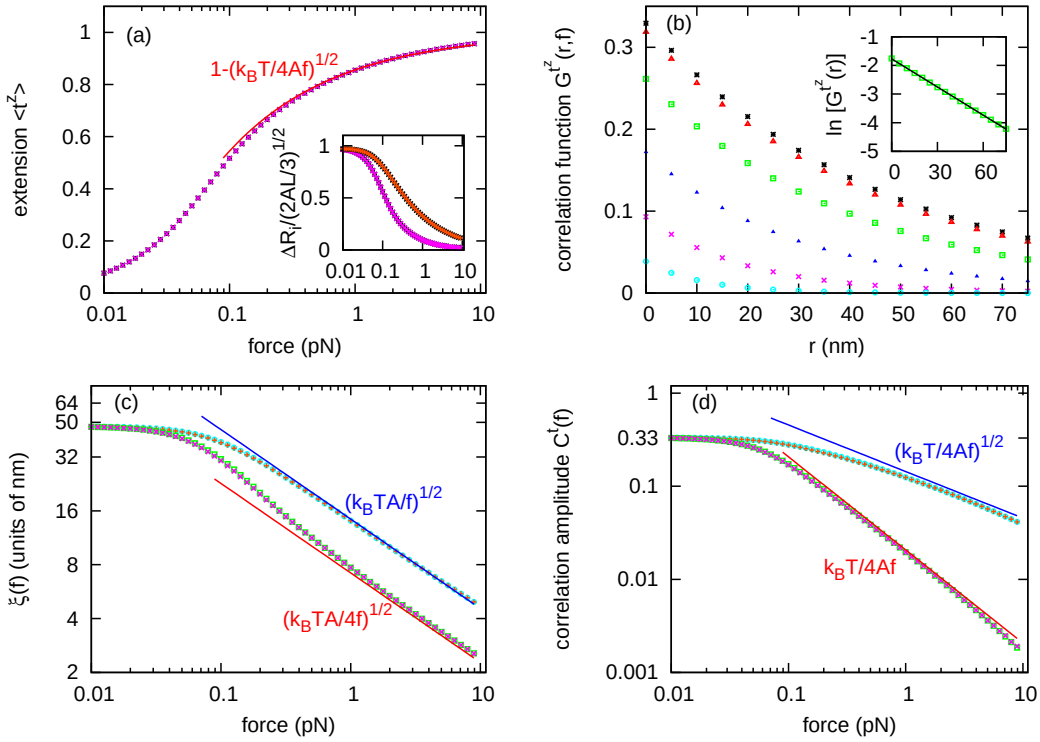

FIG. 1. (color online) Tangent vector correlation functions and correlation lengths for naked DNA. (a) The average z-component of the tangent vector, $\langle \hat{\mathbf{t}}^z \rangle$, as a function of force (pink crosses), agreeing with the normalized force-extension curve for naked DNA from Ref. [1, 2] (black asterisks) and the high-force limit of the extension from continuous WLC model [3] (red curve). Inset shows the extension fluctuation curves: z-direction extension fluctuation calculated from correlation function $G^{\hat{\mathbf{t}}^z}(r,f)$ (bottom pink crosses) and the corresponding result from Ref. [1] (bottom black asterisks, overlapping with bottom pink crosses); x (or y)-direction extension fluctuation calculated from $G^{\hat{\mathbf{t}}^x}(r,f)$ (top orange pluses) and the corresponding result calculated by Eq. A6 of Ref. [4] (top black asterisks, overlapping with top orange pluses). (b) The correlation function for the z-component of tangent vector $G^{\hat{\mathbf{t}}^z}(r,f)$ as a function of contour length $r = jb$ for a series of fixed forces (different curves). The top curve has force $f = 0.01$ pN (black asterisks), then in decreasing order the curves correspond to $f = 0.02$ (red triangles), 0.05 (green squares), 0.1 (blue filled triangles), 0.2 (pink crosses), and 0.5 (aqua circles) pN. Inset depicts $\ln(G^{\hat{\mathbf{t}}^z}(r))$ versus the distance at $f = 0.1$ pN (green squares) and a linear fit (black line), showing that the correlation function decays exponentially. (c) The longitudinal correlation length $\xi_\parallel$ versus applied force: extracted from correlation function $G^{\hat{\mathbf{t}}^z}(r,f)$ (bottom pink crosses), asymptotic result from eigenvalues (bottom green squares, overlapping with bottom pink crosses), and high-force limit from continuous WLC model [3] (bottom red line); the transverse correlation length $\xi_\perp$ versus applied force: extracted from correlation function $G^{\hat{\mathbf{t}}^x}(r,f)$ (top orange pluses), asymptotic result from eigenvalues (top aqua circles, overlapping with top orange pluses), and high-force limit from continuous WLC model [3] (top blue line). In high-force range, $\xi_\perp = 2\xi_\parallel = \sqrt{k_B TA/f}$. (d) The longitudinal correlation function amplitude $C_\parallel(f)$: extracted from correlation function $G^{\hat{\mathbf{t}}^z}(r,f)$ (bottom pink crosses), $G^{\hat{\mathbf{t}}^z}(0,f)$ (bottom green squares, overlapping with bottom pink crosses), and high-force limit from continuous WLC model [3] (bottom red line); the transverse correlation function amplitude $C_\perp(f)$: extracted from correlation function $G^{\hat{\mathbf{t}}^x}(r,f)$ (top orange pluses), $G^{\hat{\mathbf{t}}^x}(0,f)$ (top aqua circles, overlapping with top orange pluses), and high-force limit from continuous WLC model [3] (top blue line).

S4. HIGH-FORCE LIMIT OF CONTINUOUS WLC MODEL

For the continuous WLC model, the effective energy of a stretched naked DNA molecule is expressed as [3, 4],

$$\beta E = \int_0^L ds \left[\frac{A}{2} (\partial_s \hat{t})^2 - \beta f \hat{t} \cdot \hat{z} \right] , \tag{4}$$

where L is the contour length of the DNA polymer and $\hat{t}(s)$ is the unit tangent vector describing the curve conformation. In the large-force limit ($f \gg k_B T/A \approx 0.1$ pN), t_x and t_y are small quantities, and $t_z = 1 - (t_x^2 + t_y^2)/2 + \mathcal{O}(t_x^4 + t_y^4)$. Then we obtain the Gaussian approximation to this model, which is valid in the large force limit $f >> k_B T/A$:

$$\beta E \approx \frac{1}{2} \int_0^L ds \left\{ A[(\partial_s t_x)^2 + (\partial_s t_y)^2] + \beta f [t_x^2 + t_y^2] \right\} - \beta f L = \frac{1}{2} \int \frac{dq}{2\pi} (Aq^2 + \beta f)[t_x^2(q) + t_y^2(q)] - \beta f L , \tag{5}$$

where $t_x(q) = \int_0^L t_x e^{-iqs}\, ds$ and $t_y(q) = \int_0^L t_y e^{-iqs}\, ds$.

We calculate the x-direction correlation function,

$$G^{\hat{\mathbf{t}}^x}(s, f) = \langle t_x(0) t_x(s) \rangle = \int_{-\infty}^{\infty} \frac{dq_1}{2\pi} \frac{dq_2}{2\pi} \langle t_x(q_1) t_x(q_2) \rangle e^{iq_2 s} = \int_{-\infty}^{\infty} \frac{dq}{2\pi} \frac{e^{iqs}}{Aq^2 + \beta f} = \sqrt{k_B T/4Af}\; e^{-s/\sqrt{k_B T A/f}}. \tag{6}$$

Similarly, we calculate the z-direction correlation function,

$$\begin{aligned}
G^{\hat{\mathbf{t}}^z}(s, f) &= \langle t_z(0) t_z(s) \rangle - \langle t_z(0) \rangle \langle t_z(s) \rangle \\
&\approx \langle \left[1 - t_x^2(0)/2 - t_y^2(0)/2\right] \left[1 - t_x^2(s)/2 - t_y^2(s)/2\right] \rangle - \langle 1 - t_x^2(0)/2 - t_y^2(0)/2 \rangle \langle 1 - t_x^2(s)/2 - t_y^2(s)/2 \rangle \\
&= \frac{1}{4} \Big[\langle t_x^2(0) t_x^2(s) \rangle + \langle t_x^2(0) t_y^2(s) \rangle + \langle t_y^2(0) t_x^2(s) \rangle + \langle t_y^2(0) t_y^2(s) \rangle \\
&\quad - \langle t_x^2(0) \rangle \langle t_x^2(s) \rangle - \langle t_x^2(0) \rangle \langle t_y^2(s) \rangle - \langle t_y^2(0) \rangle \langle t_x^2(s) \rangle - \langle t_y^2(0) \rangle \langle t_y^2(s) \rangle \Big] \\
&= \frac{1}{2} \left[\langle t_x^2(0) t_x^2(s) \rangle - \langle t_x^2(0) \rangle \langle t_x^2(s) \rangle \right] \\
&= \frac{1}{2} \left[\int_{-\infty}^{\infty} \frac{dq_1}{2\pi} \frac{dq_2}{2\pi} \frac{dq_3}{2\pi} \frac{dq_4}{2\pi} \langle t_x(q_1) t_x(q_2) t_x(q_3) t_x(q_4) \rangle e^{iq_3 s} e^{iq_4 s} - \langle t_x^2(0) \rangle \langle t_x^2(s) \rangle \right] \\
&= \left(\int_{-\infty}^{\infty} \frac{dq}{2\pi} \frac{e^{iqs}}{Aq^2 + \beta f} \right)^2 = (k_B T/4Af)\; e^{-s/\sqrt{k_B T A/4f}} .
\end{aligned} \tag{7}$$

From the forms of the correlation functions we observe that the longitudinal correlation length is exactly half of the transverse correlation length. The amplitudes of these leading contributions to the two correlation functions are also exactly related, with the longitudinal amplitude being exactly the square of the transverse one. Thus, in the high-force limit, the leading (slowest decaying) contribution to the longitudinal correlation function (7) is just the square of the leading contribution to the transverse correlation function (6).

S5. THE LARGE-N LIMIT OF THE CORRELATION FUNCTIONS $G^{\hat{\mathbf{t}}^x}(r,f)$ AND $G^{\hat{\mathbf{t}}^z}(r,f)$ FOR NAKED DNA

For naked DNA ($\mu=-\infty$), $\mathbf{T}=\begin{pmatrix}\mathbf{A}e^{\mu+\eta} & \mathbf{A}e^{\mu}\\ \mathbf{B} & \mathbf{B}\end{pmatrix}=\begin{pmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{B} & \mathbf{B}\end{pmatrix}$. So,

$$\begin{aligned}
G^{\hat{\mathbf{t}}^z}(r,f) &= \frac{\mathrm{Tr}(\hat{\mathbf{t}}^z\mathbf{B}^j\hat{\mathbf{t}}^z\mathbf{B}^{N-j})}{\mathrm{Tr}(\mathbf{B}^N)}-\left(\frac{\mathrm{Tr}(\hat{\mathbf{t}}^z\mathbf{B}^N)}{\mathrm{Tr}(\mathbf{B}^N)}\right)^2\\
&= \frac{\sum_{m,k_1,k_2}\left(\mathbf{t}_m^z\right)_{k_1,k_2}\left(\lambda_{k_2}^m\right)^j\left(\mathbf{t}_m^z\right)_{k_2,k_1}\left(\lambda_{k_1}^m\right)^{N-j}}{\sum_{m',k'}\left(\lambda_{k'}^{m'}\right)^N}-\left(\frac{\sum_{m,k}\left(\mathbf{t}_m^z\right)_{k,k}\left(\lambda_k^m\right)^N}{\sum_{m',k'}\left(\lambda_{k'}^{m'}\right)^N}\right)^2\\
&\approx \sum_{m,k_1,k_2}\left(\mathbf{t}_m^z\right)_{k_1,k_2}\left(\mathbf{t}_m^z\right)_{k_2,k_1}\left(\frac{\lambda_{k_2}^m}{\lambda_{k_1}^m}\right)^j\left(\frac{\lambda_{k_1}^m}{\lambda_0^0}\right)^N-\left(\sum_{m,k}\left(\mathbf{t}_m^z\right)_{k,k}\left(\frac{\lambda_k^m}{\lambda_0^0}\right)^N\right)^2\\
&\approx \sum_{k_2=0}\left(\mathbf{t}_0^z\right)_{0,k_2}\left(\mathbf{t}_0^z\right)_{k_2,0}\left(\frac{\lambda_{k_2}^0}{\lambda_0^0}\right)^j-\left[\left(\mathbf{t}_0^z\right)_{0,0}\right]^2\\
&= \sum_{k_2=1}\left(\mathbf{t}_0^z\right)_{0,k_2}\left(\mathbf{t}_0^z\right)_{k_2,0}\left(\frac{\lambda_{k_2}^0}{\lambda_0^0}\right)^j
\end{aligned}\tag{8}$$

where the λ_k^m ($m=0,\pm1,\pm2,\cdots$; $k=|m|,|m|+1,|m|+2,\cdots$) are the eigenvalues of the m-th diagonal block of the naked DNA matrix $\mathbf{B}$ and $\lambda_{|m|}^m>\lambda_{|m|+1}^m>\lambda_{|m|+2}^m>\cdots$. The S6 of the Supplemental Material will show how applied force f affects the eigenvalues $\{\lambda_k^m\}$, *i.e.*, the transfer matrix spectrum of naked DNA, where λ_0^0 is the largest eigenvalue. In the large-N limit, only the terms with $m=0$ and $k_1=0$ survive. Here $\left(\mathbf{t}_m^z\right)_{k_1,k_2}$ is the matrix element of the m-th diagonal block of the matrix $\mathbf{t}^z$ (ignoring n-dependence) expressed by taking the eigenvectors of the m-th diagonal block of the naked DNA matrix $\mathbf{B}$ ($\{|\lambda_k^m\rangle\}$) as the basis, $\left(\mathbf{t}_m^z\right)_{k_1,k_2}=\langle\lambda_{k_1}^m|\hat{\mathbf{t}}_m^z|\lambda_{k_2}^m\rangle$. Similarly,

$$\begin{aligned}
G^{\hat{\mathbf{t}}^x}(r,f) &= \frac{\mathrm{Tr}(\hat{\mathbf{t}}^+\mathbf{B}^j\hat{\mathbf{t}}^-\mathbf{B}^{N-j})}{2\mathrm{Tr}(\mathbf{B}^N)}=\frac{\sum_m\left(\mathbf{t}^+_{\mathbf{m},\mathbf{m-1}}\right)\left(\mathbf{B}_{\mathbf{m-1}}\right)^j\left(\mathbf{t}^-_{\mathbf{m-1},\mathbf{m}}\right)\left(\mathbf{B}_{\mathbf{m}}\right)^{N-j}}{2\sum_{m'}\left(\mathbf{B}_{\mathbf{m}'}\right)^N}\\
&= \frac{\sum_{m,k_1,k_2}\left(\mathbf{t}^+_{m,m-1}\right)_{k_1,k_2}\left(\lambda_{k_2}^{m-1}\right)^j\left(\mathbf{t}^-_{m-1,m}\right)_{k_2,k_1}\left(\lambda_{k_1}^m\right)^{N-j}}{2\sum_{m',k'}\left(\lambda_{k'}^{m'}\right)^N}\\
&\approx \frac{1}{2}\sum_{m,k_1,k_2}\left(\mathbf{t}^+_{m,m-1}\right)_{k_1,k_2}\left(\mathbf{t}^-_{m-1,m}\right)_{k_2,k_1}\left(\frac{\lambda_{k_2}^{m-1}}{\lambda_{k_1}^m}\right)^j\left(\frac{\lambda_{k_1}^m}{\lambda_0^0}\right)^N\\
&\approx \frac{1}{2}\sum_{k_2=1}\left(\mathbf{t}^+_{0,-1}\right)_{0,k_2}\left(\mathbf{t}^-_{-1,0}\right)_{k_2,0}\left(\frac{\lambda_{k_2}^1}{\lambda_0^0}\right)^j
\end{aligned}\tag{9}$$

where $\left(\mathbf{t}^+_{m,m-1}\right)_{k_1,k_2}=\langle\lambda_{k_1}^m|\hat{\mathbf{t}}^+_{m,m-1}|\lambda_{k_2}^{m-1}\rangle$ and $\left(\mathbf{t}^-_{m,m-1}\right)_{k_1,k_2}=\langle\lambda_{k_1}^{m-1}|\hat{\mathbf{t}}^-_{m-1,m}|\lambda_{k_2}^m\rangle$. Here we also taked use of a symmetry $\lambda_k^m=\lambda_k^{-m}$ (see Eq. 2 in S1 of the Supplemental Material).

We note that according to the notation of this paper, Eq.(20) and Eq.(21) of Ref. [1] should be written as,

$$\Delta F(f,n)=-k_BT\frac{\lambda_0^0}{\lambda_1^0}\frac{\mathbf{A}_{01}^0\mathbf{A}_{10}^0}{\left(\mathbf{A}_{00}^0\right)^2}e^{-n\ln(\lambda_0^0/\lambda_1^0)}\tag{10}$$

and

$$\begin{aligned}
C(f) &= k_BT\frac{\lambda_0^0}{\lambda_1^0}\frac{\mathbf{A}_{01}^0\mathbf{A}_{10}^0}{\left(\mathbf{A}_{00}^0\right)^2}\\
\xi(f) &= \frac{b}{\ln(\lambda_0^0/\lambda_1^0)}
\end{aligned}\tag{11}$$

where $\mathbf{A}_{k_1k_2}^m$ is the matrix element of the m-th diagonal block of the matrix $\mathbf{A}$ expressed by taking the eigenvectors of the m-th diagonal block of the naked matrix $\mathbf{B}$ ($\{|\lambda_k^m\rangle\}$) as the basis, $\mathbf{A}_{k_1,k_2}^m=\langle\lambda_{k_1}^m|\mathbf{A}_m|\lambda_{k_2}^m\rangle$; λ_0^0 and λ_1^0 are the leading and next-to-leading eigenvalues of the 0-th ($m=0$) diagonal block of the naked matrix $\mathbf{B}$, respectively.

S6. TRANSFER MATRIX SPECTRUM FOR NAKED DNA

We calculated the force-dependence of the eigenvalues $\lambda_k^m(f)$ of the naked DNA matrix $\mathbf{B}$, $\langle\hat{t}|\mathbf{B}|\hat{t}'\rangle = e^{-\frac{a}{2}|\hat{t}'-\hat{t}|^2+\beta bf\hat{t}\cdot\hat{z}}$ (see Fig. 2). The free energy spectrum of this naked DNA is:

$$\mathcal{F}_{mk} = -k_BT(L/b)\ln(\lambda_k^m)\ , \tag{12}$$

where L is the contour length of the DNA polymer and b is the segment length. This is also the spectrum of the 1-d classical Heisenberg model. We note that in the low-force limit, the spectrum approaches that of a free rotor (up to a constant, $\propto \ell(\ell+1)$, $\ell = 0, 1, 2, \cdots$) while in the large-force limit, the spectrum approaches that of a harmonic oscillator (again, up to a constant, $\propto n$, $n = 0, 1, 2, \cdots$) [3].

