The Journal of Chemical Physics

RESEARCH ARTICLE | OCTOBER 02 2024

In silico study of DNA mononucleotide self-assembly

Special Collection: Monte Carlo methods, 70 years after Metropolis et al. (1953)

Mattia Trapella ; Tommaso Bellini ; Cristiano De Michele

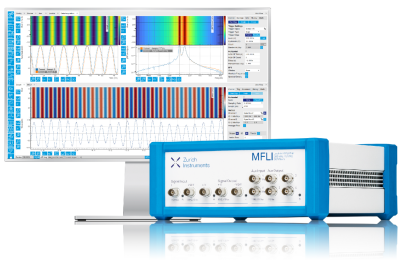
Check for updates

J. Chem. Phys. 161, 134905 (2024)
https://doi.org/10.1063/5.0226019

View Online
Export Citation

05 October 2024 08:02:28

Challenge us.
What are your needs for periodic signal detection?
Zurich Instruments
Find out more



AIP Publishing

The Journal of Chemical Physics ARTICLE pubs.aip.org/aip/jcp

In silico study of DNA mononucleotide self-assembly

Cite as: J. Chem. Phys. 161, 134905 (2024); doi: 10.1063/5.0226019
Submitted: 28 June 2024 • Accepted: 18 September 2024 •
Published Online: 2 October 2024

View Online Export Citation CrossMark

Mattia Trapella,[1] Tommaso Bellini,[2] and Cristiano De Michele[3,a)]

AFFILIATIONS

[1] Dipartimento di Fisica e Geologia, Università di Perugia, Perugia, Italy
[2] Dipartimento di Biotecnologie Mediche e Medicina Traslazionale, Università Degli Studi di Milano, Milano, Italy
[3] Dipartimento di Fisica, "Sapienza" Università di Roma, Roma, Italy

Note: This paper is part of the JCP Special Topic on Monte Carlo Methods, 70 Years after Metropolis *et al.* (1953).
[a)]**Author to whom correspondence should be addressed:** cristiano.demichele@uniroma1.it

ABSTRACT

Recent experiments have demonstrated the self-assembly and long-range ordering of concentrated aqueous solutions of DNA and RNA mononucleotides. These are found to form Watson–Crick pairs that stack into columns that become spatially organized into a columnar liquid-crystalline phase. In this work, we numerically investigate this phase behavior by adopting an extremely coarse-grained model in which nucleotides are represented as semi-disk-like polyhedra decorated with attractive (patchy) sites that mimic the stacking and pairing interactions. We carry out Monte Carlo simulations of these patchy polyhedra by adapting algorithms borrowed from computer graphics. This model reproduces the features of the experimental phase behavior, which essentially depends on the combination of pairing and stacking interactions.

Published under an exclusive license by AIP Publishing. https://doi.org/10.1063/5.0226019

05 October 2024 08:02:28

I. INTRODUCTION

The fundamental role of nucleic acids in biology arises from their capacity for mutual selective interactions. This feature governs all the key steps in the maintenance, transmission, and translation of the genetic code. The knowledge and control of such interactions have enabled a relentless flowering of new capacities that range from the development of RNA-based drugs[1] to nanotechnological innovation.[2] Watson–Crick selectivity results from a combination of pairing and stacking interactions between the nucleotides. While the combined effect of these two molecular processes is quantitatively known, the mechanism governing their interplay is still under discussion and experimental scrutiny.[3–5]

The standard thermodynamic description of DNA is built on the notion that the largest contribution to the binding strength of complementary sequences is given by stacking forces, while pairing mainly acts as a gatekeeper, providing a selective free energy penalty when non-complementary bases are in contact.[5,6] The relevance of stacking in the interactions between distinct DNA and RNA duplexes has emerged in the observation of spontaneous collective ordering in concentrated solutions of oligomer duplexes provided by well-paired blunt-ended terminals. In these systems, the attractive interaction of the duplex terminals favors their linear aggregation and, at sufficiently large concentrations and low temperatures, their self-assembly into liquid crystal phases, in which long chains of aggregated duplexes are equally oriented while preserving their 3D liquid structure (nematic phase) or else adopting a hexagonal 2D packing of the columns (columnar phases).[7] Recent direct measurements[8] confirmed that inter-duplex end-to-end stacking is equal to or even larger than intra-duplex stacking.

In previous works, the authors of the present manuscript and other groups have developed coarse-grained models in which DNA duplexes (DNADs) are approximated by hard cylinders (HCs) and quasi-cylinders with the stacking interactions represented by two attractive potentials at the two terminals.[9–12] These models well describe the phase behavior of DNADs (e.g., the B-DNA CGCGAATTCGCG known as the Dickerson dodecamer), exhibiting isotropic, nematic, and columnar phases.

Recent experimental research has demonstrated that self-assembly and collective ordering are also present in solutions of ultrashort DNA and RNA oligomers[13,14] and even in solutions of nucleotides.[15] Specifically, aqueous solutions of DNA and RNA

Published under an exclusive license by AIP Publishing

mononucleotide triphosphates (dNTPs) at sufficiently high concentrations and low temperatures undergo a phase transition to a columnar liquid crystal phase.[15] This finding marks a difference with previous observations of liquid crystal ordering, in that here there is no duplex pre-existing aggregation and ordering. In this case, the linear aggregation has to rely on the simultaneous occurrence of pairing and stacking, a behavior still not investigated via computer simulation.

Atomistic simulations of nucleic acids[16–18] are typically used to study nucleic acids but they are computationally demanding for investigating liquid crystalline ordering at high concentrations (e.g. columnar phase). Moreover, atomistic force fields[19–22] are known to not accurately quantify the strength of base stacking interactions.[23–27] Other approaches based on a more coarse-grained representation of DNA have also been proposed,[28–33] but they are also rather demanding in terms of computer time.

