Helical Structure Determines Different Susceptibilities of dsDNA, dsRNA, and tsDNA to Counterion-Induced Condensation

Alexei A. Kornyshev[†] and Sergey Leikin[‡*]
[†]Department of Chemistry, Imperial College London, London, United Kingdom; and [‡]Section of Physical Biochemistry, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland

ABSTRACT Recent studies of counterion-induced condensation of nucleic acid helices into aggregates produced several puzzling observations. For instance, trivalent cobalt hexamine ions condensed double-stranded (ds) DNA oligomers but not their more highly charged dsRNA counterparts. Divalent alkaline earth metal ions condensed triple-stranded (ts) DNA oligomers but not dsDNA. Here we show that these counterintuitive experimental results can be rationalized within the electrostatic zipper model of interactions between molecules with helical charge motifs. We report statistical mechanical calculations that reveal dramatic and nontrivial interplay between the effects of helical structure and thermal fluctuations on electrostatic interaction between oligomeric nucleic acids. Combining predictions for oligomeric and much longer helices, we also interpret recent experimental studies of the role of counterion charge, structure, and chemistry. We argue that an electrostatic zipper attraction might be a major or even dominant force in nucleic acid condensation.

INTRODUCTION

Interactions of nucleic acids at close separations play crucial role in some of the most fundamental biological reactions, including packing of genetic material in cells and viruses, homologous pairing, and RNA folding. Nucleic acid helices are among the most highly charged biological molecules (see Table 1). The ability of nucleic acids to overcome the repulsion caused by their extremely high charge density is generally attributed to counterion-induced (or mediated) attractive interactions, because specific counterions are needed to condense them into compact aggregates or structures (1).

The physics underlying these attractive interactions is still debated (1–7). The proposed ideas include attractive hydration forces (8,9), bridging by shared counterions (10), correlated fluctuations in the density of condensed counterions (11), attraction between Wigner-crystal-like lattices of counterions (12,13), and zipper-like alignment of negatively charged strands of the sugar-phosphate backbone opposite to positively charged strands of counterions bound in nucleic acid grooves (14). For detailed analysis of these and other ideas, see, e.g., Kornyshev et al. (7) and references therein. Resolving which mechanisms of counterion-induced attraction contribute the most to interactions between nucleic acids is important not only for understanding the principles of nucleic acid behavior but also for rapidly expanding bioengineering applications of these molecules (15).

One of the most straightforward approaches to this problem is to examine how the propensity of nucleic acids to condense depends on their structure as well as on the charge and chemistry of the counterions. Experimental studies along these lines produced a number of challenging observations. Among the most puzzling: Oligomeric double-strand (ds) RNA helices, expected to be more susceptible to counterion-induced condensation due to their higher surface charge density, were found to resist condensation (16). Oligomeric triple-strand (ts) DNA molecules were found to be susceptible to condensation by Mg^{2+} and Ca^{2+} whereas their dsDNA analogs were not (17). No evidence of attractive interactions was detected between oligomeric dsDNA helices at counterion concentrations below the condensation threshold (18). These observations might be explained by different molecular mechanisms, as proposed in the original studies. However, the Occam's Razor principle of simplicity (entities are not to be multiplied beyond necessity; see, e.g., Baker (19)) provides a compelling argument for seeking a more universal interpretation.

To examine whether these observations might indeed be different facets of a more general mechanism, here we extend an earlier theory of DNA-DNA interactions to oligomeric nucleic acids with different number and configuration of strands. We take into account that the total interaction energy for oligomers may not exceed the thermal energy k_BT by much, even at short distances between the molecules (k_B is the Boltzmann constant and T is the absolute temperature). In this case, thermal motions may produce important qualitative effects, necessitating a thorough statistical mechanical analysis, which is described in this article.

Our calculations rationalize different susceptibilities of oligomeric dsRNA, dsDNA, and tsDNA to condensation based on known differences in the helical structure of the molecules and preferential counterion adsorption in their grooves rather than possible differences in the underlying

Submitted December 10, 2012, and accepted for publication March 18, 2013.

*Correspondence: leikins@mail.nih.gov

Editor: Nathan Baker.

0006-3495/13/05/2031/11 $2.00

http://dx.doi.org/10.1016/j.bpj.2013.03.033

TABLE 1 Structural parameters and charge density of dsDNA (B-form), dsRNA, and tsDNA

Parameter	B-DNA	dsRNA	tsDNA
Helical pitch (Å)	33.8	30.9	38.4
Number of basepairs per helical turn	10	11	12
Groove half-width (rad/π)	0.4, 0.6	0.64, 0.36	0.36, 0.35, 0.29
Radius at the center of phosphate groups (Å)	9	8.5	9.5
Axial distance per elementary charge, l_c (Å)	1.7	1.4	1.1
Average surface charge density of phosphates, σ ($\mu C/cm^2$)	17	21	25

interactions. These results support the idea of zipper-like alignment of negatively charged phosphate strands with positively charged counterions bound in grooves of the opposing molecule as a mechanism of counterion-induced attraction between the molecules. We argue that this electrostatic zipper mechanism might also explain recently observed counterion charge, structure, and chemistry effects on condensation of longer nucleic acids. Other interactions might contribute to the condensation, but they are less universal in nature and cannot explain the entire data set.

To facilitate understanding of our results by readers who may be less interested in mathematical aspects of the theory than in the nature of the main results, we formulate the model and basic equations, and then concentrate on the physics of the predicted and observed effects. All mathematical derivations are appended in the Supporting Material. Because the statistical mechanics is different for oligomeric and longer nucleic acids and because relevant calculations for long dsDNA have been published already (20), here we focus on the calculations for oligomers. In Discussion, however, we consider recent experimental data for both oligomeric and longer molecules and evaluate more general principles for the electrostatic zipper mechanism of nucleic acid condensation.

MODEL

Basic approach

In this study, we calculate the free energy of electrostatic interactions between helical oligonucleotides in solution and in aggregates as a function of their interaxial separation. For simplicity, we focus on parallel-oriented oligomers and use the corresponding free energy for qualitatively evaluating whether differences in their helical structure and preferential binding of counterions might explain the experimental observations. This approach relies on previous studies of longer molecules, which suggest that chiral interactions between helices at small interaxial angles do not change the qualitative conclusions on requirements for intermolecular attraction and aggregation (7). Analysis of interactions at nonzero interaxial angles is much more complicated and lies beyond the scope of this study.

To clarify the underlying molecular mechanisms and physical principles, we develop an analytical theory that follows the general approach recently reviewed in Kornyshev et al. (7). Alternatively, this problem could be addressed by computer simulations based on explicit, all-atom solvent models or Poisson-Boltzmann solvent description. Such simulations are realistic for short oligonucleotide helices. Yet, they provide less clear delineation of different interaction mechanisms and they are still imprecise. They also utilize multiple simplifications and assumptions (interaction potentials for the atoms, summation of long-range interactions, effects of box size and boundary conditions, validity of the nonlinear Poisson-Boltzmann model in the immediate vicinity of the nucleic acid surface, etc.). Still, they are complementary to our approach and might be interesting to pursue.

Structure and charge state of nucleic acid helices

We model the helices as dielectric cylinders with charged groups at the cylinder/water interface (Fig. 1). To elucidate effects of the nucleic-acid structure and preferential counterion binding at specific sites that may play a crucial role in intermolecular interactions (7), we account not only for the average surface charge density but also for helical patterning of the charges (Fig. 1). In Table 1, we gather pertinent parameters of nucleic acid structure and charge distribution that appear in the theory. We describe negatively charged phosphate groups as point charges located along two helical strands in dsDNA and dsRNA or along three helical strands in tsDNA. We assume that the fraction θ of the total phosphate charge is compensated by bound counterions, which are treated as part of the nucleic acid charge. The net surface charge density of the molecule is $(1–\theta)\sigma$, where σ is the surface change density of phosphates. The fraction f_1 of bound counterions is located in the center of the minor groove, the fraction f_2 is located in the center of the major groove in dsDNA and dsRNA or distributed equally between the centers of the two other grooves in tsDNA, and the remaining $(1–f_1–f_2)$ counterions are bound at random locations. We describe the effects of other, free electrolyte ions within the Debye-Hückel (linearized Poisson-Boltzmann) theory. The advantages and drawbacks of this approach were discussed in detail in Kornyshev et al. (7).

Pair interaction energy

We base our calculations on the theory of electrostatic interactions between molecules with helical charge motifs (7). Within this theory, the electrostatic interaction energy between two parallel helices of length L at large interaxial distances R may be approximated with

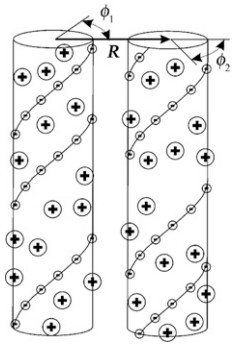

FIGURE 1 Model of electrostatic interaction between two parallel, double-stranded nucleic acids at close separations. The molecules are represented by dielectric cylinders with helical strands of point-like negative charges at the surface (*small circles* with the *minus sign* connected by *helical lines*). Bound counterions are modeled as point-like charges (*large circles* with the *plus sign*) at the cylinder surface, located randomly or in the middle between the strands of negative charges. Counterions responsible for the nonlinear screening of the nucleic acid charge are considered to be bound (7). Roughness of nucleic acid surfaces and poorly known dielectric properties of water in their grooves preclude evaluation of the effective cylinder radius with better than 2–3 Å accuracy from known molecular structures. For B-DNA, the effective radius that provides the best fit for the measured intermolecular forces is 11.2 Å, which is within the expected range between the 9 Å radius at centers and 12 Å radius at outer surfaces of phosphate groups (20). Because radii at the centers of phosphate groups in dsRNA and tsDNA are within 0.5 Å of that for B-DNA (Table 1), we use the same 11.2 Å effective radius for all three nucleic acids. It is important to emphasize that separation between nucleic acid surfaces in aggregates condensed by counterions is typically <10 Å (8,9), which is smaller than the separation between the strands of negative charges (Table 1). Therefore, it is essential to account for the discreteness and helical arrangement of the strands in any theory of such aggregates (7). Zipper-like juxtaposition of positively charged counterions bound in the grooves with negatively charged phosphate strands on the opposing molecule leads to electrostatic zipper attraction between the molecules, provided that bound counterions balance a sufficient fraction of the phosphate charges (14).

$$\frac{E_{\text{int}}}{k_{\text{B}}T} \approx L\frac{2l_{\text{B}}}{l_c^2}\sum_{n=0}^{\infty}(2-\delta_{n,0})\zeta_n^2\left\{c_n^{im}\frac{\exp(-2\kappa_n R)}{\kappa_n R} + (-1)^n c_n\frac{\exp(-\kappa_n R)}{\sqrt{\kappa_n R}}\frac{1}{L}\int_0^L dz\,\cos\left(n[\Delta\phi(z)]\right)\right\}. \quad (1)$$

Here, $l_{\text{B}} \approx 7$ Å is the Bjerrum length in water, and l_c is the axial length per phosphate group in the nucleic acid helix (see Table 1). The summation is performed over helical harmonics n of the interaction. $\delta_{n,m} = 1$ at $n = m$ $\delta_{n,m} = 0$ at $n \neq m$, ζ_n are helical harmonics of the density of the charged groups and bound counterions on nucleic acid surface (see Eq. S9 in the Supporting Material), e.g.,

$$\zeta_n = [f_1 + f_2(-1)^n + (1 - f_1 - f_2)\delta_{n,0}]\theta - \cos\left(n\tilde{\phi}_s\right) \quad (2)$$

for double-stranded helices with the azimuthal half-width of the minor groove $\tilde{\phi}_s$, and $\Delta\phi(z)$ is the difference between the azimuthal orientations of the helices at point z along their length.

$$\kappa_n = \sqrt{\kappa^2 + \left(\frac{2\pi n}{H}\right)^2} \quad (3)$$

is the reciprocal decay length for each helical harmonic of the direct charge-charge interaction, κ is the reciprocal screening length in electrolyte solution, H is the helical pitch of the molecules, and c_n and c_n^{im} are coefficients that depend only on κ_n and radius of the molecules, which are defined by Eqs. S7 and S8 of the Supporting Material. Note that Eq. 1 is a more transparent, simplified expression obtained by asymptotic expansion of Bessel functions at large R. More accurate Eqs. S1–S5, presented in the Supporting Material, were used in all of our calculations.

Each interaction harmonic consists of two contributions given by the two terms in the curly brackets in Eq. 1. The first one describes the energy of the dielectric core of each helix in the electric field created by the other helix, hereafter referred to as "image-charge interactions" (interaction of a charge with a dielectric may be represented by a sum of interaction of this charge with a set of images of this charge across the dielectric interface). The second one describes the energy of charges on one helix in the electric field created by the other helix, hereafter referred to as "direct charge-charge interactions". It depends on mutual azimuthal orientation of the molecules which varies over the length of the molecules due to torsional fluctuations, as discussed in the next section.