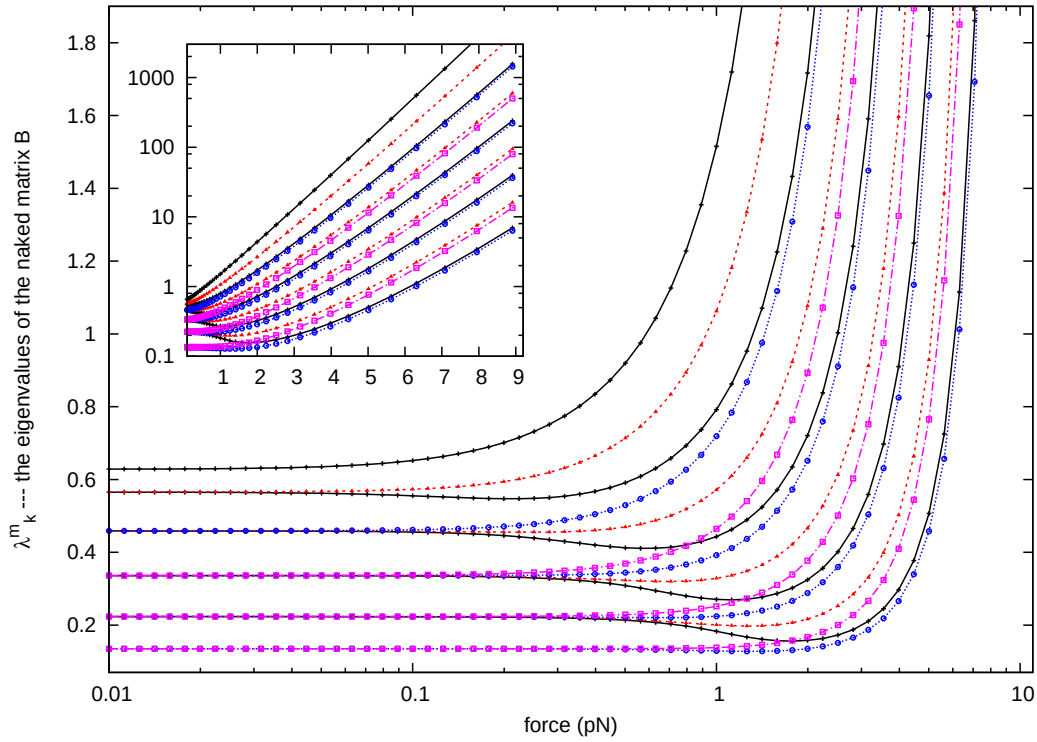

FIG. 2. (color online) Force-dependence of the eigenvalues $\lambda_k^m(f)$ of the "naked" matrix $\mathbf{B}$. Here we only show the largest several eigenvalues for the $m = 0$ block, λ_k^0 (black line and pluses), for the $m = 1$ block, λ_k^1 (red dashed line and filled triangles), for the $m = 2$ block, λ_k^2 (blue dashed line and circles), and for the $m = 3$ block, λ_k^3 (pink line and diamonds). Note that $\lambda_1^1 > \lambda_1^0$. Inset shows that in the high-force range, the spectrum approaches that of a harmonic oscillator, which is closely related to the relation $\xi_\parallel = \xi_\perp/2$.

S7. PROTEIN OCCUPATION CORRELATION FUNCTIONS AND CORRELATION LENGTHS FOR THE CASE $\gamma = 0$ AND $a' = 50$

To focus on the force-generated cooperativity, we turn off the intrinsic protein-protein binding cooperativity (*i.e.* set $\eta = 0$) and calculate the protein occupation correlation functions and correlation lengths along a DNA molecule in contact with dilute DNA-bending protein solution (see Fig. 3 for the case $\mu = -5$ and Fig. 4 for the case $\mu = -3$). Here, the parameters $\gamma = 0$ and $a' = 50$ correspond to a rather stiff protein-DNA complex with a preferred bending angle of 90°.

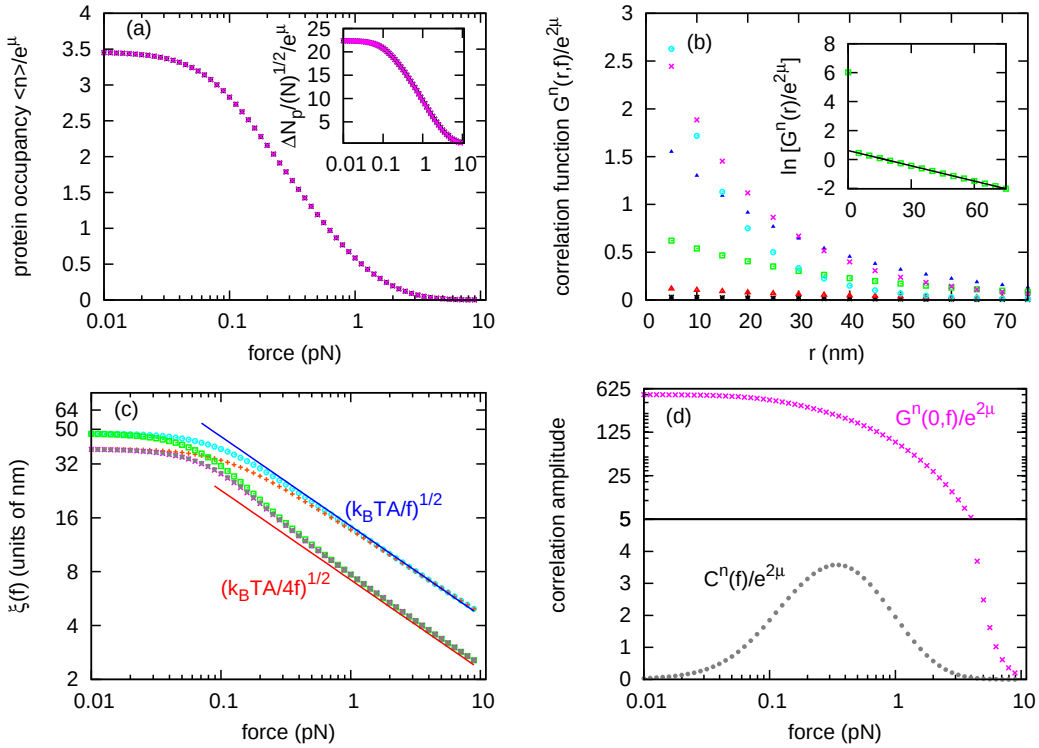

FIG. 3. (color online) Protein occupation correlation functions and correlation lengths along DNA polymer in dilute protein solution ($\mu = -5$) without intrinsic protein-protein binding cooperativity ($\eta = 0$). (a) The protein occupation $\langle n \rangle / e^{\mu}$ as a function of force (pink crosses), agreeing with the normalized protein binding number from Ref.[1] (black asterisks). Inset shows the protein occupation fluctuation curve calculated from protein occupation correlation functions (pink crosses), which matches the corresponding result from Ref. [1] (black asterisks). (b) The protein occupation correlation function $G^{\hat{\mathbf{n}}}(r, f)/e^{2\mu}$ as a function of contour length $r = jb$ for a series of fixed forces (different curves). At the left, bottom curve has force $f = 0.01$ pN (black asterisks), then in increasing order the curves show forces $f = 0.02$ (red triangles), 0.05 (green squares), 0.1 (blue filled triangles), 0.2 (pink crosses), and 0.4 (aqua circles) pN. Inset depicts $\ln[G^{\mathbf{n}}(r)/e^{2\mu}]$ versus the distance at $f = 0.1$ pN (green squares) and a linear fit (black line), showing that the non-zero distance protein occupation correlation function $G^{\mathbf{n}}(r, f)$ decays exponentially and that the zero distance protein occupation correlation function $G^{\mathbf{n}}(0, f)$ is bigger than the non-zero distance protein occupation correlation function decaying amplitude. (c) The protein occupation correlation length as a function of applied force extracted from correlation function $G^{\hat{\mathbf{n}}}(r, f)$ (bottom grey filled circles). As reference, here we also show the longitudinal tangent vector correlation length $\xi_{\parallel}(f)$ extracted from correlation function $G^{\hat{\mathbf{t}}^z}(r, f)$ (bottom pink crosses, overlapping with bottom grey filled circles), the transverse tangent vector correlation length $\xi_{\perp}(f)$ extracted from correlation function $G^{\hat{\mathbf{t}}^x}(r, f)$ (top orange pluses), the correlation lengths for naked DNA (bottom green squares for longitudinal and top aqua circles for transverse), and the high-force limit for naked DNA from continuous WLC model [3] (bottom red line for longitudinal and top blue line for transverse). We find the protein occupation correlation length to be (numerically) equal to the longitudinal tangent vector correlation length $\xi_{\parallel}(f)$. (d) The protein occupation correlation amplitude $C^{n}(f)/e^{2\mu}$ (bottom grey filled circles) and the zero-distance protein occupation correlation function $G^{\mathbf{n}}(0, f)/e^{2\mu}$ (top pink crosses).

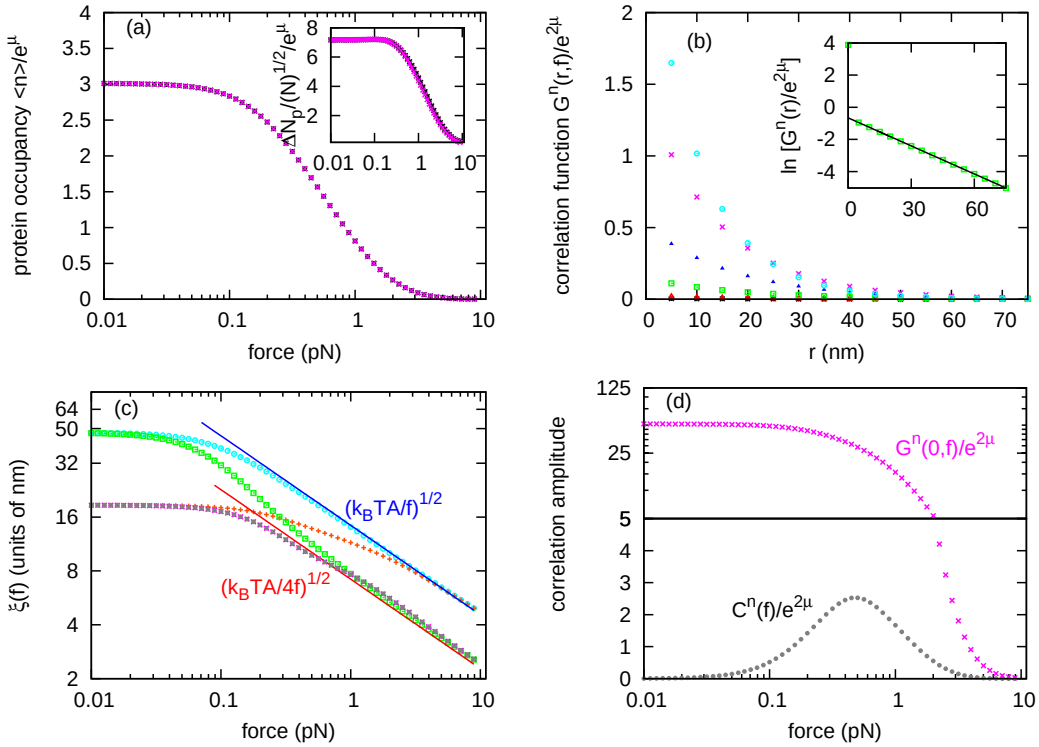

FIG. 4. (color online) Protein occupation correlation functions and correlation lengths along DNA polymer in dilute protein solution as in Fig. 3 except for the use of a different chemical potential value ($\mu = -3$).

From the results we find: (i) the protein occupancy correlation function $G^{\mathbf{n}}(r, f)$ decays exponentially with distance; the zero-distance protein occupation correlation function $G^{\mathbf{n}}(0, f)$ is bigger than the non-zero distance protein occupation correlation function decaying amplitude; (ii) the protein occupancy correlation length, (*i.e.* the decay length) is (numerically) equal to the longitudinal tangent vector correlation length $\xi_{\parallel}(f, \mu)$; (iii) the protein occupancy correlation is small for low force and high force and has a peak value at around 0.5 pN, and the peak value position shifts with increasing of the protein concentration (see panels (d) of Fig. 3 and 4).

We calculated properties of the model for the case of nonzero intrinsic cooperative interaction between DNA-bending proteins. Results are shown in Fig. 5 for $\mu = -5$, $\eta = 1$ and in Fig. 6 for $\mu = -3$, $\eta = 1$.

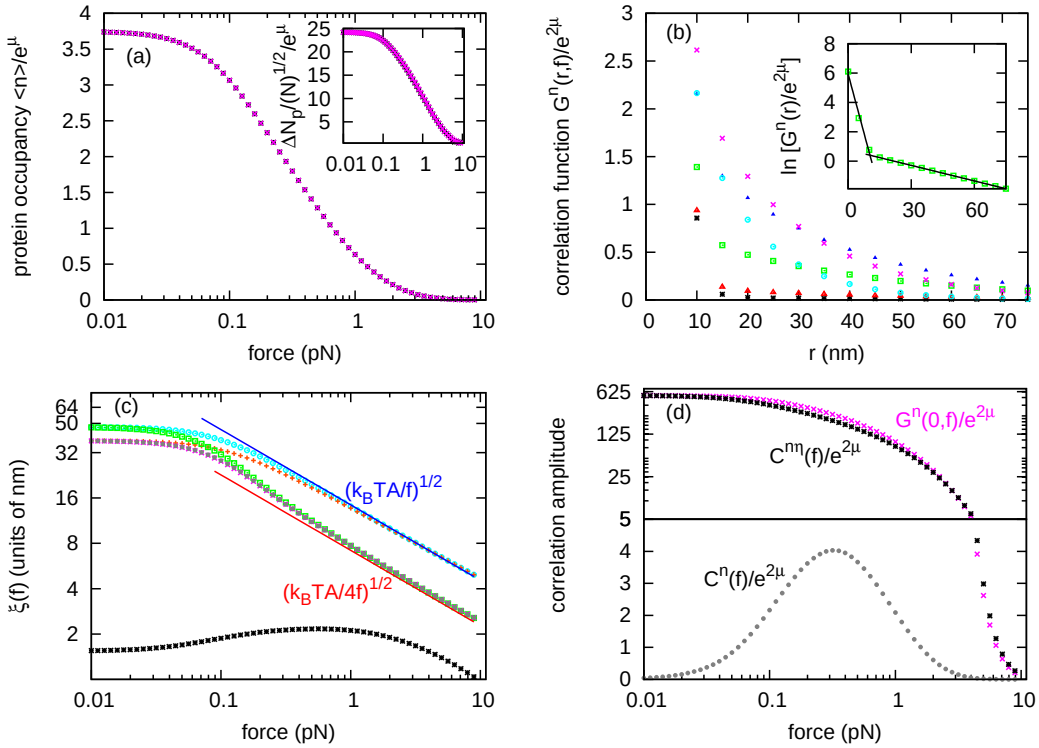

FIG. 5. (color online) Protein occupation correlation functions and correlation lengths along DNA polymer in dilute protein solution ($\mu = -5$) with intrinsic protein-protein binding cooperativity ($\eta = 1$). (a) The protein occupation $\langle n \rangle / e^{\mu}$ as a function of force (pink crosses), agreeing with the normalized protein binding number from Ref.[1] (black asterisks). Inset shows the protein occupation fluctuation curve calculated from protein occupation correlation functions (pink crosses), which matches the corresponding result from Ref. [1] (black asterisks). (b) The protein occupation correlation function $G^{\hat{\mathbf{n}}}(r, f)/e^{2\mu}$ as a function of contour length $r = jb$ for a series of fixed forces (different curves). At the left, bottom curve has force $f = 0.01$ pN (black asterisks), then in increasing order the curves show forces $f = 0.02$ (red triangles), 0.05 (green squares), 0.1 (blue filled triangles), 0.4 (aqua circles), and 0.2 (pink crosses) pN. Inset depicts $\ln[G^{\mathbf{n}}(r)/e^{2\mu}]$ versus the distance at $f = 0.1$ pN (green squares) and a linear fit (black line), showing that the protein occupation correlation function $G^{\mathbf{n}}(r, f)$ includes two kinds of different correlations, and that both of them decay exponentially. (c) The two different protein occupation correlation lengths as a function of applied force: the short-range correlation length (bottom black asterisks) which comes from the intrinsic protein-protein binding cooperativity, and the long-range correlation length (middle grey filled circles) which comes from the longitudinal bending fluctuations. As reference, here we also show the longitudinal tangent vector correlation length $\xi_{\parallel}(f)$ extracted from correlation function $G^{\hat{\mathbf{t}}^z}(r, f)$ (middle pink crosses, overlapping with middle grey filled circles), the transverse tangent vector correlation length $\xi_{\perp}(f)$ extracted from correlation function $G^{\hat{\mathbf{t}}^x}(r, f)$ (top orange pluses), the correlation lengths for naked DNA (middle green squares for longitudinal and top aqua circles for transverse), and the high-force limit for naked DNA from continuous WLC model [3] (middle red line for longitudinal and top blue line for transverse). We find that the long-range protein occupation correlation length is (numerically) equal the longitudinal tangent vector correlation length $\xi_{\parallel}(f)$. (d) The long-range protein occupation correlation amplitude $C^{n}(f)/e^{2\mu}$ (bottom grey filled circles), the short-range protein occupation correlation amplitude $C^{n\eta}(f)/e^{2\mu}$ (top black asterisks), and the zero-distance protein occupation correlation function $G^{\mathbf{n}}(0, f)/e^{2\mu}$ (top pink crosses).

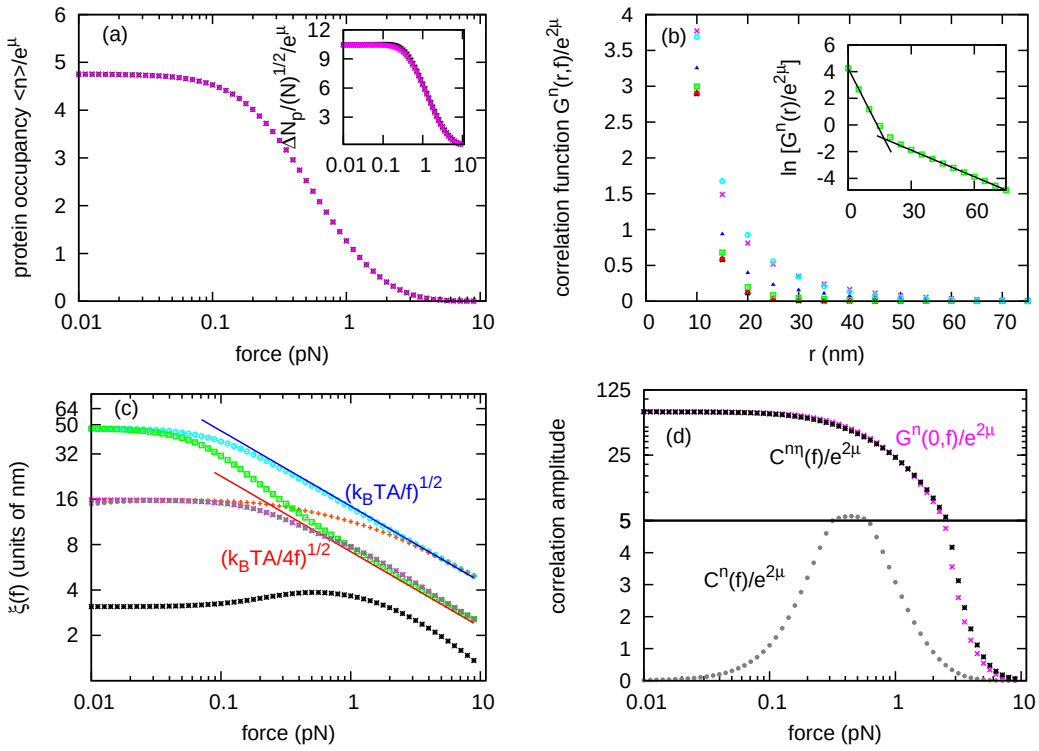

FIG. 6. (color online) Protein occupation correlation functions and correlation lengths along DNA polymer in dilute protein solution as in Fig. 5 except for the use of a different chemical potential value ($\mu = -3$).

