Biophysical Journal, Volume *120*

Supplemental Information

Impact of Self-Association on the Architectural Properties of Bacterial Nucleoid Proteins

Marc Joyeux

Impact of self-association on the architectural properties of bacterial nucleoid proteins

- Supporting Material –

M. Joyeux

Laboratoire Interdisciplinaire de Physique,
CNRS and Université Grenoble Alpes, Grenoble, France

MODEL AND SIMULATIONS

Temperature T is assumed to be 298 K throughout the study. The DNA molecule is modeled as a circular chain of $n = 2880$ beads with radius $a = 1.0$ nm separated at equilibrium by a distance $l_0 = 2.5$ nm and enclosed in a sphere with radius $R_0 = 120$ nm. Each bead represents 7.5 DNA base pairs (bp). The contour length of the DNA molecule and the cell volume correspond approximately to 1/200th of the values for *E. coli* cells, so that the nucleic acid concentration of the model is close to the physiological one (≈10 mM). The potential energy of the DNA chain consists of 4 terms, namely, the stretching energy $V_{\mathrm{s}}^{\mathrm{DNA}}$, the bending energy $V_{\mathrm{b}}^{\mathrm{DNA}}$, the electrostatic repulsion $V_{\mathrm{e}}^{\mathrm{DNA}}$, and a confinement term $V_{\mathrm{w}}^{\mathrm{DNA}}$

$$E^{\mathrm{DNA}} = V_{\mathrm{s}}^{\mathrm{DNA}} + V_{\mathrm{b}}^{\mathrm{DNA}} + V_{\mathrm{e}}^{\mathrm{DNA}} + V_{\mathrm{w}}^{\mathrm{DNA}} \ . \tag{S1}$$

The stretching and bending contributions write

$$V_{\mathrm{s}}^{\mathrm{DNA}} = \frac{h}{2}\sum_{k=1}^{n}(l_k - l_0)^2 \tag{S2}$$

$$V_{\mathrm{b}}^{\mathrm{DNA}} = \frac{g}{2}\sum_{k=1}^{n}\theta_k^2 \ , \tag{S3}$$

where $\mathbf{r}_k$ denotes the position of DNA bead k (with the convention that $\mathbf{r}_{n+k} = \mathbf{r}_k$), $l_k = \|\mathbf{r}_k - \mathbf{r}_{k+1}\|$ the distance between two successive beads, and $\theta_k = \arccos((\mathbf{r}_{k-1} - \mathbf{r}_k)(\mathbf{r}_k - \mathbf{r}_{k+1}) / (\|\mathbf{r}_{k-1} - \mathbf{r}_k\| \, \|\mathbf{r}_k - \mathbf{r}_{k+1}\|))$ the angle formed by three successive beads. The stretching energy $V_{\mathrm{s}}^{\mathrm{DNA}}$ is a computational device without biological meaning, which is aimed at avoiding a rigid rod description. The stretching force constant h is set to $h = 100\, k_{\mathrm{B}}T / l_0^2$, which ensures that the variations of the distance between successive beads

remain small enough (1). In contrast, the bending rigidity constant is obtained from the known persistence length of the DNA, $\xi = 50$ nm, according to $g = \xi\, k_{\mathrm{B}}T / l_0 = 20\, k_{\mathrm{B}}T$.

Moreover, it is assumed that the repulsion between DNA beads that are not close neighbours along the chain is driven by electrostatics and can be expressed as a sum of repulsive Debye-Hückel potentials with hard core

$$V_{\mathrm{e}}^{\mathrm{DNA}} = e_{\mathrm{DNA}}^2 \sum_{k=1}^{n-4} \sum_{K=k+4}^{n} H(\|\mathbf{r}_k - \mathbf{r}_K\| - 2a) \ , \tag{S4}$$

where

$$H(r) = \frac{1}{4\pi\varepsilon r} \exp\left(-\frac{r}{r_D}\right) \ . \tag{S5}$$

In Eq. (S4), e_{DNA} denotes the electric charge placed at the centre of each DNA bead. The numerical value $e_{\mathrm{DNA}} = -3.525\,\bar{e}$, where $\bar{e}$ is the absolute charge of the electron, is the product of l_0 and the net linear charge density along a DNA molecule immersed in a buffer with monovalent cations derived from Manning's counterion condensation theory ($-\bar{e}/\ell_B \approx -1.41\ \bar{e}/\mathrm{nm}$) (2,3). In Eq. (S5), $\varepsilon = 80\,\varepsilon_0$ denotes the dielectric constant of the buffer and $r_D = 1.07$ nm the Debye length, whose value corresponds to a concentration of monovalent salt of 100 mM, which includes the (implicit) cationic counterions that are required for the global electroneutrality of the investigated systems. Interactions between close neighbours ($1 \le |k-K| \le 3$) are not included in Eq. (S4) because it is considered that they are already accounted for in the stretching and bending terms. The repulsive interaction between two DNA beads is shown as a blue short-dashed line in Fig. S1. Note that the equilibrium distance between two successive beads ($l_0 = 2.5$ nm) is small enough to ensure that different DNA segments do not cross in spite of the small value of r_D.

The confinement term $V_{\mathrm{w}}^{\mathrm{DNA}}$ is taken as a sum of repulsive terms

$$V_{\mathrm{w}}^{\mathrm{DNA}} = 10\, k_{\mathrm{B}}T \sum_{k=1}^{n} f(\|\mathbf{r}_k\|) \ , \tag{S6}$$

where the function f is defined according to

$$\begin{aligned} &f(r) = 0, \quad \text{if} \quad r \le R_0 \\ &f(r) = \frac{r^6}{R_0^6} - 1, \quad \text{if} \quad r > R_0 . \end{aligned} \tag{S7}$$

In addition to the DNA chain, the confining sphere contains $P = 200$ DNA-binding protein chains, which corresponds to a protein concentration approximately twice the concentration of H-NS dimers during the cell growth phase and six times the concentration during the stationary phase (4). The number of protein chains was chosen to be as small as possible, but still large enough for the effects discussed in the present paper (compaction and/or stiffening) to be clearly seen in the simulations. $P = 200$ turns out to be an adequate choice in this respect. We note in passing that most experiments dealing with the architectural properties of nucleoid proteins were similarly performed at protein concentrations much larger than physiological ones (5-7). Each DNA-binding protein *j* is modeled as a chain of 7 beads (indexes $m = 1,2,..,7$), where the two terminal beads $m = 1$ and $m = 7$ represent the DNA-binding sites (see Fig. 1). Protein beads have the same radius $a = 1.0$ nm as DNA beads and are separated at equilibrium by the same distance $l_0 = 2.5$ nm. The internal energy of each protein chain *j*, E^j, consists of 4 terms

$$E^j = V_{\mathrm{s}}^j + V_{\mathrm{b}}^j + V_{\mathrm{ev}}^j + V_{\mathrm{w}}^j \ , \tag{S8}$$

where the stretching, bending, and confinement contributions are very similar to their DNA counterparts

$$V_{\mathrm{s}}^j = \frac{h}{2}\sum_{m=1}^{6}(L_{jm} - l_0)^2 \tag{S9}$$

$$V_{\mathrm{b}}^j = \frac{g}{2}\sum_{m=3}^{m=5}\Theta_{jm}^2 \tag{S10}$$

$$V_{\mathrm{w}}^j = 10\, k_{\mathrm{B}}T \sum_{m=1}^{7} f(\|\mathbf{R}_{jm}\|) \ . \tag{S11}$$

In Eqs. (S9)-(S11), $\mathbf{R}_{jm}$ denotes the position of bead m $(1 \le m \le 7)$ of protein chain *j* $(1 \le j \le P)$, L_{jm} the distance between beads *m* and $m+1$ of protein chain *j*, and Θ_{jm} the angle formed by beads $m-1$, *m*, and $m+1$ of protein chain *j*. The values of the stretching force constant *h* and the bending force constant *g* are identical to those of the DNA chain. Note, however, that the sum in the expression of V_{b}^j runs from $m = 3$ to $m = 5$, which means that the DNA-binding beads $m = 1$ and $m = 7$ can rotate without energy penalty around beads $m = 2$ and $m = 6$, respectively. The free rotation of terminal beads mimics the flexible linkers, which connect the C-terminal DNA-binding domains of H-NS to the main body of the dimer (8).