Here, we introduce a coarse-grained model of monomers, representing each a mononucleotide, in which we mimic both pairing and stacking through attractive patches. We show that with this approach, we can retrieve a phase diagram matching observations in conditions where both pairing and stacking are essential to the findings. Pairs of DNA mononucleotides can stack one on top of each other, forming longer strands and thus inducing the formation of liquid crystal phases. Unlike DNAD and their modeling with HC,[9,11,12] the cylinder-like self-assembled dNTP aggregates do not exhibit a nematic phase, featuring instead a direct phase transition from the isotropic phase into the columnar phase as the temperature is lowered and the concentration is increased.

The combination of the different mechanisms of assembly, in which both pairing and stacking are present, and the different resulting phase diagrams, missing the nematic phase, call for the development of computational models able to capture these features.

II. MODEL AND SIMULATION DETAILS

Based on the model previously developed by some of the authors of the present paper for DNADs, we model dNTPs as semi-disk-like polyhedra (SDPs) of diameter $D = 2.0$ nm and height $h = 0.2$ nm decorated with four attractive patches, as shown in Fig. 1(a). The value of D is a standard approximation for B-DNA duplexes.[10,34] Here, the duplexes might better map into a somewhat larger cylinder because of the two additional phosphates at each nucleotide. However, since the internal structure of the observed columns of pairs is undetermined and since the choice of $D = 2$ nm is by itself an approximation balancing backbone ridges and grooves, we opted for the minimalist approach of holding D to the standard value to keep our model as simple as possible. Any attempts to account for the additional steric hindrance and electrostatic interactions carried by the triphosphate groups of dNTPs would require careful atomistic simulations beyond the scope of this work. More specifically, two patches are placed along the cut edge of the semi-disk to replicate the base–base pairing interaction, while two additional patches are located at the center of the two bases of the semi-disk to mimic the stacking attraction between two dNTPs. In the following, we will use reduced temperature $T^* = k_BT/U_s$ and reduced pressure $P^* = Pv_0/(k_BT)$, where T and P are temperature and pressure, respectively, v_0 is the volume

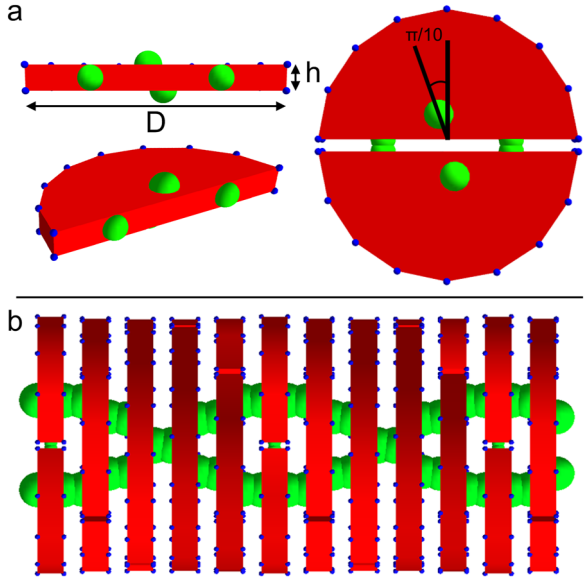


FIG. 1. (a) Patchy semi-disk model of a DNA mononucleotides. Each semi-disk-like polyhedron, with $D = 2.0$ nm and $h = 0.2$ nm, has four attractive patches, two for stacking and two for pairing. (b) Side view of a 12 base pair long DNA double strand modeled with 24 semi-disks. Small blue spheres highlight the vertices of the polyhedra.

of an SDP, and U_s is the stacking energy. To take into account the twist typical of B-DNA duplexes and, thus, properly account for its flexibility, the two stacking patches are placed on each monomer so as to give rise to a complete revolution of their columnar assembly every ten bases. This is achieved by rotating the two patches belonging to a single semi-disk (SD) by $\pi/10$, as shown in Fig. 1(a).

For the overlap detection needed in Monte Carlo (MC) simulations, we employed the *Xenocollide* algorithm,[35] a very efficient and robust (i.e., ensuring that no overlaps are missed) method that has been already used in computer graphics engines. Since self-assembled aggregates of DNA mononucleotides resemble double-stranded DNA duplexes (DNADs), our SDPs are designed in such a way that they self-assemble into cylinder-like aggregates [see Fig. 1(b)]. We adjusted the model parameters to obtain reasonable values for the persistence and contour length of the self-assembled DNA duplexes (see Appendix B). These adjustments include a fine-tuning of the spatial placement and size of the pairing patches, as well as a refinement of the geometry and position of the stacking patches.

The phase behavior of DNADs has been successfully reproduced by a model based on patchy HCs.[10] For example, Dickerson dodecamers have been successfully modeled as HCs decorated with two sticky patches on their bases.[12] It is expected that if dNTPs are modeled by exact patchy semi-disks that self-assemble into cylinder-like aggregates (CLAs), the correct phase behavior can be obtained. Anyway, we model a dNTP as an SDP composed of a finite number of vertices, thus differing from an exact semi-disk (i.e., the solid obtained by taking half of a disk). Hence, an issue to address is to

05 October 2024 08:02:28

Published under an exclusive license by AIP Publishing

estimate the minimum number of vertices of the SDP, which provides a physical behavior of the system equivalent to that of exact semi-disks. In Sec. II A, we will discuss a procedure to establish this minimal number of vertices. Another important issue concerns equilibration times. Starting from a configuration of fully unbonded SDPs, they have to self-assemble into cylinder-like aggregates and under the right conditions organize into a liquid crystalline phase. This process can be rather demanding in terms of computing time; hence, in Sec. II B, we will discuss a procedure to speed up equilibration.

A. Semi-disks as convex polyhedra

The *Xenocollide* algorithm's time complexity for checking the overlap between two polyhedra scales linearly with the number of vertices. Therefore, selecting the appropriate number of vertices involves balancing efficiency and accuracy.

