ChemPhysChem

Articles
doi.org/10.1002/cphc.202000630

Chem
Europ

European C
Societies P

Check for
updates

Binding Mode Multiplicity and Multiscale Chirality in the Supramolecular Assembly of DNA and a π-Conjugated Polymer

Mathieu Fossépré,*[a] Inbal Tuvi-Arad,[b] David Beljonne,[a] Sébastien Richeter,[c] Sébastien Clément,[c] and Mathieu Surin*[a]

Water-soluble π-conjugated polymers are increasingly considered for DNA biosensing. However, the conformational rearrangement, supramolecular organization and dynamics upon interaction with DNA have been overlooked, which prevents the rational design of such detection tools. To elucidate the binding of a cationic polythiophene (CPT) to DNA with atomistic resolution, we performed molecular simulations of their supramolecular assembly. Comparison of replicated simulations show a multiplicity of CPT binding geometries that contribute to the wrapping of CPT around DNA. The different binding geometries are stabilized by both electrostatic interactions between CPT lateral cations and DNA phosphodiesters and van der Waals interactions between the CPT backbone and the DNA grooves. Simulated circular dichroism (CD) spectra show that the induced CD signal stems from a conserved geometrical feature across the replicated simulations, i.e. the presence of segments of syn configurations between thiophene units along the CPT chain. At the macromolecular scale, we inspected the different shapes related to the CPT binding modes around the DNA through symmetry metrics. Altogether, molecular dynamics (MD) simulations, model Hamiltonian calculations of the CD spectra, and symmetry indices provide insights into the origin of induced chirality from the atomic to the macromolecular scale. Our multidisciplinary approach points out the hierarchical aspect of CPT chiral organization induced by DNA.

1. Introduction

Molecular recognition of DNA by molecules is fundamental in regards to the crucial role of DNA in medical diagnosis, food safety, gene expression profiling, drug screening, among many other biochemical or medical studies.[1–3] Diverse strategies for designing DNA biosensors have been proposed during the last decades, using for instance electrochemical or fluorescence signals, the latter requiring most often to covalently label the DNA with a fluorophore (dye).[4] Covalently-bound labeling dyes that transduce a recognition event into an optically-detected signal involve chemical modifications of DNA, which often affect the chemical specificity of the targeted DNA. For instance, the specific recognition between antigen and DNA is perturbed when the oligonucleotide is tagged with a large fluorophore.[5]

Supramolecular approaches for detection of DNA are thus appealing to maintain the chemical integrity and specificity of the biomolecular target.[6,7] The chemical nature of DNA offers a multitude of potential intermolecular interactions for the design of biosensors, i.e., from non-directional but strong electrostatic interactions to weak but directional hydrogen bonds, dipole-dipole or π-stacking interactions.[8,9] The polyanionic nature of DNA is often considered to yield strong electrostatic interactions between a given cationic molecule and the negatively-charged phosphodiester backbone, leading to the formation of stable supramolecular complexes.[10–12] In a previous study, we used a cationic π-conjugated polymer, i.e., poly[3-(6′-(trimethylphosphonium)hexyl)thiophene-2,5-diyl] (P3HT–PMe3), to interact with the anionic phosphodiester backbone along DNA. We showed that the achiral P3HT–PMe3 polymer presents an induced CD (ICD) signal upon interaction with DNA, this signal being sensitive to the DNA sequence. Nucleotides can thus be optically detected without covalent modification of the DNA, through the formation of polyplexes with P3HT–PMe3.[13] Such polyplexes were used to monitor in real-time the cleavage of DNA by HpaI (Haemophilus parainfluenzae), an endonuclease enzyme, via chiroptical signals.[14]

Thanks to the tremendous progress of computer performance, the accessible size and timescale for performing more 'realistic' simulations of biomolecular systems have increased.[15] In our abovementioned study,[14] experimental ICD results were complemented by MD simulations to bring atomistic insights into the organization and the dynamics of P3HT–PMe3 before and after the cleavage of DNA. Simulations of drug-DNA complexes are now becoming routine to propose atomistic

[a] Dr. M. Fossépré, Prof. D. Beljonne, Prof. M. Surin
Laboratory for Chemistry of Novel Materials
Centre of Innovation and Research in Materials and Polymers (CIRMAP)
University of Mons - UMONS
20 Place du Parc, Mons 7000 (Belgium)
E-mail: mathieu.fossepre@umons.ac.be
        mathieu.surin@umons.ac.be
[b] Prof. I. Tuvi-Arad
Department of Natural Sciences
The Open University of Israel
Raanana (Israel)
E-mail: inbaltu@openu.ac.il
[c] Dr. S. Richeter, Prof. S. Clément
ICGM, Université de Montpellier
CNRS, ENSCM, Montpellier (France)

ChemPhysChem

Articles
doi.org/10.1002/cphc.202000630

Chemistry Europe
European Chemical
Societies Publishing

models of DNA-ligand binding and a deeper understanding of the identification, dynamics, and the multiplicity of DNA binding sites.[16] Multiple binding modes were observed for small DNA ligands and for larger polycationic DNA binders.[17] It is therefore worth investigating the possible binding modes of P3HT–PMe$_3$ to DNA and its impact on ICD signals, to understand the origin of the chirality induction in P3HT–PMe$_3$/DNA complexes. We have carried out replicated MD simulations of the supramolecular assembly between P3HT–PMe$_3$ and a DNA fragment on a long MD timescale (up to 10 microseconds). In our simulations, we have considered a P3HT–PMe$_3$ polymer composed of 47 monomer units, which correspond to the average degree of polymerization of the polymer from $<M_n>$.[13] The selected DNA (referred to as dsR$_{43}$) is composed of 43 base pairs, as reported in Ref. 14. We first compare the stability and the dynamics of the various P3HT–PMe$_3$ binding modes around dsR$_{43}$. Based on the MD snapshots, we then simulate the CD spectra of the conjugated polymer by solving a model Hamiltonian including through-space and through-bond interactions between successive monomer units. Our findings highlight the relationship between the local conformational effects of the polymer and the magnitude and sign of the CD response.