The excluded volume term $V_{\rm ev}^{j}$ ensures that beads belonging to the same chain *j* repel each other at short distances and do not overlap. $V_{\rm ev}^{j}$ is expressed in terms of the repulsive part of a Lennard-Jones 3-6 potential

$$V_{\rm ev}^{j} = 8\,k_{\rm B}T \sum_{m=1}^{5} \sum_{M=m+2}^{7} F(\left\| \mathbf{R}_{jm} - \mathbf{R}_{jM} \right\| | \, r_0) \ , \tag{S12}$$

where $r_0 = 3$ nm and $F(r \,|\, r_0)$ is defined according to

$$\begin{aligned} &F(r \,|\, r_0) = \frac{r_0^6}{r^6} - 2\frac{r_0^3}{r^3} + 1, \quad \text{if} \quad r \le r_0 \\ &F(r \,|\, r_0) = 0, \quad \text{if} \quad r > r_0 \, . \end{aligned} \tag{S13}$$

The repulsive interaction between two protein beads belonging to the same chain is shown as a thin red solid line in Fig. S1.

Interactions between the DNA chain and protein chain *j*, $E^{{\rm DNA}/j}$, are taken as the sum of (attractive or repulsive) Debye-Hückel potentials with hard core, which are complemented with repulsive excluded volume terms for DNA-binding beads $m = 1$ and $m = 7$

$$E^{{\rm DNA}/j} = V_{\rm e}^{{\rm DNA}/j} + V_{\rm ev}^{{\rm DNA}/j} \ , \tag{S14}$$

where

$$V_{\rm e}^{{\rm DNA}/j} = e_{\rm DNA} \sum_{k=1}^{n} \sum_{m=1}^{7} e_m H(\left\| \mathbf{r}_k - \mathbf{R}_{jm} \right\| - a) \tag{S15}$$

and

$$V_{\rm ev}^{{\rm DNA}/j} = 4\,k_{\rm B}T \sum_{k=1}^{n} \sum_{m=1,7} G(\left\| \mathbf{r}_k - \mathbf{R}_{jm} \right\| | \, s_0) \ . \tag{S16}$$

Positive charges $e_m = 2.4\,\overline{e}$ are placed at the centre of DNA-binding beads $m = 1$ and $m = 7$, and negative charges $e_m = -1.2\,\overline{e}$ at the centre of the other beads ($2 \le m \le 6$). The values of these effective charges are compatible with those obtained from a naive counting of the number of positively and negatively charged residues in published crystallographic structures of H-NS (9), except that H-NS is globally neutral, whereas the protein chains have here a total slightly negative charge of $-1.2\,\overline{e}$, in agreement with the fact that most proteins encoded in the genome of *E. coli* are anionic (10). In Eq. (S16), $s_0 = 2$ nm and $G(r \,|\, s_0)$ is the repulsive part of a Lennard-Jones 1-2 potential

$$\begin{aligned} &G(r \,|\, s_0) = \frac{s_0^2}{r^2} - 2\frac{s_0}{r} + 1, \quad \text{if} \quad r \le s_0 \\ &G(r \,|\, s_0) = 0, \quad \text{if} \quad r > s_0 \, . \end{aligned} \tag{S17}$$

The interaction between DNA-binding protein beads ($m = 1$ and $m = 7$) and DNA beads is shown as a thick green long-dashed line in Fig. S1 and the repulsive interaction between other protein beads ($2 \le m \le 6$) and DNA beads as a thin green long-dashed line. The most stable DNA/protein complex is shown in Fig. S2(a) : each DNA-binding protein bead binds simultaneously to two successive DNA beads with a total binding energy of $-7.8\, k_{\mathrm{B}}T$, which is comparable to experimentally determined values for complexes of DNA and H-NS ($\approx -11.0\, k_{\mathrm{B}}T$ (11)). It is seen in this figure that the terminal protein bead and the two DNA beads overlap to some extent for the most stable DNA/protein complex. The reason is that the radius of the beads $a = 1.0$ nm is essentially used to compute the translational diffusion coefficient $D_t = (k_B T)/(6\pi\eta a)$ which governs the Langevin equations (see Eq. (S25) below). Nonetheless, the distance between the centers of a DNA bead and a protein bead can become smaller than $2a$, which reflects the fact that proteins may insert loops in the major or minor groove of the DNA molecule. The sum of the attractive Debye-Hückel potential with hard core at distance *a* (Eq. (S15)) and the repulsive part of a Lennard-Jones 1-2 potential without hard core (Eq. (S16)) results in a maximum binding energy of $-3.9\, k_{\mathrm{B}}T$ at a center-center distance of 1.60 nm (thick green long-dashed line in Fig. S1), which is indeed smaller than $2a = 2.0$ nm and is responsible for the overlaps in Fig. S2(a).

The fact that DNA/protein interactions are mediated uniquely by effective electrostatic charges placed at the center of each bead is certainly a strong approximation, because it is known that proteins and cationic counterions compete for binding to the DNA, and that binding of a protein is consequently accompanied by the release of counterions in the buffer (12-14). Since the released counterions regain translational entropy, the net energy balance for the binding of ligands to the DNA results from subtle enthalpy-entropy compensations (15,16), a point which is overlooked in the present models. There is however no reason, why this approximation should affect the validity of the results discussed in the main text, because these results depend essentially on the ability of the proteins to bind to the DNA and on the bond energy, not on the nature of their interactions.

The two models discussed in the main text differ only in the expression of the interaction $E^{j/J}$ between two protein chains *j* and *J*. The two models agree on the fact that each protein chain contains two isomerization sites, but the properties of these sites differ from Model I to Model II.

In Model I, each isomerization site is made up of a single bead, namely bead $m = 2$ for one site and bead $m = 6$ for the other one (see Fig. 1). The isomerization beads of protein

chain j may bind to the isomerization beads of chain J but repel all other beads through an excluded volume term. The other beads of chain j ($m \neq 2$ and $m \neq 6$) repel all the beads of chain J through the same excluded volume term. More explicitly, for Model I the interaction energy between protein chains j and J, $E^{j/J}$, writes

$$E^{j/J} = V_{\text{iso}}^{j/J} + V_{\text{ev}}^{j/J} \ , \tag{S18}$$

where

$$V_{\text{iso}}^{j/J} = \varepsilon_{\text{LJ}} \sum_{\{m,M\}\in\Im} W(\left\| \mathbf{R}_{jm} - \mathbf{R}_{JM} \right\| | u_0) \tag{S19}$$

and

$$V_{\text{ev}}^{j/J} = 8\, k_{\text{B}}T \sum_{\{m,M\}\notin\Im} F(\left\| \mathbf{R}_{jm} - \mathbf{R}_{JM} \right\| | r_0) \ . \tag{S20}$$

In Eqs. (S19) and (S20), $\Im$ stands for the ensemble

$$\Im = \{\{2,2\},\{2,6\},\{6,6\}\} \ . \tag{S21}$$

In Eq. (S19), the condition $\{m,M\} \in \Im$ indicates that the sum applies to pairs of isomerization beads. In contrast, the restriction $\{m,M\} \notin \Im$ in Eq. (S20) indicates that pairs of isomerization beads do not contribute to the sum. Moreover, in Eq. (S19), $W(r \,|\, u_0)$ is a Lennard-Jones 3-6 potential

$$W(r \,|\, u_0) = \frac{u_0^6}{r^6} - 2\frac{u_0^3}{r^3} \ , \tag{S22}$$

with $u_0 = 1.0$ nm. The interaction potential of two isomerization protein beads is shown for $\varepsilon_{\text{LJ}} = 8\, k_{\text{B}}T$ as a thick red solid line in Fig. S1. $W(r \,|\, u_0)$ is minimum for $r = u_0$, with $W(u_0 \,|\, u_0) = -1$, so that for Model I the isomerization binding energy is simply $-\varepsilon_{\text{LJ}}$. The excluded volume term in Eq. (S19) is similar to that in Eq. (S12) and the function $F(r \,|\, r_0)$ is defined in Eq. (S13).