Here, we discuss how to determine the minimal number of vertices of each SDP to provide an accurate description of the physical properties of the system. Our strategy was to build the equation of state (EOS) of hard convex polyhedra that have a cylinder-like shape (HCPs) and to compare this EOS with that of real hard cylinders. We carried out MC simulations in the isobaric–isothermal (NPT) ensemble for HCs of aspect ratio $X_0 = L/D = 3.7$, where L is their length and D is their diameter. We also simulated HCPs with the same aspect ratio and composed of 16, 24, and 32 vertices. Before the production runs, we equilibrated the system for at least 4×10^6 MC steps and we checked the equilibration by inspecting system density.

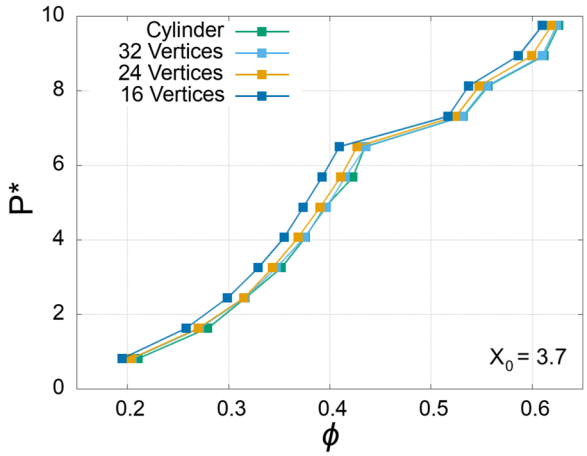


FIG. 2. Equation of state of HCs with an aspect ratio of $L/D = 3.7$. The equation of state of HCs is compared with the one obtained for HCPs with 16, 24, and 32 vertices.

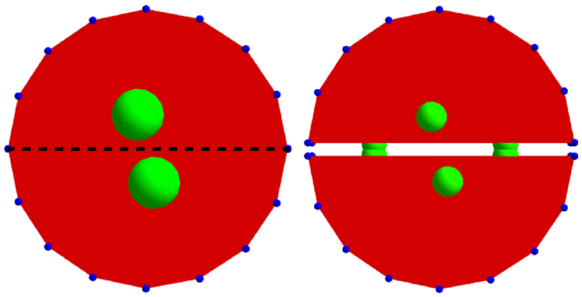


FIG. 3. Top view of HCPs and two SDPs with a mesh of vertices providing the same accuracy in their physical behavior. The HCP comprises a total of 32 vertices (16 on the bottom and 16 on the top), while an "equivalent" SDP is composed of 18 vertices.

Figure 2 shows the resulting EOS (i.e., P^* vs volume fraction ϕ) obtained for HCs and HCPs with a variable number of vertices. It can be seen that a very good approximation of the EOS of HCs is achieved with $V = 32$ vertices and that already with 16 vertices, the qualitative physical behavior is obtained, where in this case, there is just a slight shift of EOS toward smaller volume fractions. Hence, in our numerical simulations of SDPs, to achieve the same level of approximation obtained for HCs with 32 vertices, we use 18 vertices for SDPs, as shown in Fig. 3.

B. From patchy hard cylinders to patchy semi-disks

In this section, we provide details of the procedure that we used to speed up the equilibration. The key idea is that if patchy SDPs self-assemble into cylinder-like aggregates one can first simulate patchy HCPs and then use equilibrium configurations of patchy HCPs—where patchy HCPs are replaced with patchy SDPs—to investigate the phase behavior of patchy SDPs. According to this strategy, one has to first check that patchy SDPs fully self-assemble into CLAs.

To verify that this is the case, we performed simulations of a system composed of a small number of SDPs (i.e., small enough to have affordable computational times) and we checked that they almost fully self-assemble into CLAs. In Fig. 4, we show an initial isotropic configuration of SDPs and the resulting final (equilibrium) configuration where all SDPs self-assembled into CLAs.

We simulate patchy HCPs with $X_0 = 2$ decorated with two patches per base (where the choice of two patches will be justified later) to obtain N and Col phases. Patchy HCs with $X_0 = 2$ and one patch per base[11,12] exhibit a phase behavior compatible with experiments if a suitable stacking energy is chosen. In order to have a phase behavior of patchy HCPs (which have two patches per base) identical to one of patchy HCs (which have one patch per base), it is sufficient to ensure that the stacking free energy of the aggregation of two patchy HCPs is identical to that of two patchy HCs. The procedure to achieve this is discussed in detail in Appendix A.

To obtain N and Col configurations, we build the whole EOS of $N = 180$ patchy HCPs, which exhibits isotropic (I), nematic (N), and columnar (Col) phases, as shown in Fig. 5 We verified that by employing $N = 1440$ HCPs, the EOSs do not change appreciably, thus ruling out possible finite-size effects.

To distinguish the various phases, we inspect the radial distribution function and we make use of the nematic S and hexagonal ψ_6 order parameters. The nematic order parameter S is defined as the largest eigenvalue S of the order tensor $\mathbf{Q}$, i.e.,

$$\mathbf{Q}_{\alpha\beta} = \frac{1}{N}\sum_i^N \left[\frac{3}{2}(\mathbf{u}_i)_\alpha(\mathbf{u}_i)_\beta - \frac{1}{2}\delta_{\alpha\beta}\right], \quad (1)$$

05 October 2024 08:02:28

Published under an exclusive license by AIP Publishing

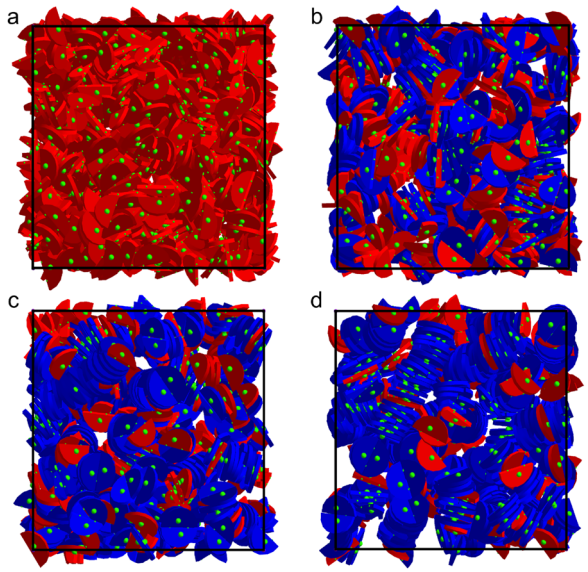


FIG. 4. (a) Starting configuration of the simulation with N = 300 SDPs. Simulation snapshots after two (b), four (c), and six (d) million MC steps. The SDPs that are fully paired are shown in blue.