Following conformational changes during dynamical processes requires the use of good molecular descriptors. These are expected to be normalized global functions of the coordinates that allow for comparison of different structures on the same scale, while capturing both major and minor geometrical changes during the process. One such set of molecular descriptors are the continuous symmetry measures (CSMs)[18–20] and their related continuous chirality measure (CCM).[21] These measures determine the distance of a given molecular structure from its nearest symmetric (or achiral) structure, on a continuous, normalized scale. They can also be used as measures of shape,[22,23] and as measures of linearity or planarity even in the lack of perfect symmetry. Since their original developments, these metrics have been proven useful for various studies of reactive[24] and catalytic processes,[25,26] as well as physical and chemical processes.[27–32] Here, we used the CSM to study the various global shapes of the P3HT–PMe$_3$ conformations in order to underline how local organizational effects influence the macromolecular shapes of CPT around DNA.

2. Results and Discussion

2.1. Structural Aspects of DNA/Polymer Assemblies

In a previous study, we proposed a stable binding mode of P3HT–PMe$_3$ around a double-stranded DNA, with preferential adsorption of the thiophene units along the DNA minor-groove, and with the lateral phosphonium groups complexed by the DNA phosphodiester groups.[14] To uncover other stable binding modes, we performed four new MD replicas (named R1 to R4), from 2 μs to 10 μs, see the Methods Section. We follow the time-evolution of the distance between the centers of mass

(COM) of both partners (Figure 1), as an average indicator of DNA/polymer proximity within the complexes.

While important oscillations over a range of ~5 Å were observed during the first μs of the MD simulations (especially for the R1 and R2 cases), a plateau is reached for each replica after ~1 μs. The dsR$_{43}$/P3HT–PMe$_3$ distances are confined within a range of few Å along the second part of the MD production phase. Although the complex is dynamic, as observed by the oscillations of the DNA-polymer COM distances versus the MD time (Figure 1), a statistical analysis of distance profiles show that these oscillations are confined within a range of few Å, especially in the second part of the MD simulations (see distributions and box plots in Supplementary S2). No dissociation between dsR$_{43}$ and P3HT–PMe$_3$ is observed, irrespective of the profile. The dsR$_{43}$/P3HT–PMe$_3$ complexes, once formed, are thus stable on the μs timescale thanks to the strong electrostatic interactions. Hence, the weights of electrostatic and van der Waals interactions to the total non-bonding interactions (around 90% and 10%, respectively, see Supplementary S3) show that the dsR$_{43}$/P3HT–PMe$_3$ assembly is mainly electrostatically-driven. We also report values of this ratio for the final MD snapshots in the same figure, showing that the latter are statistically representative in terms of the balance of interactions at play. However, a visual inspection of the final MD snapshots for the four runs (Figure 2) show that the geometry of P3HT–PMe$_3$ around dsR$_{43}$ are quite different. P3HT–PMe$_3$ conformations depict different levels of coiling, where conformations of R2 and R4 conformations are more extended than those of R1 and R3 runs.

Note that the P3HT–PMe$_3$ end-to-end distance (Figure 3) varies from around 60 to 70 Å for replicas R1 and R3 (highly coiled polymers) to ~100 Å for replicas R2 and R4 (extended polymers), far from the end-to-end distances of the P3HT–PMe$_3$ without DNA, which are centered around ~140 Å (Figure 3). In addition, variations of end-to-end distances are much less important when P3HT–PMe$_3$ is complexed with DNA (see distributions and box plots of end-to-end profiles in Supplementary S4). This means that DNA constricts the polymer
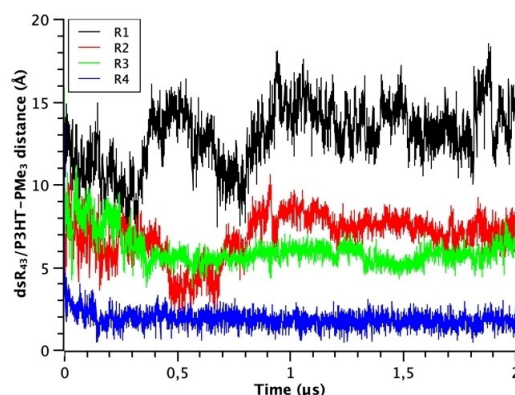

Figure 1. Distance profiles between dsR$_{43}$ and P3HT–PMe$_3$ centers of mass for the four replicas MD simulations. For the R4 replica, the complete profile (from 0 to 10 μs) is available in Supplementary S1 in the Supporting Information.
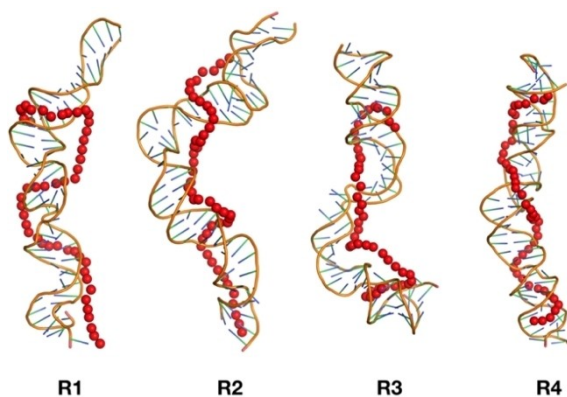
ChemPhysChem

Articles
doi.org/10.1002/cphc.202000630

**Chemistry
Europe**
European Chemical
Societies Publishing

Figure 2. Final MD snapshots of dsR$_{43}$/P3HT—PMe$_3$ complexes (replicas R1, R2, R3 and R4). The DNA fragments are depicted in cartoon representation (backbones in orange; bases in green/blue sticks). The polymer (red) is depicted in low resolution, each red particle being a thiophene unit centered on the sulfur atom.
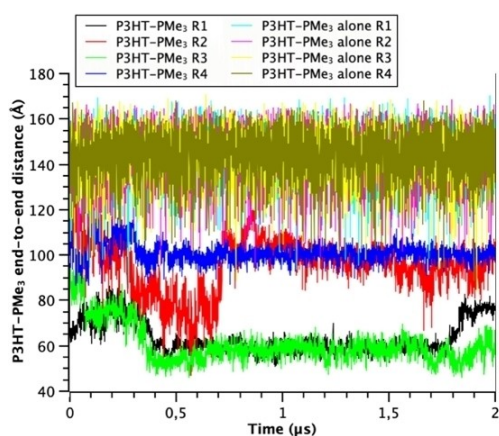

Figure 3. Top end-to-end distance profiles of DNA-bound (four replicas) and unbound P3HT—PMe$_3$ (four replicas). Distances were calculated from the centers of mass of terminal monomers of the polymer chain. For R4 replicas, the complete profiles (from 0 to 10 μs) are available in Supplementary S1 in the Supporting Information.

dynamics to a smaller conformational ensemble, although the polymer is still presenting some dynamical behavior. In the last MD snapshots, shown in Figure 2, we observe that some segments of the coiled polymer conformations (in particular for R1 and R3) do not interact closely all along the DNA axis, leaving a part of the DNA fragment free from polymer interactions. The 'wrapping' of the polymer around the DNA is also varying in terms of positioning along the DNA helical axis.