In Model II, each isomerization site is instead made up of two beads, namely beads $m = 2$ and $m = 3$ for one site and beads $m = 5$ and $m = 6$ for the other one (see Fig. 1). An isomerization site of one protein chain can bind to an isomerization site of another protein chain, but only in a "head-to-tail" fashion. This is obtained by imposing that beads $m = 2$ and $m = 6$ of one chain can bind to beads $m = 3$ and $m = 5$ of the other chain, but repel all other beads. More explicitly, the interaction energy between protein chains j and J, $E^{j/J}$, has the same expression for Model II as for Model I, except that the ensemble $\Im$ is defined according to

$$\Im = \{\{2,3\},\{2,5\},\{3,6\},\{5,6\}\} \ . \qquad \text{(S23)}$$

$\varepsilon_{\mathrm{LJ}}$ is the only variable parameter of the model. It was varied from $4\,k_{\mathrm{B}}T$ to $12\,k_{\mathrm{B}}T$, in order to check the influence of the isomerization binding energy on the equilibrium properties of the system. As already mentioned, for Model I the isomerization binding energy is just $-\varepsilon_{\mathrm{LJ}}$. For Model II, the isomerization binding energy of two protein chains was estimated by letting two chains frozen in minimum energy conformations (Π-shaped) slide parallel to each other and computing their interaction energy as a function of the coordinates of the central bead of one chain relative to the central bead of the other one. The Π-shaped conformation of protein chains was used for the calculation of the isomerization binding energy, because terminal protein beads rotate without energy penalty around beads $m = 2$ and $m = 6$ (Eq. (S10)) and the 90° angle ensures minimum repulsion between these beads. As shown for $\varepsilon_{\mathrm{LJ}} = 8\,k_{\mathrm{B}}T$ in Fig. S2(b), the interaction energy displays a sharp minimum around $x = 7.6$ nm and $y = 1.1$ nm, whose depth is taken as the isomerization binding energy. As illustrated in Fig. S3, this energy varies almost linearly with $\varepsilon_{\mathrm{LJ}}$, from about $-5\,k_{\mathrm{B}}T$ for $\varepsilon_{\mathrm{LJ}} = 4\,k_{\mathrm{B}}T$ to about $-21\,k_{\mathrm{B}}T$ for $\varepsilon_{\mathrm{LJ}} = 12\,k_{\mathrm{B}}T$. For the sake of comparison, it is reminded that the experimentally determined value of the enthalpy change upon forming a complex between two H-NS dimers is $-10.2\,k_{\mathrm{B}}T$ (17). For Model II, two protein chains cannot bind simultaneously to the same isomerization site of a third protein chain, which ensures that protein chains isomerize in the form of filaments rather than clusters.

As for DNA/protein interactions, it is stressed that modeling protein/protein interactions as sums of Lennard-Jones potentials is a strong approximation, because protein self-association is governed by a variety of factors, including electrostatic interactions, hydrogen bonds, geometric frustrations, and hydrophobic interactions (18-20). There is however again no reason, why this approximation should affect the validity of the results discussed in the main text, because these results depend essentially on the geometry of protein assemblies and on the strength of the bonds, not on the nature of protein/protein interactions.

The total potential energy of the system, E_{pot}, is the sum

$$E_{\mathrm{pot}} = E^{\mathrm{DNA}} + \sum_{j=1}^{P} E^{j} + \sum_{j=1}^{P} E^{\mathrm{DNA}/j} + \sum_{j=1}^{P-1} \sum_{J=j+1}^{P} E^{j/J} \ . \qquad \text{(S24)}$$

The dynamics of the model was investigated by integrating numerically the Langevin equations of motion with kinetic energy terms neglected. Practically, the updated position

vector for each bead (whether DNA or protein), $\mathbf{r}_j^{(n+1)}$, was computed from the current position vector, $\mathbf{r}_j^{(n)}$, according to

$$\mathbf{r}_j^{(n+1)} = \mathbf{r}_j^{(n)} + \frac{D_t\,\Delta t}{k_B T}\mathbf{F}_j^{(n)} + \sqrt{2\,D_t\,\Delta t}\;\xi^{(n)}\ , \tag{S25}$$

where the translational diffusion coefficient D_t is equal to $(k_B T)/(6\pi\eta a)$ and $\eta = 0.00089$ Pa s is the viscosity of the buffer at $T = 298$ K. $\mathbf{F}_j^{(n)}$ is the vector of inter-particle forces arising from the potential energy E_{pot}, $\xi^{(n)}$ a vector of random numbers extracted at each step n from a Gaussian distribution of mean 0 and variance 1, and Δt the integration time step, which was set to 1.0 ps. After each integration step, the position of the center of the confining sphere was slightly adjusted so as to coincide with the center of mass of the DNA molecule.

SUPPORTING REFERENCES

1. Jian, H., A. Vologodskii, and T. Schlick. 1997. A combined wormlike-chain and bead model for dynamic simulations of long linear DNA. *J. Comp. Phys.* 136**:**168-179

2. Manning, G. S. 1969. Limiting laws and counterion condensation in polyelectrolyte solutions. I. Colligative properties. *J. Chem. Phys.* 51:924-933.

3. Oosawa, F. 1971. Polyelectrolytes. Marcel Dekker, New York.

4. Azam, T. A., A. Iwata, A. Nishimura, S. Ueda, and A. Ishihama. 1999. Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J. Bacteriol.* 181:6361-6370.

5. Dame, R. T. 2005. The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Mol. Microbiol.* 56:858-870.

6. Song, D., and J. J. Loparo. 2015. Building bridges within the bacterial chromosome. *Trends Genet.* 31:164-173.

7. Liu Y. J., H. Chen, L. J. Kenney, and J. Yan. 2010. A divalent switch drives H-NS/DNA-binding conformations between stiffening and bridging modes. *Genes Dev.* 24:339-344.

8. Dorman, C. J., J. C. D. Hinton, and A. Free. 1999. Domain organization and oligomerization among H-NS like nucleoid-associated proteins in bacteria. *Trends Microbiol.* 7:124-128.

9. Arold, S. T., P. G. Leonard, G. N. Parkinson, and J. E. Ladbury. 2010. H-NS forms a superhelical protein scaffold for DNA condensation. *P. Natl. Acad. Sci. USA.* 107:15728-15732.

10. Link, A. J., K. Robison, and G. M. Church. 1997. Comparing the predicted and observed properties of proteins encoded in the genome of *Escherichia coli* K-12. *Electrophoresis*. 18:1259-1313.

11. Ono, S., M. D. Goldberg, T. Olsson, D. Esposito, J. C. D. Hinton, and J. E. Ladbury. 2005. H-NS is a part of a thermally controlled mechanism for bacterial gene regulation. *Biochem. J.* 391:203-213.