Because κ_n linearly increases with n at large n and the energy associated with each harmonic n exponentially decreases with increasing κ_n, the sum over n in Eq. 1 rapidly converges. For dsDNA it is usually sufficient to retain only the terms with $n \leq 2$ (7). For tsDNA, however, the term with $n = 3$ is also important. Thus, for consistency, here we retain the terms with $n \leq 3$ for all nucleic acid helices.

Strictly speaking, Eq. 1 was derived for long helices ($L >> R,H$) and it neglects edge effects. Most of the experiments discussed in this study were performed with 25-basepair (bp) or longer oligonucleotides. For 25-bp helices $L/R \approx L/H \sim 2$, so that the predictions of the model based on this equation might be expected to be more qualitative than quantitative, which is important to keep in mind. Note also that the discreteness of the surface charges leads to additional harmonics in the electrostatic pair interaction energy that are not included into Eq. 1 (21,22). Because contribution of these harmonics is negligible for helices that have 10–12 charges per helical turn (7), we do not include them into our calculations.

Nonideality of the helical structure and thermal fluctuations

For ideal helices, the difference in the azimuthal orientations is the same along the entire length of the molecules,

i.e., $\Delta\phi$ is independent of z. In real nucleic acid molecules, however, sequence-related variations and thermal fluctuations in the stacking of basepairs result in variation of $\Delta\phi$ with z. Because the interaction energy depends on $\Delta\phi(z)$, the energetic cost of this variation contributes to the interaction free energy (23). This cost is described by

$$\frac{E_{\Delta\phi}}{k_BT} = \frac{l_p^h}{4}\int_0^L \left(\frac{d[\Delta\phi(z) - \Delta\phi_0(z)]}{dz}\right)^2 dz, \quad (4)$$

where $\Delta\phi_0(z)$ describes intrinsic misalignment of nonidentical helices associated with sequence-related differences in their basepair stacking.

$$l_p^h = \frac{C_tC_s}{\left(C_s + (2\pi/H)^2C_t\right)k_BT} \quad (5)$$

is the helical persistence length of the molecules, and C_t and C_s are the torsional and stretching elasticity moduli, respectively (20). For B-DNA, $k_BT/C_t \approx k_BT(2\pi/H)^2/C_s \approx 1.4 \times 10^5$ cm^{-1} (24,25) and $l_p^h \approx$ 350 Å. Here we consider only interactions between parallel, identical helices, for which $\Delta\phi_0(z) = 0$. This is the case in most experimental studies of oligonucleotide nucleic acids.

Interaction free energy

The free energy of pair interaction between two parallel helical oligonucleotides at an interaxial separation R may thus be calculated as

$$\frac{F(R)}{k_BT} = -\ln\left[\frac{\int D\{\Delta\phi(z)\}\exp\left(-\frac{E_{\text{int}} + E_{\Delta\phi}}{k_BT}\right)}{\int D\{\Delta\phi(z)\}\exp\left(-\frac{E_{\Delta\phi}}{k_BT}\right)}\right], \quad (6)$$

where $D\{\Delta\phi(z)\}$ indicates functional integration over all possible trajectories of $\Delta\phi(z)$ and the reference state for the free energy is at infinite R ($E_{\text{int}} = 0$). Calculation of these functional integrals, which follows the variational trial-function approach proposed in Lee et al. (20), is described in the Supporting Material.

To calculate the free energy of aggregates we utilize the approach proposed in Cherstvy et al. (26), which accounts for both pairwise-additive and nonadditive interactions (see Eqs. S40–S42 in the Supporting Material). In this case, it is also important to account for renormalization of the screening length κ due to increased concentration of free counterions in the aggregate compared to the surrounding solution (see Eq. S41 in the Supporting Material); we do this renormalization here using the scheme suggested in Cherstvy et al. (26).

RESULTS

B-DNA

Fig. 2 illustrates main features of electrostatic interactions between nucleic acid oligomers predicted by our theory, using the example of 50-bp, double-stranded B-DNA. The interaction becomes attractive once a sufficient fraction of phosphate charge is neutralized by counterions bound in the grooves. The attraction results from zipper-like alignment of negatively charged phosphate strands with positively charged counterions bound in grooves of the opposing molecule (Fig. 1), as predicted earlier for long DNA double helices (7,14). Preferential counterion binding in the major groove favors the attraction between B-DNA helices (Fig. 2 *B*). Depending on the extent of phosphate charge neutralization and counterion partitioning between the grooves, the attraction may become strong enough to drive aggregation of the molecules (Fig. 2 *D*).

Thermal fluctuations in the helical pitch (twisting and stretching of the helices) diminish the attraction by disrupting the zipper alignment. For long helices, this effect of

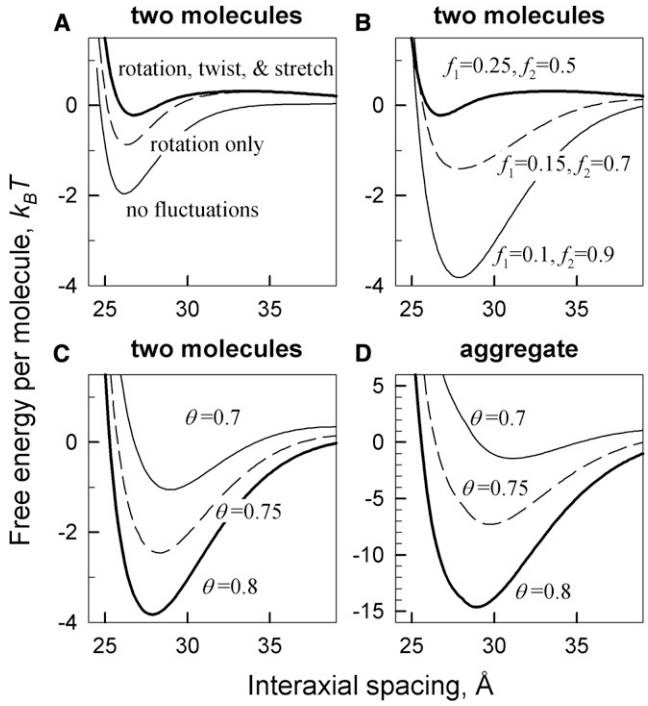

FIGURE 2 Free energy of interaction between 50 bp B-DNA double helices at 7 Å Debye screening length. (*A*) Effects of thermal rotation, twisting, and stretching of the helix on the pair interaction energy ($\theta = 0.8, f_1 = 0.25, f_2 = 0.5$). (*B*) Effects of bound counterions partitioning between the middle of the minor groove (f_1), middle of the major groove (f_2), and random locations ($1–f_1–f_2$) on the pair interaction energy ($\theta = 0.8$). (*C* and *D*) Interaction free energy per molecule in DNA pairs (*C*) and hexagonal, columnar aggregates (*D*) at different neutralization of DNA charge (θ) by bound counterions ($f_1 = 0.1, f_2 = 0.9$). (*B–D*) We accounted for thermal rotation, twisting, and stretching of the helix, using $l_p^h = 350$ Å helical persistence length.

pitch fluctuations is small (7). For oligonucleotides, it may be large (Fig. 2 *A*). Depending on the oligonucleotide length and partitioning of bound counterions, it may completely wipe out the attraction. Because pitch fluctuations affect higher helical harmonics of the surface charge density more than lower harmonics, their effect depends on counterion partitioning between the grooves.

Thermal rotations of oligonucleotides as a whole about their long axes may diminish the attraction even more than pitch fluctuations (Fig. 2 *A*). Unlike pitch fluctuations, they do not involve costly elastic deformations. The cost of rotations is associated only with intermolecular interaction. It is proportional to the length of the molecules and becomes prohibitively high for long DNA molecules, but for shorter oligonucleotides it is comparable to or even smaller than the thermal energy, k_BT.

The conditions for energetically favorable formation of multimolecular aggregates are similar to those at which two molecules begin experiencing net attraction in solution (see Fig. 2, *C* and *D*). Hence, a simpler calculation of the threshold for pairwise attraction in solution is sufficient for qualitative evaluation of the conditions for counterion-induced aggregation. Yet, above this attraction threshold, the energy gain per molecule is dramatically larger in hexagonal, columnar aggregates, suggesting that oligonucleotide pairing should not occur below the aggregation threshold. The underlying reason is that six pairwise interactions result in larger net interaction energy per molecule and stronger suppression of thermal fluctuations in aggregates. This effect is partially counterbalanced by shorter effective screening length, $1/\kappa$, due to counterion accumulation inside aggregates (26). The shorter screening length might significantly weaken attractive electrostatic interactions when a smaller fraction, θ, of DNA charge, is neutralized by bound counterions.

dsRNA

Fig. 3 compares the interaction free energy for a pair of B-DNA oligonucleotides with that for a pair of dsRNA, calculated for oligonucleotide lengths and electrolyte concentrations corresponding to the conditions of experiments reported in Li et al. (16). Based on the observed preferential binding of cobalt-hexamine (27–31), we plotted the predictions of our theory for 90% localization of the bound counterions in the major groove. Overall, B-DNA condensation is strongly favored whereas dsRNA condensation is not, as long as 70% or more counterions bind in the major groove. In contrast, dsRNA condensation is favored whereas B-DNA condensation is not by preferential counterion binding in the minor groove.

The opposite behavior of B-DNA and dsRNA results primarily from the difference in the relative width of their minor and major grooves. The major groove of B-DNA is ~50% wider than the minor one. The major groove of

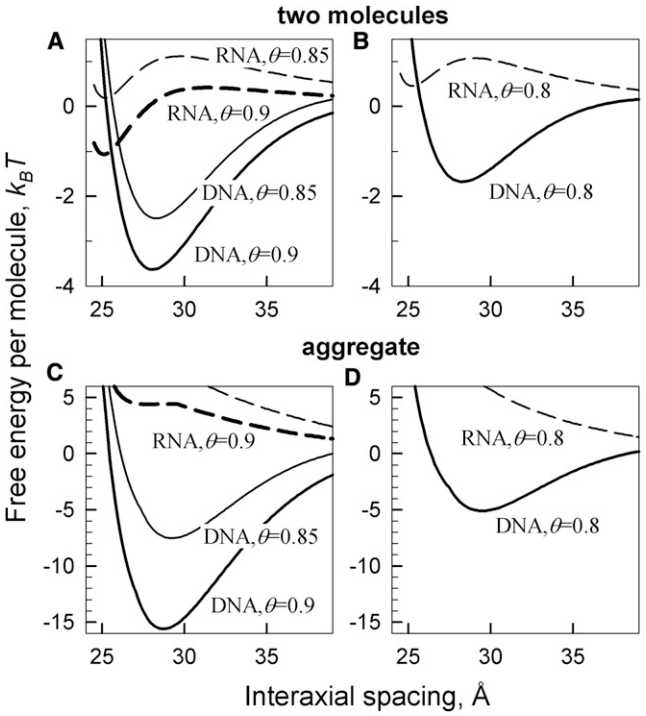

FIGURE 3 Preferential counterion binding in the major groove favors condensation of B-DNA but not double stranded A-RNA. (*A*) Interaction free energy per molecule in pairs of 25-bp A-RNA (*dashed lines*) and 25 bp B-DNA (*solid lines*) in 20 mM NaCl, 5 mM cobalt hexamine (cohex) calculated at 85% ($\theta = 0.85$) and 90% ($\theta = 0.9$) of DNA and RNA charge neutralization by cohex ions preferentially bound in the major groove (90% in the major groove and 10% in the minor groove). (*B*) Interaction free energy per molecule for the same pairs in 100 mM NaCl and 0.8 mM cohex, calculated assuming the same 90/10% partitioning of the ions between the major and minor grooves but lower θ. (*C* and *D*) Free energy per molecule in hexagonal, columnar aggregates of 25-bp A-RNA (*dashed lines*) and 25-bp B-DNA (*solid lines*) at the same conditions as in panels *A* and *B*, respectively.

dsRNA is almost 50% narrower than the minor groove. Preferential localization of bound counterions in the middle of the wider groove results in a larger separation between positive and negative charges along the molecule, which leads to a stronger electrostatic zipper attraction. In addition, such counterion localization favors an azimuthal alignment of the molecules that is more compatible with the symmetry of hexagonal aggregates (7). The combination of these two factors promotes condensation of double helices upon counterion binding in the wider groove and inhibits the condensation upon counterion binding in the narrower groove. Counterion adsorption in the major grove thus induces condensation of double stranded B-DNA, whereas dsRNA resists condensation by such counterions.

tsDNA

Binding of the third oligonucleotide strand in the middle of the major groove of B-DNA, which produces tsDNA

molecules, has two major effects on intermolecular interactions (Fig. 4):

One effect is that the third strand increases the axial charge density by >50%, enhancing counterion condensation/binding to DNA surface. For instance, the value of θ predicted by the Onsawa-Manning theory at low ionic strength is $\theta = 1-l_c/ql_B$, yielding $\theta \approx 0.88$ for B-DNA ($l_c \approx 1.7$ Å) and $\theta \approx 0.92$ for tsDNA ($l_c \approx 1.1$ Å) in the case of divalent counterions ($q = 2$). Stronger counterion condensation/binding promotes intermolecular attraction and aggregation of the molecules.