Comparing with previous results, we find when there is intrinsic protein-protein binding cooperativity, the interaction gives rise to a new protein occupation correlation which has a bigger amplitude but a shorter correlation length. The short-range protein occupation correlation length increases as the protein concentration (μ) is increased or as cooperativity strength (η) is increased.

S8. PROTEIN OCCUPATION CORRELATION LENGTHS FOR OTHER VALUES OF γ AND a'

We calculated the correlation lengths for varied values of preferred angle θ and bending rigidity a'. Fig. 7 shows results for $\theta = 30°$ and $a' = 50$, Fig. 8 shows results for $\theta = 150°$ and $a' = 50$, and Fig. 9 shows results for $\theta = 90°$ and $a' = 2$.

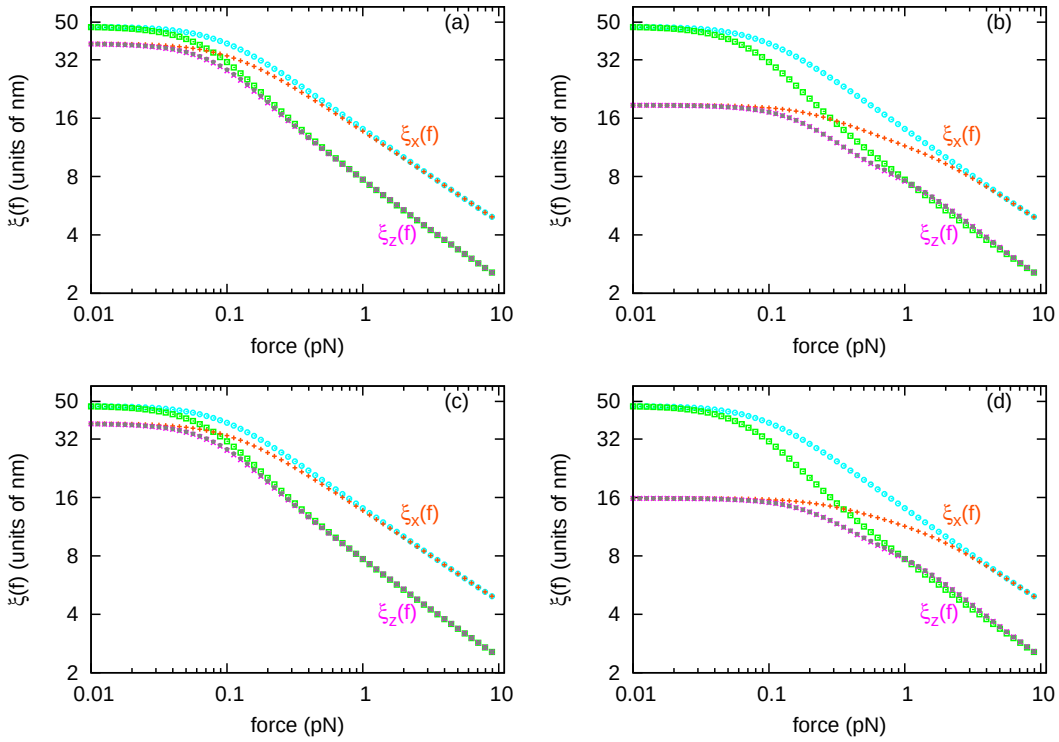

FIG. 7. The longitudinal, transverse, and protein occupation correlation lengths along DNA polymer for $\theta = 30^o$ ($\gamma = \cos 30^o$) and $a' = 50$. (a)-(d) show results for $\mu = -5$ and $\eta = 0$, $\mu = -3$ and $\eta = 0$, $\mu = -5$ and $\eta = 1$, and $\mu = -3$ and $\eta = 1$, respectively. In each panel, we show the longitudinal tangent vector correlation length $\xi_{\|}$ (bottom pink crosses), the transverse tangent vector correlation length $\xi_{\perp}$ (middle orange pluses), and the long-range protein occupation correlation length (bottom grey filled circles, overlapping with bottom pink crosses). As reference, we also show the longitudinal and transverse tangent vector correlation lengths for naked DNA (middle green squares for longitudinal and top aqua circles for transverse).

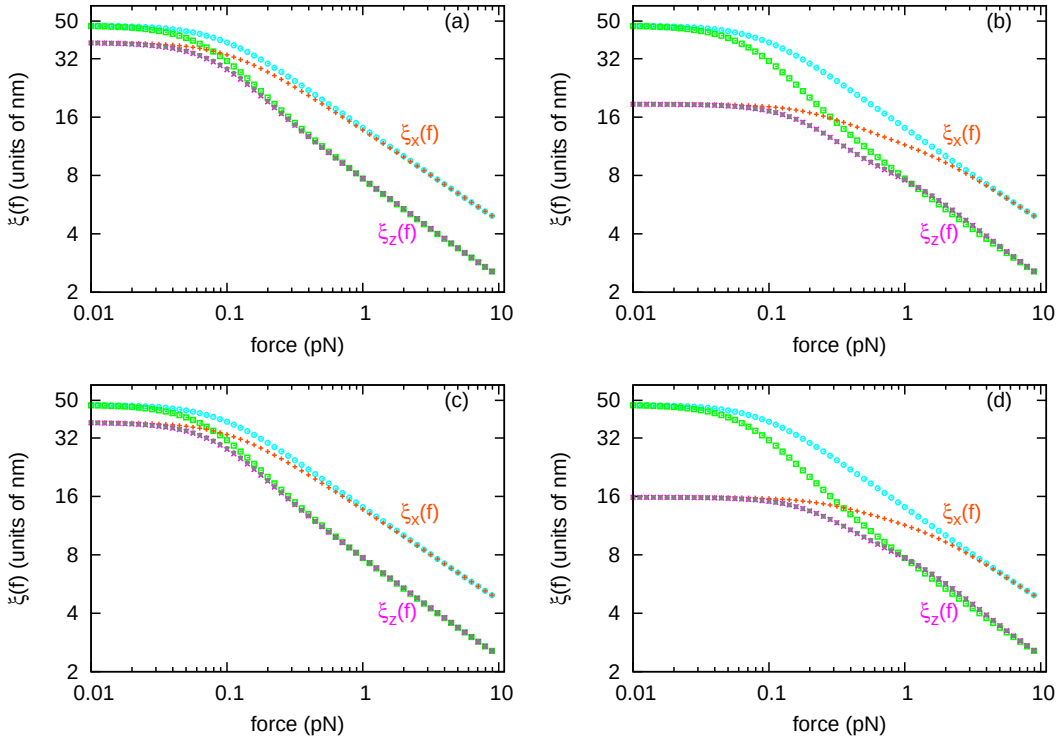

FIG. 8. The longitudinal, transverse, and protein occupation correlation lengths along DNA polymer for $\theta = 150^o$ ($\gamma = \cos 150^o$) and $a' = 50$. (a)-(d) show results for $\mu = -5$ and $\eta = 0$, $\mu = -3$ and $\eta = 0$, $\mu = -5$ and $\eta = 1$, and $\mu = -3$ and $\eta = 1$, respectively; all other parameters and color-coding are the same as in Fig. 7.

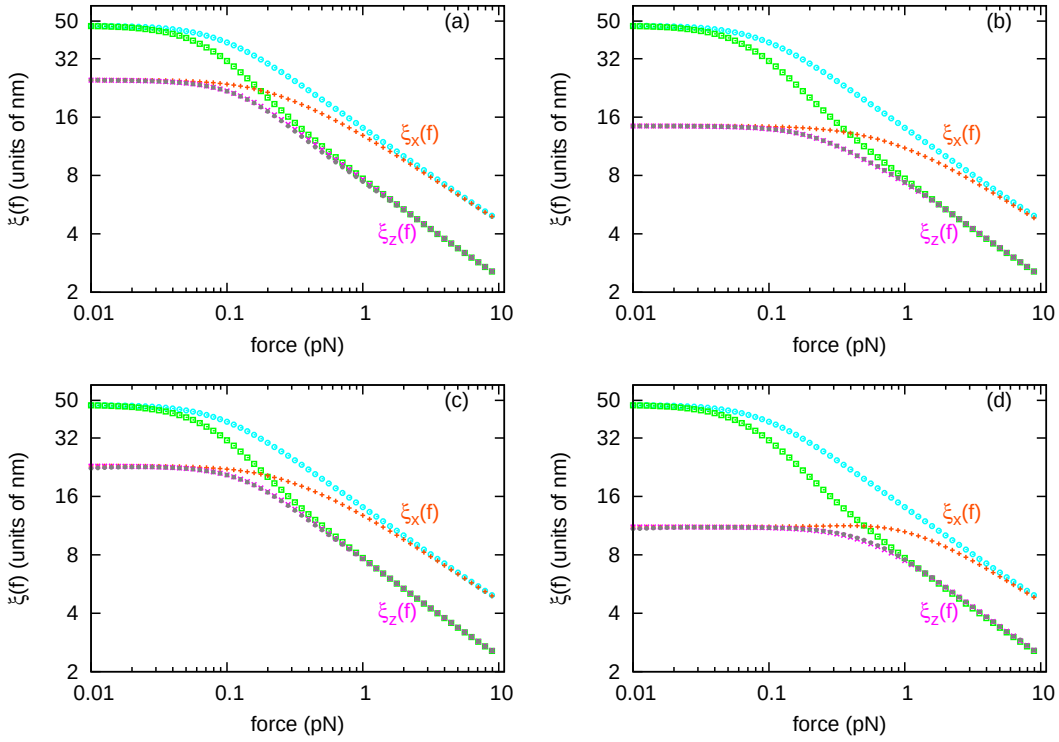

FIG. 9. The longitudinal, transverse, and protein occupation correlation lengths along DNA polymer for $\theta = 90°$ ($\gamma = \cos 90° = 0$) and $a' = 2$. (a)-(d) show results for $\mu = -5$ and $\eta = 0$, $\mu = -4$ and $\eta = 0$, $\mu = -5$ and $\eta = 1$, and $\mu = -4$ and $\eta = 1$, respectively; all other parameters and color-coding are the same as in Fig. 7.

From these results we find that different γ and a' values affect numerical details of the calculation results, but do not affect the main conclusion of the manuscript, *i.e.*, that the range of interaction between DNA-bending proteins is controlled by the second-longest correlation length for bending fluctuations. This is to be expected since different choices of γ and a' do not affect the axial rotation symmetry of the interaction between the proteins and DNA.

[1] H. Zhang and J.F. Marko, Phys. Rev. E **82**, 051906 (2010).
[2] J. Yan and J.F. Marko, Phys. Rev. E **68**, 011905 (2003).
[3] J.F. Marko and E.D. Siggia, Macromolecules **28**, 8759 (1995).
[4] J. Yan, R. Kawamura, and J.F. Marko, Phys. Rev. E **71**, 061905 (2005).

Supplemental Material for:

The range of interaction between DNA-bending proteins is controlled by the second-longest correlation length for bending fluctuations

Houyin Zhang and John F. Marko

S1. TRANSFER MATRIX IN SPHERICAL HARMONIC BASIS

By using symmetries, the definition of the Wigner 3J symbols, and partial-wave decomposition of $e^{\beta b f \hat{t}\cdot\hat{z}}$, we re-expressed the two matrices $\mathbf{A}$ and $\mathbf{B}$ in a space with the spherical harmonics $\{|l,m\rangle\}$ as the basis,

$$\begin{aligned}\langle lm|\mathbf{A}|l'm'\rangle = (-1)^m\delta_{mm'}\sqrt{(2l+1)(2l'+1)}\sqrt{\frac{8\pi}{a'}} \\ \times\int_0^{+\infty} dq[j_{l'}(q)e^{-\frac{q^2}{2a'}}(e^{i(\frac{\pi}{2}l'-q\gamma)}+e^{-i(\frac{\pi}{2}l'-q\gamma)})] \\ \times\sum_{l_1} i_{l_1}(\beta bf)(2l_1+1)\begin{pmatrix} l & l_1 & l' \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l & l_1 & l' \\ -m & 0 & m\end{pmatrix}\ ,\end{aligned} \tag{1}$$

and

$$\begin{aligned}\langle lm|\mathbf{B}|l'm'\rangle = (-1)^m\delta_{mm'}\sqrt{(2l+1)(2l'+1)}[4\pi e^{-a}i_{l'}(a)] \\ \times\sum_{l_1} i_{l_1}(\beta bf)(2l_1+1)\begin{pmatrix} l & l_1 & l' \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l & l_1 & l' \\ -m & 0 & m\end{pmatrix}\ ,\end{aligned} \tag{2}$$

where the $j_l(x)$ are spherical Bessel functions of the first kind, and the $i_l(x)$ are modified spherical Bessel functions of the first kind. Both of these two matrices are block diagonal.

S2. TANGENT VECTOR OPERATORS IN SPHERICAL HARMONIC BASIS

Choosing spherical harmonics $\{|l,m\rangle\}$ as the basis, by making use of Wigner 3J symbols, we obtain the matrix elements of the three tangent vector component operators,

$$\begin{aligned}\langle l,m,n|\hat{\mathbf{t}}^+|l',m',n'\rangle &= \delta_{nn'}(-1)^{m+1}\sqrt{2(2l+1)(2l'+1)}\begin{pmatrix} l & 1 & l' \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l & 1 & l' \\ -m & 1 & m'\end{pmatrix}\ , \\ \langle l,m,n|\hat{\mathbf{t}}^-|l',m',n'\rangle &= \delta_{nn'}(-1)^{m}\sqrt{2(2l+1)(2l'+1)}\begin{pmatrix} l & 1 & l' \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l & 1 & l' \\ -m & -1 & m'\end{pmatrix}\ , \\ \langle l,m,n|\hat{\mathbf{t}}^z|l',m',n'\rangle &= \delta_{nn'}(-1)^{m+1}\sqrt{(2l+1)(2l'+1)}\begin{pmatrix} l & 1 & l' \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l & 1 & l' \\ -m & 0 & m'\end{pmatrix}\ .\end{aligned} \tag{3}$$

S3. TANGENT VECTOR CORRELATION FUNCTIONS FOR NAKED DNA

Figure 1 shows the transverse and longitudinal tangent vector correlation functions and corresponding correlation lengths for naked DNA (DNA in the absence of protein).

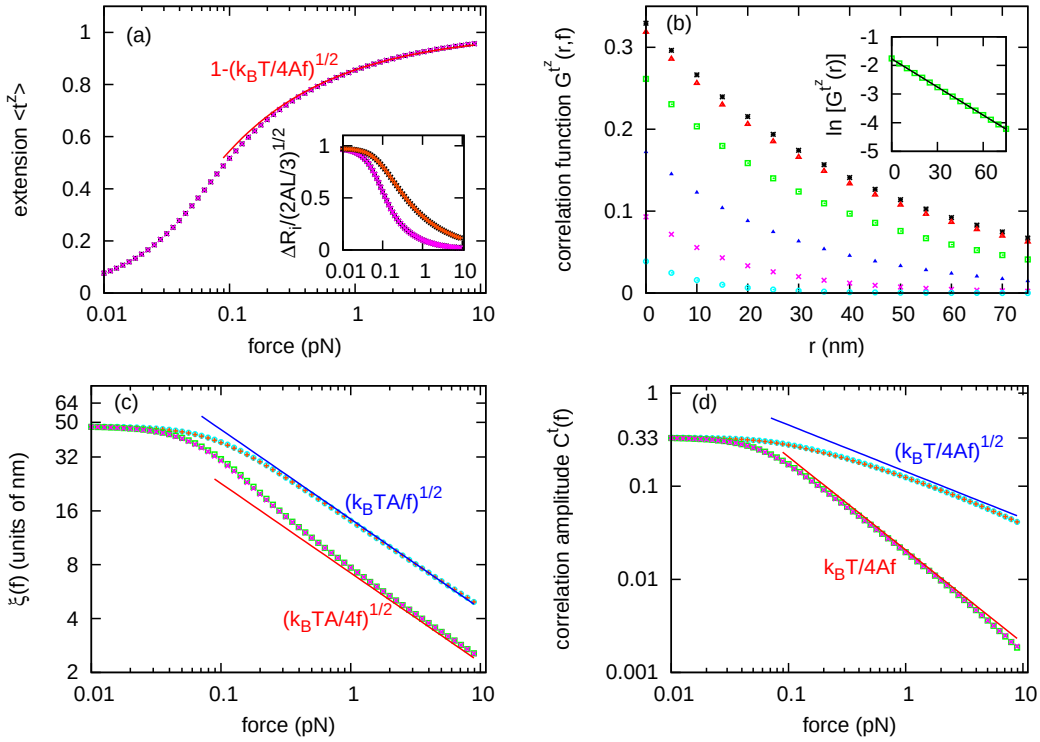

FIG. 1. (color online) Tangent vector correlation functions and correlation lengths for naked DNA. (a) The average z-component of the tangent vector, $\langle \hat{\mathbf{t}}^z \rangle$, as a function of force (pink crosses), agreeing with the normalized force-extension curve for naked DNA from Ref. [1, 2] (black asterisks) and the high-force limit of the extension from continuous WLC model [3] (red curve). Inset shows the extension fluctuation curves: z-direction extension fluctuation calculated from correlation function $G^{\hat{\mathbf{t}}^z}(r,f)$ (bottom pink crosses) and the corresponding result from Ref. [1] (bottom black asterisks, overlapping with bottom pink crosses); x (or y)-direction extension fluctuation calculated from $G^{\hat{\mathbf{t}}^x}(r,f)$ (top orange pluses) and the corresponding result calculated by Eq. A6 of Ref. [4] (top black asterisks, overlapping with top orange pluses). (b) The correlation function for the z-component of tangent vector $G^{\hat{\mathbf{t}}^z}(r,f)$ as a function of contour length $r = jb$ for a series of fixed forces (different curves). The top curve has force $f = 0.01$ pN (black asterisks), then in decreasing order the curves correspond to $f = 0.02$ (red triangles), 0.05 (green squares), 0.1 (blue filled triangles), 0.2 (pink crosses), and 0.5 (aqua circles) pN. Inset depicts $\ln(G^{\hat{\mathbf{t}}^z}(r))$ versus the distance at $f = 0.1$ pN (green squares) and a linear fit (black line), showing that the correlation function decays exponentially. (c) The longitudinal correlation length $\xi_\parallel$ versus applied force: extracted from correlation function $G^{\hat{\mathbf{t}}^z}(r,f)$ (bottom pink crosses), asymptotic result from eigenvalues (bottom green squares, overlapping with bottom pink crosses), and high-force limit from continuous WLC model [3] (bottom red line); the transverse correlation length $\xi_\perp$ versus applied force: extracted from correlation function $G^{\hat{\mathbf{t}}^x}(r,f)$ (top orange pluses), asymptotic result from eigenvalues (top aqua circles, overlapping with top orange pluses), and high-force limit from continuous WLC model [3] (top blue line). In high-force range, $\xi_\perp = 2\xi_\parallel = \sqrt{k_B TA/f}$. (d) The longitudinal correlation function amplitude $C_\parallel(f)$: extracted from correlation function $G^{\hat{\mathbf{t}}^z}(r,f)$ (bottom pink crosses), $G^{\hat{\mathbf{t}}^z}(0,f)$ (bottom green squares, overlapping with bottom pink crosses), and high-force limit from continuous WLC model [3] (bottom red line); the transverse correlation function amplitude $C_\perp(f)$: extracted from correlation function $G^{\hat{\mathbf{t}}^x}(r,f)$ (top orange pluses), $G^{\hat{\mathbf{t}}^x}(0,f)$ (top aqua circles, overlapping with top orange pluses), and high-force limit from continuous WLC model [3] (top blue line).