12. Record, M. T., C. F. Anderson, and T. M. Lohman. 1978. Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening, and ion effects on water activity. *Q. Rev. Biophys.* 11:103-178.

13. Mascotti, D. P., and T. M. Lohman. 1990. Thermodynamic extent of counterion release upon binding oligolysines to single-stranded nucleic acids. *P. Natl. Acad. Sci. USA.* 87:3142-3146.

14. Fenley, M. O., C. Russo, and G. S. Manning. 2011. Theoretical assessment of the oligolysine model for ionic interactions in protein-DNA complexes. *J. Phys. Chem. B* 115:9864-9872.

15. Breslauer, K. J., D. P. Remeta, W.-Y. Chou, R. Ferrante, J. Curry, D. Zaunczkowski, J. G. Snyder, and L. A. Marky. 1987. Enthalpy-entropy compensations in drug-DNA binding studies. *P. Natl. Acad. Sci. USA.* 84:8922-8926.

16. Wang, S., A. Kumar, K. Aston, B. Nguyen, J. K. Bashkin, D. W. Boykin, and W. D. Wilson. 2013. Different thermodynamic signatures for DNA minor groove binding with changes in salt concentration and temperature. *Chem. Commun.* 49:8543-8545.

17. Ceschini, S., G. Lupidi, M. Coletta, C. L. Pon, E. Fioretti, and M. Angeletti. 2000. Multimeric self-assembly equilibria involving the histone-like protein H-NS. A thermodynamic study. *J. Biol. Chem.* 275:729-734.

18. Meyer, E. E., K. J. Rosenberg, and J. Israelachvili. 2006. Recent progress in understanding hydrophobic interactions. *P. Natl. Acad. Sci. USA.* 103:15739–15746.

19. Lin, M. S., N. L. Fawzi, and T. Head-Gordon. 2007. Hydrophobic potential of mean force as a solvation function for protein structure prediction. *Structure*. 15:727–740.

20. Makowski, M., C. Czaplewski, A. Liwo, and H. A. Scheraga. 2010. Potential of mean force of association of large hydrophobic particles: Toward the nanoscale limit. *J. Phys. Chem. B.* 114:993–1003.

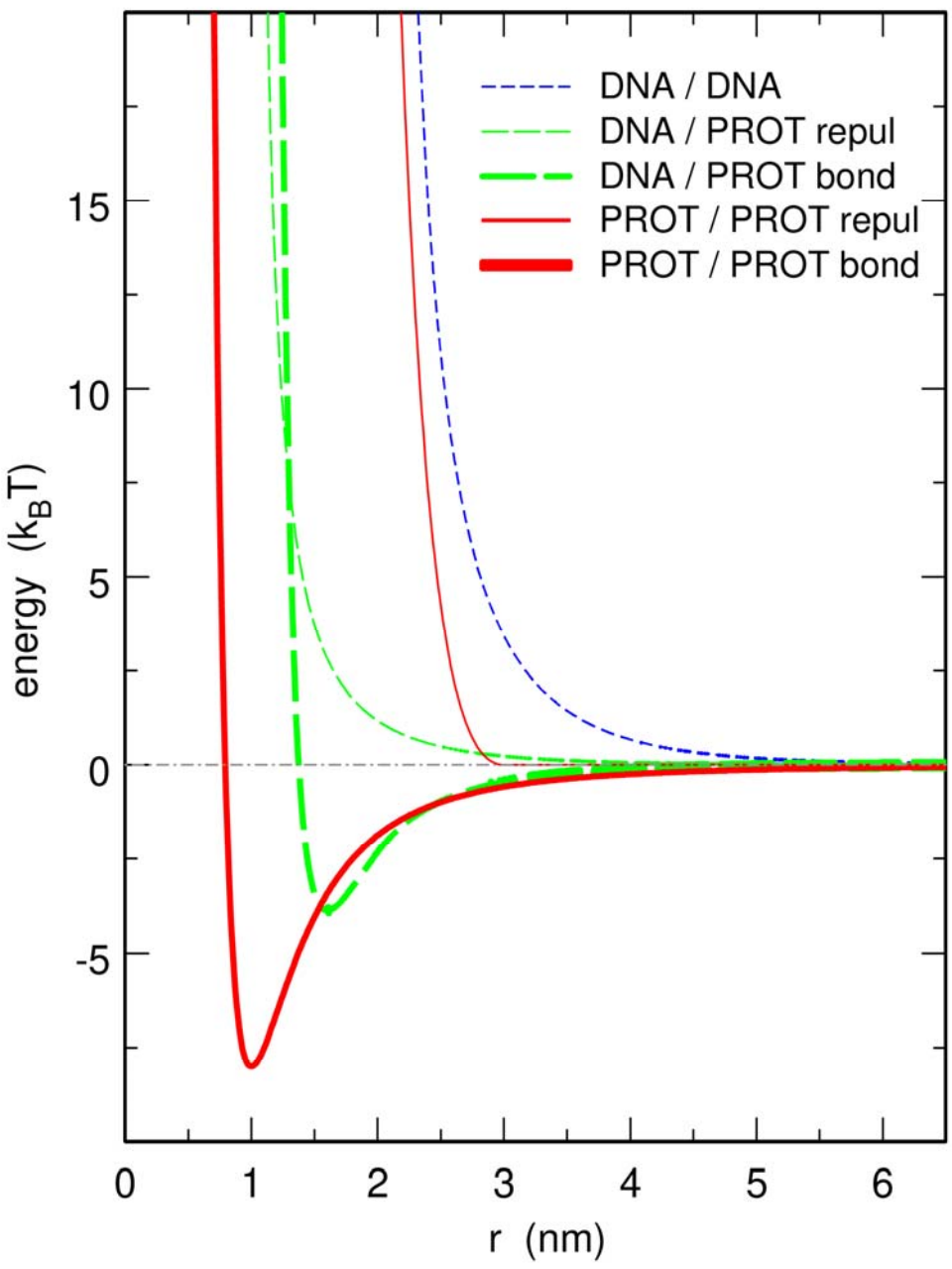

Figure S1 : Plot, as a function of the distance r between the centers of the beads, of the various bead-bead interaction potentials of the model. ***Blue short-dashed line*** : DNA-DNA repulsive interaction (Eqs. (S4)-(S5)). ***Thin green long-dashed line*** : repulsive potential between a DNA bead and a protein bead with $2 \le m \le 6$ (Eq. (S15)). ***Thick green long-dashed line*** : binding potential between a DNA bead and a DNA-binding protein bead with $m = 1$ or $m = 7$ (Eqs. (S15)-(S17)). ***Thin red solid line*** : repulsive potential between two protein beads, which do not belong to the ensemble $\Im$ (Eqs. (S20), (S21) and (S23)). ***Thick red solid line*** : binding potential between two protein beads, which belong to the ensemble $\Im$ (Eqs. (S19), (S21), (S22) and (S23)). r is expressed in nm, energy values in units of $k_B T$.