The other effect is that the third strand nearly equalizes the widths of the grooves (Table 1). In this case, preferential counterion adsorption in any one of the grooves favors intermolecular attraction and similar azimuthal orientation of all molecules, which can be easily accommodated in columnar aggregates.

These two effects might lead, e.g., to condensation of tsDNA by divalent counterions that preferentially bind in the minor groove, as discussed below.

DISCUSSION

From the perspective of experimentally distinguishable predictions, it is useful to categorize proposed condensation mechanisms based on the following concepts of attractive interactions between nucleic acids that might be induced by counterions:

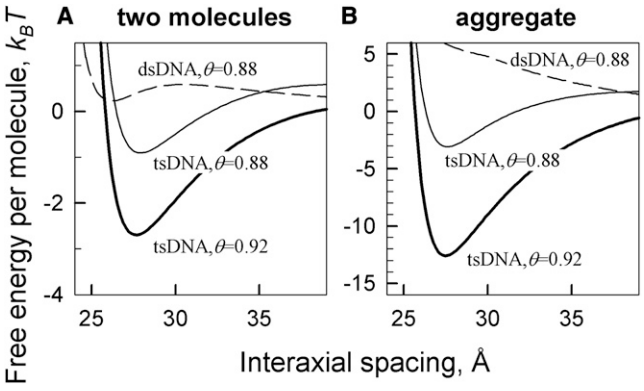

FIGURE 4 Preferential binding of Mg^{2+}, Ba^{2+} and Ca^{2+} in the minor groove promotes condensation of triple-stranded but not double-stranded DNA. (*A*) Free energy of pair interaction per one molecule calculated for 10 mM divalent ion concentrations in 10 mM Tris, 1 mM EDTA, assuming that 60% of bound divalent counterions are located in the middle of the minor groove with the rest located at random sites on DNA surface. The ionic strength of the Tris/EDTA buffer has only a small effect on the interaction energy. The values of θ were calculated from the Onsager-Manning theory for divalent ions, $\theta = 1 - l_c/2l_B$, yielding $\theta \approx 0.88$ for double-stranded DNA ($l_c \approx 1.7$ Å) and $\theta \approx 0.92$ for triple-stranded DNA ($l_c \approx 1.1$ Å). The curve for tsDNA calculated at the value of θ expected for dsDNA is shown to demonstrate the relative contribution of the phosphate charge pattern in tsDNA condensation, but for practical purposes only the dashed and bold lines should be compared. (*B*) Free energy per molecule in hexagonal, columnar aggregates calculated at the same conditions as in panel *A*.

Bridging by counterions that are bound to two or more nucleic acid molecules (10) or form a condensed, liquid-like layer shared by these molecules (32). This interaction is most susceptible to changes in counterion dimensions. It is likely to be dominant in the case of very large counterions, such as poly(amido amine) dendrimers (33).

Counterion-correlation attraction between juxtaposed negatively charged phosphates and positively charged bound counterions that are spatially separated on the surface of each molecule due to repulsion between the counterions (12,13) or fluctuations in their density (11,34,35). This attraction originates from and depends strongly on interactions and resulting correlations between bound counterions that move freely along nucleic acid surfaces. Theoretically, it is expected to be dominant for point-like multivalent counterions that do not have strong preferential binding sites (7,36).

Electrostatic zipper attraction between juxtaposed helical phosphate strands and counterions in nucleic acid grooves (14). This attraction requires efficient zipper-like alignment of the strands with opposing grooves. It depends strongly on the nucleic acid structure and is expected to be dominant in the case of significant preferential counterion binding in nucleic acid grooves (7).

Our previous analysis of the published data suggested that preferential counterion binding in grooves might be the most prominent observed feature, likely responsible for differences in the ability of similarly charged but chemically or structurally different counterions to induce DNA condensation (7). Comparison of calculated and measured forces between B-DNA molecules supported the dominant role of this mechanism in DNA condensation by many counterions (20). Yet, interpretations of the same experiments based on the other two mechanisms could not be excluded (2–4,7).

Below we discuss studies of oligomeric dsDNA, tsDNA, and dsRNA as well as studies of counterion charge, structure, and chemistry effects on condensation of longer dsDNA, which are more recent and provide more data for discriminating between predictions of different theories. In our analysis, these studies show that zipper-like juxtaposition of phosphate strands with groove counterions is a major factor in nucleic acid condensation, although they do not resolve whether electrostatic or hydration forces, for which the predictions are similar (7,37), cause the zipper attraction.

Interaction between B-DNA oligonucleotides linked into pairs by flexible linkers at counterion concentrations below DNA condensation threshold

The authors of this study investigated effects of spermidine^{3+} on interactions between double helices within linked pairs of 12-bp oligonucleotides and linked pairs of

80-bp oligonucleotides (18). They observed aggregation of the pairs above a threshold spermidine^{3+} concentration, but no significant association between the helices within each pair below the threshold. For 12-bp oligonucleotides, they estimated that the lack of any observable association below the aggregation threshold means that the attraction between the helices within each pair must be $<$0.1–0.2 k_BT/bp. Bai et al. (18) hypothesized that "more than two neighboring helices are required to generate a pronounced attraction" needed for DNA aggregation observed in their own and multiple other studies.

Our predictions for the free energy per molecule in a pair (Fig. 2, *B* and *C*) and columnar aggregate (Fig. 2 *D*) provide a theoretical foundation for this idea, revealing the nature of complex many-body effects in aggregates. The aggregation occurs once the energy benefit per molecule exceeds several k_BT. Because the attractive energy per molecule is several times lower in a pair than in an aggregate (Fig. 2, *C* and *D*), the predicted pairwise attraction below the aggregation threshold is <1 k_BT per total length of the double helices (50 bp in the case illustrated in Fig. 2). This energy is consistent with the maximum attractive energy estimated in Fig. 6 of Bai et al. (18) and it is not sufficient to overcome the entropy loss upon the association of either 12- or 80-bp oligonucleotide pairs connected by flexible linkers.

Our calculations predict stable pairing below or near the aggregation threshold only for much longer DNA molecules, which are capable of forming braids that optimize chiral electrostatic interactions (38). For oligomeric molecules, stable association requires multiple pairwise attractions per molecule in columnar assemblies. Our calculations show that the cumulative effect of these attractions overcomes the entropy loss and the osmotic pressure of free counterions that make the aggregate electroneutral, provided that at least 80–90% of the DNA's negative charge is compensated by bound counterions (as required for the condensation (1)).

Resistance of dsRNA to condensation at cobalt-hexamine^{3+} concentrations that condense dsDNA

This observation cannot be explained by theories of electrostatic interactions that do not account for the structure of nucleic acids and approximate them as featureless, homogeneously charged cylinders (16). In such theories, nucleic acid condensation propensity is determined by counterions, which are expected to exhibit stronger binding and correlations at surfaces of more charged molecules (see, e.g., Kornyshev et al. (7) and Grosberg et al. (36) and references therein). Both axial and surface densities of phosphate charge are higher in dsRNA than double-stranded B-DNA (Table 1), and bound counterions do appear to screen dsRNA charge more efficiently than B-DNA (39). Therefore, the observation of the resistance of dsRNA to condensation by cobalt-hexamine^{3+} prompted its authors to seek an alternative explanation. In particular, they suggested that this resistance might be related to cobalt-hexamine^{3+} binding inside the narrow cleft in the major groove of dsRNA, so that the counterions are buried deep inside the core of the molecules and cannot effectively balance the surface charge of phosphates (16). However, in view of the analysis presented above, there is no longer a need in this assumption.

The depth of cobalt-hexamine^{3+} location in the major groove may affect interactions between dsRNAs, but our calculations suggest an immediate, simpler explanation for the differences in the condensation propensities of dsRNA and B-DNA. This explanation is based on assuming that the stereochemistry of cobalt-hexamine^{3+} interaction with nucleic acid bases leads to its preferential binding in the major groove of both dsRNA, as reported in Davis et al. (27), Kieft and Tinoco (28), and Rüdisser and Tinoco (29), and B-DNA, as reported in Ouameur and Tajmir-Riahi (30) and Robinson and Wang (31). The difference in the condensation of these molecules by cobalt-hexamine^{3+} is then caused by the difference in the relative widths of their major grooves (Fig. 3). The condensation of oligomers (as well as longer nucleic acids) is promoted by counterion binding in the wider groove, which enhances the electrostatic zipper attraction. In B-DNA, the major groove is wider than the minor one. In dsRNA, the major groove is narrower than the minor one. Less favorable electrostatic zipper interaction upon cobalt-hexamine^{3+} binding in the narrower groove rationalizes why dsRNA resists the condensation (Fig. 3).

Condensation of tsDNA by Mg^{2+}, Ba^{2+}, and Ca^{2+} ions that do not condense dsDNA

This observation might be explained by several different contributing factors, as noted by its authors (17). Our calculations point to two such factors, which might contribute almost equally (Fig. 4):

The first factor is higher charge of tsDNA leads to stronger binding of divalent ions to its surface, weakening the repulsion associated with the net charge (charge of the molecules plus bound counterions) and enhancing the electrostatic zipper attraction associated with helical patterning of this charge. Enhancement of the zipper attraction is a straightforward consequence of more bound counterions, which are responsible for this force. The reason behind weakening of the repulsion between higher charged molecules is less intuitive. Onsager-Manning theory predicts counterion binding that ensures the same net axial charge density (phosphates plus bound counterions) regardless of the phosphate density (40). The same axial density means that the surface density of the net charge should be lower on tsDNA compared to dsDNA, because the former has a larger radius. Lower surface charge density leads to weaker electrostatic repulsion at the same surface-to-surface separation between the molecules.

The second factor is geometry of the three tsDNA grooves, which have similar yet not identical widths (Table 1). Counterion binding in the middle of one of these grooves is likely to provide more favorable salt-bridge interactions with phosphates, with which alkaline earth metal ions form poorly soluble salts. Provided that these ions do bind preferentially in the minor groove of B-DNA (41) and in one of the three grooves of tsDNA, we expect stronger electrostatic zipper attraction and condensation of tsDNA but not dsDNA oligonucleotides (Fig. 4). Interestingly, putrescine^{2+} ions do not condense either tsDNA (17) or B-DNA (42). Within our theory, this might be explained by insufficient preferential groove partitioning of these organic ions, which have very different chemical natures, structures, and interactions with phosphates than alkaline earth metal ions. Yet, we agree with the tsDNA study authors that further experiments are needed to unambiguously resolve the condensation mechanism.

We should note that bridging by divalent alkaline earth metal ions and side-by-side aggregation of 10-bp dsDNA oligonucleotides was predicted by computer simulations in Luan and Aksimentiev (43). However, experimental studies provide evidence only for end-to-end attraction and stacking of dsDNA oligonucleotides in solutions of these ions (44–46).

Condensation of short versus long nucleic acids

The only distinction of oligonucleotides in our theory is more pronounced independent thermal rotation of each molecule about its axis compared to longer nucleic acid molecules. Such rotations affect the zipper-like alignment of phosphates and counterions, altering the requirements for counterion-induced condensation. In the case of longer nucleic acids, they are suppressed by prohibitively large energetic cost associated with the loss of electrostatic zipper attraction.

Because the cost of thermal rotations depends not only on the length of the molecules but also on counterion binding and partitioning, the resulting differences in predictions for shorter and longer molecules are significant in some but not all cases. For instance, our calculations predict weak zipper attraction upon preferential counterion binding in the narrow groove, resulting in low energetic cost and a strong effect of thermal rotations on interactions between short double helices. The rotations further suppress the attraction, making 25-bp dsRNA molecules resistant to condensation by cobalt hexamine (Fig. 3). In contrast, preferential counterion binding in the wider groove strengthens the zipper attraction, increasing the cost and suppressing thermal rotations. We expect interactions between 25-bp B-DNA double helices and interactions between much longer B-DNA molecules to be similar in the presence of such counterions, e.g., consistent with observations for B-DNA condensation by spermine and spermidine (8,47,48).

It is also important to note that interactions between molecules shorter or comparable in length with the helical pitch (~10 or fewer basepairs or triplets) might be affected by edge effects not incorporated into the theory. Nevertheless, our theory might work qualitatively even for these molecules, because of their tendency toward end-to-end stacking (44–46,49).

Effects of counterion charge, chemistry, and structure on condensation of long dsDNA molecules

Preferential counterion binding seems to explain not only how structures of different nucleic acid molecules affect their condensation but also how counterion structure and chemistry affect condensation of the same dsDNA molecules. In particular, we previously argued that preferential binding in the major groove might be responsible for the ability of divalent transition (versus alkaline earth) metal ions and some (but not all) polyamine ions to induce dsDNA condensation (7). A number of more recent observations lend further support to this hypothesis.