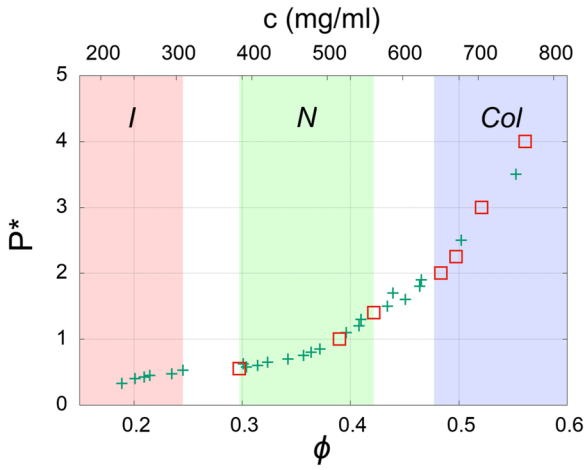


FIG. 5. Equation of state for patchy HCPs with aspect-ratio $X_0 = 2$ and two attractive patches per base. The simulations employ $N = 180$ particles and evolve over 4×10^6 MC steps. The squares in red represent the simulations whose final configurations are used for simulating patchy SDPs after the replacement of patchy HCPs.

where N is the number of particles, $\alpha\beta \in \{x, y, z\}$, and the unit vector $\mathbf{u}_{i\alpha}$ is the component α of the orientation (i.e., the symmetry axis) of particle i. If S is significantly larger than 0, particles are aligned along a given axis called the nematic director.

The hexagonal order parameter ψ_6 is defined as

$$\psi_6 = \left\langle \frac{1}{N}\sum_{i=1}^{N} \frac{1}{n(i)} \sum_{j=1}^{n(i)} e^{6i\theta_{ij}} \right\rangle, \tag{2}$$

where $n(i)$ is the number of first neighbors to the molecule i on the same layer and θ_{ij} is the angle between the two molecules i and j for an axis orthogonal to the nematic axis. Finally, the radial distribution function is defined as follows:

$$g(\mathbf{r}) = \frac{1}{\rho N}\left\langle \sum_{i=1}^{N} \sum_{j\neq i} \delta(\mathbf{r} - (\mathbf{r}_i - \mathbf{r}_j)) \right\rangle, \tag{3}$$

where ρ is the density of the system, $\mathbf{r}_i$ is the position of the ith particle, and $\delta(\mathbf{r})$ is the Dirac delta function.

From these patchy HCP simulations, we selected the state points shown as red squares in Fig. 5 to obtain nematic and columnar configurations.

In these configurations, we replaced each patchy HCP with a set of 24 patchy SDPs, as illustrated in Fig. 6.The replacement of patchy HCPs with CLAs of patchy SDPs has to be done without introducing overlaps between SDPs; hence, the SDPs must be entirely contained within the original cylinder. Moreover, the distance between the paired bases of the CLA is chosen in such a way that the final height matches that of the original cylinder. Additionally, the CLA is constructed by introducing a rotation of $\pi/5$, between successive base pairs, to ensure that the stacking patches are all bonded. Note that the adoption of four-patch HCPs, where two patches on one base are rotated with respect to the ones on the other base, ensures that after having replaced them with patchy SDPs, the bonds between two patchy HCPs are preserved.

Figure 7 shows a nematic configuration before [Fig. 7(a)] and after [Fig. 7(b)] the replacement of patchy HCPs with patchy SDPs for $P^* = 0.55$.

To verify the stability of N and Col, obtained by substituting patchy HCPs with CLAs composed of paired patchy SDPs, we performed NVT simulations. Since we used $N = 180$ patchy HCPs to build the initial configurations, the total number of SDPs in these simulations was equal to 4320.

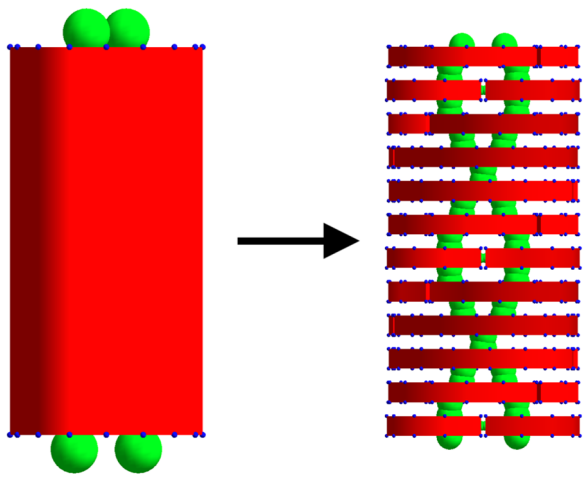


FIG. 6. A patchy HCP is replaced by 24 patchy SDPs. The patches of the HCP are rotated in such a way that they coincide with the ones of the two terminal SDPs once the replacement has taken place.

05 October 2024 08:02:28

Published under an exclusive license by AIP Publishing

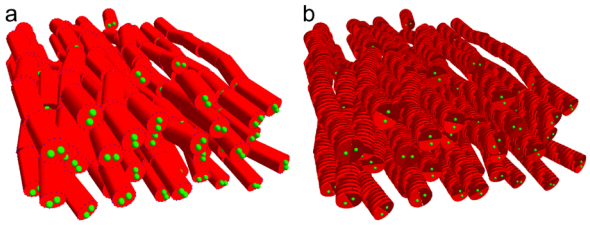


FIG. 7. (a) Nematic phase with cylinders at pressure $P^* = 0.55$. (b) Nematic phase after replacement of patchy HCPs with patchy SDPs.