S4. HIGH-FORCE LIMIT OF CONTINUOUS WLC MODEL

For the continuous WLC model, the effective energy of a stretched naked DNA molecule is expressed as [3, 4],

$$\beta E = \int_0^L ds \left[\frac{A}{2} (\partial_s \hat{t})^2 - \beta f \hat{t} \cdot \hat{z} \right] , \tag{4}$$

where L is the contour length of the DNA polymer and $\hat{t}(s)$ is the unit tangent vector describing the curve conformation. In the large-force limit ($f \gg k_B T/A \approx 0.1$ pN), t_x and t_y are small quantities, and $t_z = 1 - (t_x^2 + t_y^2)/2 + \mathcal{O}(t_x^4 + t_y^4)$. Then we obtain the Gaussian approximation to this model, which is valid in the large force limit $f >> k_B T/A$:

$$\beta E \approx \frac{1}{2} \int_0^L ds \left\{ A[(\partial_s t_x)^2 + (\partial_s t_y)^2] + \beta f [t_x^2 + t_y^2] \right\} - \beta f L = \frac{1}{2} \int \frac{dq}{2\pi} (Aq^2 + \beta f)[t_x^2(q) + t_y^2(q)] - \beta f L , \tag{5}$$

where $t_x(q) = \int_0^L t_x e^{-iqs}\, ds$ and $t_y(q) = \int_0^L t_y e^{-iqs}\, ds$.

We calculate the x-direction correlation function,

$$G^{\hat{\mathbf{t}}^x}(s, f) = \langle t_x(0) t_x(s) \rangle = \int_{-\infty}^{\infty} \frac{dq_1}{2\pi} \frac{dq_2}{2\pi} \langle t_x(q_1) t_x(q_2) \rangle e^{iq_2 s} = \int_{-\infty}^{\infty} \frac{dq}{2\pi} \frac{e^{iqs}}{Aq^2 + \beta f} = \sqrt{k_B T/4Af}\; e^{-s/\sqrt{k_B T A/f}}. \tag{6}$$

Similarly, we calculate the z-direction correlation function,

$$\begin{aligned}
G^{\hat{\mathbf{t}}^z}(s, f) &= \langle t_z(0) t_z(s) \rangle - \langle t_z(0) \rangle \langle t_z(s) \rangle \\
&\approx \langle \left[1 - t_x^2(0)/2 - t_y^2(0)/2\right] \left[1 - t_x^2(s)/2 - t_y^2(s)/2\right] \rangle - \langle 1 - t_x^2(0)/2 - t_y^2(0)/2 \rangle \langle 1 - t_x^2(s)/2 - t_y^2(s)/2 \rangle \\
&= \frac{1}{4} \Big[\langle t_x^2(0) t_x^2(s) \rangle + \langle t_x^2(0) t_y^2(s) \rangle + \langle t_y^2(0) t_x^2(s) \rangle + \langle t_y^2(0) t_y^2(s) \rangle \\
&\quad - \langle t_x^2(0) \rangle \langle t_x^2(s) \rangle - \langle t_x^2(0) \rangle \langle t_y^2(s) \rangle - \langle t_y^2(0) \rangle \langle t_x^2(s) \rangle - \langle t_y^2(0) \rangle \langle t_y^2(s) \rangle \Big] \\
&= \frac{1}{2} \left[\langle t_x^2(0) t_x^2(s) \rangle - \langle t_x^2(0) \rangle \langle t_x^2(s) \rangle \right] \\
&= \frac{1}{2} \left[\int_{-\infty}^{\infty} \frac{dq_1}{2\pi} \frac{dq_2}{2\pi} \frac{dq_3}{2\pi} \frac{dq_4}{2\pi} \langle t_x(q_1) t_x(q_2) t_x(q_3) t_x(q_4) \rangle e^{iq_3 s} e^{iq_4 s} - \langle t_x^2(0) \rangle \langle t_x^2(s) \rangle \right] \\
&= \left(\int_{-\infty}^{\infty} \frac{dq}{2\pi} \frac{e^{iqs}}{Aq^2 + \beta f} \right)^2 = (k_B T/4Af)\; e^{-s/\sqrt{k_B T A/4f}} .
\end{aligned} \tag{7}$$

From the forms of the correlation functions we observe that the longitudinal correlation length is exactly half of the transverse correlation length. The amplitudes of these leading contributions to the two correlation functions are also exactly related, with the longitudinal amplitude being exactly the square of the transverse one. Thus, in the high-force limit, the leading (slowest decaying) contribution to the longitudinal correlation function (7) is just the square of the leading contribution to the transverse correlation function (6).

S5. THE LARGE-N LIMIT OF THE CORRELATION FUNCTIONS $G^{\hat{\mathbf{t}}^x}(r,f)$ AND $G^{\hat{\mathbf{t}}^z}(r,f)$ FOR NAKED DNA

For naked DNA ($\mu=-\infty$), $\mathbf{T}=\begin{pmatrix}\mathbf{A}e^{\mu+\eta} & \mathbf{A}e^{\mu}\\ \mathbf{B} & \mathbf{B}\end{pmatrix}=\begin{pmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{B} & \mathbf{B}\end{pmatrix}$. So,

$$\begin{aligned}
G^{\hat{\mathbf{t}}^z}(r,f) &= \frac{\mathrm{Tr}(\hat{\mathbf{t}}^z\mathbf{B}^j\hat{\mathbf{t}}^z\mathbf{B}^{N-j})}{\mathrm{Tr}(\mathbf{B}^N)}-\left(\frac{\mathrm{Tr}(\hat{\mathbf{t}}^z\mathbf{B}^N)}{\mathrm{Tr}(\mathbf{B}^N)}\right)^2\\
&= \frac{\sum_{m,k_1,k_2}\left(\mathbf{t}_m^z\right)_{k_1,k_2}\left(\lambda_{k_2}^m\right)^j\left(\mathbf{t}_m^z\right)_{k_2,k_1}\left(\lambda_{k_1}^m\right)^{N-j}}{\sum_{m',k'}\left(\lambda_{k'}^{m'}\right)^N}-\left(\frac{\sum_{m,k}\left(\mathbf{t}_m^z\right)_{k,k}\left(\lambda_k^m\right)^N}{\sum_{m',k'}\left(\lambda_{k'}^{m'}\right)^N}\right)^2\\
&\approx \sum_{m,k_1,k_2}\left(\mathbf{t}_m^z\right)_{k_1,k_2}\left(\mathbf{t}_m^z\right)_{k_2,k_1}\left(\frac{\lambda_{k_2}^m}{\lambda_{k_1}^m}\right)^j\left(\frac{\lambda_{k_1}^m}{\lambda_0^0}\right)^N-\left(\sum_{m,k}\left(\mathbf{t}_m^z\right)_{k,k}\left(\frac{\lambda_k^m}{\lambda_0^0}\right)^N\right)^2\\
&\approx \sum_{k_2=0}\left(\mathbf{t}_0^z\right)_{0,k_2}\left(\mathbf{t}_0^z\right)_{k_2,0}\left(\frac{\lambda_{k_2}^0}{\lambda_0^0}\right)^j-\left[\left(\mathbf{t}_0^z\right)_{0,0}\right]^2\\
&= \sum_{k_2=1}\left(\mathbf{t}_0^z\right)_{0,k_2}\left(\mathbf{t}_0^z\right)_{k_2,0}\left(\frac{\lambda_{k_2}^0}{\lambda_0^0}\right)^j
\end{aligned}\tag{8}$$

where the λ_k^m ($m=0,\pm1,\pm2,\cdots$; $k=|m|,|m|+1,|m|+2,\cdots$) are the eigenvalues of the m-th diagonal block of the naked DNA matrix $\mathbf{B}$ and $\lambda_{|m|}^m>\lambda_{|m|+1}^m>\lambda_{|m|+2}^m>\cdots$. The S6 of the Supplemental Material will show how applied force f affects the eigenvalues $\{\lambda_k^m\}$, *i.e.*, the transfer matrix spectrum of naked DNA, where λ_0^0 is the largest eigenvalue. In the large-N limit, only the terms with $m=0$ and $k_1=0$ survive. Here $\left(\mathbf{t}_m^z\right)_{k_1,k_2}$ is the matrix element of the m-th diagonal block of the matrix $\mathbf{t}^z$ (ignoring n-dependence) expressed by taking the eigenvectors of the m-th diagonal block of the naked DNA matrix $\mathbf{B}$ ($\{|\lambda_k^m\rangle\}$) as the basis, $\left(\mathbf{t}_m^z\right)_{k_1,k_2}=\langle\lambda_{k_1}^m|\hat{\mathbf{t}}_m^z|\lambda_{k_2}^m\rangle$. Similarly,

$$\begin{aligned}
G^{\hat{\mathbf{t}}^x}(r,f) &= \frac{\mathrm{Tr}(\hat{\mathbf{t}}^+\mathbf{B}^j\hat{\mathbf{t}}^-\mathbf{B}^{N-j})}{2\mathrm{Tr}(\mathbf{B}^N)}=\frac{\sum_m\left(\mathbf{t}^+_{\mathbf{m},\mathbf{m-1}}\right)\left(\mathbf{B}_{\mathbf{m-1}}\right)^j\left(\mathbf{t}^-_{\mathbf{m-1},\mathbf{m}}\right)\left(\mathbf{B}_{\mathbf{m}}\right)^{N-j}}{2\sum_{m'}\left(\mathbf{B}_{\mathbf{m}'}\right)^N}\\
&= \frac{\sum_{m,k_1,k_2}\left(\mathbf{t}^+_{m,m-1}\right)_{k_1,k_2}\left(\lambda_{k_2}^{m-1}\right)^j\left(\mathbf{t}^-_{m-1,m}\right)_{k_2,k_1}\left(\lambda_{k_1}^m\right)^{N-j}}{2\sum_{m',k'}\left(\lambda_{k'}^{m'}\right)^N}\\
&\approx \frac{1}{2}\sum_{m,k_1,k_2}\left(\mathbf{t}^+_{m,m-1}\right)_{k_1,k_2}\left(\mathbf{t}^-_{m-1,m}\right)_{k_2,k_1}\left(\frac{\lambda_{k_2}^{m-1}}{\lambda_{k_1}^m}\right)^j\left(\frac{\lambda_{k_1}^m}{\lambda_0^0}\right)^N\\
&\approx \frac{1}{2}\sum_{k_2=1}\left(\mathbf{t}^+_{0,-1}\right)_{0,k_2}\left(\mathbf{t}^-_{-1,0}\right)_{k_2,0}\left(\frac{\lambda_{k_2}^1}{\lambda_0^0}\right)^j
\end{aligned}\tag{9}$$

where $\left(\mathbf{t}^+_{m,m-1}\right)_{k_1,k_2}=\langle\lambda_{k_1}^m|\hat{\mathbf{t}}^+_{m,m-1}|\lambda_{k_2}^{m-1}\rangle$ and $\left(\mathbf{t}^-_{m,m-1}\right)_{k_1,k_2}=\langle\lambda_{k_1}^{m-1}|\hat{\mathbf{t}}^-_{m-1,m}|\lambda_{k_2}^m\rangle$. Here we also taked use of a symmetry $\lambda_k^m=\lambda_k^{-m}$ (see Eq. 2 in S1 of the Supplemental Material).

We note that according to the notation of this paper, Eq.(20) and Eq.(21) of Ref. [1] should be written as,

$$\Delta F(f,n)=-k_BT\frac{\lambda_0^0}{\lambda_1^0}\frac{\mathbf{A}_{01}^0\mathbf{A}_{10}^0}{\left(\mathbf{A}_{00}^0\right)^2}e^{-n\ln(\lambda_0^0/\lambda_1^0)}\tag{10}$$

and

$$\begin{aligned}
C(f) &= k_BT\frac{\lambda_0^0}{\lambda_1^0}\frac{\mathbf{A}_{01}^0\mathbf{A}_{10}^0}{\left(\mathbf{A}_{00}^0\right)^2}\\
\xi(f) &= \frac{b}{\ln(\lambda_0^0/\lambda_1^0)}
\end{aligned}\tag{11}$$

where $\mathbf{A}_{k_1k_2}^m$ is the matrix element of the m-th diagonal block of the matrix $\mathbf{A}$ expressed by taking the eigenvectors of the m-th diagonal block of the naked matrix $\mathbf{B}$ ($\{|\lambda_k^m\rangle\}$) as the basis, $\mathbf{A}_{k_1,k_2}^m=\langle\lambda_{k_1}^m|\mathbf{A}_m|\lambda_{k_2}^m\rangle$; λ_0^0 and λ_1^0 are the leading and next-to-leading eigenvalues of the 0-th ($m=0$) diagonal block of the naked matrix $\mathbf{B}$, respectively.

S6. TRANSFER MATRIX SPECTRUM FOR NAKED DNA

We calculated the force-dependence of the eigenvalues $\lambda_k^m(f)$ of the naked DNA matrix $\mathbf{B}$, $\langle\hat{t}|\mathbf{B}|\hat{t}'\rangle = e^{-\frac{a}{2}|\hat{t}'-\hat{t}|^2+\beta bf\hat{t}\cdot\hat{z}}$ (see Fig. 2). The free energy spectrum of this naked DNA is:

$$\mathcal{F}_{mk} = -k_BT(L/b)\ln(\lambda_k^m)\ , \tag{12}$$

where L is the contour length of the DNA polymer and b is the segment length. This is also the spectrum of the 1-d classical Heisenberg model. We note that in the low-force limit, the spectrum approaches that of a free rotor (up to a constant, $\propto \ell(\ell+1)$, $\ell = 0, 1, 2, \cdots$) while in the large-force limit, the spectrum approaches that of a harmonic oscillator (again, up to a constant, $\propto n$, $n = 0, 1, 2, \cdots$) [3].

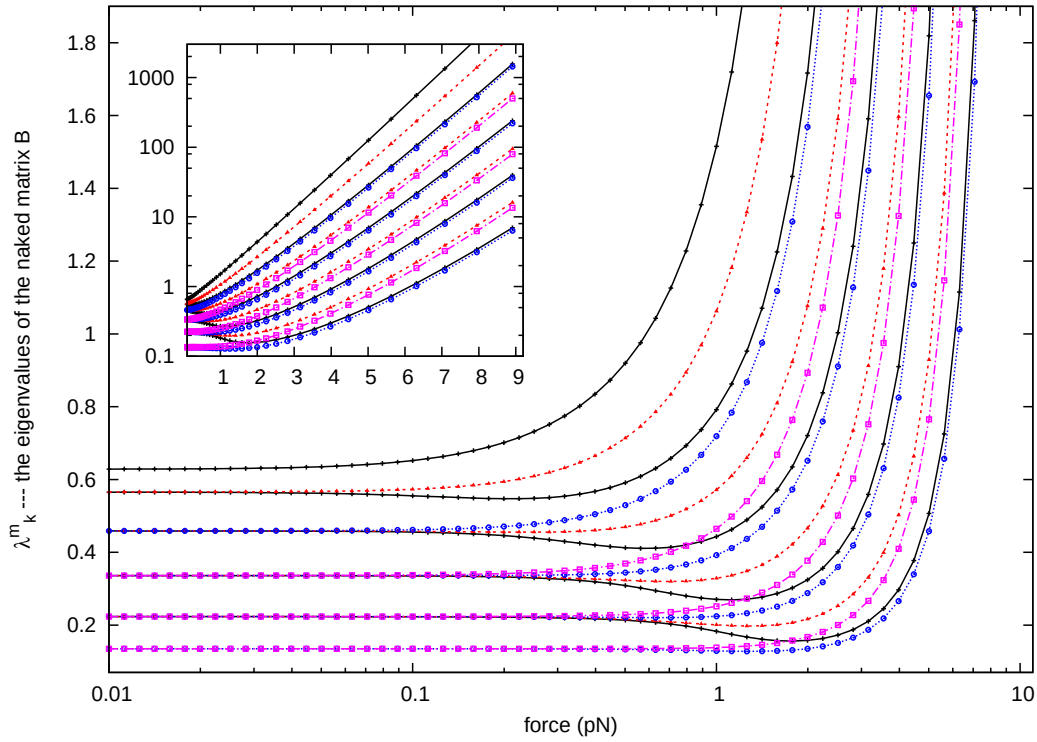

FIG. 2. (color online) Force-dependence of the eigenvalues $\lambda_k^m(f)$ of the "naked" matrix $\mathbf{B}$. Here we only show the largest several eigenvalues for the $m = 0$ block, λ_k^0 (black line and pluses), for the $m = 1$ block, λ_k^1 (red dashed line and filled triangles), for the $m = 2$ block, λ_k^2 (blue dashed line and circles), and for the $m = 3$ block, λ_k^3 (pink line and diamonds). Note that $\lambda_1^1 > \lambda_1^0$. Inset shows that in the high-force range, the spectrum approaches that of a harmonic oscillator, which is closely related to the relation $\xi_\parallel = \xi_\perp/2$.

S7. PROTEIN OCCUPATION CORRELATION FUNCTIONS AND CORRELATION LENGTHS FOR THE CASE $\gamma = 0$ AND $a' = 50$

To focus on the force-generated cooperativity, we turn off the intrinsic protein-protein binding cooperativity (*i.e.* set $\eta = 0$) and calculate the protein occupation correlation functions and correlation lengths along a DNA molecule in contact with dilute DNA-bending protein solution (see Fig. 3 for the case $\mu = -5$ and Fig. 4 for the case $\mu = -3$). Here, the parameters $\gamma = 0$ and $a' = 50$ correspond to a rather stiff protein-DNA complex with a preferred bending angle of 90°.