In particular, different effects of various chiral isoforms of methylated spermidine^{3+} and spermine^{4+} on DNA condensation (50) are difficult to explain without assuming preferential adsorption of these counterions at specific sites. Similar critical fraction of charge neutralization required for DNA condensation by ε-oligolysines with different length and charge (51) is consistent with the electrostatic zipper attraction, assuming similar partitioning of these ions between DNA grooves. It is not consistent with the other attraction mechanisms, because ε-oligolysines that vary from 3 to 31 residue length and from 4+ to 31+ charge are unlikely to exhibit similar counterion correlations and ability to bridge adjacent DNA molecules. Forces measured between DNA molecules condensed by polyarginines are also more consistent with predictions for zipper-like alignment and resulting electrostatic or hydration attraction between phosphate strands and counterions bound in the major groove (52,53).

Effects of intermolecular interaction on counterion binding pattern

Because preferential counterion adsorption significantly affects intermolecular interactions, the interactions might affect the counterion binding pattern as well (7,26). Here, we do not explicitly describe the latter effect, because we expect the interaction energy to be insufficient for overcoming the cost of relocating the counterions away from their preferential binding sites in most cases of nucleic acid condensation into hydrated aggregates. For instance, our calculations predict more favorable interactions between oligonucleotide double helices in hydrated aggregates when counterions are located in the wider

groove (Fig. 3). Yet, they show that the interaction energy benefit of relocating to the wider minor groove of dsRNA is $<1\ k_BT$ per counterion and likely insufficient for overcoming preferential cobalt-hexamine binding in the major groove.

Counterions with weak or no preference for binding in a specific groove might relocate upon nucleic acid aggregation, e.g., such relocation might contribute to condensation of B-DNA by Mn^{2+} at elevated temperatures (26). However, detailed analysis of such counterion relocation is more complicated and beyond the scope of this study.

More importantly, intermolecular interactions might significantly affect counterion binding locations in nucleic acid crystals, in which these interactions are much stronger and favor different counterion binding patterns. In particular, our calculations show that interactions between very closely packed and immobilized B-DNA molecules in crystals are optimized by counterion binding in the narrower rather than wider groove, in contrast to hydrated aggregates. Smaller separation between the negative and positive charges upon narrow groove binding produces a shorter-range electric field that leads to more favorable electrostatic interactions in crystals. The corresponding energy benefit per ion is particularly large when longer polypeptides with multiple positively charged residues bind in the minor rather than major groove of B-DNA. It might explain, e.g., why arginine-containing polypeptides expected to bind in the major groove of fully hydrated DNA (52,53) are found in the minor groove of deformed DNA in nucleosome crystals (54).

Hybrid condensation mechanisms

X-ray structures of nucleosome crystals also illustrate that different mechanisms of counterion-mediated attraction between nucleic acids might sometimes represent different facets of the same interaction. These structures reveal zipper-like alignment of positively charged histone tails bound in the minor groove and phosphate strands on the opposing double helix surface (55,56). Yet, DNA packing in nucleosome crystals is so close that it is difficult to distinguish whether histone tails bind in the groove between two phosphate strands of one double helix and attract phosphate strands on the other helix or whether they bridge all strands together.

Overall, separation of distinct nucleic acid condensation mechanisms is certainly useful for conceptual clarity. However, depending on specific circumstances, several mechanisms might operate at the same time and even blend together. In another example, point-like 3+ or higher valence ions might mediate attraction by binding along helical strands of phosphate charges in a manner that blends together counterion-correlation attraction and zipper-like alignment of charges on opposing surfaces, as suggested by recent calculations (57–59).

CONCLUSIONS

1. Because of inaccuracies inherent in any theory, simulation, or measurement of mesoscopic phenomena, comparison of conceptual predictions and observations in a wider variety of nucleic acid condensation studies is more revealing than parsing of finer details of any specific calculation or experiment.
2. The common theme emerging from experimental studies is that condensation efficiency depends strongly on the nucleic acid structure and preferential counterion binding at specific sites, which are the distinguishing features of the electrostatic zipper attraction.
3. The concept of electrostatic zipper attraction appears to explain a wide variety of observations, including the most puzzling features of oligonucleotide condensation that are difficult to interpret otherwise.
4. The major role of this interaction in nucleic acid condensation does not exclude contributions from bridging by counterions or counterion-correlation attraction, which might become significant under some circumstances.

Revisiting the concept of electrostatic zipper attraction between helical molecules was motivated by the revival of interest in counterion-induced condensation of different nucleic acids, particularly oligomeric nucleic acids whose interactions can be better characterized experimentally and computationally. We believe that the statistical mechanical theory of interaction and aggregation of such molecules developed in this study provides a useful basis for interpreting the results of such studies.

SUPPORTING MATERIAL

Three figures and 57 equations, and reference (60) are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(13)00372-X.

The authors thank Dominic Lee, Adrian Parsegian, Lois Pollack, and Don Rau for useful discussions.

This work was supported by funding from the Engineering and Physical Sciences Research Council, UK (grant No. EP/H004319/1) and Human Frontiers Science Foundation (grant No. RGP0049/2010-C102) to A.A.K. as well as funding from the Intramural Research Program of the National Institute of Child Health and Human Development, National Institutes of Health, to S.L.

REFERENCES

1. Bloomfield, V. A. 1997. DNA condensation by multivalent cations. *Biopolymers.* 44:269–282.
2. Carrivain, P., A. Cournac, …, M. Barbi. 2012. Electrostatics of DNA compaction in viruses, bacteria and eukaryotes: functional insights and evolutionary perspective. *Soft Matter.* 8:9285–9301.
3. Teif, V. B., and K. Bohinc. 2011. Condensed DNA: condensing the concepts. *Prog. Biophys. Mol. Biol.* 105:208–222.
4. Wong, G. C. L., and L. Pollack. 2010. Electrostatics of strongly charged biological polymers: ion-mediated interactions and self-organization in nucleic acids and proteins. *Annu. Rev. Phys. Chem.* 61:171–189.

5. Cherstvy, A. G. 2011. Electrostatic interactions in biological DNA-related systems. *Phys. Chem. Chem. Phys.* 13:9942–9968.

6. Kornyshev, A. A. 2010. Physics of DNA: unraveling hidden abilities encoded in the structure of 'the most important molecule'. *Phys. Chem. Chem. Phys.* 12:12352–12378.

7. Kornyshev, A. A., D. J. Lee, …, A. Wynveen. 2007. Structure and interactions of biological helices. *Rev. Mod. Phys.* 79:943–996.

8. Rau, D. C., and V. A. Parsegian. 1992. Direct measurement of the intermolecular forces between counterion-condensed DNA double helices. Evidence for long range attractive hydration forces. *Biophys. J.* 61:246–259.

9. Rau, D. C., and V. A. Parsegian. 1992. Direct measurement of temperature-dependent solvation forces between DNA double helices. *Biophys. J.* 61:260–271.

10. Allison, S. A., J. C. Herr, and J. M. Schurr. 1981. Structure of viral ϕ29 DNA condensed by simple triamines: a light-scattering and electron-microscopy study. *Biopolymers.* 20:469–488.

11. Marquet, R., and C. Houssier. 1991. Thermodynamics of cation-induced DNA condensation. *J. Biomol. Struct. Dyn.* 9:159–167.

12. Rouzina, I., and V. A. Bloomfield. 1996. Macroion attraction due to electrostatic correlation between screening counterions. 1. Mobile surface-adsorbed ions and diffuse ion cloud. *J. Phys. Chem. US.* 100:9977–9989.

13. Shklovskii, B. I. 1999. Wigner crystal model of counterion induced bundle formation of rodlike polyelectrolytes. *Phys. Rev. Lett.* 82:3268–3271.

14. Kornyshev, A. A., and S. Leikin. 1999. Electrostatic zipper motif for DNA aggregation. *Phys. Rev. Lett.* 82:4138–4141.

15. Seeman, N. C. 2010. Nanomaterials based on DNA. *Annu. Rev. Biochem.* 79:65–87.

16. Li, L., S. A. Pabit, …, L. Pollack. 2011. Double-stranded RNA resists condensation. *Phys. Rev. Lett.* 106:108101.

17. Qiu, X., V. A. Parsegian, and D. C. Rau. 2010. Divalent counterion-induced condensation of triple-strand DNA. *Proc. Natl. Acad. Sci. USA.* 107:21482–21486.

18. Bai, Y., R. Das, …, S. Doniach. 2005. Probing counterion modulated repulsion and attraction between nucleic acid duplexes in solution. *Proc. Natl. Acad. Sci. USA.* 102:1035–1040.

19. Baker, A. 2011. Simplicity. *In* The Stanford Encyclopedia of Philosophy. E. N. Zalta, editor. Stanford University Press, Stanford, CA.

20. Lee, D. J., A. Wynveen, …, S. Leikin. 2010. Undulations enhance the effect of helical structure on DNA interactions. *J. Phys. Chem. B.* 114:11668–11680.

21. Kornyshev, A. A., and S. Leikin. 1998. Symmetry laws for interaction between helical macromolecules. *Biophys. J.* 75:2513–2519.

22. Kornyshev, A. A., and S. Leikin. 1998. Electrostatic interaction between helical macromolecules in dense aggregates: an impetus for DNA poly- and mesomorphism. *Proc. Natl. Acad. Sci. USA.* 95:13579–13584.

23. Cherstvy, A. G., A. A. Kornyshev, and S. Leikin. 2004. Torsional deformation of double helix in interaction and aggregation of DNA. *J. Phys. Chem. B.* 108:6508–6518.

24. Crothers, D. M., J. Drak, …, S. D. Levene. 1992. DNA bending, flexibility, and helical repeat by cyclization kinetics. *Methods Enzymol.* 212:3–29.

25. Wenner, J. R., M. C. Williams, …, V. A. Bloomfield. 2002. Salt dependence of the elasticity and overstretching transition of single DNA molecules. *Biophys. J.* 82:3160–3169.

26. Cherstvy, A. G., A. A. Kornyshev, and S. Leikin. 2002. Temperature-dependent DNA condensation triggered by rearrangement of adsorbed cations. *J. Phys. Chem. B.* 106:13362–13369.

27. Davis, J. H., T. R. Foster, …, S. E. Butcher. 2007. Role of metal ions in the tetraloop-receptor complex as analyzed by NMR. *RNA.* 13:76–86.

28. Kieft, J. S., and I. Tinoco, Jr. 1997. Solution structure of a metal-binding site in the major groove of RNA complexed with cobalt (III) hexamine. *Structure.* 5:713–721.

29. Rüdisser, S., and I. Tinoco, Jr. 2000. Solution structure of Cobalt(III) hexamine complexed to the GAAA tetraloop, and metal-ion binding to G•A mismatches. *J. Mol. Biol.* 295:1211–1223.

30. Ouameur, A. A., and H. A. Tajmir-Riahi. 2004. Structural analysis of DNA interactions with biogenic polyamines and cobalt(III)hexamine studied by Fourier transform infrared and capillary electrophoresis. *J. Biol. Chem.* 279:42041–42054.

31. Robinson, H., and A. H. Wang. 1996. Neomycin, spermine and hexamine cobalt (III) share common structural motifs in converting B- to A-DNA. *Nucleic Acids Res.* 24:676–682.

32. Ray, J., and G. S. Manning. 1994. An attractive force between two rodlike polyions mediated by the sharing of condensed counterions. *Langmuir.* 10:2450–2461.

33. Ainalem, M. L., A. M. Carnerup, …, K. Schillen. 2009. Condensing DNA with poly(amido amine) dendrimers of different generations: means of controlling aggregate morphology. *Soft Matter.* 5:2310–2320.

34. Arenzon, J. J., J. F. Stilck, and Y. Levin. 1999. Simple model for attraction between like-charged polyions. *Eur. Phys. J. B.* 12:79–82.

35. Ha, B. Y., and A. J. Liu. 1999. Counterion-mediated, non-pairwise-additive attractions in bundles of like-charged rods. *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics.* 60:803–813.

36. Grosberg, A. Y., T. T. Nguyen, and B. I. Shklovskii. 2002. Colloquium: The physics of charge inversion in chemical and biological systems. *Rev. Mod. Phys.* 74:329–345.

37. Kornyshev, A. A., and S. Leikin. 1997. Theory of interaction between helical molecules. *J. Chem. Phys.* 107:3656–3674.

38. Cortini, R., A. A. Kornyshev, …, S. Leikin. 2011. Electrostatic braiding and homologous pairing of DNA double helices. *Biophys. J.* 101:875–884.

39. Pabit, S. A., X. Qiu, …, L. Pollack. 2009. Both helix topology and counterion distribution contribute to the more effective charge screening in dsRNA compared with dsDNA. *Nucleic Acids Res.* 37:3887–3896.

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

41. Sines, C. C., L. McFail-Isom, …, L. D. Williams. 2000. Cations mediate B-DNA conformational heterogeneity. *J. Am. Chem. Soc.* 122:11048–11056.

42. Gosule, L. C., and J. A. Schellman. 1978. DNA condensation with polyamines. I. Spectroscopic studies. *J. Mol. Biol.* 121:311–326.

43. Luan, B., and A. Aksimentiev. 2008. DNA attraction in monovalent and divalent electrolytes. *J. Am. Chem. Soc.* 130:15754–15755.

44. Qiu, X., L. W. Kwok, …, L. Pollack. 2006. Measuring inter-DNA potentials in solution. *Phys. Rev. Lett.* 96:138101.

45. Li, L., S. A. Pabit, …, L. Pollack. 2008. Closing the lid on DNA end-to-end stacking interactions. *Appl. Phys. Lett.* 92:223901.