The values of plateaus reached in Figure 1 are an indicator of the proximity between the COM of both dsR$_{43}$ and P3HT—PMe$_3$. A higher distance means that the COM of both components are far from each other, as a result of the sliding motion of the polymer along the helical axis of the DNA. Plateaus reported in Figure 1 extend from 2 Å for replica R4 to 13 Å for replica R1, with intermediate values around 6–7 Å for R2 and R3. Both dsR$_{43}$/P3HT—PMe$_3$ COM and P3HT—PMe$_3$ end-to-end distances permit to classify P3HT—PMe$_3$ binding modes:

off-centered (with respect to the DNA helical axis) and coiled in R1, centered and elongated in R2, centered and coiled in R3, and centered and elongated in R4. Importantly, the different possibilities of 'wrapping' the P3HT—PMe$_3$ around DNA cannot be obviously discriminated with unique energy criteria. The difference between average potential energies, reported in Supplementary S5, are indeed in the order of their corresponding standard deviations. In the next section, we analyze how the monomer units are interacting with the DNA helix throughout the four MD replicas.

2.2. Multiple Adsorption Modes in DNA/Polymer Assemblies

A visual comparison of final MD snapshots in Figure 2 clearly depicts that thiophene units along the polymer backbone are differently accommodated around the DNA helix. To further understand the adsorption modes, non-bonded interactions involving thiophene moieties were detected by considering contacts between sulfur atoms of thiophene units and DNA atoms (called 'S-DNA' contacts hereafter). The number of S-DNA contacts are similar for the final MD snapshots of the four replicas (Supplementary S6), regardless of the cut-off distance used to record S-DNA contacts. However, the repartition of S-DNA contacts along the polymer backbone is different (Figure 4). Figure 4 bottom shows the number of S-DNA contacts for the four replicas. The larger domains of close contacts between thiophene units and DNA are mainly related to minor-groove binding domains, as highlighted in pink in Figure 4 top. Major-groove binding of thiophene units is also encountered but related to a smaller number of S-contacts or shorter domains of thiophene units. Besides these groove-binding domains, other segments of P3HT—PMe$_3$ units are not within grooves (see monomer units in yellow in Figure 4 top), but rather involved in electrostatic interactions between phosphonium groups and DNA phosphodiester groups, as highlighted in black in Figure 4 top (referred to as 'DNA-close phosphonium'). Concerning these electrostatic interactions, we have determined the close contacts between the P3HT—PMe$_3$ cationic phosphonium groups and DNA anionic phosphodiester groups through the P-P distances (denoted as 'P-P' contacts hereafter). The number of such contacts is similar for the four final MD snapshots (Supplementary S6), regardless of the cut-off distance chosen to record P-P contacts. It is worth mentioning that the phosphonium side groups of the polymer can be close to several DNA phosphodiesters, explaining the many P-P contacts. Hence, the proximity between the phosphonium groups and the various DNA phosphodiesters groups can change with time, a motion facilitated by the flexibility of the polymer alkyl side-chains. Altogether, the MD replicas show diverse adsorption modes of CPT monomer units to DNA. Most of the thiophene units of P3HT—PMe$_3$ are located within DNA minor grooves and major grooves. The thiophene units that are not groove-bound have their phosphonium groups interacting with the DNA phosphodiester backbone, occurring often in major grooves. The wrapping of P3HT—PMe$_3$ around DNA exhibits a sliding motion as well as different
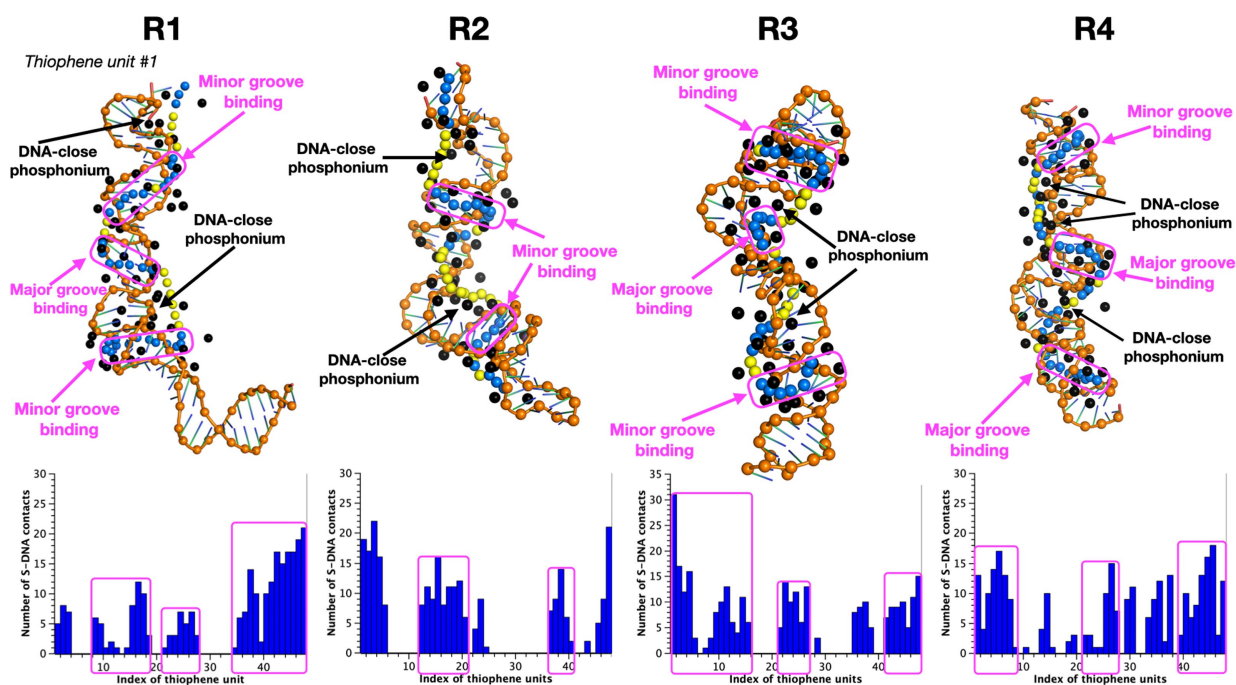
ChemPhysChem