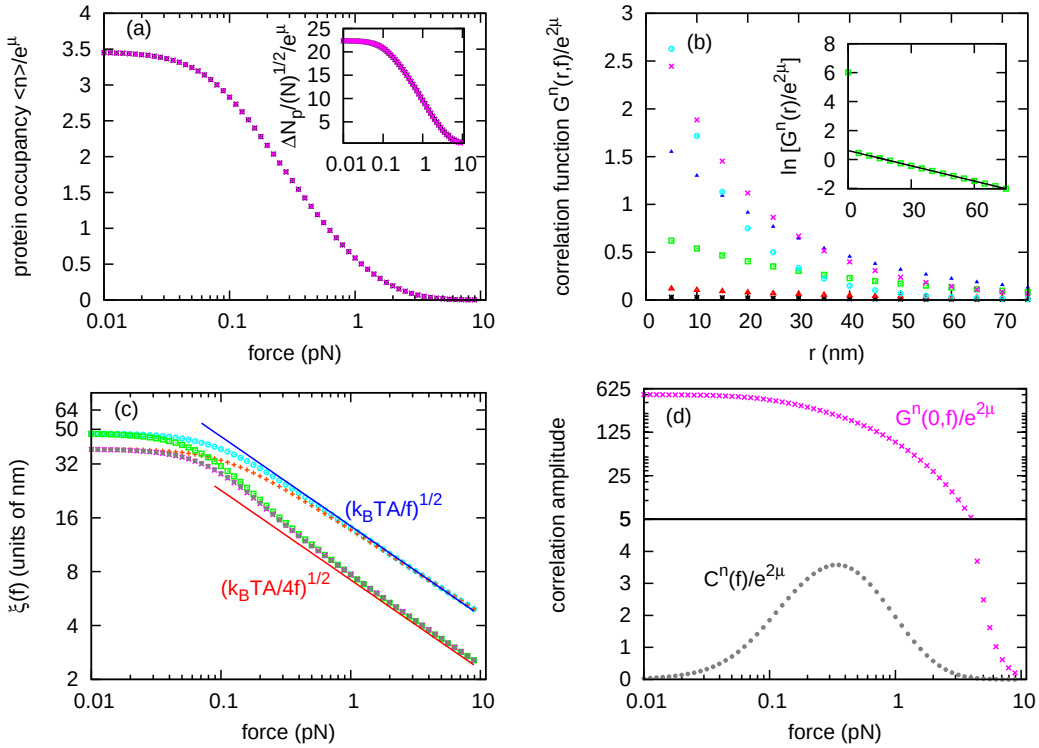

FIG. 3. (color online) Protein occupation correlation functions and correlation lengths along DNA polymer in dilute protein solution ($\mu = -5$) without intrinsic protein-protein binding cooperativity ($\eta = 0$). (a) The protein occupation $\langle n \rangle / e^{\mu}$ as a function of force (pink crosses), agreeing with the normalized protein binding number from Ref.[1] (black asterisks). Inset shows the protein occupation fluctuation curve calculated from protein occupation correlation functions (pink crosses), which matches the corresponding result from Ref. [1] (black asterisks). (b) The protein occupation correlation function $G^{\hat{\mathbf{n}}}(r, f)/e^{2\mu}$ as a function of contour length $r = jb$ for a series of fixed forces (different curves). At the left, bottom curve has force $f = 0.01$ pN (black asterisks), then in increasing order the curves show forces $f = 0.02$ (red triangles), 0.05 (green squares), 0.1 (blue filled triangles), 0.2 (pink crosses), and 0.4 (aqua circles) pN. Inset depicts $\ln[G^{\mathbf{n}}(r)/e^{2\mu}]$ versus the distance at $f = 0.1$ pN (green squares) and a linear fit (black line), showing that the non-zero distance protein occupation correlation function $G^{\mathbf{n}}(r, f)$ decays exponentially and that the zero distance protein occupation correlation function $G^{\mathbf{n}}(0, f)$ is bigger than the non-zero distance protein occupation correlation function decaying amplitude. (c) The protein occupation correlation length as a function of applied force extracted from correlation function $G^{\hat{\mathbf{n}}}(r, f)$ (bottom grey filled circles). As reference, here we also show the longitudinal tangent vector correlation length $\xi_{\parallel}(f)$ extracted from correlation function $G^{\hat{\mathbf{t}}^z}(r, f)$ (bottom pink crosses, overlapping with bottom grey filled circles), the transverse tangent vector correlation length $\xi_{\perp}(f)$ extracted from correlation function $G^{\hat{\mathbf{t}}^x}(r, f)$ (top orange pluses), the correlation lengths for naked DNA (bottom green squares for longitudinal and top aqua circles for transverse), and the high-force limit for naked DNA from continuous WLC model [3] (bottom red line for longitudinal and top blue line for transverse). We find the protein occupation correlation length to be (numerically) equal to the longitudinal tangent vector correlation length $\xi_{\parallel}(f)$. (d) The protein occupation correlation amplitude $C^{n}(f)/e^{2\mu}$ (bottom grey filled circles) and the zero-distance protein occupation correlation function $G^{\mathbf{n}}(0, f)/e^{2\mu}$ (top pink crosses).

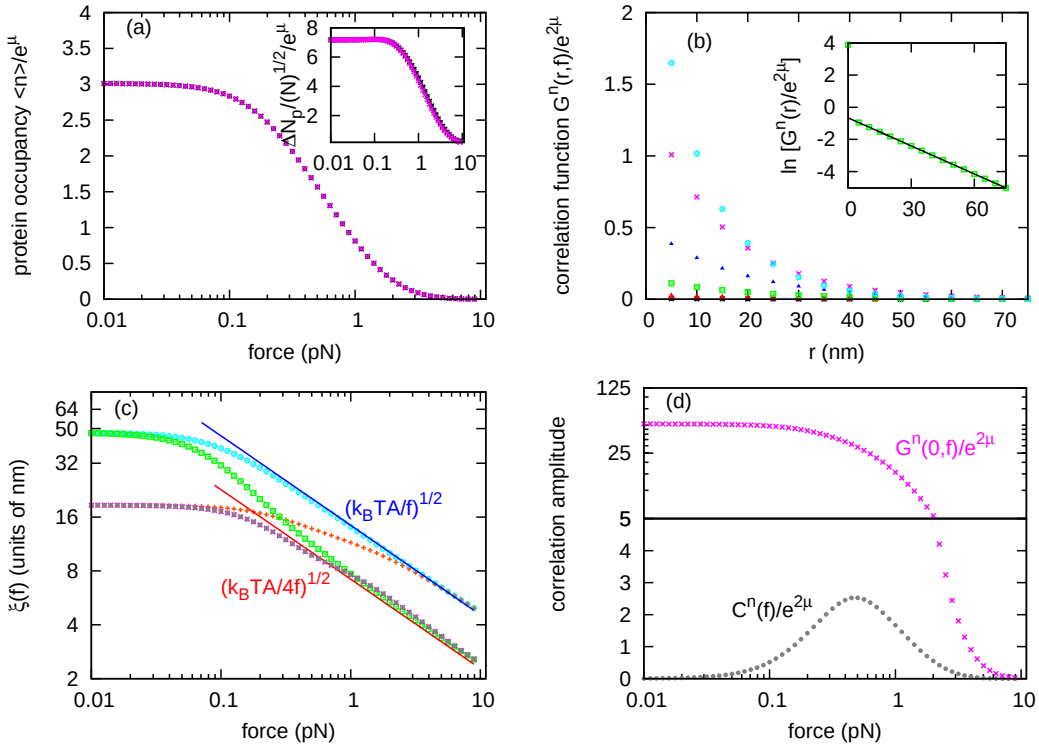

FIG. 4. (color online) Protein occupation correlation functions and correlation lengths along DNA polymer in dilute protein solution as in Fig. 3 except for the use of a different chemical potential value ($\mu = -3$).

From the results we find: (i) the protein occupancy correlation function $G^{\mathbf{n}}(r, f)$ decays exponentially with distance; the zero-distance protein occupation correlation function $G^{\mathbf{n}}(0, f)$ is bigger than the non-zero distance protein occupation correlation function decaying amplitude; (ii) the protein occupancy correlation length, (*i.e.* the decay length) is (numerically) equal to the longitudinal tangent vector correlation length $\xi_{\parallel}(f, \mu)$; (iii) the protein occupancy correlation is small for low force and high force and has a peak value at around 0.5 pN, and the peak value position shifts with increasing of the protein concentration (see panels (d) of Fig. 3 and 4).

We calculated properties of the model for the case of nonzero intrinsic cooperative interaction between DNA-bending proteins. Results are shown in Fig. 5 for $\mu = -5$, $\eta = 1$ and in Fig. 6 for $\mu = -3$, $\eta = 1$.

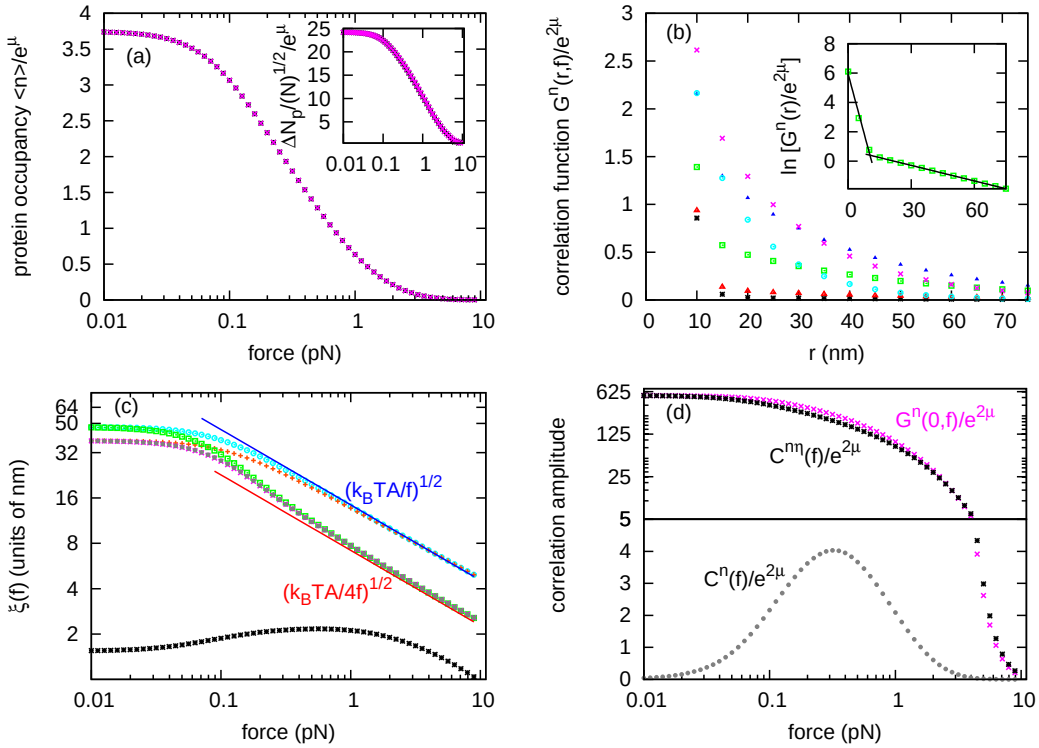

FIG. 5. (color online) Protein occupation correlation functions and correlation lengths along DNA polymer in dilute protein solution ($\mu = -5$) with intrinsic protein-protein binding cooperativity ($\eta = 1$). (a) The protein occupation $\langle n \rangle / e^{\mu}$ as a function of force (pink crosses), agreeing with the normalized protein binding number from Ref.[1] (black asterisks). Inset shows the protein occupation fluctuation curve calculated from protein occupation correlation functions (pink crosses), which matches the corresponding result from Ref. [1] (black asterisks). (b) The protein occupation correlation function $G^{\hat{\mathbf{n}}}(r, f)/e^{2\mu}$ as a function of contour length $r = jb$ for a series of fixed forces (different curves). At the left, bottom curve has force $f = 0.01$ pN (black asterisks), then in increasing order the curves show forces $f = 0.02$ (red triangles), 0.05 (green squares), 0.1 (blue filled triangles), 0.4 (aqua circles), and 0.2 (pink crosses) pN. Inset depicts $\ln[G^{\mathbf{n}}(r)/e^{2\mu}]$ versus the distance at $f = 0.1$ pN (green squares) and a linear fit (black line), showing that the protein occupation correlation function $G^{\mathbf{n}}(r, f)$ includes two kinds of different correlations, and that both of them decay exponentially. (c) The two different protein occupation correlation lengths as a function of applied force: the short-range correlation length (bottom black asterisks) which comes from the intrinsic protein-protein binding cooperativity, and the long-range correlation length (middle grey filled circles) which comes from the longitudinal bending fluctuations. As reference, here we also show the longitudinal tangent vector correlation length $\xi_{\parallel}(f)$ extracted from correlation function $G^{\hat{\mathbf{t}}^z}(r, f)$ (middle pink crosses, overlapping with middle grey filled circles), the transverse tangent vector correlation length $\xi_{\perp}(f)$ extracted from correlation function $G^{\hat{\mathbf{t}}^x}(r, f)$ (top orange pluses), the correlation lengths for naked DNA (middle green squares for longitudinal and top aqua circles for transverse), and the high-force limit for naked DNA from continuous WLC model [3] (middle red line for longitudinal and top blue line for transverse). We find that the long-range protein occupation correlation length is (numerically) equal the longitudinal tangent vector correlation length $\xi_{\parallel}(f)$. (d) The long-range protein occupation correlation amplitude $C^{n}(f)/e^{2\mu}$ (bottom grey filled circles), the short-range protein occupation correlation amplitude $C^{n\eta}(f)/e^{2\mu}$ (top black asterisks), and the zero-distance protein occupation correlation function $G^{\mathbf{n}}(0, f)/e^{2\mu}$ (top pink crosses).

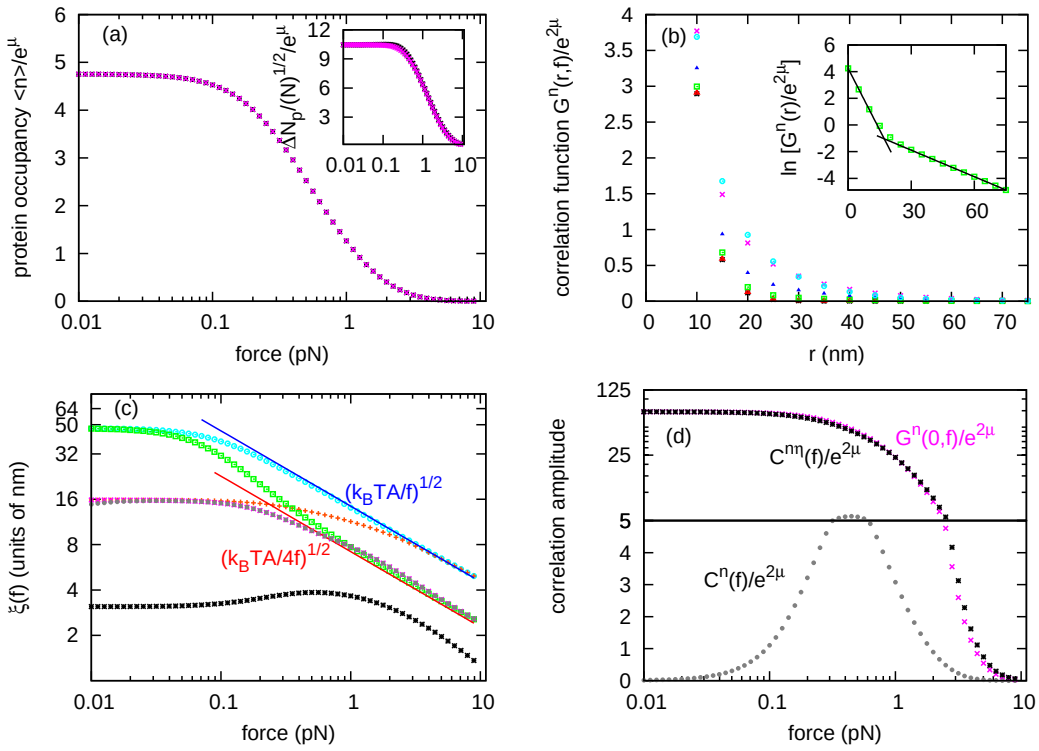

FIG. 6. (color online) Protein occupation correlation functions and correlation lengths along DNA polymer in dilute protein solution as in Fig. 5 except for the use of a different chemical potential value ($\mu = -3$).

Comparing with previous results, we find when there is intrinsic protein-protein binding cooperativity, the interaction gives rise to a new protein occupation correlation which has a bigger amplitude but a shorter correlation length. The short-range protein occupation correlation length increases as the protein concentration (μ) is increased or as cooperativity strength (η) is increased.

S8. PROTEIN OCCUPATION CORRELATION LENGTHS FOR OTHER VALUES OF γ AND a'

We calculated the correlation lengths for varied values of preferred angle θ and bending rigidity a'. Fig. 7 shows results for $\theta = 30°$ and $a' = 50$, Fig. 8 shows results for $\theta = 150°$ and $a' = 50$, and Fig. 9 shows results for $\theta = 90°$ and $a' = 2$.

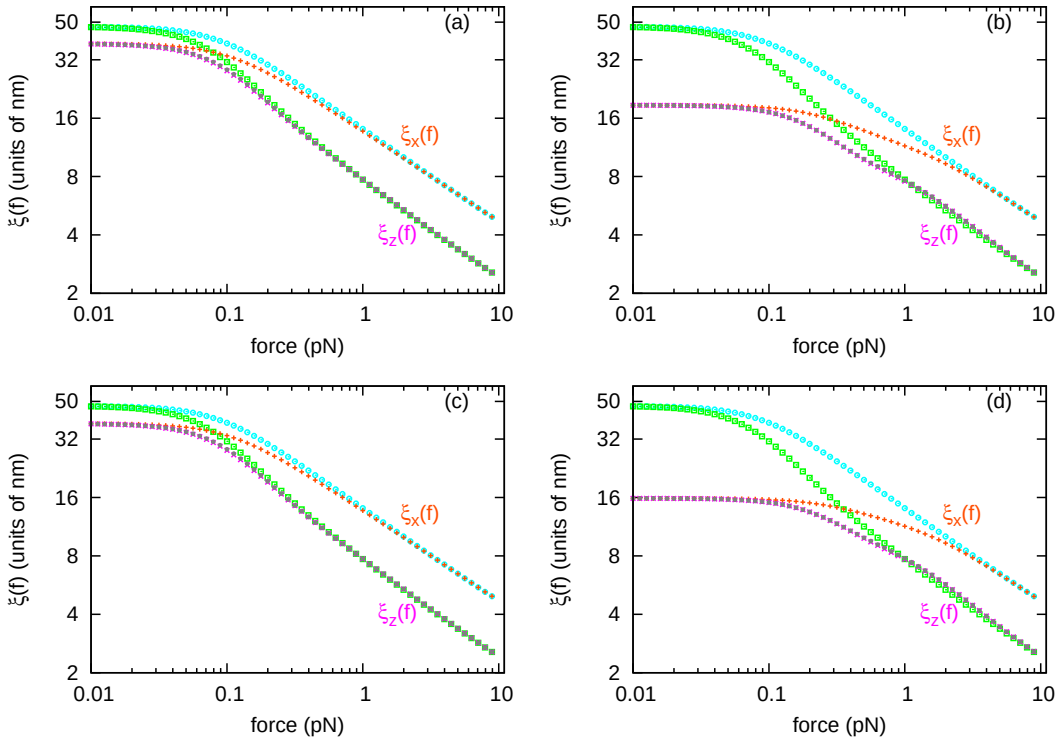

FIG. 7. The longitudinal, transverse, and protein occupation correlation lengths along DNA polymer for $\theta = 30^o$ ($\gamma = \cos 30^o$) and $a' = 50$. (a)-(d) show results for $\mu = -5$ and $\eta = 0$, $\mu = -3$ and $\eta = 0$, $\mu = -5$ and $\eta = 1$, and $\mu = -3$ and $\eta = 1$, respectively. In each panel, we show the longitudinal tangent vector correlation length $\xi_{\|}$ (bottom pink crosses), the transverse tangent vector correlation length $\xi_{\perp}$ (middle orange pluses), and the long-range protein occupation correlation length (bottom grey filled circles, overlapping with bottom pink crosses). As reference, we also show the longitudinal and transverse tangent vector correlation lengths for naked DNA (middle green squares for longitudinal and top aqua circles for transverse).