46. De Michele, C., L. Rovigatti, …, F. Sciortino. 2012. Self-assembly of short DNA duplexes: from a coarse-grained model to experiments through a theoretical link. *Soft Matter.* 8:8388–8398.

47. Qiu, X., K. Andresen, …, L. Pollack. 2008. Abrupt transition from a free, repulsive to a condensed, attractive DNA phase, induced by multivalent polyamine cations. *Phys. Rev. Lett.* 101:228101.

48. Raspaud, E., D. Durand, and F. Livolant. 2005. Interhelical spacing in liquid crystalline spermine and spermidine-DNA precipitates. *Biophys. J.* 88:392–403.

49. Zanchetta, G., F. Giavazzi, …, T. Bellini. 2010. Right-handed double-helix ultrashort DNA yields chiral nematic phases with both right- and left-handed director twist. *Proc. Natl. Acad. Sci. USA.* 107:17497–17502.

50. Nayvelt, I., M. T. Hyvönen, …, T. J. Thomas. 2010. DNA condensation by chiral alpha-methylated polyamine analogues and protection of cellular DNA from oxidative damage. *Biomacromolecules.* 11:97–105.

51. Korolev, N., N. V. Berezhnoy, …, L. Nordenskiöld. 2012. A universal description for the experimental behavior of salt-(in)dependent oligocation-induced DNA condensation. *Nucleic Acids Res.* 40:2808–2821.

52. DeRouchey, J., V. A. Parsegian, and D. C. Rau. 2010. Cation charge dependence of the forces driving DNA assembly. *Biophys. J.* 99:2608–2615.

53. DeRouchey, J. E., and D. C. Rau. 2011. Role of amino acid insertions on intermolecular forces between arginine peptide condensed DNA helices: implications for protamine-DNA packaging in sperm. *J. Biol. Chem.* 286:41985–41992.

54. West, S. M., R. Rohs, …, B. Honig. 2010. Electrostatic interactions between arginines and the minor groove in the nucleosome. *J. Biomol. Struct. Dyn.* 27:861–866.

55. Luger, K., A. W. Mäder, …, T. J. Richmond. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*. 389:251–260.

56. Schalch, T., S. Duda, …, T. J. Richmond. 2005. X-ray structure of a tetranucleosome and its implications for the chromatin fiber. *Nature.* 436:138–141.

57. Kanduc, M., J. Dobnikar, and R. Podgornik. 2009. Counterion-mediated electrostatic interactions between helical molecules. *Soft Matter.* 5:868–877.

58. Lee, D. J. 2010. Charge renormalization of helical macromolecules. *J. Phys. Condens. Mat.* 22:414101.

59. Lee, D. J. 2011. Correlation forces between helical macro-ions in the weak coupling limit. *J. Phys. Condens. Mat.* 23:105102.

60. Wynveen, A., D. J. Lee, and A. A. Kornyshev. 2005. Statistical mechanics of columnar DNA assemblies. *Eur. Phys. J. E. Soft Matter.* 16:303–318.

SUPPORTING MATERIAL:

Helical structure determines different susceptibilities of ds-DNA, ds-RNA, and ts-DNA to counterion-induced condensation

Alexei A. Kornyshev[1], **and Sergey Leikin**[2,*]

[1]Deparment of Chemistry, Imperial College London, SW7 2AZ, London, UK and [2]Section of Physical Biochemistry, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892

Methods

This Supporting Material provides derivations and complete expressions for all equations used in the present study. It is structured as follows. In section I, we provide previously published expressions for pair interaction potentials and derive coefficients for the simplified Eq. 1 of the main text. In section II, we derive the free energy of pair interaction between two parallel oligomeric helices that takes into account helical pitch fluctuations and rotations of the helices about their axes. In section III, we calculate the same free energy using a variational approximation for the rotations and demonstrate that this approximation works almost as well as the exact calculation of the partition function. In section IV, we utilize the variational approximation to derive expressions for the free energy of hexagonal, columnar aggregates. In section V, we discuss average azimuthal alignment of molecules in pairs and in aggregates.

I. Pair interaction potentials

I.a. *Asymptotic expansion at large interaxial distances.*

For two long, parallel helices, the electrostatic interaction energy may be represented by a sum over helical harmonics n (1)

$$E_{\text{int}} = k_B T \int_0^L dz \sum_{n=0}^{\infty} \left\{ (-1)^n \tilde{u}_n(R) \cos\left(n\left[\Delta\phi(z)\right]\right) + \tilde{u}_n^{im}(R) \right\} \tag{S1}$$

where L is the length of the helices, $\Delta\phi(z)$ is the difference between the azimuthal orientations of the helices at point z along their length, R is the interaxial distance between them,

$$\tilde{u}_n(R) = \frac{2l_B}{l_c^2}\left(2-\delta_{n,0}\right)\zeta_n^2 \frac{K_0(\kappa_n R)}{(\kappa_n a)^2\left[K_n'(\kappa_n a)\right]^2}, \tag{S2}$$

$$\tilde{u}_n^{im}(R) = \frac{2l_B}{l_c^2}\left(2-\delta_{n,0}\right)\zeta_n^2 \frac{\Omega_{n,n}(\kappa_n R, \kappa_n a)}{(\kappa_n a)^2\left[K_n'(\kappa_n a)\right]^2}, \tag{S3}$$

l_B is the Bjerrum length (≈ 7 Å at room temperature), $1/l_c$ is the linear density of the fixed charged groups on the molecular surface ($l_c \approx 1.7$ Å for *B*-DNA), ζ_n are normalized helical moments of the surface charge density, a is the radius of the helices, κ_n is the reciprocal screening length for the helical harmonic n,

$$\kappa_n = \sqrt{\kappa^2 + \left(\frac{2\pi n}{H}\right)^2}\ , \tag{S4}$$

$K_n(x)$ is the modified Bessel function of the second kind (sometimes referred to as the Macdonald function); $K_n'(x) \equiv dK_n(x)/dx$,

$$\Omega_{n,m}(x,y) = -\sum_{j=-\infty}^{\infty} \left[K_{n-j}(x)\right]\left[K_{j-m}(x)\right]\frac{I_j'(y)}{K_j'(y)}, \tag{S5}$$

is the Ω-function introduced in (2), $I_n(y)$ is the modified Bessel function of the first kind, and $I_n'(x) \equiv dI_n(x)/dx$.

At κR>>1, we may utilize the following asymptotic expansion

$$K_n(\kappa_n R) \approx \sqrt{\frac{\pi}{2\kappa_n R}}\exp(-\kappa_n R), \tag{S6}$$

where we took into account that $\kappa_0=\kappa$ and $\kappa_n>\kappa$. After substitution of Eq. S6 into Eqs. S2 and S5, we may rewrite Eq. S1 in the form presented in the main text as Eq. 1, where the coefficients c_n and c_n^{im} are given by

$$c_n = \frac{\sqrt{\pi/2}}{(\kappa_n a)^2\left[K_n'(\kappa_n a)\right]^2} \tag{S7}$$

and

$$c_n^{im} = \frac{\pi}{2(\kappa_n a)^2\left[K_n'(\kappa_n a)\right]^2}\sum_{j=-\infty}^{\infty}\frac{I_j'(\kappa_n a)}{-K_j'(\kappa_n a)}. \tag{S8}$$

Note, however, that Eqs. S1-S5 are just as convenient for practical calculations. Because the latter equations are more accurate at smaller R, we have utilized them in all of our calculations. The simplified, less accurate Eq. 1 is provided in the main text for illustrating the exponential asymptotic dependence of the pair interaction energy for each helical harmonic on $\kappa_n R$. Because $\kappa_n \approx 2\pi n/H$ at large n, this exponential dependence leads to very rapid convergence of the sums in Eqs. 1 and S1. For nucleic acids, it is usually sufficient to account only for the terms with $n \le 2$. Higher helical harmonics become relevant only for special surface charge patterns, for which $\zeta_1 \approx 0$ and $\zeta_2 \approx 0$.

I.b. *Helical moments of the surface charge density.*

Normalized helical moments of the surface density, ζ_n can be calculated from known distribution of fixed charges and bound counterions as described in (1). Here, we utilize a simplified, qualitative model proposed in (3, 4). Specifically, we treat all charges as point-like. We assume that all fixed charges are located along helical lines (strands, two in dsDNA and dsRNA or three in tsDNA) on the surface of cylindrical dielectric core of each molecule. The fraction θ of the fixed charge is compensated by bound

counterions that are located either in the middle of the space (groove) between the helical strands of fixed charges or randomly. For a nucleic acid helix that has N similar strands of fixed phosphate charges, which are separated by grooves of different widths, this model yields

$$\zeta_n = \left| \theta\delta_{n,0} + \theta\sum_{j=1}^{N} f_j \left[\exp(in\varphi_j) - \delta_{n,0} \right] - \frac{1}{N}\sum_{j=1}^{N} \exp\left[in\left(\varphi_j + \frac{\omega_j}{2} \right) \right] \right|, \tag{S9}$$

where φ_j is the azimuthal coordinate of the middle of groove j in the ground state relative to a selected reference point (e.g., middle of the minor groove in double stranded helices), f_j is the fraction of bound counterions located in the middle of this groove, ω_j is the azimuthal width of this groove, and $|z|$ indicates modulus of the complex number z. For double helices, this expression may be rewritten in the more simple form of Eq. 2. Benefits and shortcomings of this model were discussed in (1).

II. Two parallel helices

Consider interaction between two straight, parallel oligonucleotide helices, which have identical sequences and are separated by the interaxial distance R. To account for thermal twisting, stretching, and rotations of these helices, we parameterize fluctuations in the azimuthal orientation $\phi_\mu(z)$ of the helix μ (=1,2) by introducing the helical phase $\Phi_\mu(z)$,

$$\Phi_\mu(z) \equiv \phi_\mu(z) - \bar{g}z \equiv \Psi_\mu + \int_0^z \left(\Omega_\mu(\tau) - \bar{g}h_\mu(\tau) \right) \frac{d\tau}{\langle h \rangle} \equiv \Psi_\mu + \delta\Phi_\mu^0(z) + \delta\Phi_\mu(z) \ . \tag{S10}$$

Here $\Psi_\mu \equiv \phi_\mu(0)$ is the azimuthal orientation of the helix at z=0; $\Omega_\mu(z)$ and $h_\mu(z)$ are the twist and axial rise per base pair (triplet in ts-DNA), respectively; $\bar{g} = \langle \Omega_\mu \rangle / \langle h_\mu \rangle$ is the average reciprocal pitch of the helix ($2\pi/H$); and $\langle\Omega_\mu\rangle$ and $\langle h_\mu\rangle$ are the average twist and rise per base pair/triplet. In an ideal, rigid helix, Ω_μ and h_μ are constant and $\Phi_\mu(z) \equiv \Psi_\mu$. In nucleic acid helices, which are not ideal, deviations from this value have two principal components: $\delta\Phi_\mu^0(z)$ describes intrinsic, sequence-related variations in the stacking of base pairs/triplets and $\delta\Phi_\mu(z)$ describe deformations caused by thermal stretching and twisting.

The cost of stretching and twisting may be approximated within the elastic rod model as (5)

$$E_\mu^{EL} = k_B T \frac{l_p^h}{2} \int_0^L \left(\frac{d\left[\delta\Phi_\mu(z) \right]}{dz} \right)^2 dz \ , \tag{S11}$$

where l_p^h is the helical persistence length (Eq. 5 of the main text). Since we are interested in calculating the interaction free energy and the interaction energy, E_{int} depends only on $\Delta\phi(z) \equiv \phi_1(z) - \phi_2(z)$, for molecules with identical sequences ($\delta\Phi_1^0(z) = \delta\Phi_2^0(z)$) we may use that

$$\Delta\phi(z) = \Psi_1 - \Psi_2 + \delta\Phi_1(z) - \delta\Phi_2(z) \equiv \Delta\Psi + \Delta\delta\Phi(z) \ , \tag{S12}$$

and represent the total elastic cost of helical pitch fluctuations of the two molecules as

$$E^{EL} = E_1^{EL} + E_2^{EL} = E_{\Delta\phi} + E_{\Sigma\Phi} \, , \tag{S13}$$

where

$$E_{\Delta\phi} = k_B T \frac{l_p^h}{4} \int_0^L \left(\frac{d\left[\Delta\delta\Phi(z)\right]}{dz} \right)^2 dz = k_B T \frac{l_p^h}{4} \int_0^L \left(\frac{d\left[\Delta\phi(z)\right]}{dz} \right)^2 dz \tag{S14}$$

determines the elastic cost of fluctuations in $\Delta\phi$ (c.f., Eq. 4 of the main text), while

$$E_{\Sigma\Phi} = k_B T \frac{l_p^h}{4} \int_0^L \left(\frac{d\left[\delta\Phi_1(z) + \delta\Phi_2(z)\right]}{dz} \right)^2 dz \tag{S15}$$

is the elastic cost of fluctuations that preserve $\Delta\phi$. The latter fluctuations may be omitted from our analysis, since they do not alter intermolecular interactions.