III. RESULTS

Experiments indicate that the stacking interaction strength is larger than the pairing interaction strength although a clear analysis of the interplay, and possibly cooperativity, of pairing and stacking is still missing. According to Ref. 8, the stacking free energy for A–T interactions is about three times larger than the pairing free energy. In our model, the entropic penalties associated with stacking and pairing—which are of comparable values—are determined by patch geometry and positioning, in turn, chosen to approximate the persistence length of the SDP aggregates. Thus, by construction, our model does not enable an easy modulation of entropy variations. We instead differentiated the enthalpic values by setting a stacking energy two times larger than the pairing energy. This "moderate" choice, motivated by the need to avoid lowering too much the pairing free energy, appears, to further checks, not to affect significantly the phase diagram (see Sec. III A for a further discussion). As an initial configuration for simulating patchy SDPs, we used those corresponding to the red squares in Fig. 5.

The patchy SDPs and their stability are studied by carrying out an MC-NVT simulation at various reduced temperatures $T^* = [0.10, 0.15]$ for each starting configuration. We first note that for all temperatures investigated, the initially nematic phases are never stable and the system always becomes isotropic, whereas the columnar configurations are stable at low temperatures. Figure 8 shows the nematic and columnar order parameters for two starting configurations, one nematic and the other one columnar. In this figure, we also show the pair distribution functions of the columnar case on the plane orthogonal to the nematic axis; note the decrease in hexagonal ordering as the temperature increases.

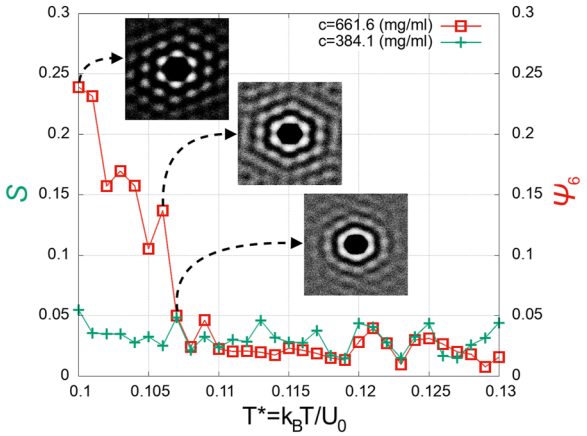


FIG. 8. Nematic (green crosses) and hexagonal (red squares) order parameters of patchy SDPs at $\phi = x0.357$. The radial distribution function is shown for the columnar phase at the temperature $T^* = 0.100$, $T^* = 0.106$ and $T^* = 0.110$.

A. Comparison with experiments

Here, we compare simulation results with experimental findings on dTTP/dATP mononucleotides.[15] We expressed system concentration c of computer simulations in mg/ml, where the conversion from volume fraction ϕ to concentration c in mg/ml can be obtained as follows:

$$c = \frac{N m_N}{V}, \tag{4}$$

where N represents the number of semi-disks within the simulation box's volume V and m_N denotes the molecular mass of a single nucleotide, measured in Daltons (Da). For dTTP and dATP filaments, the molecular masses are $m_N = 482.2$ Da and $m_N = 491.2$ Da, respectively. Since the experimental results refer to a mixture of dTTP and dATP filaments, we consider for the conversion the average molecular masses of the two species, i.e., $m_N = 486.7$ Da. Figure 9, shows the experimental findings of dATP/dTTP filaments[15] together with MC simulation results of patchy SDPs. Above a concentration of about 600 mg/ml, simulations exhibit a phase transition from *I* to *Col*, whereas at lower concentrations, the system is isotropic for all temperatures investigated.

This phase behavior well matches experimental observations even in the shape of the thermal stability of the *Col* phase. To check the stability of these results with respect to alternative choices of pairing energy, we carried out simulations for $c = 670$ mg/ml at $T^* = 0.1$ and 0.104 (the largest ones at which a stable *Col* phase is found) with stacking energy 10 times larger than pairing energy (i.e., $U_p = U_s/10$, where U_p is the pairing energy), finding that the *Col* phase indeed remains stable.

The absence of a nematic phase in the phase diagram of patchy SDPs can be rationalized through Wertheim-like theory developed in Ref. 10. If we consider self-assembling bifunctional HCs of length L and diameter D, they will form a polydisperse set of polymers (e.g., see Ref. 11). We start noting that in our present case, the self-assembled structure composed of patchy SDPs can be regarded as a set of 12 disks formed each by two paired patchy SDPs, where the stacking free energy of two disks is equal to that of two patchy HCPs.

According to the theory, the average number M of HCs belonging to a polymer at a given concentration and temperature in the isotropic phase is

$$M(\phi, T) = \frac{1}{2}\left(1 + \sqrt{1 + 4\phi e^{k_I \phi \eta(\phi) + \beta \Delta F_b}}\right), \tag{5}$$

where k_I is a geometric factor that does not depend on HC length L, $\eta(\phi) = \frac{1}{4}\frac{4-3\phi}{(1-\phi)^2}$ is the Parsons–Lee factor, and ΔF_b is the stacking free energy. If we replace each HC with a set of N_{sd} disks of diameter D and length L/N_{sd}, since the stacking free energy between two disks is still ΔF_b (since both stacking energy and bonding entropy are the same), we can expect M to be unchanged. Anyway, the average con-