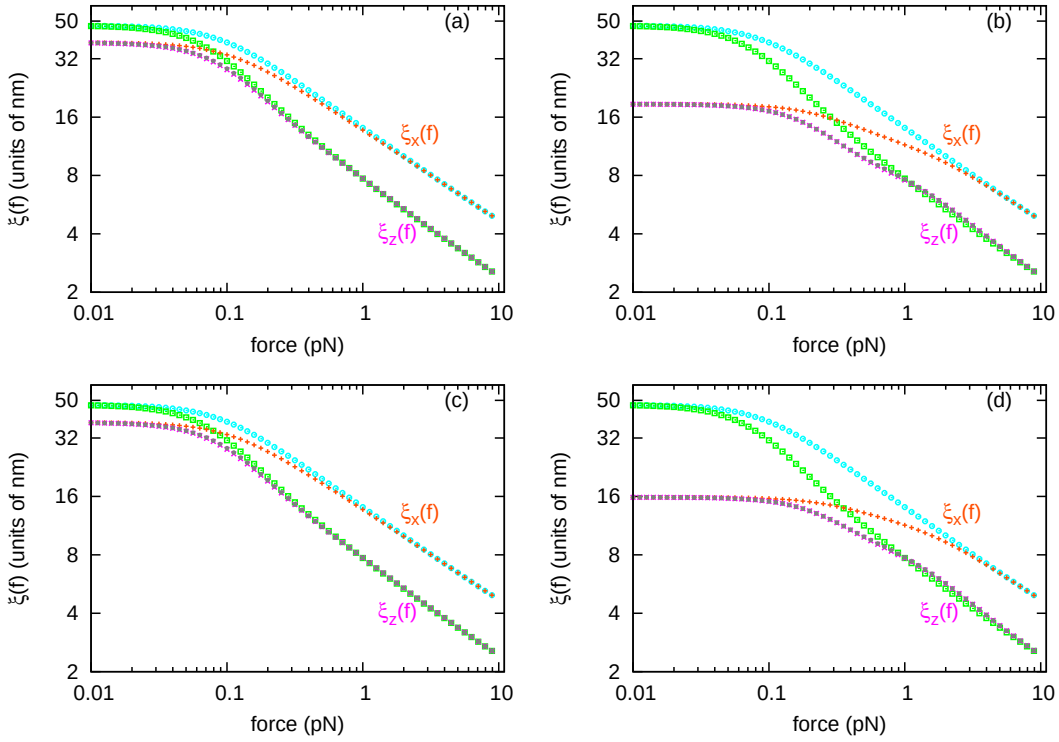

FIG. 8. The longitudinal, transverse, and protein occupation correlation lengths along DNA polymer for $\theta = 150^o$ ($\gamma = \cos 150^o$) and $a' = 50$. (a)-(d) show results for $\mu = -5$ and $\eta = 0$, $\mu = -3$ and $\eta = 0$, $\mu = -5$ and $\eta = 1$, and $\mu = -3$ and $\eta = 1$, respectively; all other parameters and color-coding are the same as in Fig. 7.

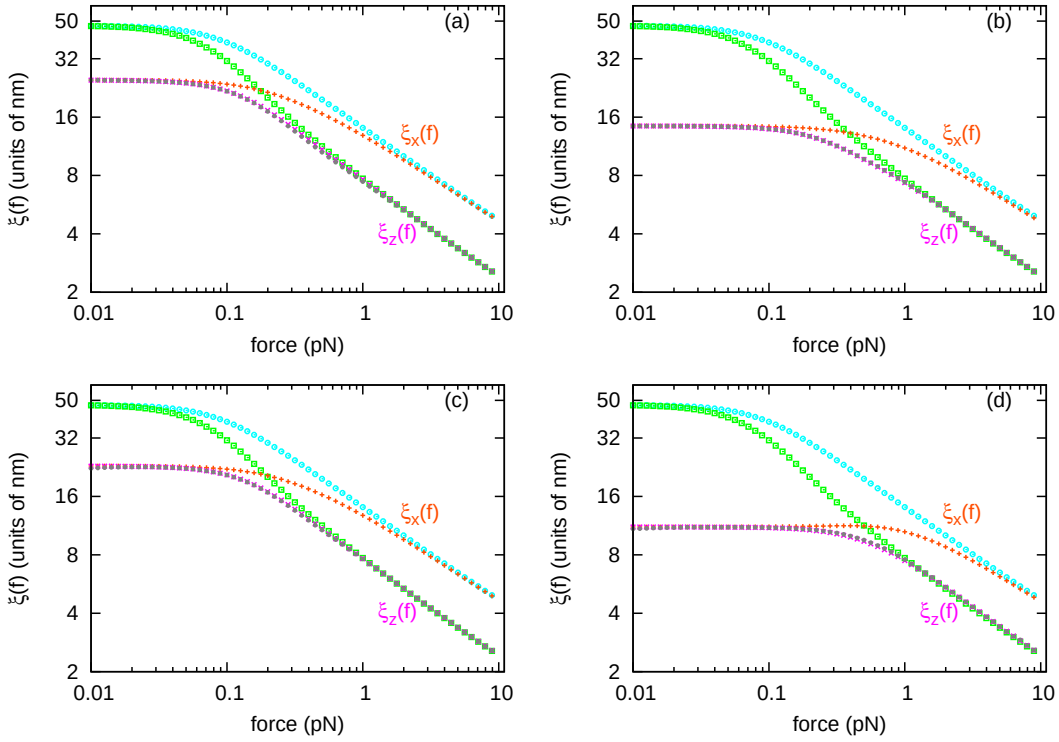

FIG. 9. The longitudinal, transverse, and protein occupation correlation lengths along DNA polymer for $\theta = 90°$ ($\gamma = \cos 90° = 0$) and $a' = 2$. (a)-(d) show results for $\mu = -5$ and $\eta = 0$, $\mu = -4$ and $\eta = 0$, $\mu = -5$ and $\eta = 1$, and $\mu = -4$ and $\eta = 1$, respectively; all other parameters and color-coding are the same as in Fig. 7.

From these results we find that different γ and a' values affect numerical details of the calculation results, but do not affect the main conclusion of the manuscript, *i.e.*, that the range of interaction between DNA-bending proteins is controlled by the second-longest correlation length for bending fluctuations. This is to be expected since different choices of γ and a' do not affect the axial rotation symmetry of the interaction between the proteins and DNA.

[1] H. Zhang and J.F. Marko, Phys. Rev. E **82**, 051906 (2010).
[2] J. Yan and J.F. Marko, Phys. Rev. E **68**, 011905 (2003).
[3] J.F. Marko and E.D. Siggia, Macromolecules **28**, 8759 (1995).
[4] J. Yan, R. Kawamura, and J.F. Marko, Phys. Rev. E **71**, 061905 (2005).

Supplemental Material for:

The range of interaction between DNA-bending proteins is controlled by the second-longest correlation length for bending fluctuations

Houyin Zhang and John F. Marko

S1. TRANSFER MATRIX IN SPHERICAL HARMONIC BASIS

By using symmetries, the definition of the Wigner 3J symbols, and partial-wave decomposition of $e^{\beta b f \hat{t}\cdot\hat{z}}$, we re-expressed the two matrices $\mathbf{A}$ and $\mathbf{B}$ in a space with the spherical harmonics $\{|l,m\rangle\}$ as the basis,

$$\begin{aligned}\langle lm|\mathbf{A}|l'm'\rangle = (-1)^m\delta_{mm'}\sqrt{(2l+1)(2l'+1)}\sqrt{\frac{8\pi}{a'}} \\ \times\int_0^{+\infty} dq[j_{l'}(q)e^{-\frac{q^2}{2a'}}(e^{i(\frac{\pi}{2}l'-q\gamma)}+e^{-i(\frac{\pi}{2}l'-q\gamma)})] \\ \times\sum_{l_1} i_{l_1}(\beta bf)(2l_1+1)\begin{pmatrix} l & l_1 & l' \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l & l_1 & l' \\ -m & 0 & m\end{pmatrix}\ ,\end{aligned} \tag{1}$$

and

$$\begin{aligned}\langle lm|\mathbf{B}|l'm'\rangle = (-1)^m\delta_{mm'}\sqrt{(2l+1)(2l'+1)}[4\pi e^{-a}i_{l'}(a)] \\ \times\sum_{l_1} i_{l_1}(\beta bf)(2l_1+1)\begin{pmatrix} l & l_1 & l' \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l & l_1 & l' \\ -m & 0 & m\end{pmatrix}\ ,\end{aligned} \tag{2}$$

where the $j_l(x)$ are spherical Bessel functions of the first kind, and the $i_l(x)$ are modified spherical Bessel functions of the first kind. Both of these two matrices are block diagonal.

S2. TANGENT VECTOR OPERATORS IN SPHERICAL HARMONIC BASIS

Choosing spherical harmonics $\{|l,m\rangle\}$ as the basis, by making use of Wigner 3J symbols, we obtain the matrix elements of the three tangent vector component operators,

$$\begin{aligned}\langle l,m,n|\hat{\mathbf{t}}^+|l',m',n'\rangle &= \delta_{nn'}(-1)^{m+1}\sqrt{2(2l+1)(2l'+1)}\begin{pmatrix} l & 1 & l' \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l & 1 & l' \\ -m & 1 & m'\end{pmatrix}\ , \\ \langle l,m,n|\hat{\mathbf{t}}^-|l',m',n'\rangle &= \delta_{nn'}(-1)^{m}\sqrt{2(2l+1)(2l'+1)}\begin{pmatrix} l & 1 & l' \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l & 1 & l' \\ -m & -1 & m'\end{pmatrix}\ , \\ \langle l,m,n|\hat{\mathbf{t}}^z|l',m',n'\rangle &= \delta_{nn'}(-1)^{m+1}\sqrt{(2l+1)(2l'+1)}\begin{pmatrix} l & 1 & l' \\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l & 1 & l' \\ -m & 0 & m'\end{pmatrix}\ .\end{aligned} \tag{3}$$

S3. TANGENT VECTOR CORRELATION FUNCTIONS FOR NAKED DNA

Figure 1 shows the transverse and longitudinal tangent vector correlation functions and corresponding correlation lengths for naked DNA (DNA in the absence of protein).

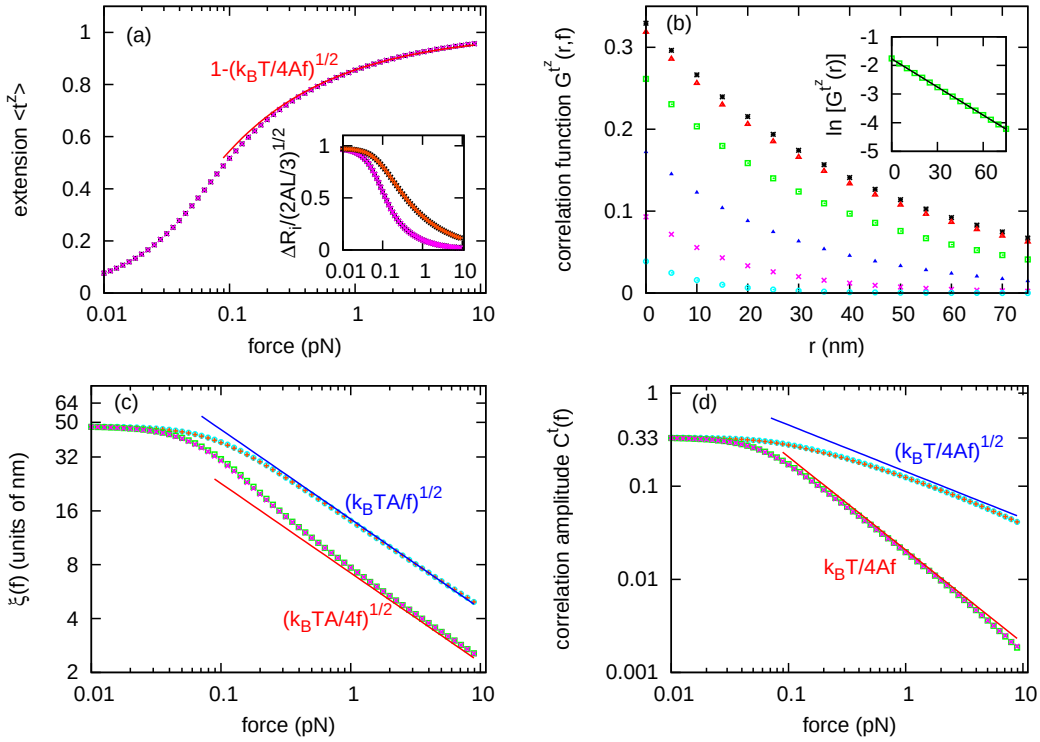

FIG. 1. (color online) Tangent vector correlation functions and correlation lengths for naked DNA. (a) The average z-component of the tangent vector, $\langle \hat{\mathbf{t}}^z \rangle$, as a function of force (pink crosses), agreeing with the normalized force-extension curve for naked DNA from Ref. [1, 2] (black asterisks) and the high-force limit of the extension from continuous WLC model [3] (red curve). Inset shows the extension fluctuation curves: z-direction extension fluctuation calculated from correlation function $G^{\hat{\mathbf{t}}^z}(r,f)$ (bottom pink crosses) and the corresponding result from Ref. [1] (bottom black asterisks, overlapping with bottom pink crosses); x (or y)-direction extension fluctuation calculated from $G^{\hat{\mathbf{t}}^x}(r,f)$ (top orange pluses) and the corresponding result calculated by Eq. A6 of Ref. [4] (top black asterisks, overlapping with top orange pluses). (b) The correlation function for the z-component of tangent vector $G^{\hat{\mathbf{t}}^z}(r,f)$ as a function of contour length $r = jb$ for a series of fixed forces (different curves). The top curve has force $f = 0.01$ pN (black asterisks), then in decreasing order the curves correspond to $f = 0.02$ (red triangles), 0.05 (green squares), 0.1 (blue filled triangles), 0.2 (pink crosses), and 0.5 (aqua circles) pN. Inset depicts $\ln(G^{\hat{\mathbf{t}}^z}(r))$ versus the distance at $f = 0.1$ pN (green squares) and a linear fit (black line), showing that the correlation function decays exponentially. (c) The longitudinal correlation length $\xi_\parallel$ versus applied force: extracted from correlation function $G^{\hat{\mathbf{t}}^z}(r,f)$ (bottom pink crosses), asymptotic result from eigenvalues (bottom green squares, overlapping with bottom pink crosses), and high-force limit from continuous WLC model [3] (bottom red line); the transverse correlation length $\xi_\perp$ versus applied force: extracted from correlation function $G^{\hat{\mathbf{t}}^x}(r,f)$ (top orange pluses), asymptotic result from eigenvalues (top aqua circles, overlapping with top orange pluses), and high-force limit from continuous WLC model [3] (top blue line). In high-force range, $\xi_\perp = 2\xi_\parallel = \sqrt{k_B TA/f}$. (d) The longitudinal correlation function amplitude $C_\parallel(f)$: extracted from correlation function $G^{\hat{\mathbf{t}}^z}(r,f)$ (bottom pink crosses), $G^{\hat{\mathbf{t}}^z}(0,f)$ (bottom green squares, overlapping with bottom pink crosses), and high-force limit from continuous WLC model [3] (bottom red line); the transverse correlation function amplitude $C_\perp(f)$: extracted from correlation function $G^{\hat{\mathbf{t}}^x}(r,f)$ (top orange pluses), $G^{\hat{\mathbf{t}}^x}(0,f)$ (top aqua circles, overlapping with top orange pluses), and high-force limit from continuous WLC model [3] (top blue line).

S4. HIGH-FORCE LIMIT OF CONTINUOUS WLC MODEL

For the continuous WLC model, the effective energy of a stretched naked DNA molecule is expressed as [3, 4],

$$\beta E = \int_0^L ds \left[\frac{A}{2} (\partial_s \hat{t})^2 - \beta f \hat{t} \cdot \hat{z} \right] , \tag{4}$$

where L is the contour length of the DNA polymer and $\hat{t}(s)$ is the unit tangent vector describing the curve conformation. In the large-force limit ($f \gg k_B T/A \approx 0.1$ pN), t_x and t_y are small quantities, and $t_z = 1 - (t_x^2 + t_y^2)/2 + \mathcal{O}(t_x^4 + t_y^4)$. Then we obtain the Gaussian approximation to this model, which is valid in the large force limit $f >> k_B T/A$:

$$\beta E \approx \frac{1}{2} \int_0^L ds \left\{ A[(\partial_s t_x)^2 + (\partial_s t_y)^2] + \beta f [t_x^2 + t_y^2] \right\} - \beta f L = \frac{1}{2} \int \frac{dq}{2\pi} (Aq^2 + \beta f)[t_x^2(q) + t_y^2(q)] - \beta f L , \tag{5}$$

where $t_x(q) = \int_0^L t_x e^{-iqs}\, ds$ and $t_y(q) = \int_0^L t_y e^{-iqs}\, ds$.

We calculate the x-direction correlation function,

$$G^{\hat{\mathbf{t}}^x}(s, f) = \langle t_x(0) t_x(s) \rangle = \int_{-\infty}^{\infty} \frac{dq_1}{2\pi} \frac{dq_2}{2\pi} \langle t_x(q_1) t_x(q_2) \rangle e^{iq_2 s} = \int_{-\infty}^{\infty} \frac{dq}{2\pi} \frac{e^{iqs}}{Aq^2 + \beta f} = \sqrt{k_B T/4Af}\; e^{-s/\sqrt{k_B T A/f}}. \tag{6}$$

Similarly, we calculate the z-direction correlation function,

$$\begin{aligned}
G^{\hat{\mathbf{t}}^z}(s, f) &= \langle t_z(0) t_z(s) \rangle - \langle t_z(0) \rangle \langle t_z(s) \rangle \\
&\approx \langle \left[1 - t_x^2(0)/2 - t_y^2(0)/2\right] \left[1 - t_x^2(s)/2 - t_y^2(s)/2\right] \rangle - \langle 1 - t_x^2(0)/2 - t_y^2(0)/2 \rangle \langle 1 - t_x^2(s)/2 - t_y^2(s)/2 \rangle \\
&= \frac{1}{4} \Big[\langle t_x^2(0) t_x^2(s) \rangle + \langle t_x^2(0) t_y^2(s) \rangle + \langle t_y^2(0) t_x^2(s) \rangle + \langle t_y^2(0) t_y^2(s) \rangle \\
&\quad - \langle t_x^2(0) \rangle \langle t_x^2(s) \rangle - \langle t_x^2(0) \rangle \langle t_y^2(s) \rangle - \langle t_y^2(0) \rangle \langle t_x^2(s) \rangle - \langle t_y^2(0) \rangle \langle t_y^2(s) \rangle \Big] \\
&= \frac{1}{2} \left[\langle t_x^2(0) t_x^2(s) \rangle - \langle t_x^2(0) \rangle \langle t_x^2(s) \rangle \right] \\
&= \frac{1}{2} \left[\int_{-\infty}^{\infty} \frac{dq_1}{2\pi} \frac{dq_2}{2\pi} \frac{dq_3}{2\pi} \frac{dq_4}{2\pi} \langle t_x(q_1) t_x(q_2) t_x(q_3) t_x(q_4) \rangle e^{iq_3 s} e^{iq_4 s} - \langle t_x^2(0) \rangle \langle t_x^2(s) \rangle \right] \\
&= \left(\int_{-\infty}^{\infty} \frac{dq}{2\pi} \frac{e^{iqs}}{Aq^2 + \beta f} \right)^2 = (k_B T/4Af)\; e^{-s/\sqrt{k_B T A/4f}} .
\end{aligned} \tag{7}$$

From the forms of the correlation functions we observe that the longitudinal correlation length is exactly half of the transverse correlation length. The amplitudes of these leading contributions to the two correlation functions are also exactly related, with the longitudinal amplitude being exactly the square of the transverse one. Thus, in the high-force limit, the leading (slowest decaying) contribution to the longitudinal correlation function (7) is just the square of the leading contribution to the transverse correlation function (6).