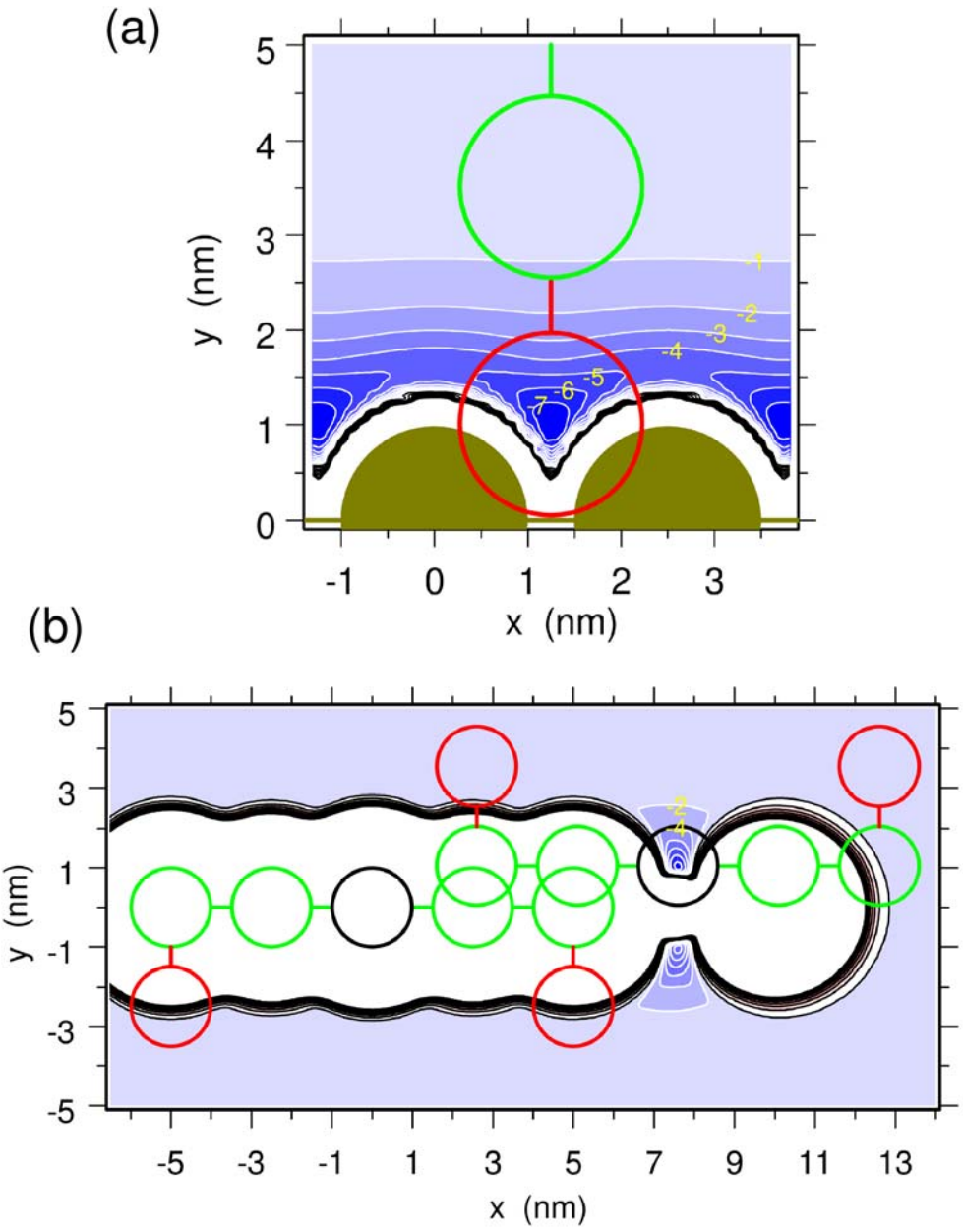

Figure S2 : **(a)** Diagram showing the most stable DNA-protein complex and how energy evolves when the protein is displaced from this position. Filled brown disks represent DNA beads and circles represent protein beads $m = 1$ (red) and $m = 2$ (green). Contour lines are separated by $1\,k_{\mathrm{B}}T$. Energy values expressed in units of $k_{\mathrm{B}}T$ are shown in yellow for a few contours. Binding energy is $-7.8\,k_{\mathrm{B}}T$. **(b)** Diagram showing the most stable complex formed by two proteins for Model II and $\varepsilon_{\mathrm{LJ}} = 8\,k_{\mathrm{B}}T$, and how energy evolves when the upper chain is displaced from this position. Contour lines are separated by $2\,k_{\mathrm{B}}T$. Energy values expressed in units of $k_{\mathrm{B}}T$ are shown in yellow for two contours. Binding energy is $-12.7\,k_{\mathrm{B}}T$. Contour lines display rotational symmetry around the *x* axis (the long axis of the fixed protein chain) and point symmetry with respect to the origin of the plot (the center of the central bead of the fixed protein chain).

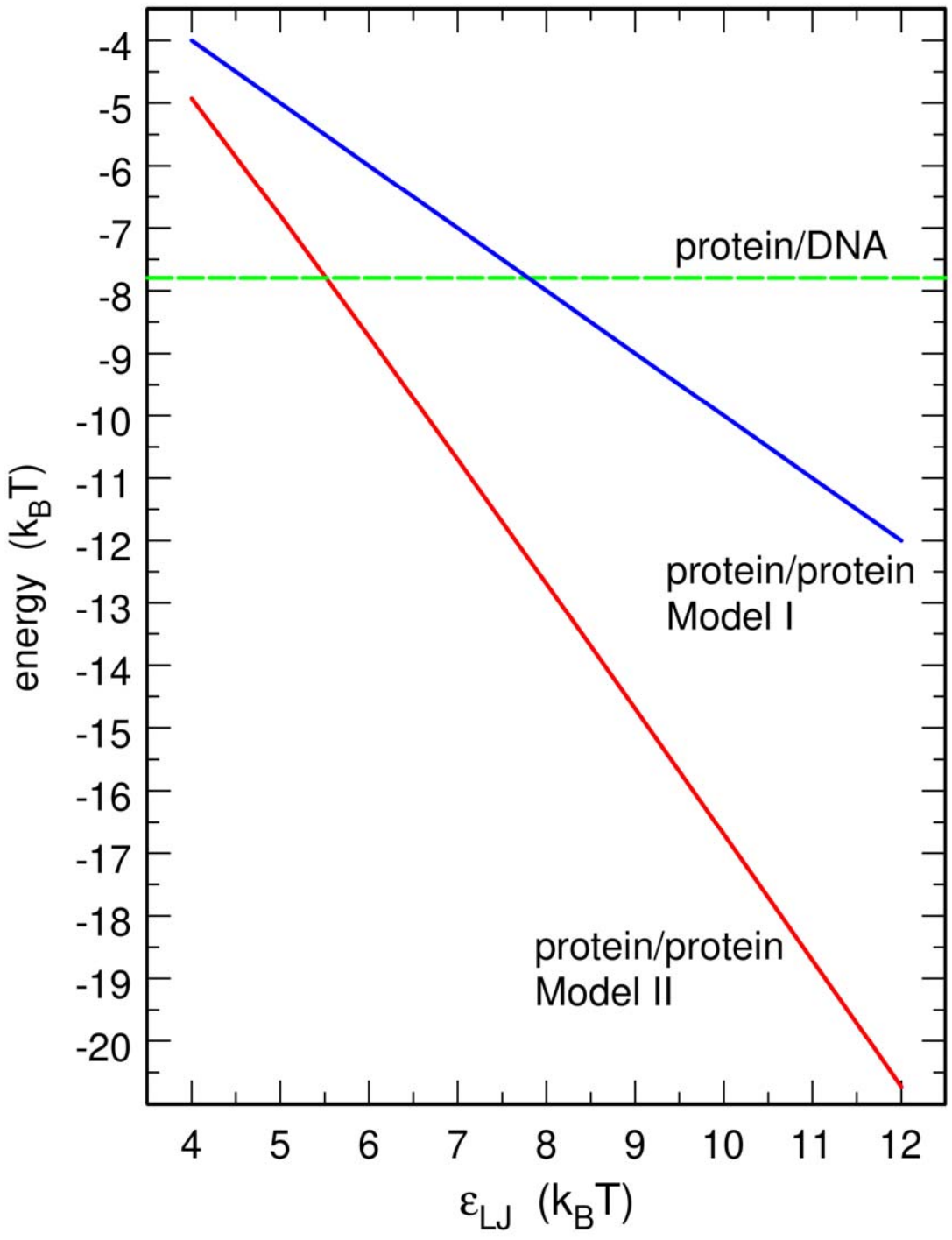

Figure S3 : The blue and red lines show the evolution, as a function of $\varepsilon_{\mathrm{LJ}}$, of the binding energy of protein/protein complexes for Models I and II, respectively. For comparison, the green dashed line shows the bond energy of DNA/protein complexes ($-7.8 k_{\mathrm{B}}T$). $\varepsilon_{\mathrm{LJ}}$ and energy values are expressed in units of $k_{\mathrm{B}}T$.