Articles
doi.org/10.1002/cphc.202000630

Chemistry Europe
European Chemical
Societies Publishing

Figure 4. Mapping of thiophene moities that interact with the DNA double helix on final MD snapshots of dsR$_{43}$/P3HT—PMe$_3$ complexes (replicas R1, R2, R3 and R4). The DNA fragments are depicted in cartoon representation (backbones in orange; bases in green/blue sticks). The polymer is depicted in low resolution (thiophene units centered on the sulfur atom). Blue particles are related to thiophenes that closely interact with DNA double helix (contacts between sulfur atoms of thiophene units and DNA atoms, S-DNA contacts > 0), yellow particles are related to non-interacting thiophene moieties (no S-DNA contact). The major interacting domains are framed in the S-DNA contacts profiles. Phosphorus atoms of DNA phosphodiesters groups and of polymer phosphonium groups are in orange and black spheres, respectively. Non-interacting thiophene units (yellow) are due to DNA-close phosphonium moieties (P-P contacts in the text).

degrees of coiling around the DNA helix axis. However, a common feature to all complexes is that the polymer wraps around the DNA double helix, with coiled sections of various sizes that occur along the minor and major DNA grooves. In the next section, we examine how the dynamics of both DNA and polymer are influenced by the different binding modes.

2.3. Dynamics of DNA and Polymer Within the Complexes

The Root Mean Square Fluctuations (RMSF, see Methods) measuring the local amplitude of the displacements with respect to the average conformation for each partner are shown in Figure 5. These estimates allow us to assess how the polymer binding geometry influences the flexibility of both dsR$_{43}$ and P3HT—PMe$_3$. The dsR$_{43}$ RMSF profiles are different for the four MD replicas. The dynamics of dsR$_{43}$ is locally influenced with portions of rigid nucleotide residues due to local coiling of the polymer. Hence, the RMSF profile of DNA in R1 (Figure 5 top) can be cut in three regions: an unbound segment of dsR$_{43}$ (residues 1 to 15) that is highly flexible, a second segment (residues 15–35) that is more rigid because of the entanglement of the polymer around DNA, and a third segment (residues 36 to 43), corresponding to intrinsically flexible DNA extremities. In contrast, the dsR$_{43}$ RMSF profile in R2 shows an important flexibility around residue-20, as the polymer is less accommodated to the DNA helix (see final MD snapshots in Figure 2 and

4). The RMSF profile of dsR$_{43}$ in R3 shows important oscillations between flexible and rigid residues. RMSF profile of dsR$_{43}$ in R4 has some similarity with the R2 one (both replicas present elongated polymer conformations), with peaks being encountered in the central region of DNA, i.e., around the residue 25. Nevertheless, the flexibility profiles of bounded dsR$_{43}$ preserve some similarities with the usual flexibility profile of unbounded DNA. Hence, the maximum flexibility (peaks in Figure 5) is reached for the residues located at the extremities of dsR$_{43}$ in each replica. These peaks are, however, more intense for R1 and R3 (coiled polymers, lower end-to-end distances) than for R2 and R4 (elongated polymers, higher end-to-end distances). The correlation coefficients between the RMSF profiles of the bound (four replicas) and unbound dsR$_{43}$ (in Supplementary S7) are always above 0.50, the minimum being 0.58 for the pair of replicas R1 and R3 and the maximum 0.87 for the pair R2 and unbound DNA. The intrinsic flexibility and dynamics of dsR$_{43}$ is thus partially conserved, even when DNA is trapped by P3HT—PMe$_3$. In contrast, correlation coefficients between the RMSF profiles of DNA-bound (four replicas) and unbound P3HT—PMe$_3$ (Supplementary S8) are weaker, in general below 0.50 and even negative in some cases. We thus conclude that, whereas the dynamics of P3HT—PMe$_3$ is extremely influenced by the type of binding modes around dsR$_{43}$, the dsR$_{43}$ flexibility is partially conserved with only local impact compared to pure DNA. The intrinsic flexibility of P3HT—PMe$_3$ is such that it could adopt many types of conformations around a DNA fragment, at

ChemPhysChem

Articles
doi.org/10.1002/cphc.202000630

Chemistry Europe
European Chemical
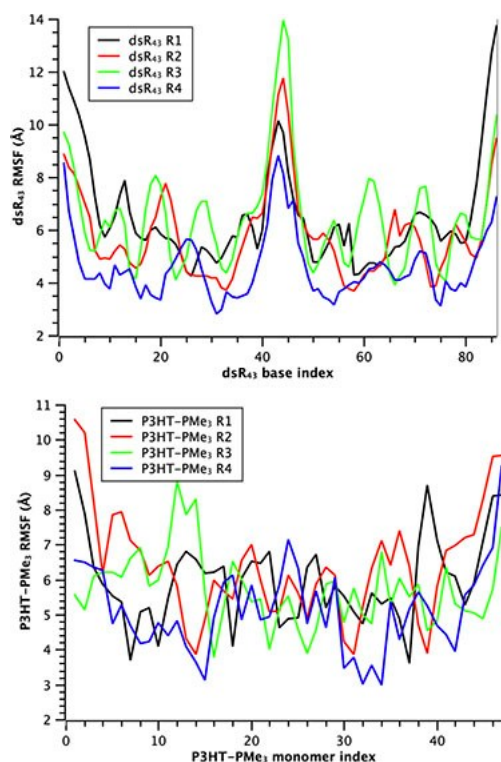Societies Publishing

Figure 5. Top: Root Mean Square Fluctuations (RMSF) of dsR_{43} for replicas R1, R2, R3 and R4, as a function of the residue number. Bottom: RMSF of P3HT—PMe$_3$ for replicas R1, R2, R3 and R4. Only heavy atoms were considered for RMSF calculations.

Table 1. Percentage of *syn* dihedral angles along P3HT—PMe$_3$ backbone for the four replicas MD simulations of the complex (DNA/polymer) and for the polymer alone.