S5. THE LARGE-N LIMIT OF THE CORRELATION FUNCTIONS $G^{\hat{\mathbf{t}}^x}(r,f)$ AND $G^{\hat{\mathbf{t}}^z}(r,f)$ FOR NAKED DNA

For naked DNA ($\mu=-\infty$), $\mathbf{T}=\begin{pmatrix}\mathbf{A}e^{\mu+\eta} & \mathbf{A}e^{\mu}\\ \mathbf{B} & \mathbf{B}\end{pmatrix}=\begin{pmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{B} & \mathbf{B}\end{pmatrix}$. So,

$$\begin{aligned}
G^{\hat{\mathbf{t}}^z}(r,f) &= \frac{\mathrm{Tr}(\hat{\mathbf{t}}^z\mathbf{B}^j\hat{\mathbf{t}}^z\mathbf{B}^{N-j})}{\mathrm{Tr}(\mathbf{B}^N)}-\left(\frac{\mathrm{Tr}(\hat{\mathbf{t}}^z\mathbf{B}^N)}{\mathrm{Tr}(\mathbf{B}^N)}\right)^2\\
&= \frac{\sum_{m,k_1,k_2}\left(\mathbf{t}_m^z\right)_{k_1,k_2}\left(\lambda_{k_2}^m\right)^j\left(\mathbf{t}_m^z\right)_{k_2,k_1}\left(\lambda_{k_1}^m\right)^{N-j}}{\sum_{m',k'}\left(\lambda_{k'}^{m'}\right)^N}-\left(\frac{\sum_{m,k}\left(\mathbf{t}_m^z\right)_{k,k}\left(\lambda_k^m\right)^N}{\sum_{m',k'}\left(\lambda_{k'}^{m'}\right)^N}\right)^2\\
&\approx \sum_{m,k_1,k_2}\left(\mathbf{t}_m^z\right)_{k_1,k_2}\left(\mathbf{t}_m^z\right)_{k_2,k_1}\left(\frac{\lambda_{k_2}^m}{\lambda_{k_1}^m}\right)^j\left(\frac{\lambda_{k_1}^m}{\lambda_0^0}\right)^N-\left(\sum_{m,k}\left(\mathbf{t}_m^z\right)_{k,k}\left(\frac{\lambda_k^m}{\lambda_0^0}\right)^N\right)^2\\
&\approx \sum_{k_2=0}\left(\mathbf{t}_0^z\right)_{0,k_2}\left(\mathbf{t}_0^z\right)_{k_2,0}\left(\frac{\lambda_{k_2}^0}{\lambda_0^0}\right)^j-\left[\left(\mathbf{t}_0^z\right)_{0,0}\right]^2\\
&= \sum_{k_2=1}\left(\mathbf{t}_0^z\right)_{0,k_2}\left(\mathbf{t}_0^z\right)_{k_2,0}\left(\frac{\lambda_{k_2}^0}{\lambda_0^0}\right)^j
\end{aligned}\tag{8}$$

where the λ_k^m ($m=0,\pm1,\pm2,\cdots$; $k=|m|,|m|+1,|m|+2,\cdots$) are the eigenvalues of the m-th diagonal block of the naked DNA matrix $\mathbf{B}$ and $\lambda_{|m|}^m>\lambda_{|m|+1}^m>\lambda_{|m|+2}^m>\cdots$. The S6 of the Supplemental Material will show how applied force f affects the eigenvalues $\{\lambda_k^m\}$, *i.e.*, the transfer matrix spectrum of naked DNA, where λ_0^0 is the largest eigenvalue. In the large-N limit, only the terms with $m=0$ and $k_1=0$ survive. Here $\left(\mathbf{t}_m^z\right)_{k_1,k_2}$ is the matrix element of the m-th diagonal block of the matrix $\mathbf{t}^z$ (ignoring n-dependence) expressed by taking the eigenvectors of the m-th diagonal block of the naked DNA matrix $\mathbf{B}$ ($\{|\lambda_k^m\rangle\}$) as the basis, $\left(\mathbf{t}_m^z\right)_{k_1,k_2}=\langle\lambda_{k_1}^m|\hat{\mathbf{t}}_m^z|\lambda_{k_2}^m\rangle$. Similarly,

$$\begin{aligned}
G^{\hat{\mathbf{t}}^x}(r,f) &= \frac{\mathrm{Tr}(\hat{\mathbf{t}}^+\mathbf{B}^j\hat{\mathbf{t}}^-\mathbf{B}^{N-j})}{2\mathrm{Tr}(\mathbf{B}^N)}=\frac{\sum_m\left(\mathbf{t}^+_{\mathbf{m},\mathbf{m-1}}\right)\left(\mathbf{B}_{\mathbf{m-1}}\right)^j\left(\mathbf{t}^-_{\mathbf{m-1},\mathbf{m}}\right)\left(\mathbf{B}_{\mathbf{m}}\right)^{N-j}}{2\sum_{m'}\left(\mathbf{B}_{\mathbf{m}'}\right)^N}\\
&= \frac{\sum_{m,k_1,k_2}\left(\mathbf{t}^+_{m,m-1}\right)_{k_1,k_2}\left(\lambda_{k_2}^{m-1}\right)^j\left(\mathbf{t}^-_{m-1,m}\right)_{k_2,k_1}\left(\lambda_{k_1}^m\right)^{N-j}}{2\sum_{m',k'}\left(\lambda_{k'}^{m'}\right)^N}\\
&\approx \frac{1}{2}\sum_{m,k_1,k_2}\left(\mathbf{t}^+_{m,m-1}\right)_{k_1,k_2}\left(\mathbf{t}^-_{m-1,m}\right)_{k_2,k_1}\left(\frac{\lambda_{k_2}^{m-1}}{\lambda_{k_1}^m}\right)^j\left(\frac{\lambda_{k_1}^m}{\lambda_0^0}\right)^N\\
&\approx \frac{1}{2}\sum_{k_2=1}\left(\mathbf{t}^+_{0,-1}\right)_{0,k_2}\left(\mathbf{t}^-_{-1,0}\right)_{k_2,0}\left(\frac{\lambda_{k_2}^1}{\lambda_0^0}\right)^j
\end{aligned}\tag{9}$$

where $\left(\mathbf{t}^+_{m,m-1}\right)_{k_1,k_2}=\langle\lambda_{k_1}^m|\hat{\mathbf{t}}^+_{m,m-1}|\lambda_{k_2}^{m-1}\rangle$ and $\left(\mathbf{t}^-_{m,m-1}\right)_{k_1,k_2}=\langle\lambda_{k_1}^{m-1}|\hat{\mathbf{t}}^-_{m-1,m}|\lambda_{k_2}^m\rangle$. Here we also taked use of a symmetry $\lambda_k^m=\lambda_k^{-m}$ (see Eq. 2 in S1 of the Supplemental Material).

We note that according to the notation of this paper, Eq.(20) and Eq.(21) of Ref. [1] should be written as,

$$\Delta F(f,n)=-k_BT\frac{\lambda_0^0}{\lambda_1^0}\frac{\mathbf{A}_{01}^0\mathbf{A}_{10}^0}{\left(\mathbf{A}_{00}^0\right)^2}e^{-n\ln(\lambda_0^0/\lambda_1^0)}\tag{10}$$

and

$$\begin{aligned}
C(f) &= k_BT\frac{\lambda_0^0}{\lambda_1^0}\frac{\mathbf{A}_{01}^0\mathbf{A}_{10}^0}{\left(\mathbf{A}_{00}^0\right)^2}\\
\xi(f) &= \frac{b}{\ln(\lambda_0^0/\lambda_1^0)}
\end{aligned}\tag{11}$$

where $\mathbf{A}_{k_1k_2}^m$ is the matrix element of the m-th diagonal block of the matrix $\mathbf{A}$ expressed by taking the eigenvectors of the m-th diagonal block of the naked matrix $\mathbf{B}$ ($\{|\lambda_k^m\rangle\}$) as the basis, $\mathbf{A}_{k_1,k_2}^m=\langle\lambda_{k_1}^m|\mathbf{A}_m|\lambda_{k_2}^m\rangle$; λ_0^0 and λ_1^0 are the leading and next-to-leading eigenvalues of the 0-th ($m=0$) diagonal block of the naked matrix $\mathbf{B}$, respectively.

S6. TRANSFER MATRIX SPECTRUM FOR NAKED DNA

We calculated the force-dependence of the eigenvalues $\lambda_k^m(f)$ of the naked DNA matrix $\mathbf{B}$, $\langle\hat{t}|\mathbf{B}|\hat{t}'\rangle = e^{-\frac{a}{2}|\hat{t}'-\hat{t}|^2+\beta bf\hat{t}\cdot\hat{z}}$ (see Fig. 2). The free energy spectrum of this naked DNA is:

$$\mathcal{F}_{mk} = -k_BT(L/b)\ln(\lambda_k^m)\ , \tag{12}$$

where L is the contour length of the DNA polymer and b is the segment length. This is also the spectrum of the 1-d classical Heisenberg model. We note that in the low-force limit, the spectrum approaches that of a free rotor (up to a constant, $\propto \ell(\ell+1)$, $\ell = 0, 1, 2, \cdots$) while in the large-force limit, the spectrum approaches that of a harmonic oscillator (again, up to a constant, $\propto n$, $n = 0, 1, 2, \cdots$) [3].

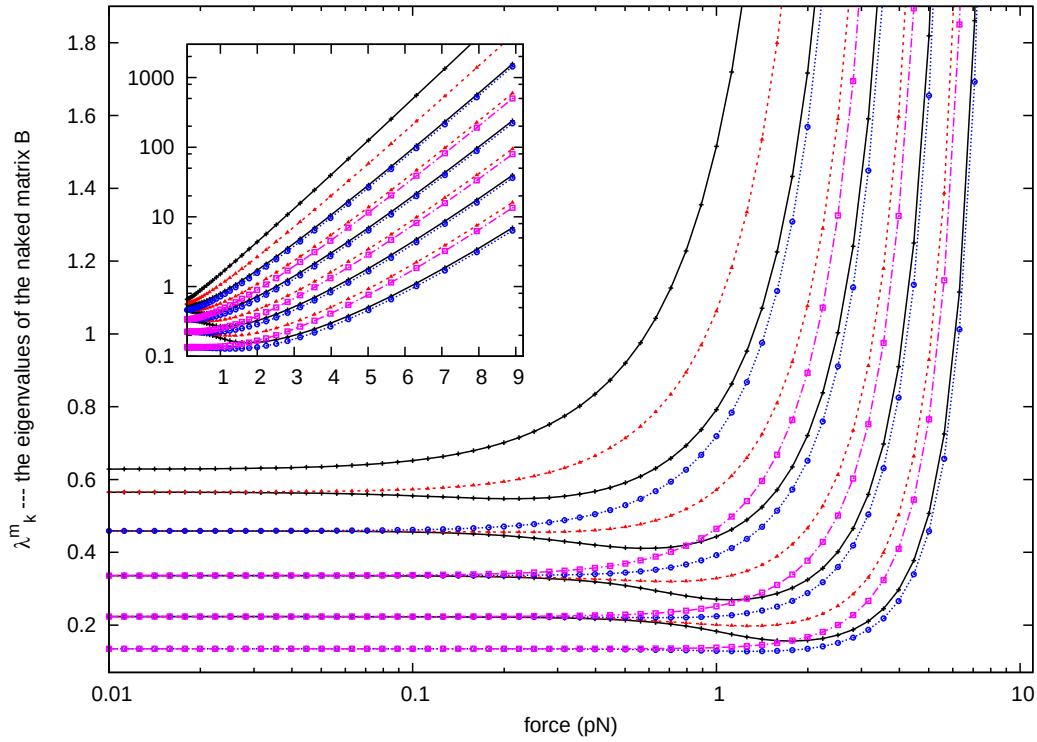

FIG. 2. (color online) Force-dependence of the eigenvalues $\lambda_k^m(f)$ of the "naked" matrix $\mathbf{B}$. Here we only show the largest several eigenvalues for the $m = 0$ block, λ_k^0 (black line and pluses), for the $m = 1$ block, λ_k^1 (red dashed line and filled triangles), for the $m = 2$ block, λ_k^2 (blue dashed line and circles), and for the $m = 3$ block, λ_k^3 (pink line and diamonds). Note that $\lambda_1^1 > \lambda_1^0$. Inset shows that in the high-force range, the spectrum approaches that of a harmonic oscillator, which is closely related to the relation $\xi_\parallel = \xi_\perp/2$.

S7. PROTEIN OCCUPATION CORRELATION FUNCTIONS AND CORRELATION LENGTHS FOR THE CASE $\gamma = 0$ AND $a' = 50$

To focus on the force-generated cooperativity, we turn off the intrinsic protein-protein binding cooperativity (*i.e.* set $\eta = 0$) and calculate the protein occupation correlation functions and correlation lengths along a DNA molecule in contact with dilute DNA-bending protein solution (see Fig. 3 for the case $\mu = -5$ and Fig. 4 for the case $\mu = -3$). Here, the parameters $\gamma = 0$ and $a' = 50$ correspond to a rather stiff protein-DNA complex with a preferred bending angle of 90°.

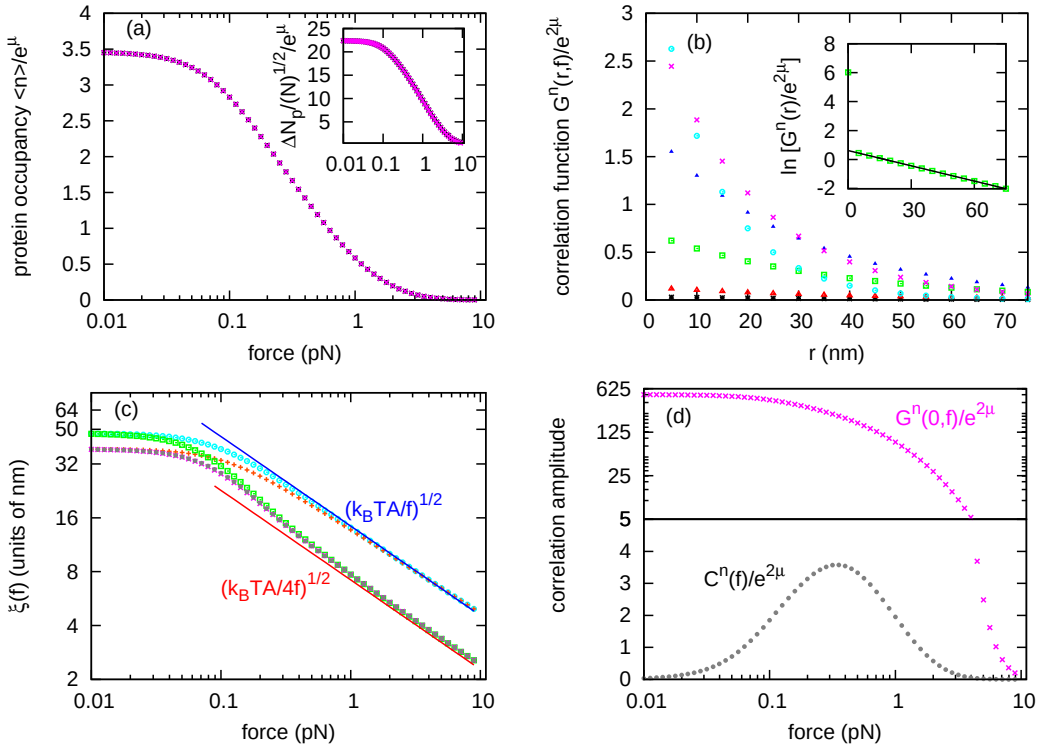

FIG. 3. (color online) Protein occupation correlation functions and correlation lengths along DNA polymer in dilute protein solution ($\mu = -5$) without intrinsic protein-protein binding cooperativity ($\eta = 0$). (a) The protein occupation $\langle n \rangle / e^{\mu}$ as a function of force (pink crosses), agreeing with the normalized protein binding number from Ref.[1] (black asterisks). Inset shows the protein occupation fluctuation curve calculated from protein occupation correlation functions (pink crosses), which matches the corresponding result from Ref. [1] (black asterisks). (b) The protein occupation correlation function $G^{\hat{\mathbf{n}}}(r, f)/e^{2\mu}$ as a function of contour length $r = jb$ for a series of fixed forces (different curves). At the left, bottom curve has force $f = 0.01$ pN (black asterisks), then in increasing order the curves show forces $f = 0.02$ (red triangles), 0.05 (green squares), 0.1 (blue filled triangles), 0.2 (pink crosses), and 0.4 (aqua circles) pN. Inset depicts $\ln[G^{\mathbf{n}}(r)/e^{2\mu}]$ versus the distance at $f = 0.1$ pN (green squares) and a linear fit (black line), showing that the non-zero distance protein occupation correlation function $G^{\mathbf{n}}(r, f)$ decays exponentially and that the zero distance protein occupation correlation function $G^{\mathbf{n}}(0, f)$ is bigger than the non-zero distance protein occupation correlation function decaying amplitude. (c) The protein occupation correlation length as a function of applied force extracted from correlation function $G^{\hat{\mathbf{n}}}(r, f)$ (bottom grey filled circles). As reference, here we also show the longitudinal tangent vector correlation length $\xi_{\parallel}(f)$ extracted from correlation function $G^{\hat{\mathbf{t}}^z}(r, f)$ (bottom pink crosses, overlapping with bottom grey filled circles), the transverse tangent vector correlation length $\xi_{\perp}(f)$ extracted from correlation function $G^{\hat{\mathbf{t}}^x}(r, f)$ (top orange pluses), the correlation lengths for naked DNA (bottom green squares for longitudinal and top aqua circles for transverse), and the high-force limit for naked DNA from continuous WLC model [3] (bottom red line for longitudinal and top blue line for transverse). We find the protein occupation correlation length to be (numerically) equal to the longitudinal tangent vector correlation length $\xi_{\parallel}(f)$. (d) The protein occupation correlation amplitude $C^{n}(f)/e^{2\mu}$ (bottom grey filled circles) and the zero-distance protein occupation correlation function $G^{\mathbf{n}}(0, f)/e^{2\mu}$ (top pink crosses).

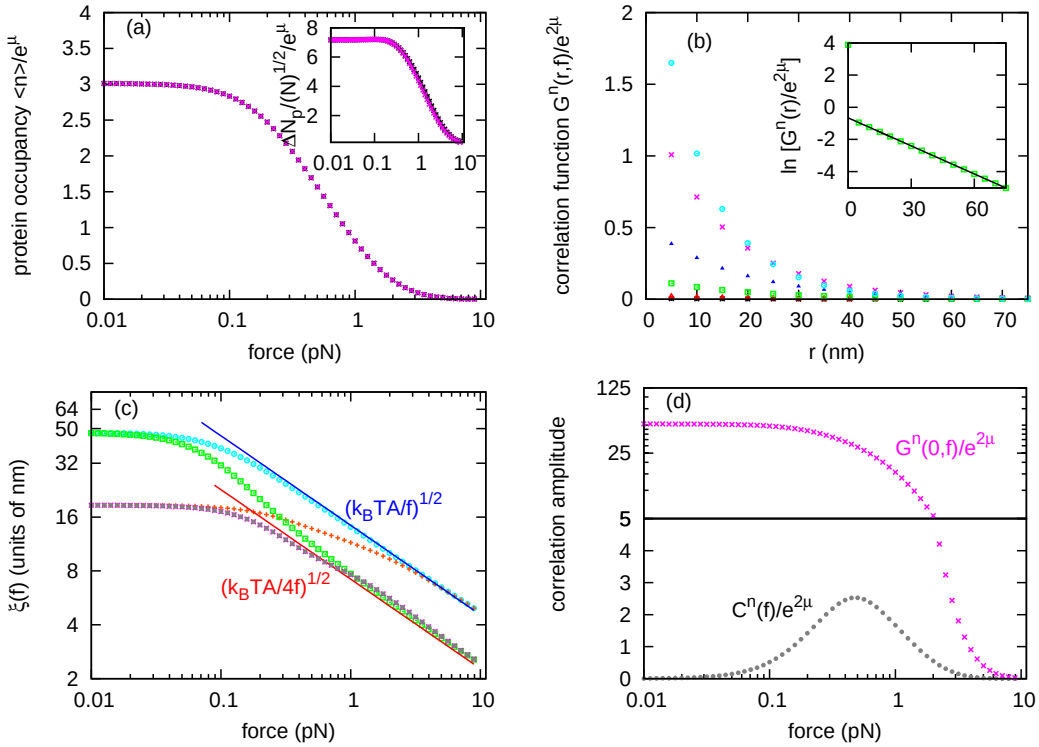

FIG. 4. (color online) Protein occupation correlation functions and correlation lengths along DNA polymer in dilute protein solution as in Fig. 3 except for the use of a different chemical potential value ($\mu = -3$).