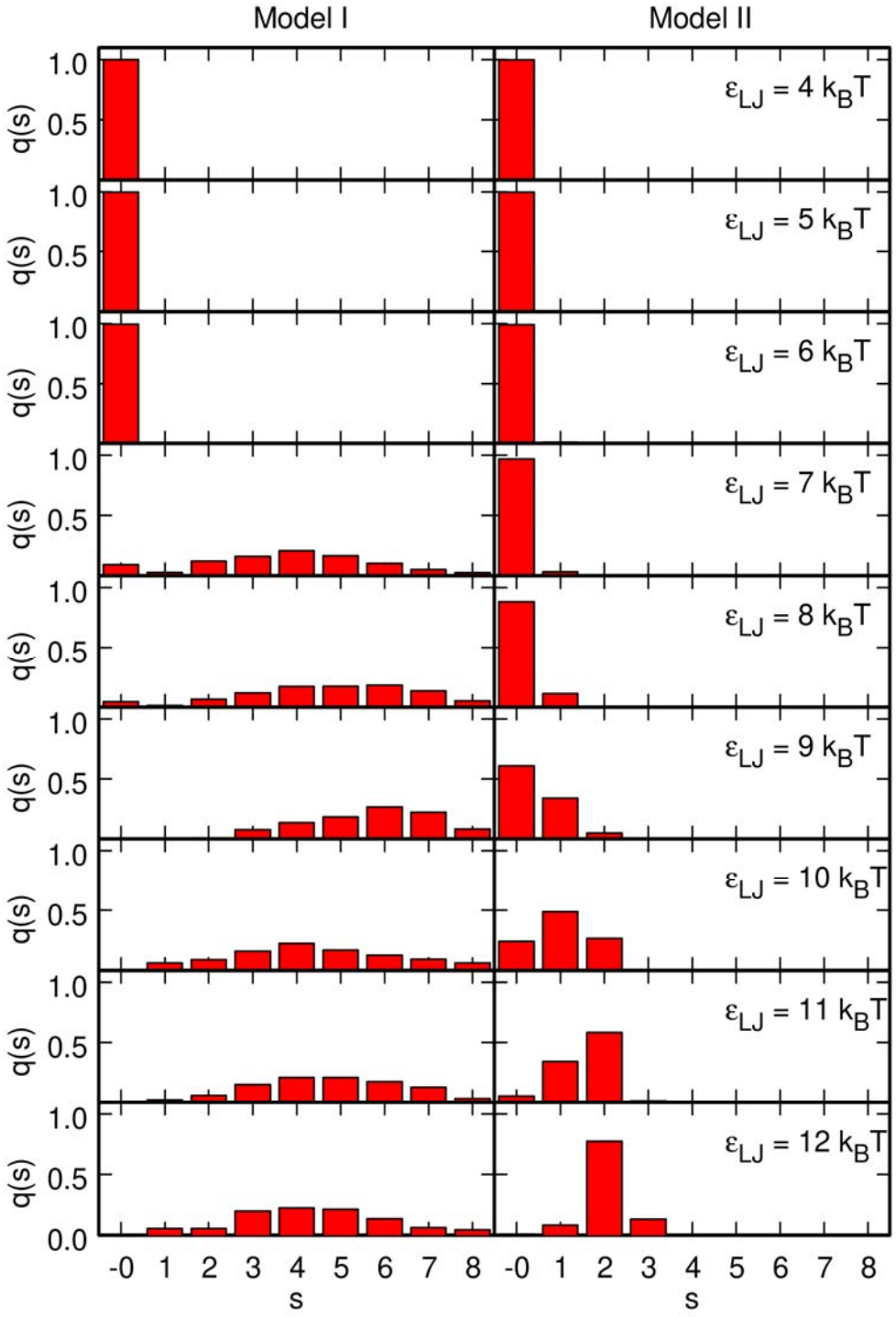

Figure S4 : Plot of the probability distribution $q(s)$ for a protein chain to bind to s other protein chains, for Model I **(left column)** and Model II **(right column),** and values of ε_{LJ} increasing from $4k_BT$ **(top)** to $12k_BT$ **(bottom)**. Each plot was obtained from a single simulation with 200 protein chains (without the DNA chain), by averaging $q(s)$ over time intervals of at least 4 ms after equilibration.

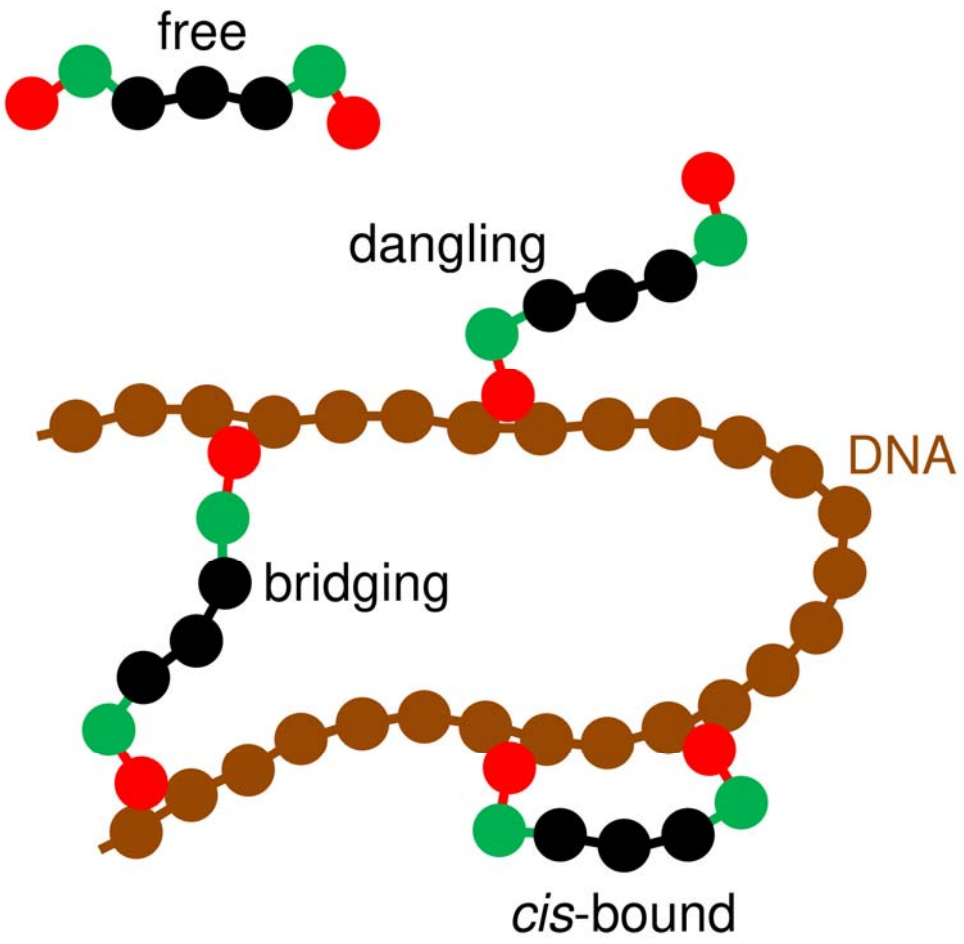

Figure S5 : Diagram showing the different modes of interaction between protein chains and a DNA chain. DNA beads are shown in brown and protein beads in red, green or black, according to the same color code as in Fig. 1.