	R1	R2	R3	R4
syn to anti proportion (*dsR$_{43}$*/P3HT—PMe$_3$)	61.4	63.2	70.0	64.6
syn to anti proportion (P3HT—PMe$_3$ alone)	43.6	43.7	43.7	43.6

least for a DNA fragment of a certain length (43 bp in the case of dsR$_{43}$). Such a result is crucial regarding the use of P3HT—PMe$_3$ in a DNA biosensor: in this supramolecular approach, the P3HT—PMe$_3$ binding with DNA does not drastically perturb the intrinsic DNA dynamics. In the next section, we inspect how the plasticity in the binding modes of P3HT—PMe$_3$ impacts the conformational properties of the different P3HT—PMe$_3$ binding modes.

2.4. A Preserved Geometrical Pattern: *Syn* or *Anti* Blocks

The coiling of P3HT—PMe$_3$ around the DNA helix obviously influences the dynamics of dihedral angles between thiophene units along the backbone. Previously, we showed that the P3HT—PMe$_3$ binding to DNA induces a rearrangement in the distribution of the dihedral angles between thiophene units, from a slight majority of *anti* configurations between thiophene units (for the pure polymer) to a majority of *syn* configurations when it interacts with DNA.[14] The change toward a *syn* P3HT—PMe$_3$ backbone is also observed in each replica of this study. Interestingly, the ratio of *syn*- to *anti*-dihedral angles is very similar between the four replicas (Table 1). The predominance of the *anti* character for the unbound polymer is also confirmed by additional MD simulations of the P3HT—PMe$_3$ alone on a longer time scale, *i.e.*, three replicas of 2 μs and a fourth replica of 10 μs (Table 1). For the polymer alone, the

repartition of the *syn* and *anti* dihedral angles along the MD run time and the sequence is completely random (Supplementary S9). In contrast, the dihedral angles profiles of P3HT—PMe$_3$ when bound to DNA (Figure 6), show distributions of *syn/anti* dihedral angles that are partitioned in large blocks of *syn* dihedral angles separated by shorter blocks of *anti* dihedral angles. Some of the large *syn* blocks (green) and smaller *anti* (red) blocks are remarkably preserved on the μs timescale. The most conserved *syn* domains are gathered in a list, see Supplementary S10. Although the different coiling and positioning of P3HT—PMe$_3$ around dsR$_{43}$ involves distinct P3HT—PMe$_3$ dihedral angle profiles, overall the *syn/anti* ratio and a blocky repartition of *syn* or *anti* angles are clearly a conserved pattern throughout the P3HT—PMe$_3$ binding modes, for the four MD replicas. In the next section, we investigate how the preserved local conformational effects influence the chiroptical properties of P3HT—PMe$_3$.

2.5. Origin of the Induced Chirality

To decipher the origin of the induced CD signals experimentally observed,[12,13] theoretical CD spectra of the polymer conformations (when bound to DNA) were calculated from MD simulations (Supplementary S11, see Methods for computational details). In addition to the handedness of MD snapshots, our simple exciton model permits a spatial mapping along the polymer backbone of the lowest exciton state contributing to CD. In Figure 6 (middle column), we have plotted the location of the first excited state along the polymer backbone (Y-axis, in number of thiophene units) versus MD time (X-axis), using a color code to label the handedness (left or right) of the polymer conformations. On the same figure, we superposed the most conserved *syn* domains (green bands), i.e., segments of five consecutive dihedral angles (pentad) that have a *syn* character more than 90% of the MD simulations. The P3HT—PMe$_3$ binding modes observed in the four replicas lead to different situations not only in terms of left- *vs.* right-handed arrangements, but also in terms of location of excited states along the polymer backbone. For replica R1, the handedness of the polymer is intermediate with around half of the conformations being left-handed. An alternating left-/right-handedness character is observed. The first part of the MD (0 to 1.0 μs) exhibits a right-handed character of the polymer. The position of the lowest excited state along the polymer backbone is definitely not well defined in this part of the MD, in agreement with oscillations in the distance between the DNA and the polymer (Figure 1) and important time- and sequence-alternation between *syn/anti*
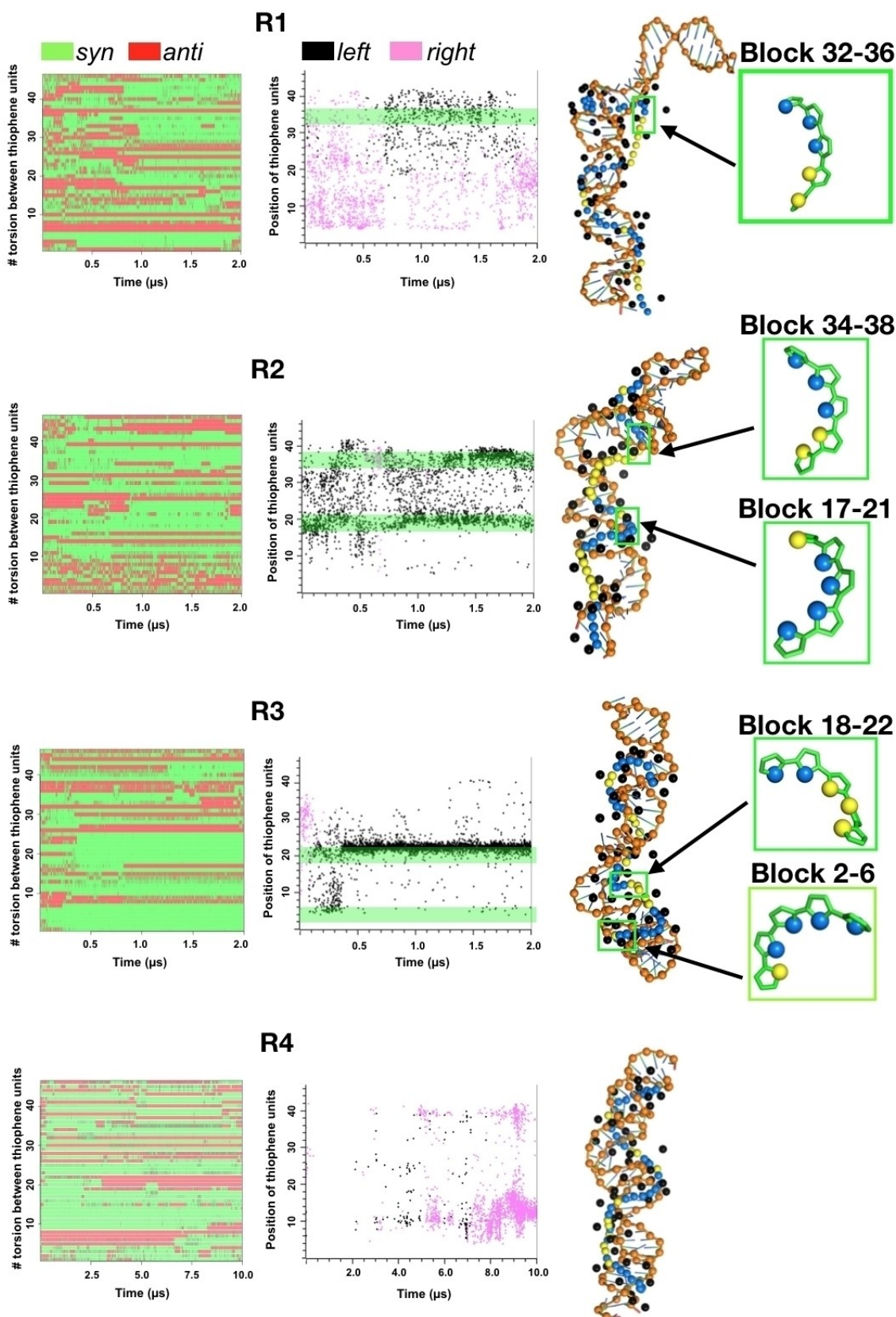
Figure 6. Position of *syn*/*anti* dihedral angles along the polymer backbone versus time (left column) and the respective location of the first excited state in our simplified excitonic model (middle column). Regions of most conserved *syn* domain are representend in green (see Supplementary S10). Right: final MD snapshots of the dsR$_{43}$/P3HT—PMe$_3$ complexes for the four replicas with expanded portion showing the pentad domain. Only the backbone atoms of the P3HT—PMe$_3$ are represented. Sulfur atoms are represented with yellow/blue spheres to label groove-bound thiophene units (blue) and non-adsorbed thiophene units (yellow).