Based on this parameterization of thermal twisting and stretching, the interaction free energy may be calculated from

$$F(R) = -k_B T \ln \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \exp\left(-\frac{F_{\Delta\Psi}(R)}{k_B T} \right) d\Delta\Psi \right] \tag{S16}$$

where

$$F_{\Delta\Psi}(R) = -k_B T \ln \left[\frac{\int D\left\{\Delta\delta\Phi(z)\right\} \exp\left(-\frac{E_{\text{int}}\left\{R, \Delta\Psi + \Delta\delta\Phi(z)\right\} + E_{\Delta\phi}\left\{\Delta\delta\Phi(z)\right\}}{k_B T} \right)}{\int D\left\{\Delta\delta\Phi(z)\right\} \exp\left(-\frac{E_{\Delta\phi}\left\{\Delta\delta\Phi(z)\right\}}{k_B T} \right)} \right] . \tag{S17}$$

is the interaction free energy at fixed $\Delta\Psi$ and $D\left\{\Delta\delta\Phi(z)\right\}$ indicates integration over all possible "trajectories" of $\Delta\delta\Phi(z)$.

$F_{\Delta\Psi}(R)$ may be calculated within the following variational ansatz (5):

$$F_{\Delta\Psi}(R) \approx F_\beta\left(\beta_{\min}, R, \Delta\Psi\right) , \tag{S18}$$

where

$$F_\beta\left(\beta, R, \Delta\Psi\right) = C_\beta - k_B T \ln Z_{eff} + \left\langle E_{\text{int}} + E_{\Delta\phi} - E_{eff} \right\rangle_{eff} , \tag{S19}$$

$$Z_{eff} = \int \exp\left(-\frac{E_{eff}}{k_B T} \right) D\left\{\Delta\delta\Phi(z)\right\} \tag{S20}$$

$$E_{eff} = E_{\Delta\phi} + k_B T \frac{\beta}{2L} \int_0^L \left[\Delta\delta\Phi(z)\right]^2 dz , \tag{S21}$$

$$\left\langle E\right\rangle_{eff}=\frac{\int E\exp\left(-\frac{E_{eff}}{k_BT}\right)D\left\{\Delta\delta\Phi\left(z\right)\right\}}{Z_{eff}}\,, \tag{S22}$$

$\beta_{min}\equiv\beta_{min}(R,\Delta\Psi)$ is the value of the variational parameter β at which $F_\beta(\beta,R,\Delta\Psi)$ is minimal, and the "integration" constant C_β ensures that $F_\beta(\beta_{min},R\to\infty,\Delta\Psi)\to 0$.

After substitution of Eq. S21 into Eqs. S19,S20,S22 and straightforward calculation of the path integrals, we find

$$\frac{F_\beta\left(\beta,R,\Delta\Psi\right)}{k_BT}=-\ln\left[\frac{\sqrt{\beta L/2l_p^h}}{\sinh\left(\sqrt{\beta L/2l_p^h}\right)}\right]-\frac{1}{2}\left[\sqrt{\frac{\beta L}{2l_p^h}}\coth\left(\sqrt{\frac{\beta L}{2l_p^h}}\right)-1\right]$$
$$+L\sum_{n=0}^{\infty}\tilde{u}_n^{im}\left(R\right)+L\sum_{n=0}^{\infty}\left(-1\right)^n\tilde{u}_n\left(R\right)\cos\left(n\Delta\Psi\right)\exp\left(-\frac{n^2}{2\beta}\left[\sqrt{\frac{\beta L}{2l_p^h}}\coth\left(\sqrt{\frac{\beta L}{2l_p^h}}\right)-1\right]\right) \tag{S23}$$

Minimization of this $F_\beta(\beta,R,\Delta\Psi)$ yields the following equation for β_{min}

$$\beta_{min}=L\sum_{n=1}^{\infty}\left(-1\right)^{n+1}\tilde{u}_n\left(R\right)n^2\cos\left(n\Delta\Psi\right)\exp\left(-\frac{n^2}{2\beta_{min}}\left[\sqrt{\frac{\beta_{min}L}{2l_p^h}}\coth\left(\sqrt{\frac{\beta_{min}L}{2l_p^h}}\right)-1\right]\right)\,, \tag{S24}$$

which can be easily solved numerically due to the rapid convergence of the sum over n. Substitution of this β_{min} into Eq. S23 and subsequently into Eqs. S18 and S16 allows us to calculate the interaction free energy at any R and $\Delta\Psi$.

Note that the variational approach used here extends beyond the Gaussian limit by utilizing a trial function, the form of which corresponds to Gaussian fluctuations but the amplitude is found by minimizing the full, non-Gaussian free energy. Thus, it is important to keep in mind that this approach may lose accuracy when the fluctuations become too large.

III. Variational approximation for thermal rotations of molecules as a whole

Consider now an alternative approximation for the free energy of pairwise interaction, which does not require numerical integration over the whole range of possible thermal rotations, $\Delta\Psi$. The reason for introducing this approximation is to lay the groundwork for analysis of columnar aggregates, in which the dependence of E_{int} on $\Delta\Psi$ for each pair of neighbor molecules leads to interdependence of all pairwise $\Delta\Psi$, making direct integration over all $\Delta\Psi$ impractical. To avoid direct integration over $\Delta\Psi$, we rely on the following two assumptions: (i) At small R, thermal rotations of the molecules about their axes are suppressed by intermolecular interactions, justifying a variational approximation for small to moderate rotational fluctuations. (ii) Once the rotational fluctuations become large at larger R and the variational approximation fails, these fluctuations should reduce the contribution of $\Delta\phi(z)$-dependent terms to $\sim 1\ k_BT$ or less, in which case inaccuracies in the calculation of this contribution become irrelevant in the context of the present study. To test the validity of this approach, we use it to calculate the same free energy of pairwise interactions as above but within a variational approximation for thermal rotations.

Within the latter approximation, we represent $\Delta\Psi$ as a sum of its average value $\Delta\Psi_0$ and small fluctuations $\Delta\delta\Psi$ near the average, i.e. we replace Eq. S12 with

$$\Delta\phi(z) = \Delta\Psi_0 + \Delta\delta\Psi + \Delta\delta\Phi(z). \tag{S25}$$

We also replace Eqs. S16, S18 with

$$F(R) \approx F_{\gamma,\beta}(\gamma_{\min}, \beta_{\min}, R, \Delta\Psi_0), \tag{S26}$$

where we introduce an additional variational parameter γ, which describes rotational fluctuations in the same way as β describes twisting and stretching. Specifically,

$$F_{\gamma,\beta}(\gamma,\beta,R,\Delta\Psi_0) = k_B T C_{\gamma,\beta} - k_B T \ln Z_{\gamma,\beta} + \left\langle E_{\text{int}} + E_{\Delta\phi} - E_{\gamma,\beta} \right\rangle_{\gamma,\beta}\ ; \tag{S27}$$

$$Z_{\gamma,\beta} = \frac{1}{2\pi}\int_{-\pi}^{\pi} d(\Delta\delta\Psi) \int \exp\left(-\frac{E_{\gamma,\beta}}{k_B T}\right) D\{\Delta\delta\Phi(z)\}\ ; \tag{S28}$$

$$E_{\gamma,\beta} = E_{\Delta\phi} + k_B T \frac{\gamma}{2}(\Delta\delta\Psi)^2 + k_B T \frac{\beta}{2L}\int_0^L [\Delta\delta\Phi(z)]^2\, dz\ ; \tag{S29}$$

the $\langle\ \rangle_{\gamma,\beta}$ averaging of energies E (E_{int}, $E_{\Delta\phi}$, or $E_{\gamma,\beta}$) is defined as

$$\left\langle E \right\rangle_{\gamma,\beta} = \frac{\frac{1}{2\pi}\int_{-\pi}^{\pi} d(\Delta\delta\Psi) \int E \exp\left(-\frac{E_{\gamma,\beta}}{k_B T}\right) D\{\Delta\delta\Phi(z)\}}{Z_{\gamma,\beta}}\ ; \tag{S30}$$

$\gamma_{\min}\equiv\gamma_{\min}(R)$ and $\beta_{\min}\equiv\beta_{\min}(R)$ are the values of the variational parameters γ and β at which $F_{\gamma,\beta}(\gamma,\beta,R,\Delta\Psi_0)$ is minimal; and the "integration" constant $C_{\gamma,\beta}$ ensures that $F_{\gamma,\beta}(\gamma_{\min},\beta_{\min},R\to\infty,\Delta\Psi_0)\to 0$. The value of $\Delta\Psi_0$ is determined by minimization of $F_{\gamma,\beta}(\gamma_{\min},\beta_{\min},R,\Delta\Psi_0)$. Note that Eq. S29 defines the effective energy $E_{\gamma,\beta}$ for $-\pi < \Delta\delta\Psi < \pi$. Unlike Eqs. S16-S24, here $-\pi/2$ is not equivalent to $3\pi/2$; and the integration in Eq. S28 has to be performed from $-\pi$ to π and not from 0 to 2π.

After calculating the integrals, we arrive at

$$\begin{aligned}
&\frac{F_{\gamma,\beta}}{k_B T} = L\sum_{n=0}^{\infty} \tilde{u}_n^{im}(R) \\
&+ C_{\gamma,\beta} - \ln\left(\frac{1}{\sqrt{2\pi\gamma}}\operatorname{erf}\left(\pi\sqrt{\gamma/2}\right)\right) - \frac{1}{2}\left(1 - \frac{\sqrt{2\pi\gamma}\exp\left(-\pi^2\gamma/2\right)}{\operatorname{erf}\left(\pi\sqrt{\gamma/2}\right)}\right) \\
&- \ln\left[\frac{\sqrt{\beta L/2l_p^h}}{\sinh\left(\sqrt{\beta L/2l_p^h}\right)}\right] - \frac{1}{2}\left[\sqrt{\frac{\beta L}{2l_p^h}}\coth\left(\sqrt{\frac{\beta L}{2l_p^h}}\right) - 1\right] \\
&+ L\sum_{n=0}^{\infty}(-1)^n \tilde{u}_n(R)\cos(n\Delta\Psi_0)\exp\left(-\frac{n^2}{2\beta}\left[\sqrt{\frac{\beta L}{2l_p^h}}\coth\left(\sqrt{\frac{\beta L}{2l_p^h}}\right) - 1\right] - \frac{n^2}{2\gamma}\left[1 - \frac{\sqrt{2\pi\gamma}\exp\left(-\pi^2\gamma/2\right)}{\operatorname{erf}\left(\pi\sqrt{\gamma/2}\right)}\right]\right)
\end{aligned} \tag{S31}$$

Minimization with respect to γ yields

$$\gamma = \beta \ . \tag{S32}$$

This is a general property of the model, which follows directly from Eqs. S25-S29 and can be derived without calculating the integrals. After substituting Eq. S32 into S31 and neglecting the terms proportional to $\exp(-\pi^2\gamma/2)$, because the whole approximation is accurate only at $\gamma \sim 1$ or larger, we find

$$\begin{aligned}\frac{F_{\gamma,\beta}\left(\gamma_{\min},\beta,R,\Delta\Phi_0\right)}{k_B T} \approx C_{\gamma,\beta} - \ln\left[\frac{\sqrt{L/l_p^h}\,\mathrm{erf}\left(\pi\sqrt{\beta/2}\right)}{2\sqrt{\pi}\sinh\left(\sqrt{\beta L/2l_p^h}\right)}\right] - \frac{1}{2}\sqrt{\frac{\beta L}{2l_p^h}}\coth\left(\sqrt{\frac{\beta L}{2l_p^h}}\right) \\ + L\sum_{n=0}^{\infty}\tilde{u}_n^{im}\left(R\right) + L\sum_{n=0}^{\infty}(-1)^n\tilde{u}_n\left(R\right)\cos\left(n\Delta\Psi_0\right)\exp\left[-\frac{n^2}{2\beta}\sqrt{\frac{\beta L}{2l_p^h}}\coth\left(\sqrt{\frac{\beta L}{2l_p^h}}\right)\right]\end{aligned} \tag{S33}$$

Minimization of this free energy leads to the following equations for $\beta_{\min}$ and $\Delta\Phi_0$:

$$\beta_{\min} = L\sum_{n=1}^{\infty}(-1)^{n+1}\tilde{u}_n\left(R\right)n^2\cos\left(n\Delta\Psi_0\right)\exp\left(-\frac{n^2}{2\beta_{\min}}\sqrt{\frac{\beta_{\min}L}{2l_p^h}}\coth\left(\sqrt{\frac{\beta_{\min}L}{2l_p^h}}\right)\right) \tag{S34}$$

and

$$\sum_{n=1}^{\infty}(-1)^{n+1}n\tilde{u}_n\left(R\right)\sin\left(n\Delta\Psi_0\right)\exp\left[-\frac{n^2}{2\beta_{\min}}\sqrt{\frac{\beta_{\min}L}{2l_p^h}}\coth\left(\sqrt{\frac{\beta_{\min}L}{2l_p^h}}\right)\right] = 0 \ . \tag{S35}$$

While Eq. S34 has to be solved numerically, Eq. S35 may be solved analytically by omitting a negligibly small contribution of the terms with $n \geq 3$. We then find that the free energy is minimal at either

$$\Delta\Phi_0 = 0 \tag{S36}$$

or

$$\cos\left(\Delta\Phi_0\right) = \frac{\tilde{u}_1\left(R\right)}{4\tilde{u}_2\left(R\right)}\exp\left[\frac{3}{2\beta_{\min}}\sqrt{\frac{\beta_{\min}L}{2l_p^h}}\coth\left(\sqrt{\frac{\beta_{\min}L}{2l_p^h}}\right)\right] \ . \tag{S37}$$

The first solution exists at any R. The second one exists only at R smaller than a critical value, at which the right hand side of Eq. S37 is equal to 1. In the latter range of R, the optimal value of $\Delta\Phi_0$ can be found by comparing the free energies calculated for both of these solutions.