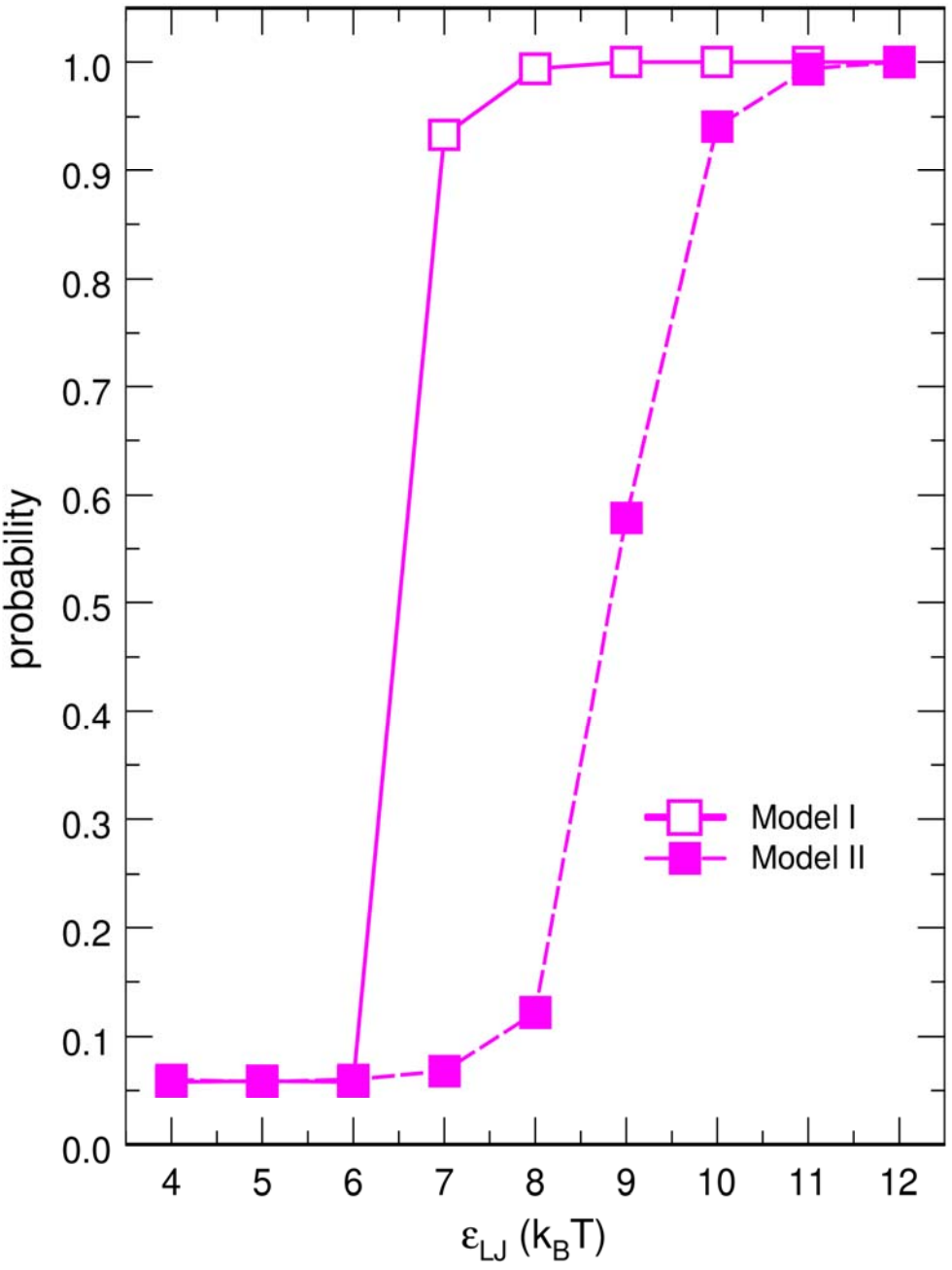

Figure S6 : Plot, as a function of ε_{LJ}, of the average fraction of protein chains which belong to a cluster bridging different DNA segments, for Model I (open symbols) and II (filled symbols). Each point was obtained from a single simulation with the DNA chain and 200 protein chains, by averaging the relevant quantity over time intervals of at least 2.5 ms after equilibration.

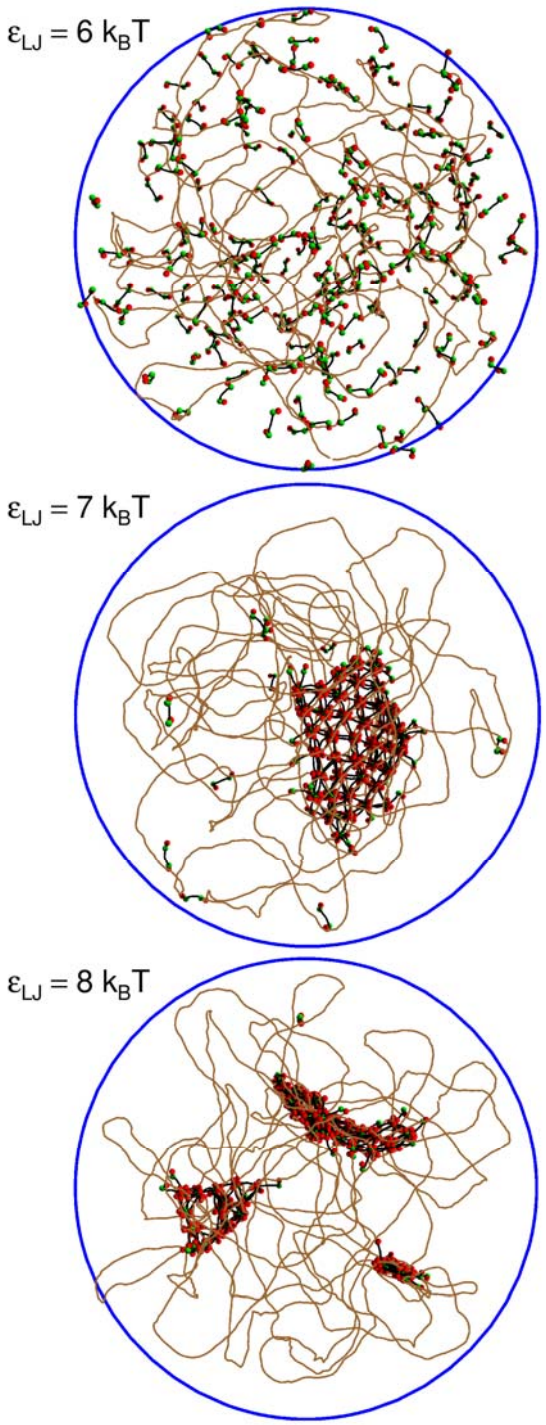

Figure S7 : Representative snapshots extracted from simulations with the DNA chain and 200 protein chains for Model I and $\varepsilon_{LJ} = 6\,k_BT$ **(top)**, $7\,k_BT$ **(middle)** and $8\,k_BT$ **(bottom).** The line joining the centers of DNA beads is shown in brown (DNA beads are not shown). DNA-binding protein beads are shown in red, isomerization beads are shown in green, other protein beads are not shown. The lines joining the centers of protein beads are shown in black. The blue circle is the trace of the confinement sphere.