dihedral angles (Figure 6 left column). From 1.0 to 1.5 µs, a left-handed character predominates, associated to a clearer location of the first excited state but also preserved *syn/anti* domains. The most conserved *syn* domain, although conserved only 80% of the time, is located in the section 32–36 (Supplementary S10), in agreement with the position of the first excited state that exhibits a left-handedness between 0.75 and 1.5 µs. A correlation between a left-handed CD signal of the DNA-bound polymer and local appearance of larger *syn* domains is suggested (see highlight in green on R1 in Figure 6). During the last part of R1 MD (1.5 to 2.0 µs), the right-handed character of the polymer is again more pronounced with a scattered location of the first excited state. Both replicas R2 and R3 exhibit a left-handed character (around 3,800 conformations over a total of 4,000 conformations are left-handed) along most of the MD simulation. For R2, the location of the first excited state is alternating between various positions of the polymer backbone during the first part of the MD simulations (from 0 to 1.0 µs), in parallel with important variations of both DNA/polymer distance and end-to-end distance (Figure 1). A highly conserved *syn* pentad is detected in the section 17–21 (in #thiophene units) along the polymer backbone (see Figure 6, structure on the right), related to the location of the first excited state in the second part of the MD (1–2 µs). The second most conserved *syn* domain is located on the section 34–38 (see Supplementary S10), covering also the location of the first exciton. In replica R3, very well-defined and conserved *syn* domains are detected in the sections 2–6 and 18–22 of the polymer, the first excitation state being located in these central and larger *syn* domains which show a left-handed character. Replica R3 presents the fastest conformational stabilization of the complex, DNA/polymer end-to-end polymer distances (Figure 3) converging after few hundred nanoseconds. Both R2 and R3 highlight a clear correlation between left-handedness of the polymer CD signal and a local conformational effect, *i.e.*, the presence of highly conserved *syn* clusters along the P3HT–PMe₃ backbone. For R4, the situation is different with only 147 out of 20,000 conformations that were left-handed. In this replica, we did not detect highly conserved *syn* clusters. The correlation between excitonic states position and handedness is therefore less evident if the polymer conformations are right-handed. Coming back to the calculated CD spectra in Supplementary S11, simulations R1, R2, and R3 show an averaged $(-/+)$ Cotton effect (i.e. left-handed), in agreement with the sign of the experimental CD signals (Supplementary S12). This signature is related to left-handed *syn* segments of the polymer where the lowest exciton state is localized, as depicted in Figure 6. In contrast, a $(+/-)$ signal is averaged for R4 simulations, for which we did not observe highly conserved *syn* clusters. It is worth stressing that left-handed character of the polymer is related to a bathochromic displacement (red-shift) of the lowest electronic transitions in our calculations, *versus* a blue-shift effect for right-handed helices. Consequently, left-handed conformations of the polymer should dominate the CD spectra at low energy. The energetics of the first exciton state that contributes to the low-energy part of the CD spectra thus reflects local geometrical conformational effects (a large

proportion of *syn* domains) and the left-handedness of the P3HT–PMe₃ conformations. The MD simulations thus shed light on the origin of the left-handed CD signals when the polymer interacts with DNA. The $(-/+)$ signature observed in the experimental induced CD signals (Supplementary S12) is therefore related to the prevalence of left-handed conformations of the polymer dominating the CD spectra at low energy.

2.6. Macromolecular Shapes of the Polymer Chain in the Complex

The different partitioning in blocks of P3HT–PMe₃ *syn/anti* dihedral angles reflects various polymer macromolecular shapes. We used the continuous symmetry measure to explore the details of these shapes. For this purpose, a 47-bead model was applied to represent the S atoms along the polymer backbone, meaning that we measured the symmetry of the polymer shape in this section. The 47-bead model was then divided by a running ruler approach[33] into segments of 5 beads. $S(C_2)$ with the identity permutation was used to measure the deviation from linearity of each segment (see Methods for computational details).

For example, the shapes of two 5-beads segments with respect to their nearest linear structures for small and high CSM values are reported in Supplementary S13. To follow the dynamics of polymer shapes, we draw a heat map of $S(C_2)$ of the 5-beads segments as a function of time, where the CSM is represented by a color scale (Figure 7) allowing the localization of elongated (low CSM) and twisted (high CSM) polymer segments.