05 October 2024 08:02:28

Published under an exclusive license by AIP Publishing

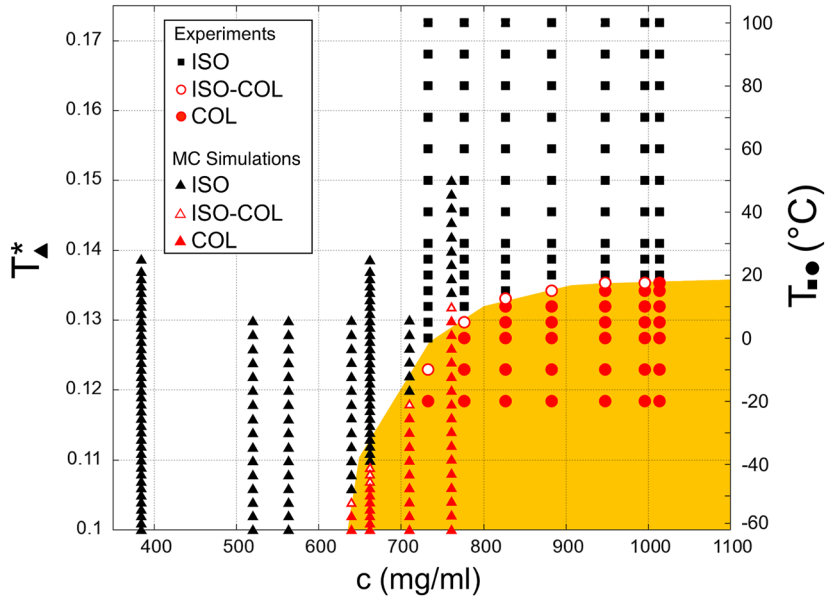


FIG. 9. Phase diagram of dTTP/dATP. The square and round symbols refer to the experimental results, while the triangular symbols refer to the results obtained from the simulations. The orange area indicates where the columnar phase is present.

tour length in the system is now $\frac{L}{N_{sd}M}$, while in the previous case, it was $\frac{L}{M}$. Hence, the average contour length of polymers becomes N_{sd} times shorter. This theoretical result implies that at a given temperature, the I–N transition of patchy SDPs shifts to higher concentrations, thus making it hindered by the columnar phase, whose phase boundaries would not change much since they are mostly driven by packing rather than elongation.

IV. CONCLUSIONS

Our study is aimed at investigating the interplay between pairing and stacking interactions in driving the self-assembly and phase behavior of DNA mononucleotides in concentrated aqueous solutions. In our extremely coarse-grained model, dNTPs are represented as SDPs decorated with attractive patches designed to mimic both base pairing and stacking interactions. With this approach, we obtained a phase diagram that, when properly scaled on the energy axis, well matches the experimental observations. In particular, the phase diagram of SDPs exhibits the liquid crystal columnar phase but not the nematic phase, possibly because of the insufficient contour length of the aggregates. Our study enlightens the cooperative role of pairing and stacking interactions and enables providing quantitative estimates of them.

Overall, our approach provides a very affordable and efficient computational tool for studying the self-assembly of DNA mononucleotides in which the interaction parameters are easily changed. This approach, based on MC simulations of hard polyhedra, can be straightforwardly modified to mimic, within the same extremely coarse-grained approximation, the assembly of macromolecules of different shapes, including proteins, peptides, and nucleic acids with different association modes, such as G-quadruplexes.[36]

ACKNOWLEDGMENTS

C.D.M. acknowledges financial support from the European Union—Next Generation EU (Grant No. MUR-PRIN2022 TAMeQUAD CUP:B53D23004500006) and from ICSC—Centro Nazionale di Ricerca in High Performance Computing, Big Data, and Quantum Computing, funded by the European Union–NextGenerationEU.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Mattia Trapella: Conceptualization (equal); Investigation (equal); Methodology (equal); Software (lead); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Tommaso Bellini**: Conceptualization (equal); Investigation (supporting); Methodology (equal); Supervision (equal); Writing – review & editing (equal). **Cristiano De Michele**: Funding acquisition (lead); Investigation (equal); Methodology (equal); Resources (lead); Software (supporting); Supervision (lead); Writing – original draft (equal); Writing – review & editing (lead).

05 October 2024 08:02:28

Published under an exclusive license by AIP Publishing

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

APPENDIX A: STACKING FREE ENERGY

Here, we discuss the calculation of the stacking free energy of HCPs decorated with four patches so that it is identical to the one previously used for bifunctional HCs.[11]

The stacking free energy is defined as

$$\beta\Delta F_b = \ln\left[2\frac{\Delta}{v_d}\right], \tag{A1}$$

where

$$\Delta = \frac{1}{4}\left\langle \int_{V_b} [e^{-\beta V(\mathbf{r}_{12},\Omega_1,\Omega_2)} - 1] d\mathbf{r}_{12} \right\rangle, \tag{A2}$$

with $\mathbf{r}$ denoting the vector connecting the centers of mass of particles 1 and 2, Ω_i representing the orientation of the ith particle, and $\langle\cdots\rangle$ indicating the average taken over all directions, v_d stands for the volume of the considered object, while V_b represents the bonding volume. For the calculation of Δ through an hit-and-miss MC, we use the same procedure described in Ref. 12, where one inserts with random positions and orientations two particles N times into a box of volume V. In the case of attractive patches, it simplifies, and for one patch per base, one has

$$\Delta_1 = \frac{VN_1}{4N}\left(e^{\beta u_0} - 1\right), \tag{A3}$$

where u_0 is the bonding energy and N_1 is the number of times that the two particles are bonded. For two patches per base, one has

$$\Delta_2 = \frac{V}{4N}\left(N_1(e^{\beta u_0'} - 1) + N_2(e^{2\beta u_0'} - 1)\right), \tag{A4}$$

where u_0' is the bonding energy and N_1 and N_2 represent the occurrences of particles forming one or two bonds, respectively. Given u_0, one can find u_0' such that $\Delta F_b(\Delta_1) = \Delta F_b(\Delta_2)$.