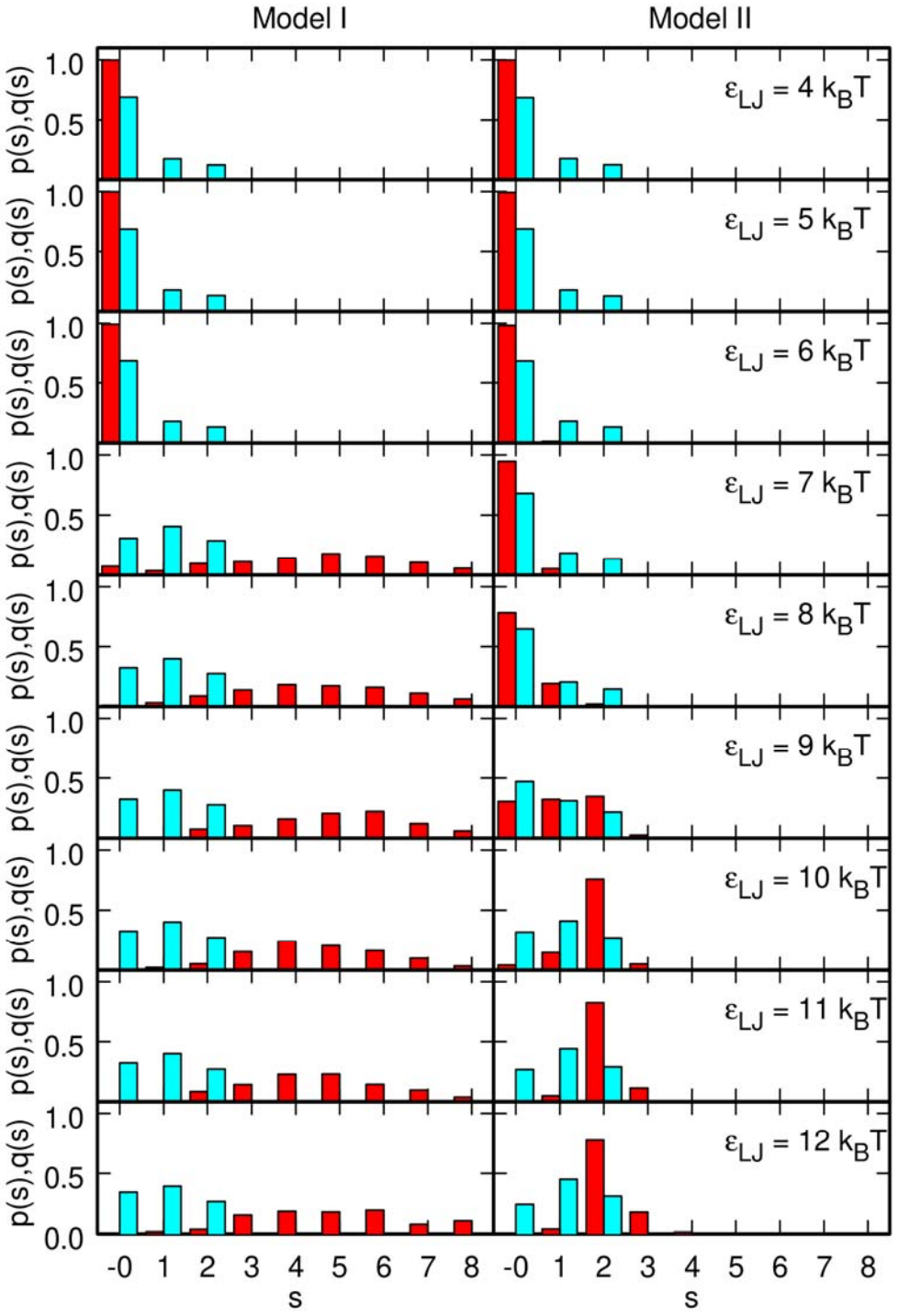

Figure S8 : Plot of the probability distribution $p(s)$ for a DNA-binding protein bead to bind to s DNA beads **(cyan)** and the probability distribution $q(s)$ for a protein chain to bind to s other protein chains **(red)**, for Model I **(left column)** and Model II **(right column),** and values of ε_{LJ} increasing from $4k_BT$ **(top)** to $12k_BT$ **(bottom)**. Each plot was obtained from a single simulation with the DNA chain and 200 protein chains, by averaging $p(s)$ and $q(s)$ over time intervals of at least 2.5 ms after equilibration.

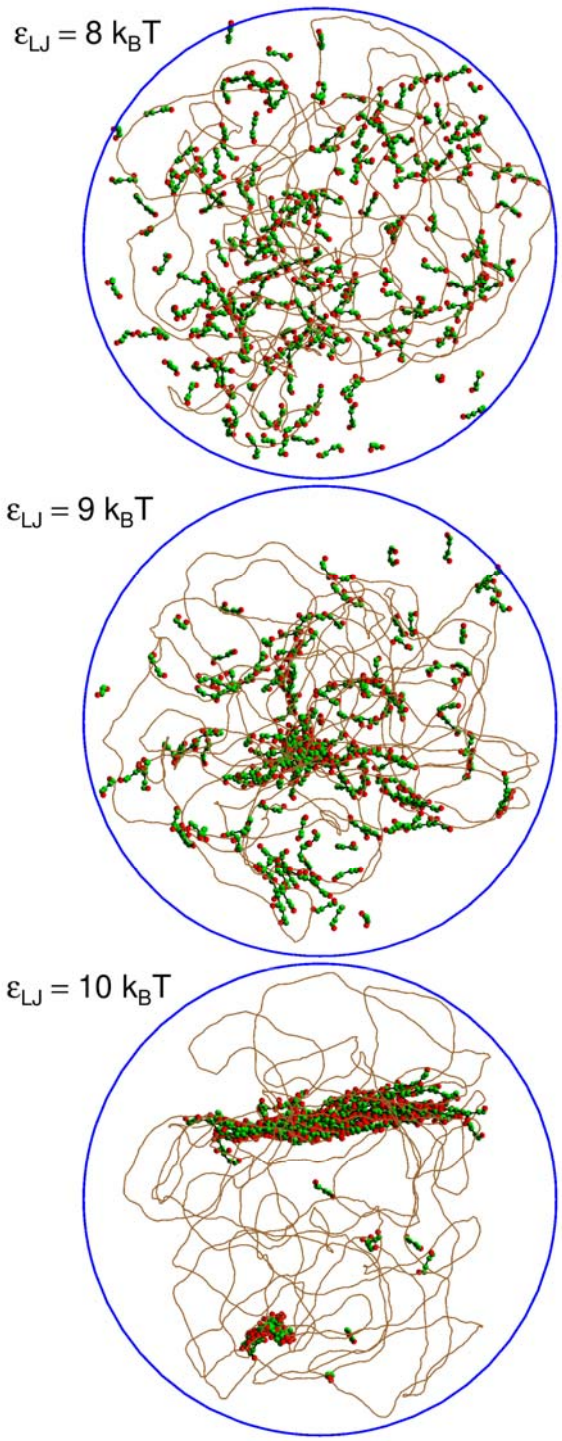

Figure S9 : Representative snapshots extracted from simulations with the DNA chain and 200 protein chains for Model II and $\varepsilon_{LJ} = 8\,k_BT$ **(top)**, $9\,k_BT$ **(middle)** and $10\,k_BT$ **(bottom).** The line joining the centers of DNA beads is shown in brown (DNA beads are not shown). DNA-binding protein beads are shown in red, isomerization beads are shown in green, other protein beads are not shown. The lines joining the centers of protein beads are shown in black. The blue circle is the trace of the confinement sphere.