A striking characteristic of these maps is the horizontal stripes of similar colors (similar CSM values). That is, while the
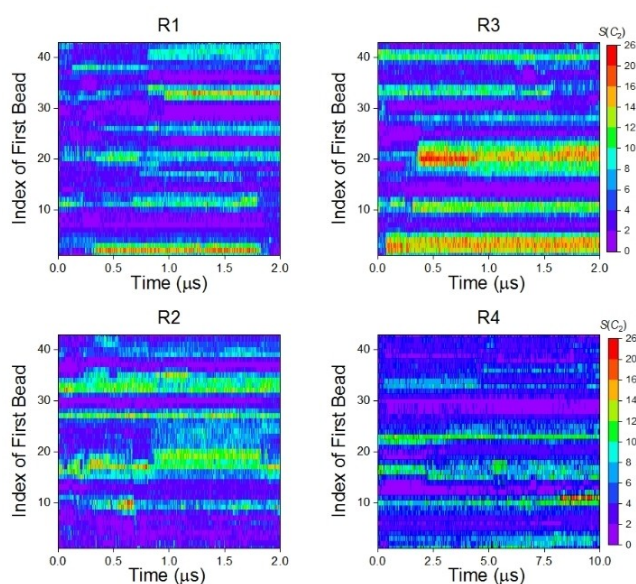

Figure 7. CSM heat map – $S(C_2)$ per 5-beads segment along the polymer as a function of time. CSM levels are represented by the color scale.

ChemPhysChem

Articles
doi.org/10.1002/cphc.202000630

Chemistry
Europe
European Chemical
Societies Publishing

process is dynamical, the structural nature of most of the segments does not change as a function of time, particularly for R3 and R4. On the maps shown in Figure 7, we can observe that twisted fragments (large and conserved green to red stripes) are encountered for replica R1 and R3. Such stripes correspond to regions of *syn* dihedral angles detected in the previous sections (see Figure 6 left and center columns). Therefore, the presence of *syn* dihedral angles is related to the most important breaking of local C_2-symmetry (according to the identity permutation), i.e., the most bent regions of the polymer, that influences the global shape of the entire polymer.

A similar CSM analysis was performed on the whole polymer bead-model of each replica. Supplementary S14 exemplifies the shape of the polymer with respect to its nearest symmetric structure for two cases: a highly twisted structure with high CSM value and a relatively elongated structure with low CSM value. As can be seen from Figure 8, the shape of the polymer for all four replicas fluctuates in time, with R2 and R4 being more linear (smaller $S(C_2)$ values) as compared with R1 and R3. Such a result can be related to the P3HT–PMe$_3$ end-to-end distance (Figure 3), varying from ~60–70 Å for replicas R1 and R3 (highly coiled polymers) to ~100 Å for replicas R2 and R4 (extended polymers). However, the end-to-end distances rely only on two points in space whereas the CSM is a global parameter of the whole geometry. The CSM analysis adds a description of the polymer shape fluctuation in the 4 replicas. In this way, R1 and R3 (coiled polymers, lower end-to-end distances) are also related to binding modes in which more intense peaks in RMSF profiles are encountered (Figure 5) in comparisons to R2 and R4 (elongated polymers, higher end-to-end distances). Supplementary S15 shows box plots of the distributions of $S(C_2)$ with the identity permutation as well as the full distributions. It can be seen that the shape of replicas R3 and R4 are relatively confined around their median CSM values while replicas R1 and especially R2 spread over a larger range of CSM values. In other words, the shape of the R2 replica

changes considerably along the MD simulation while in replica R4 it is more conserved, showing a wider range of polymer binding modes.

3. Conclusions

In this study, we show that the supramolecular self-assembly between a cationic polythiophene and double-stranded DNA oligonucleotide is mainly driven by electrostatic interactions, while van der Waals interactions play an important role by optimizing the position of thiophene units in the DNA grooves. The polymer can approach DNA in several ways, leading to a multiplicity of wrapping modes around DNA. These modes are distinguished by a sliding motion as well as different levels of coiling along the DNA helix axis. Importantly, the binding modes impact differently the dynamics of both DNA and polymer partners. The polymer can adopt many types of conformations, thanks to its intrinsic flexibility compared to DNA, whereas, the DNA backbone is more conserved. Due to the different wrapping modes of P3HT–PMe$_3$ around dsR$_{43}$, distinct P3HT–PMe$_3$ dihedral angle profiles were observed. A common feature of these profiles is apparent: a blocky repartition of *syn* or *anti* angles persists after the binding of P3HT–PMe$_3$. Such a preserved geometrical pattern of *syn* or *anti* blocks allow us to better understand the chiroptical properties of the complexes, in particular the induced CD signals in the spectral region of the polymer. The low-energy part of the CD spectra is correlated to a local geometrical effect on dihedral angles backbone (a large proportion of conserved *syn* domains) and the unexpected left-handedness of the polymer backbone. In addition, we used the continuous symmetry measures (CSM) to study the various shapes of the polymer fragments along the polymer conformations, i.e. by partitioning the polymer with 5-beads fragments at a coarse-grained resolution. We detected that the most twisted segments break the local C_2-symmetry and create bent regions that match the regions of large and conserved *syn* dihedral angles. These are responsible for the low-energy part of the CD spectra. The consequences of local C_2-symmetry breaking or, in other words, local deviation from linearity, are important regarding the overall level of elongation of the polymer's shape, as diverse levels of elongation are observed according to the binding modes. Altogether, this theoretical study points out the hierarchical aspect of the process where chiral organization of the achiral polymer is induced by DNA binding, and highlights the relationship with CD signals. In the future, our multiscale strategy could be applied to various environmental conditions (e.g., solvent, counter-ion type and salt concentrations) to understand how such parameters influence the recognition modes between DNA and achiral polymers and the chiral properties of such self-assemblies. For instance, DNA can form condensed phases at high salt concentration, a state that would impact the self-assembly between DNA and cationic polymers in comparisons to low ionic strength in dilute solutions. In addition, our approach could be extended to other supramolecular systems
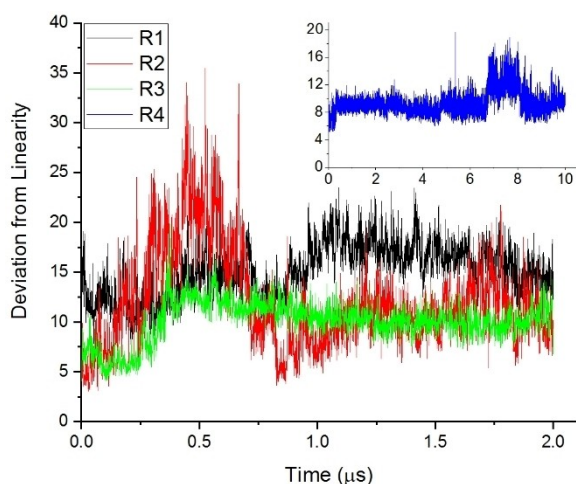

Figure 8. Deviation from linearity of the four replicas as a function of time as estimated by $S(C_2)$ with the identity permutation. Black: R1. Red: R2. Green: R3. Blue: R4 (inset).