APPENDIX B: PERSISTENCE LENGTH

The persistence length is a fundamental parameter in polymer physics that quantifies the stiffness of a polymer chain. It represents the length over which the direction of the polymer chain remains correlated. The correlation function $C(k)$ is defined as

$$C(k) = \langle \cos\ \theta_{i,i+k}\rangle = \langle \vec{u}_i \cdot \vec{u}_{i+k}\rangle, \tag{B1}$$

where $\vec{u}_i$ are unit vectors connecting each pair i of semi-disks along the filament, separated by k semi-disks. This function is typically expected to decay exponentially with the characteristic persistence length parameter, l_p,

$$C(k) = \langle \cos\ \theta_{i,i+k}\rangle = e^{-k/l_p}. \tag{B2}$$

If the polymer under consideration has a high persistence length and therefore the angle θ between two contiguous monomers is close to zero, an approximation can be made in Eq. (B2). Considering a short section of the filament, $s \ll l_p$, we have

$$\langle \cos\ \theta(k)\rangle \approx 1 - \frac{k}{l_p}. \tag{B3}$$

Since the angle θ is small, $\cos(\theta(s)) \approx 1 - \theta^2(s)/2$, leading to

$$\langle \theta^2(s)\rangle \approx \frac{2s}{l_p} \rightarrow l_p \approx \frac{2s}{\langle \theta^2(s)\rangle}. \tag{B4}$$

Therefore, the persistence length can be calculated by measuring the angle between first-neighboring monomers. This method offers a significant advantage over exponential fitting as it eliminates the need to simulate filaments close to the persistence length. We validated this approximation by comparing the results with those obtained using the exponential fit method, finding that for filaments with $l_p \geq 15$, the error is less than 1%.

The size and position of the stacking and pairing patches can be tuned to match the persistence and contour length of the CLAs with that of real DNA duplexes (DNADs). Experimentally, DNADs have a persistence length of ~50 nm or 150 base pairs of DNA,[37] while the length of a DNAD composed of 12 base pairs is about 4 nm. To estimate the persistence length in the case of CLAs of SDPs, the first neighbor method is used along with Eq. (B4). Initially, we studied how the persistence length varies with the position of the stacking patch. The position of this patch is indicated by k, where $k \cdot R_{sd}$ indicates the distance of the patch from the center of the SDP. Consequently, if $k = 0$, the patch will be at the center of the SDP, while if $k = 1$, the patch will be on its rim. The graph of the persistence length as a function of k is shown in Fig. 10. Being $r_{sd} = D/2$, the stacking patch is positioned at $k = 0.35r_{sd}$, which is the minimum distance required to prevent filament branching.

Hence, the size of the stacking patch was used as a parameter to achieve the desired persistence length, as shown in Fig. 11. By choosing a stacking patch radius of 0.14 nm, a persistence length of ~150 base pairs is achieved.

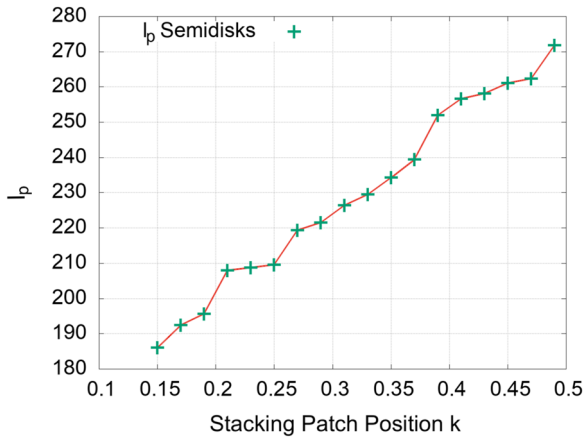


FIG. 10. Persistence length l_p in units of base pairs as a function of the position k of the stacking patch relative to the center of the SDP. The radii of stacking and pairing patches radii are 0.11 and 0.05 nm, respectively.

05 October 2024 08:02:28

Published under an exclusive license by AIP Publishing

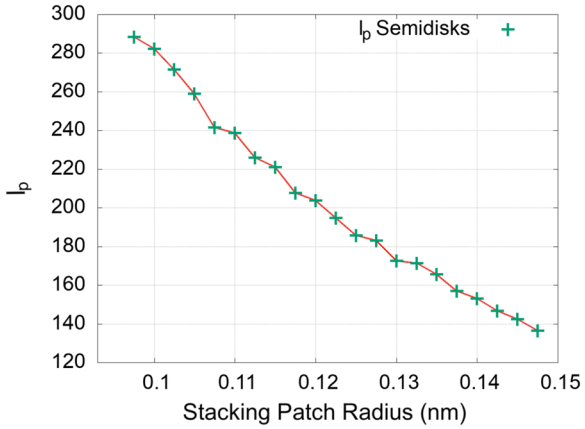


FIG. 11. Persistence length l_p in units of base pairs as a function of the radius of the stacking patch. The pairing patches radius is equal to 0.05 nm.

Then, we calculate the total length of a strand composed of 12 base pairs. Experimentally, a DNA strand of 12 base pairs measures ~4 nm in length. The strand length is examined as a function of the SDP height. Given the constraint of the maximum stacking patch dimension and the necessity of a stacking patch radius of 0.14 nm to achieve a l_p value of 150 bases, SDPs with a height of $h = 0.2$ nm are selected to ensure proper strand flexibility. This method results in a strand length of ~4.4 nm, which is reasonably close to the experimental value of 4 nm.