Together, Eqs. S26 and S33-S37 define the interaction free energy at small to moderate rotational fluctuations, i.e. at $\gamma_{\min} = 2\beta_{\min} \sim 1$ or larger. Because the amplitude of the fluctuations increases and $\beta_{\min}$ decreases with increasing R, this approximation cannot be used at large R. We are interested primarily in small R, at which this approximation is expected to work. Nevertheless, this limitation prevents us from determining $C_{\gamma,\beta}$ in Eq. S33 by requiring $F_{\gamma,\beta}\left(\gamma_{\min},\beta_{\min},R\to 0,\Delta\Phi_0\right)\to 0$, because Eq. S34 loses nonzero solutions for $\beta_{\min}$ at $R > R_*$ (Fig. S1).

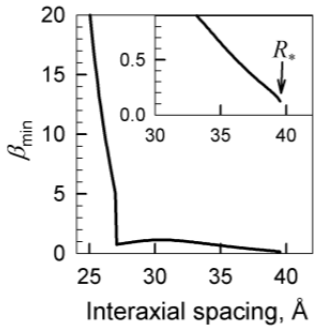

Figure S1. Variational parameter β_{min} for interaction between two 50 bp *B*-DNA double helices calculated from Eqs. S34-S37 at 7 Å Debye screening length, θ=0.8, f_1=0.1, f_2=0.9. At $R > R_*$, Eqs. S34-S37 have no nonzero solutions for β_{min}, indicating that our approximation cannot be used in this range of R. At smaller R, Eqs. S34-S37 have up to four nonzero solutions. The bold line shows the values of β_{min} that produce the lowest free energy upon substitution into Eq. S33.

To determine $C_{\gamma,\beta}$, we construct the following interpolation. We assume that $F(R)$ is described by Eqs. S26, S33-S37 with nonzero β_{min} all the way up to $R = R_*$. We use

$$F\left(R > R_*\right) \approx k_B T L \sum_{n=0}^{\infty} \tilde{u}_n^{im}\left(R\right) + k_B T L \tilde{u}_0\left(R\right) , \tag{S38}$$

which is the expected asymptotic behavior at very large R, at which rotational fluctuations completely wipe out the contribution of $\Delta\phi$-dependent interactions. We determine $C_{\gamma,\beta}$ by assuming the continuity of $F(R)$ at $R=R_*$, yielding

$$\begin{aligned} C_{\gamma,\beta} = & \ln\left[\frac{\sqrt{L/l_p^h}\,\mathrm{erf}\left(\pi\sqrt{\beta_{\min}^* /2}\right)}{2\sqrt{\pi}\,\sinh\left(\sqrt{\beta_{\min}^* L/2l_p^h}\right)}\right] + \frac{1}{2}\sqrt{\frac{\beta_{\min}^* L}{2l_p^h}}\coth\left(\sqrt{\frac{\beta_{\min}^* L}{2l_p^h}}\right) \\ & -L\sum_{n=0}^{\infty}(-1)^n \tilde{u}_n\left(R_*\right)\cos\left(n\Delta\Psi_0^*\right)\exp\left[-\frac{n^2}{2\beta_{\min}^*}\sqrt{\frac{\beta_{\min}^* L}{2l_p^h}}\coth\left(\sqrt{\frac{\beta_{\min}^* L}{2l_p^h}}\right)\right] \end{aligned} , \tag{S39}$$

where $\beta_{\min}^*$ and $\Delta\Psi_0^*$ are the solutions of Eqs. S34-S37 at $R=R_*$ calculated numerically.

Because we interpolate across the region of intermediate rotational fluctuations, in which the contribution of $\Delta\phi$-dependent interactions should be ~ 1 k_BT, we expect ~1 k_BT accuracy of this approximation. Comparison with the more accurate, direct integration over thermal rotations described in the preceding section confirms that this is indeed the case (Fig. S2). Such accuracy is sufficient for the purpose of the present study. In other words, the simplified variational approximation for thermal rotations works sufficiently well to justify its utilization for calculating the interaction free energy in hexagonal, columnar aggregates, in which case we do not have the option of direct, numerical integration over the rotations.

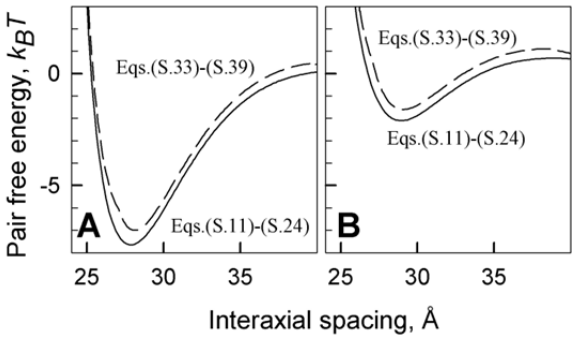

Figure S2. Comparison of the variational approximation, Eqs. S33-S39, with the partition function calculation , Eqs. S16-S24, for thermal rotations. The calculations were performed for 50 bp *B*-DNA double helices at 7 Å Debye screening length, f_1=0.1, f_2=0.9, and θ= 0.8 (**A**) or θ= 0.7 (**B**).

IV. Hexagonal, columnar aggregates

IV.a. *Electrostatic energy*

The electrostatic energy of a hexagonal, columnar aggregate may be approximated as (6)

$$E_{hex} \approx NL\sqrt{3}\int_R^\infty \left(\sum_i n_i^0 \left[\exp\left(-\frac{q_i e\varphi_s}{k_B T} \right) - 1 \right] \right) R' dR' + \frac{k_B T}{2} \sum_{\mu=1}^{N} \sum_{\nu=1}^{6} \sum_{n=1}^{\infty} \int_0^L dz \left[(-1)^n \tilde{u}_n(R) \cos\left(n\left[\Phi_\mu(z) - \Phi_\nu(z) \right] \right) + \tilde{u}_n^{im}(R) \right] \quad \text{(S40)}$$

Here N is the number of the molecules in the aggregate; n_i^0 are the concentrations of different ions (labeled by index i) in the electrolyte solution outside the aggregate; q_i are the charge numbers for these ions (e.g., q_i=2 for a 2+ ion and q_i=-1 for a 1- ion); e is the elementary charge; φ_s is the average electrostatic potential at the surface of the Wigner-Seitz cell around each molecule relative to the electrolyte solution outside the aggregate; index ν labels six nearest neighbors of each molecule μ; R is the interaxial distance between the nearest neighbors; $\Phi_\mu(z)$ is the helical phase of molecule μ at the axial coordinate z; $\tilde{u}_n(R)$ and $\tilde{u}_n^{im}(R)$ are defined by Eqs. S2-S5, in which κ_n is still defined by Eq. 3 but with the following renormalized value of κ,

$$\kappa = \sqrt{4\pi l_B \left(\sum_i n_i^0 q_i^2 \exp\left(\frac{-q_i e\varphi_s}{k_B T} \right) \right)} \quad . \quad \text{(S41)}$$

The value of φ_s in turn depends on κ, as given by the following equation (1, 6)

$$\frac{2l_B(1-\theta)}{l_c \kappa a} \frac{I_0(\kappa R_s)K_1(\kappa R_s) + K_0(\kappa R_s)I_1(\kappa R_s)}{K_1(\kappa a)I_1(\kappa R_s) - I_1(\kappa a)K_1(\kappa R_s)} = \left[\sum_i n_i^0 q_i \exp\left(\frac{-q_i e\varphi_s}{k_B T} \right) \Big/ \sum_i n_i^0 q_i^2 \exp\left(\frac{-q_i e\varphi_s}{k_B T} \right) \right] , \quad \text{(S42)}$$

where $R_s = R\sqrt{\sqrt{3}/2\pi}$.

The first term in Eq. S40 accounts for the electrostatic energy associated with the average net charge of the molecules (fixed charges plus bound counterions), which is not pairwise additive. The second term accounts for the energy associated with helical patterning of discrete charges, which can be calculated by pairwise summation of the corresponding helical harmonics for all nearest neighbor molecules. Renormalization of κ takes into account the Donnan equilibrium, which drives counterions into the interior of the aggregate from the bulk, effectively increasing the ionic strength inside the aggregate relative to its value outside. Thus, the effective value of κ increases with the increase of the aggregate density. For derivation of Eqs. S40-S42 and detailed discussion of the corresponding approximations, see (1, 6).

IV.b. *Free energy*

To calculate the aggregate free energy, we utilize an approximation that is similar to the one discussed in the preceding section, except for one important distinction. The average difference between the helical phases of two nearest neighbor molecules depends also on orientation of other molecules in the aggregate, necessitating a slightly different approach to calculation of the average helical phases and fluctuations in them. Specifically, as in the preceding sections (c.f. Eqs. S10,S25), we represent the helical phase of each molecule as

$$\Phi_\mu(z) = \Psi_\mu^0 + \delta\Psi_\mu + \delta\Phi_\mu^0(z) + \delta\Phi_\mu(z)\ , \tag{S43}$$

where Ψ_μ^0 and $\delta\Psi_\mu$ are the average value and thermal fluctuations of Ψ_μ, respectively, and $\delta\Phi_\mu(z)$ and $\delta\Phi_\mu(z)$ have exactly the same meaning as in Eq. S10. We approximate the aggregate free energy with

$$F_{hex}(R) \approx F_{\gamma,\beta}^{hex}\left(\gamma_{\min}, \beta_{\min}, R, \Psi_\mu^0\right), \tag{S44}$$

where

$$F_{\gamma,\beta}^{hex}\left(\gamma,\beta,R,\bar{\Phi}_\mu\right) = k_B T N C_{\gamma,\beta}^{hex} - k_B T \ln Z_{\gamma,\beta}^{hex} + \left\langle E_{hex} + \sum_{\mu=1}^{N} E_\mu^{ER} - E_{\gamma,\beta}^{hex} \right\rangle_{\gamma,\beta}^{hex} \tag{S45}$$

$$Z_{\gamma,\beta}^{hex} = \frac{1}{(2\pi)^N}\int_{-\pi}^{\pi}\prod_{\mu=1}^{N} d\left(\delta\Psi_\mu\right)\int\prod_{\mu=1}^{N} D\left\{\delta\Phi_\mu(z)\right\}\exp\left(-\frac{E_{\gamma,\beta}^{hex}}{k_B T}\right) \tag{S46}$$

$$E_{\gamma,\beta}^{hex} = \sum_{\mu=1}^{N} E_\mu^{ER} + k_B T\frac{\gamma}{2}\sum_{\mu=1}^{N}\left(\delta\Psi_\mu\right)^2 + k_B T\frac{\beta}{2L}\sum_{\mu=1}^{N}\int_0^L\left[\delta\Phi_\mu(z)\right]^2 dz\ , \tag{S47}$$

$$\left\langle E\right\rangle_{\gamma,\beta}^{hex} = \frac{\frac{1}{(2\pi)^N}\int_{-\pi}^{\pi}\prod_{\mu=1}^{N} d\left(\delta\hat{\Phi}_\mu\right)\int\prod_{\mu=1}^{N} D\left\{\delta\Phi_\mu(z)\right\}\left[E\exp\left(-\frac{E_{\gamma,\beta}^{hex}}{k_B T}\right)\right]}{Z_{\gamma,\beta}^{hex}}\ , \tag{S48}$$

where the meaning of the variational parameters γ and β as well as the integration constant $C_{\gamma,\beta}^{hex}$ remains the same as in the preceding section. Note that Eq. S47 presumes all thermal rotations and helical pitch fluctuations to be independent of each other.