From the results we find: (i) the protein occupancy correlation function $G^{\mathbf{n}}(r, f)$ decays exponentially with distance; the zero-distance protein occupation correlation function $G^{\mathbf{n}}(0, f)$ is bigger than the non-zero distance protein occupation correlation function decaying amplitude; (ii) the protein occupancy correlation length, (*i.e.* the decay length) is (numerically) equal to the longitudinal tangent vector correlation length $\xi_{\parallel}(f, \mu)$; (iii) the protein occupancy correlation is small for low force and high force and has a peak value at around 0.5 pN, and the peak value position shifts with increasing of the protein concentration (see panels (d) of Fig. 3 and 4).

We calculated properties of the model for the case of nonzero intrinsic cooperative interaction between DNA-bending proteins. Results are shown in Fig. 5 for $\mu = -5$, $\eta = 1$ and in Fig. 6 for $\mu = -3$, $\eta = 1$.

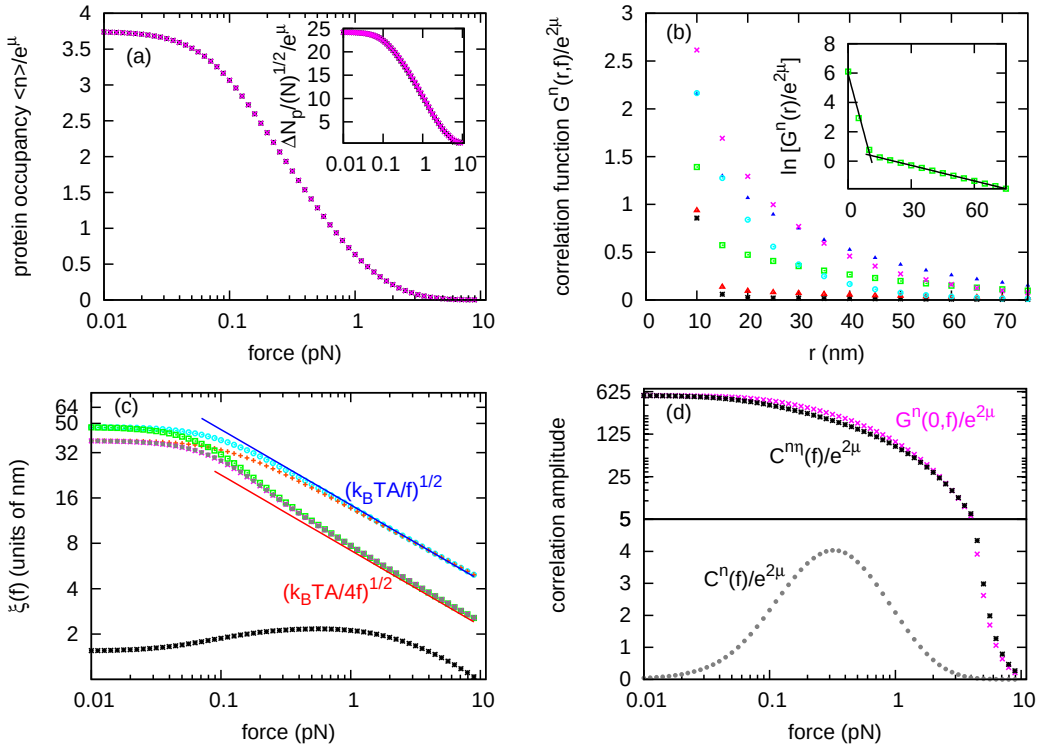

FIG. 5. (color online) Protein occupation correlation functions and correlation lengths along DNA polymer in dilute protein solution ($\mu = -5$) with intrinsic protein-protein binding cooperativity ($\eta = 1$). (a) The protein occupation $\langle n \rangle / e^{\mu}$ as a function of force (pink crosses), agreeing with the normalized protein binding number from Ref.[1] (black asterisks). Inset shows the protein occupation fluctuation curve calculated from protein occupation correlation functions (pink crosses), which matches the corresponding result from Ref. [1] (black asterisks). (b) The protein occupation correlation function $G^{\hat{\mathbf{n}}}(r, f)/e^{2\mu}$ as a function of contour length $r = jb$ for a series of fixed forces (different curves). At the left, bottom curve has force $f = 0.01$ pN (black asterisks), then in increasing order the curves show forces $f = 0.02$ (red triangles), 0.05 (green squares), 0.1 (blue filled triangles), 0.4 (aqua circles), and 0.2 (pink crosses) pN. Inset depicts $\ln[G^{\mathbf{n}}(r)/e^{2\mu}]$ versus the distance at $f = 0.1$ pN (green squares) and a linear fit (black line), showing that the protein occupation correlation function $G^{\mathbf{n}}(r, f)$ includes two kinds of different correlations, and that both of them decay exponentially. (c) The two different protein occupation correlation lengths as a function of applied force: the short-range correlation length (bottom black asterisks) which comes from the intrinsic protein-protein binding cooperativity, and the long-range correlation length (middle grey filled circles) which comes from the longitudinal bending fluctuations. As reference, here we also show the longitudinal tangent vector correlation length $\xi_{\parallel}(f)$ extracted from correlation function $G^{\hat{\mathbf{t}}^z}(r, f)$ (middle pink crosses, overlapping with middle grey filled circles), the transverse tangent vector correlation length $\xi_{\perp}(f)$ extracted from correlation function $G^{\hat{\mathbf{t}}^x}(r, f)$ (top orange pluses), the correlation lengths for naked DNA (middle green squares for longitudinal and top aqua circles for transverse), and the high-force limit for naked DNA from continuous WLC model [3] (middle red line for longitudinal and top blue line for transverse). We find that the long-range protein occupation correlation length is (numerically) equal the longitudinal tangent vector correlation length $\xi_{\parallel}(f)$. (d) The long-range protein occupation correlation amplitude $C^{n}(f)/e^{2\mu}$ (bottom grey filled circles), the short-range protein occupation correlation amplitude $C^{n\eta}(f)/e^{2\mu}$ (top black asterisks), and the zero-distance protein occupation correlation function $G^{\mathbf{n}}(0, f)/e^{2\mu}$ (top pink crosses).

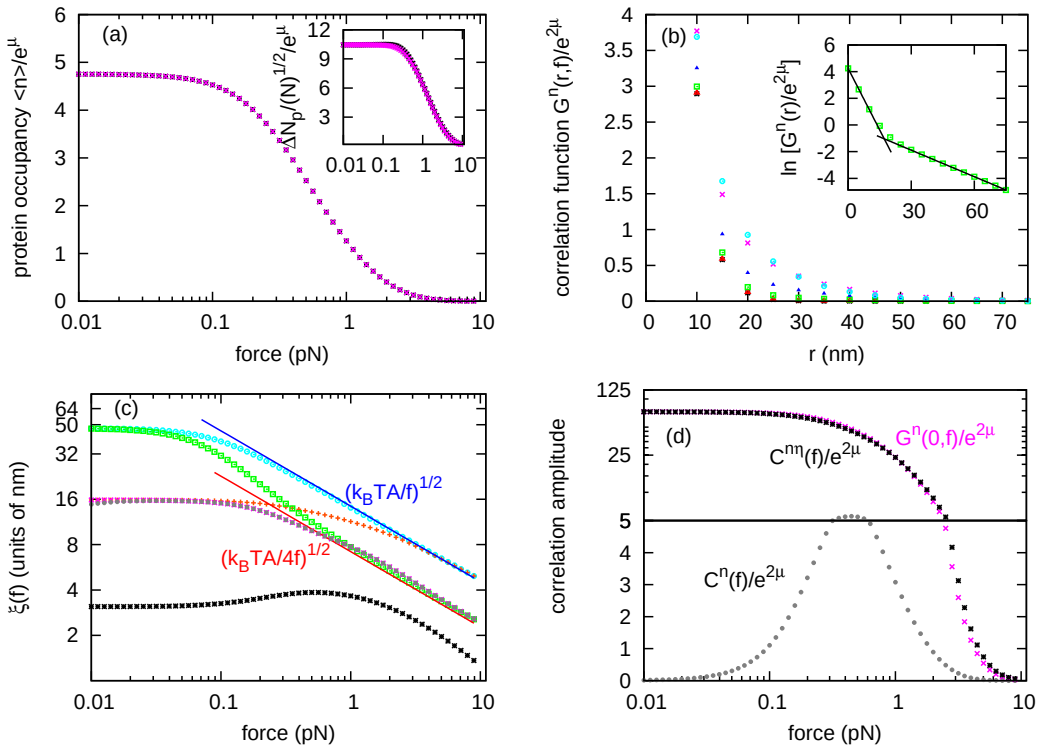

FIG. 6. (color online) Protein occupation correlation functions and correlation lengths along DNA polymer in dilute protein solution as in Fig. 5 except for the use of a different chemical potential value ($\mu = -3$).

Comparing with previous results, we find when there is intrinsic protein-protein binding cooperativity, the interaction gives rise to a new protein occupation correlation which has a bigger amplitude but a shorter correlation length. The short-range protein occupation correlation length increases as the protein concentration (μ) is increased or as cooperativity strength (η) is increased.

S8. PROTEIN OCCUPATION CORRELATION LENGTHS FOR OTHER VALUES OF γ AND a'

We calculated the correlation lengths for varied values of preferred angle θ and bending rigidity a'. Fig. 7 shows results for $\theta = 30°$ and $a' = 50$, Fig. 8 shows results for $\theta = 150°$ and $a' = 50$, and Fig. 9 shows results for $\theta = 90°$ and $a' = 2$.

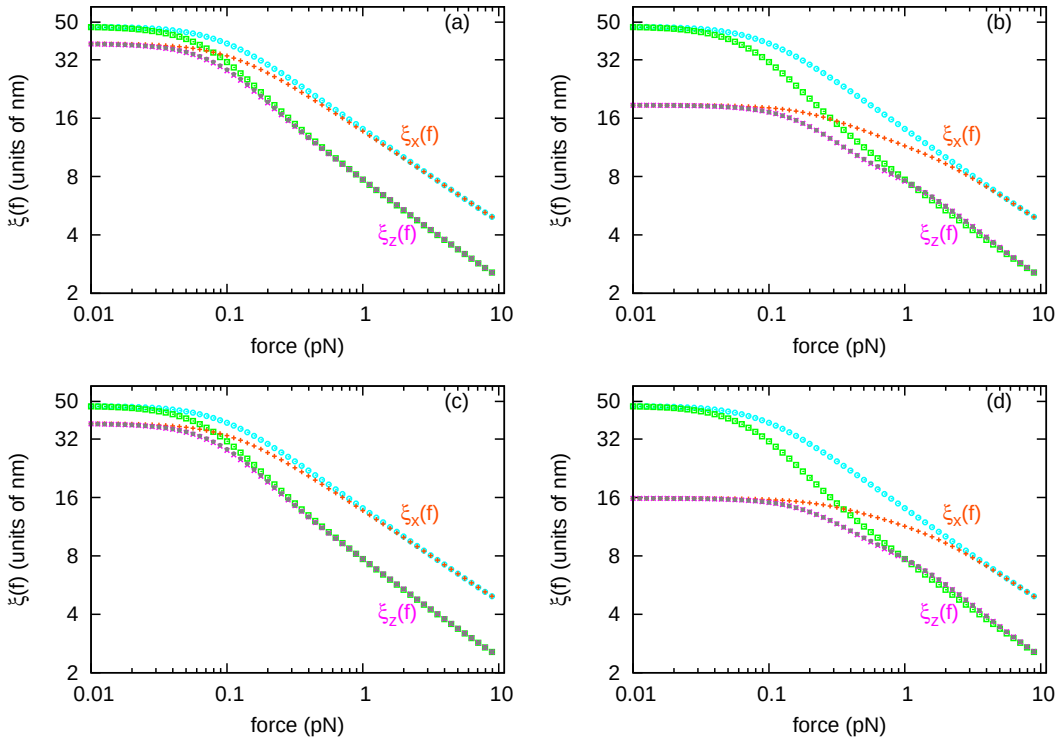

FIG. 7. The longitudinal, transverse, and protein occupation correlation lengths along DNA polymer for $\theta = 30^o$ ($\gamma = \cos 30^o$) and $a' = 50$. (a)-(d) show results for $\mu = -5$ and $\eta = 0$, $\mu = -3$ and $\eta = 0$, $\mu = -5$ and $\eta = 1$, and $\mu = -3$ and $\eta = 1$, respectively. In each panel, we show the longitudinal tangent vector correlation length $\xi_{\|}$ (bottom pink crosses), the transverse tangent vector correlation length $\xi_{\perp}$ (middle orange pluses), and the long-range protein occupation correlation length (bottom grey filled circles, overlapping with bottom pink crosses). As reference, we also show the longitudinal and transverse tangent vector correlation lengths for naked DNA (middle green squares for longitudinal and top aqua circles for transverse).

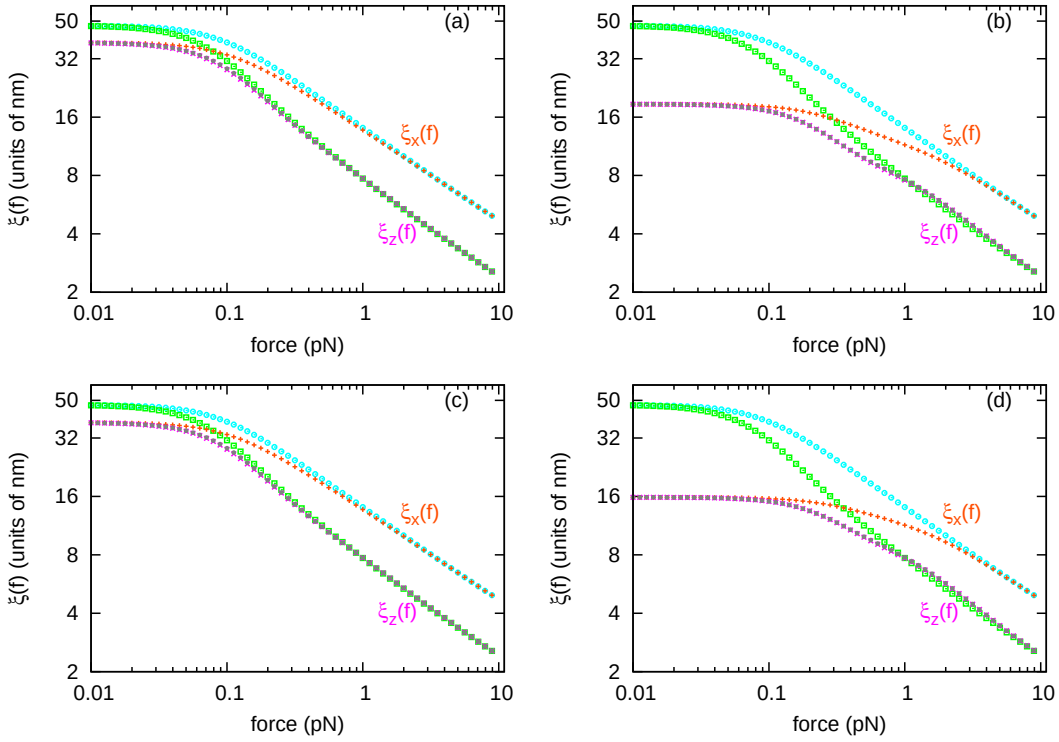

FIG. 8. The longitudinal, transverse, and protein occupation correlation lengths along DNA polymer for $\theta = 150^o$ ($\gamma = \cos 150^o$) and $a' = 50$. (a)-(d) show results for $\mu = -5$ and $\eta = 0$, $\mu = -3$ and $\eta = 0$, $\mu = -5$ and $\eta = 1$, and $\mu = -3$ and $\eta = 1$, respectively; all other parameters and color-coding are the same as in Fig. 7.

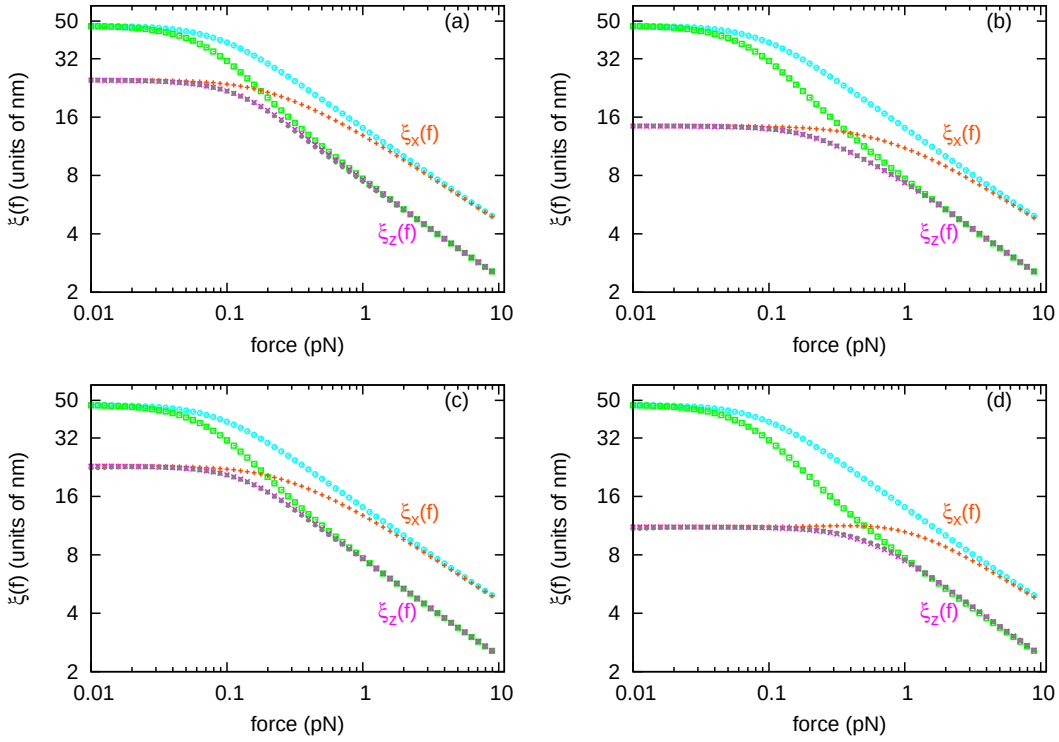

FIG. 9. The longitudinal, transverse, and protein occupation correlation lengths along DNA polymer for $\theta = 90°$ ($\gamma = \cos 90° = 0$) and $a' = 2$. (a)-(d) show results for $\mu = -5$ and $\eta = 0$, $\mu = -4$ and $\eta = 0$, $\mu = -5$ and $\eta = 1$, and $\mu = -4$ and $\eta = 1$, respectively; all other parameters and color-coding are the same as in Fig. 7.

From these results we find that different γ and a' values affect numerical details of the calculation results, but do not affect the main conclusion of the manuscript, *i.e.*, that the range of interaction between DNA-bending proteins is controlled by the second-longest correlation length for bending fluctuations. This is to be expected since different choices of γ and a' do not affect the axial rotation symmetry of the interaction between the proteins and DNA.

[1] H. Zhang and J.F. Marko, Phys. Rev. E **82**, 051906 (2010).
[2] J. Yan and J.F. Marko, Phys. Rev. E **68**, 011905 (2003).
[3] J.F. Marko and E.D. Siggia, Macromolecules **28**, 8759 (1995).
[4] J. Yan, R. Kawamura, and J.F. Marko, Phys. Rev. E **71**, 061905 (2005).