REFERENCES

[1]M. Winkle, S. M. El-Daly, M. Fabbri, and G. A. Calin, Nat. Rev. Drug Discovery **20**, 629 (2021).
[2]N. C. Seeman, *Structural DNA Nanotechnology* (Cambridge University Press, 2016).
[3]K. M. Guckian, B. A. Schweitzer, R. X.-F. Ren, C. J. Sheils, D. C. Tahmassebi, and E. T. Kool, J. Am. Chem. Soc. **122**, 2213 (2000).
[4]M. Zacharias, J. Phys. Chem. B **124**, 10345 (2020).
[5]P. Yakovchuk, E. Protozanova, and M. D. Frank-Kamenetskii, Nucleic Acids Res. **34**, 564 (2006).
[6]J. SantaLucia, Proc. Natl. Acad. Sci. U. S. A. **95**, 1460 (1998).
[7]M. Nakata, G. Zanchetta, B. D. Chapman, C. D. Jones, J. O. Cross, R. Pindak, T. Bellini, and N. A. Clark, Science **318**, 1276 (2007).
[8]F. Kilchherr, C. Wachauf, B. Pelz, M. Rief, M. Zacharias, and H. Dietz, Science **353**, aaf5508 (2016).
[9]T. Kuriabova, M. D. Betterton, and M. A. Glaser, J. Mater. Chem. **20**, 10366 (2010).
[10]C. De Michele, T. Bellini, and F. Sciortino, Macromolecules **45**, 1090 (2012).
[11]K. T. Nguyen, F. Sciortino, and C. De Michele, Langmuir **30**, 4814 (2014).
[12]C. De Michele, L. Rovigatti, T. Bellini, and F. Sciortino, Soft Matter **8**, 8388 (2012).
[13]T. P. Fraccia, G. P. Smith, L. Bethge, G. Zanchetta, G. Nava, S. Klussmann, N. A. Clark, and T. Bellini, ACS Nano **10**, 8508 (2016).
[14]M. Todisco, T. P. Fraccia, G. P. Smith, A. Corno, L. Bethge, S. Klussmann, E. M. Paraboschi, R. Asselta, D. Colombo, G. Zanchetta, N. A. Clark, and T. Bellini, ACS Nano **12**, 9750 (2018).
[15]G. P. Smith, T. P. Fraccia, M. Todisco, G. Zanchetta, C. Zhu, E. Hayden, T. Bellini, and N. A. Clark, Proc. Natl. Acad. Sci. U. S. A. **115**, E7658 (2018).
[16]J. Yoo, D. Winogradoff, and A. Aksimentiev, "Biophysical and computational methods—Cryo EM," Curr. Opin. Struct. Biol. **64**, 88 (2020).
[17]S. Naskar, S. Saurabh, Y. H. Jang, Y. Lansac, and P. K. Maiti, Soft Matter **16**, 634 (2020).
[18]J. Yoo and A. Aksimentiev, Phys. Chem. Chem. Phys. **20**, 8432 (2018).
[19]A. Pérez, I. Marchán, D. Svozil, J. Sponer, T. E. Cheatham, C. A. Laughton, and M. Orozco, Biophys. J. **92**, 3817 (2007).
[20]I. Ivani, P. D. Dans, A. Noy, A. Pérez, I. Faustino, A. Hospital, J. Walther, P. Andrio, R. Goñi, A. Balaceanu, G. Portella, F. Battistini, J. L. Gelpí, C. González, M. Vendruscolo, C. A. Laughton, S. A. Harris, D. A. Case, and M. Orozco, Nat. Methods **13**, 55 (2016).
[21]M. Zgarbová, J. Šponer, M. Otyepka, T. E. Cheatham III, R. Galindo-Murillo, and P. Jurečka, J. Chem. Theory Comput. **11**, 5723 (2015).
[22]K. Hart, N. Foloppe, C. M. Baker, E. J. Denning, L. Nilsson, and A. D. MacKerell, Jr., J. Chem. Theory Comput. **8**, 348 (2012).
[23]S. Saurabh, Y. Lansac, Y. H. Jang, M. A. Glaser, N. A. Clark, and P. K. Maiti, Phys. Rev. E **95**, 032702 (2017).
[24]C. Maffeo, J. Yoo, J. Comer, D. B. Wells, B. Luan, and A. Aksimentiev, J. Phys.: Condens. Matter **26**, 413101 (2014).
[25]C. Maffeo, B. Luan, and A. Aksimentiev, Nucleic Acids Res. **40**, 3812 (2012).
[26]V. Minhas, T. Sun, A. Mirzoev, N. Korolev, A. P. Lyubartsev, and L. Nordenskiöld, J. Phys. Chem. B **124**, 38 (2020).
[27]J. Yoo and A. Aksimentiev, J. Phys. Chem. Lett. **3**, 45 (2012).
[28]T. A. Knotts IV, N. Rathore, D. C. Schwartz, and J. J. de Pablo, J. Chem. Phys. **126**, 084901 (2007).
[29]J. J. de Pablo, Annu. Rev. Phys. Chem. **62**, 555 (2011).
[30]T. E. Ouldridge, A. A. Louis, and J. P. K. Doye, J. Chem. Phys. **134**, 085101 (2011).
[31]J. P. K. Doye, T. E. Ouldridge, A. A. Louis, F. Romano, P. Šulc, C. Matek, B. E. K. Snodin, L. Rovigatti, J. S. Schreck, R. M. Harrison, and W. P. J. Smith, Phys. Chem. Chem. Phys. **15**, 20395 (2013).
[32]J. P. K. Doye, H. Fowler, D. Prešern, J. Bohlin, L. Rovigatti, F. Romano, P. Šulc, C. K. Wong, A. A. Louis, J. S. Schreck, M. C. Engel, M. Matthies, E. Benson, E. Poppleton, and B. E. K. Snodin, "The oxDNA coarse-grained model as a tool to simulate DNA origami," in *DNA and RNA Origami: Methods and Protocols*, edited by J. Valero (Springer US, New York, NY, 2023), pp. 93–112.
[33]S. Naskar and P. K. Maiti, J. Mater. Chem. B **9**, 5102 (2021).
[34]C. R. Calladine, H. R. Drew, B. F. Luisi, and A. A. Travers, in *Understanding DNA*, 3rd ed., edited by C. R. Calladine, H. R. Drew, B. F. Luisi, and A. A. Travers (Academic Press, Oxford, 2004), pp. 1–17.
[35]D. Fiser, "XenoCollide MPR algorithm," 2012, https://github.com/danfis/libccd, copyright (c)2010-2012 Daniel Fiser (danfis@danfis.cz), Intelligent and Mobile Robotics Group, Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague. All rights reserved.
[36]B. P. Rosi, V. Libera, L. Bertini, A. Orecchini, S. Corezzi, G. Schirò, P. Pernot, R. Biehl, C. Petrillo, L. Comez, C. De Michele, and A. Paciaroni, J. Am. Chem. Soc. **145**, 16166 (2023).
[37]M. Gerald S, Biophys. J. **91**, 3607 (2006).

05 October 2024 08:02:28

Published under an exclusive license by AIP Publishing