Calculation of the integrals and minimization with respect to γ and β may be performed as described in the preceding section. Minimization with respect to Ψ_{μ}^{0} may be performed as described in (1, 5, 7). As described in the latter studies, the optimal average azimuthal orientations $\Psi_{\mu 1}^{0}$, $\Psi_{\mu 2}^{0}$, and $\Psi_{\mu 3}^{0}$ of any three nearest neighbor molecules that form the elementary triangular cell of the hexagonal lattice are related to each other as

$$\Psi_{\mu 1}^{0} - \Psi_{\mu 2}^{0} = \Psi_{\mu 3}^{0} - \Psi_{\mu 1}^{0} = \Psi_{p}, \quad \Psi_{\mu 3}^{0} - \Psi_{\mu 2}^{0} = 2\Psi_{p}. \tag{S49}$$

The free energy is then given by

$$\begin{aligned}\frac{F_{hex}(R)}{Nk_BT} \approx C_{\gamma,\beta}^{hex} - \ln\left[\frac{\sqrt{L/l_p^h}\,\mathrm{erf}\left(\pi\sqrt{\beta_{\min}/2}\right)}{2\sqrt{2\pi}\sinh\left(\sqrt{\beta_{\min}L/4l_p^h}\right)}\right] - \frac{1}{2}\sqrt{\frac{\beta_{\min}L}{4l_p^h}}\coth\left(\sqrt{\frac{\beta_{\min}L}{4l_p^h}}\right) \\ + L\sqrt{3}\int_R^{\infty}\left(\sum_i n_i^0\left[\exp\left(-\frac{q_i e\varphi_s}{k_BT}\right)-1\right]\right)R'dR' + 3L\sum_{n=1}^{\infty}\tilde{u}_n^{im}(R) \\ + 3L\sum_{n=1}^{\infty}(-1)^n\tilde{u}_n(R)\sigma\left(n\Psi_p\right)\exp\left[-\frac{n^2}{\beta_{\min}}\sqrt{\frac{\beta_{\min}L}{4l_p^h}}\coth\left(\sqrt{\frac{\beta_{\min}L}{4l_p^h}}\right)\right]\end{aligned} \tag{S50}$$

where

$$\sigma(x) = \frac{2\cos(x)+\cos(2x)}{3}, \tag{S51}$$

and $\beta_{\min}$ and Ψ_p can be found as the roots of

$$\beta_{\min} = 3L\sum_{n=1}^{\infty}(-1)^{n+1}\tilde{u}_n(R)n^2\sigma\left(n\Psi_p\right)\exp\left[-\frac{n^2}{\beta_{\min}}\sqrt{\frac{\beta_{\min}L}{4l_p^h}}\coth\left(\sqrt{\frac{\beta_{\min}L}{4l_p^h}}\right)\right] \tag{S52}$$

and

$$\sum_{n=1}^{\infty}(-1)^{n+1}n\tilde{u}_n(R)\left[\sin\left(n\Psi_p\right)+\sin\left(2n\Psi_p\right)\right]\exp\left[-\frac{n^2}{\beta_{\min}}\sqrt{\frac{\beta_{\min}L}{4l_p^h}}\coth\left(\sqrt{\frac{\beta_{\min}L}{4l_p^h}}\right)\right] = 0 \tag{S53}$$

that minimize $F_{hex}(R)$.

As in the preceding section, solutions of Eq. S53 that minimize the free energy may be found by neglecting $n \geq 3$ terms in the sum, yielding

$$\Psi_p = 0 \tag{S54}$$

or

$$\cos\left(\Psi_p\right)=\frac{1}{4}\left(1+\sqrt{1+\frac{2\tilde{u}_1\left(R\right)}{\tilde{u}_2\left(R\right)}\exp\left[\frac{3}{\beta_{\min}}\sqrt{\frac{\beta_{\min}L}{4l_p^h}}\coth\left(\sqrt{\frac{\beta_{\min}L}{4l_p^h}}\right)\right]}\right). \tag{S55}$$

at R smaller than the critical value, at which the right hand side of Eq. S55 is equal to 1. Ψ_p=0 is optimal at R larger than the latter critical value. The optimal Ψ_p at smaller R can be found by comparing the free energies for Ψ_p=0 and Ψ_p given by Eq. S55. Other roots of Eq. S53, e.g. $\cos(\Psi_p) \approx -1/2$, describe solutions with larger free energies.

As in the previous section, the integration constant C_{hex} may be found by interpolating between the free energy at small R, which is given by Eqs. S50-S55 and the limiting form of the free energy at large R, yielding

$$\begin{aligned} C_{\gamma,\beta}^{hex} \approx & \ln\left[\frac{\sqrt{L/l_p^h}\,\mathrm{erf}\left(\pi\sqrt{\beta_{\min}^*/2}\right)}{2\sqrt{2\pi}\sinh\left(\sqrt{\beta_{\min}^*L/4l_p^h}\right)}\right]+\frac{1}{2}\sqrt{\frac{\beta_{\min}^*L}{4l_p^h}}\coth\left(\sqrt{\frac{\beta_{\min}^*L}{4l_p^h}}\right) \\ & +3L\sum_{n=1}^{\infty}(-1)^{n+1}\tilde{u}_n\left(R_*^{hex}\right)\sigma\left(n\Psi_p^*\right)\exp\left[-\frac{n^2}{\beta_{\min}^*}\sqrt{\frac{\beta_{\min}^*L}{4l_p^h}}\coth\left(\sqrt{\frac{\beta_{\min}^*L}{4l_p^h}}\right)\right] \end{aligned}, \tag{S56}$$

where R_*^{hex} is the maximum value of R, at which Eqs. S52,S53 have nonzero solutions for $\beta_{\min}$, $\beta_{\min}^* = \beta_{\min}\left(R_*^{hex}\right)$, and $\Psi_p^* = \Psi_p\left(R_*^{hex}\right)$.

Eqs. S50-S56 define the approximation for the aggregate free energy used for the calculations in the main text. Another, more complicated statistical mechanical approach to calculating aggregate free energy was described in (8).

V. Azimuthal alignment of molecules in pairs and aggregates

The azimuthal alignment of molecules and their rotational fluctuations in pairs and aggregates are illustrated in Fig. S3, using the example of 50 bp *B*-DNA. The color map in Fig S3A shows the probability density for the difference in azimuthal orientation of two parallel molecules, $\Delta\Psi$, which is defined by

$$p\left(\Delta\Psi\right)=\frac{2\pi\exp\left(-F_{\Delta\Psi}\left(R\right)/k_BT\right)}{\int_0^{2\pi}\exp\left(-F_{\Delta\Psi}\left(R\right)/k_BT\right)d\Delta\Psi}. \tag{S57}$$

The yellow line shows the most probable $\Delta\Psi$ ($\Delta\Psi_{\min}$), for which $F_{\Delta\Psi}(R)$ is minimal at the given R. The magenta line shows the optimal alignment $\Delta\Psi_0(R)$ calculated within the variational model. Fig S3B shows root mean square deviations from $\Delta\Psi_{\min}$ and $\Delta\Psi_0$ calculated by direct integration and within the variational model, respectively. Note that $p(\Delta\Psi)=p(-\Delta\Psi)$.

In the case of two parallel molecules, the variational model slightly overestimates fluctuations in $\Delta\Psi$, but otherwise provides a reasonably accurate description of the alignment. At small R, $\pm\Delta\Psi_0$ represents

the most probable alignment ± $\Delta\Psi_{min}$. With increasing R, $\Delta\Psi_{min}$ decreases until it becomes zero at a critical value of R. Near this critical point, the probabilities of all $\Delta\Psi$ between $-\Delta\Psi_{min}$ and $\Delta\Psi_{min}$ are approximately equal, which is why the variational model predicts a sharp transition of the optimal average alignment to $\Delta\Psi_0$=0 before $\Delta\Psi_{min}$ becomes zero. Beyond the critical R, $\Delta\Psi_0$=$\Delta\Psi_{min}$=0.

In hexagonal, columnar aggregates, simultaneous optimization of all pairwise alignments at small R is impossible (7). The most energetically favorable alignment of the molecules in such aggregates is illustrated in the diagram under Fig 3C. Another feature of the alignment is the suppression of fluctuations in the alignment by interactions of each molecule with multiple nearest neighbors (vs. just one neighbor in a pair).

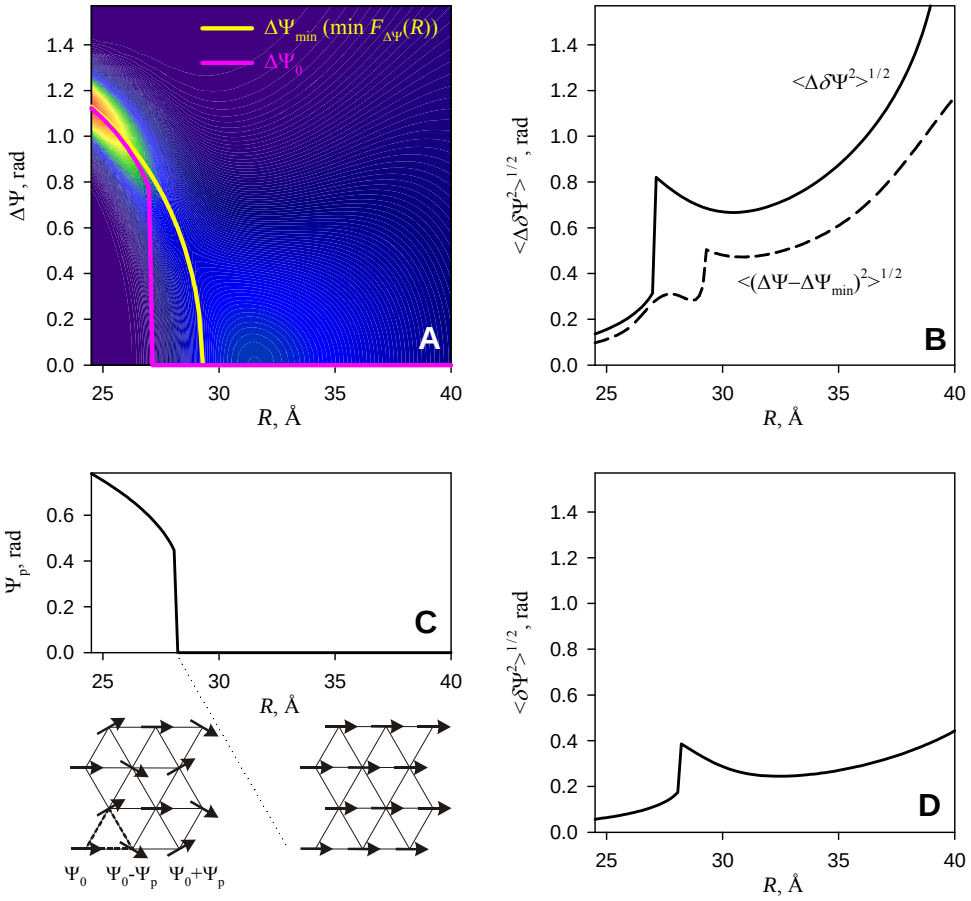

Figure S3. Azimuthal alignment of 50 bp *B*-DNA molecules in pairs (**A, B**) and aggregates (**C, D**) at θ= 0.8, f_1=0.1, f_2=0.9, and 7 Å Debye screening length. **A**. The rainbow color map represents the probability density for $\Delta\Psi(R)$ at fixed R calculated from Eq. S57, in which the violet color depicts the least and red color the most probable states; the yellow line shows the most probable $\Delta\Psi$ at a given R, for which $F_{\Delta\Psi}$ is minimal; the magenta line shows optimal $\Delta\Psi_0$ for the variational model. **B**. Root mean square fluctuations of $\Delta\Psi$ within the model defined by Eqs. S16-S24 (dashed line) and root mean square $\Delta\delta\Psi$ within the variational model defined by Eqs. S25-S39 (solid line). **C**, **D** Alignment (**C**) and azimuthal fluctuations of molecules (**D**) in hexagonal, columnar aggregates calculated within the variational model. Diagrams under the panel C illustrate optimal azimuthal orientations within each triangle of nearest neighbors above and below critical R, at which Ψ_p becomes zero.

Supporting References

1. Kornyshev, A. A., D. J. Lee, S. Leikin, and A. Wynveen. 2007. Structure and interactions of biological helices. Rev Mod Phys 79:943-996.
2. Kornyshev, A. A., and S. Leikin. 1997. Theory of interaction between helical molecules. J Chem Phys 107:3656-3674.
3. Kornyshev, A. A., and S. Leikin. 1998. Electrostatic interaction between helical macromolecules in dense aggregates: an impetus for DNA poly- and meso-morphism. Proc Natl Acad Sci U S A 95:13579-13584.
4. Kornyshev, A. A., and S. Leikin. 1999. Electrostatic zipper motif for DNA aggregation. Phys Rev Lett 82:4138-4141.
5. Lee, D. J., A. Wynveen, A. A. Kornyshev, and S. Leikin. 2010. Undulations enhance the effect of helical structure on DNA interactions. J Phys Chem B 114:11668-11680.
6. Cherstvy, A. G., A. A. Kornyshev, and S. Leikin. 2002. Temperature-dependent DNA condensation triggered by rearrangement of adsorbed cations. J Phys Chem B 106:13362-13369.
7. Cherstvy, A. G., A. A. Kornyshev, and S. Leikin. 2004. Torsional Deformation of Double Helix in Interaction and Aggregation of DNA. J Phys Chem B 108:6508-6518.
8. Wynveen, A., D. J. Lee, and A. A. Kornyshev. 2005. Statistical mechanics of columnar DNA assemblies. Eur Phys J E Soft Matter 16:303-318.