ChemPhysChem

Articles
doi.org/10.1002/cphc.202000630

**Chemistry
Europe**
European Chemical
Societies Publishing

where chiral induction processes are important, for instance to develop chiroptical (bio)sensors.

Methods Section

Molecular Dynamics Protocol

Four replicas of the same MD simulation were performed to investigate the possible various binding modes of the polymer around dsR$_{43}$. The starting conformations of P3HT–PMe$_3$ were built within the LEaP program, the molecular builder module of AMBER16.[34] The starting B-DNA conformations were built within the NAB (Nucleic Acid Builder) AmberTools16 module.[34] P3HT–PMe$_3$ was oriented perpendicular to the DNA axis, to avoid any bias of the polymer pre-orientation polymer with DNA. The starting inter-distance between the DNA and P3HT–PMe$_3$ was superior to 15 Å to prevent any bias of the DNA/CPT recognition mode due to initial intermolecular interactions before the MD simulations. All simulations were achieved with the GPU version of AMBER16 package. The parmBSC1 force field (FF) was used to consider DNA dynamics[35], and GAFF 2.1 FF was employed for the polymer.[36] A re-parameterization of the bithiophene torsion potential estimated at the MP2/cc-pvdz level and charge calculations were performed in a previous work.[37] A total number of 37 Na$^+$ counter ions were added to neutralize the net charge of the molecular systems, i.e., 84–47 = 37 ions with 47 positive charges from P3HT–PMe$_3$ and 84 negative charges from the DNA, that is composed of 84 phosphate groups, dangling phosphate for 5' terminal residues being not considered in the AMBER model of DNA.[35] Although environment conditions such as the type of counter-ion or salt concentration can strongly influence DNA dynamics[38], as well as the recognition modes between DNA and P3HT–PMe$_3$, the MD simulations were performed in an implicit solvent using the Generalized Born solvation model[39] in order to conduct a relevant conformational sampling on the microsecond regime as well as to consider several replicas with reasonable computational resources. A MD protocol in three steps (minimization-heating/equilibration-production) was applied. A first minimization was performed to relax the ionic atmosphere of the molecular systems. For this, a steepest descent minimization of 1,000 steps was carried out with harmonic positional constraints applied to the DNA and the polymer with a force constant of 25 kcal.mol^{-1}.Å^{-2}, followed by a conjugate gradient minimization in 9,000 steps. A second 10,000 steps minimization stage, still subdivided in 1,000 steepest descent steps followed by 9,000 conjugate gradient steps, was carried out without any restraints on the molecular system to relax the entire molecular system. The heating-equilibration stage was then performed for 2 ns in the NVT ensemble. A 2 fs time step was used as the SHAKE algorithm constrained the length of covalent bonds that involved hydrogen atoms. Harmonic positional constraints were applied to DNA and P3HT–PMe$_3$ with a force constant of 10.0 kcal mol^{-1} Å^{-2} to avoid any deformation of DNA or P3HT–PMe$_3$ due to the heating process. First, the system was heated from 0 to 300 K over 1 ns. The temperature was then maintained for another 1 ns with the Langevin thermostat and a collision frequency of 1.0 ps^{-1}. Finally, the MD production stage was performed for a simulation time of 2.0 μs for each replica R1, R2 and R3. A longer MD production, i.e., 10 μs, was considered for replica R4. During the production stage, a 2 fs time step was also employed as well as the Langevin thermostat with a 1.0 ps^{-1} collision frequency.

Analysis of MD Production Phase

MD snapshots were recorded each 0.5 ns, resulting in a 4,000 frames trajectory at the end of the production MD (20,000 frames for replica R4). Extraction of MD snapshots and trajectory analyses were done with the CPPTRAJ AmberTool16 module. Root-Mean-Square Deviations (RMSDs) were calculated in reference to the heavy atoms, with omissions of hydrogen atoms. RMSFs of a residue is calculated from the average displacement of the set of heavy atoms of each residue, giving a measure of the flexibility along the sequence [Eq. (1)]:

$$RMSF = \sqrt{\frac{1}{T} \sum_{t=1}^{T} (x_i(t) - \langle x_i \rangle)^2} \tag{1}$$

where T is the duration of the simulation (number of steps), $x_i(t)$ denotes the position (coordinates) of atom i at time t and $\langle x_i \rangle$ is the averaged position of atom i. For rendering of MD snapshots, we used PyMOL 2.2.0.[40] In-house R scripts were used for profiles of dihedral angle between thiophene moieties and for statistical analyses of the MD trajectories.[41]

CD Spectra Calculations

To gain insight into the origin for the chiral response of P3HT–PMe$_3$ interacting with DNA, we solved a simple Frenkel exciton model including both through-space and through-bond interactions between the thiophene monomer units of P3HT–PMe$_3$ CD spectra were then simulated on the basis of the calculated rotational strengths from the ground state to the excitonic state manifold. Results of the model have been validated against semi-empirical quantum-chemical excited-state calculations performed on selected frames, as detailed in a previous study.[14]

Continuous Symmetry Measures

The input for a CSM calculation is the original structure and the desired point group G. The algorithm is based on searching for all the structures with the same set of atoms and connectivity that belong to G, and calculating the distance of these structures to the original molecule.[18] The structure that gives the minimum distance is used as the reference structure with respect to which the measure is calculated. Mathematically, the distance between the two structures is defined by Equation (2):

$$S(G) = 100 \cdot \frac{min\left[\sum_{k=1}^{N} |\mathbf{Q}_k - \mathbf{P}_k|^2\right]}{\sum_{k=1}^{N} |\mathbf{Q}_k - \mathbf{Q}_0|^2} \tag{2}$$

where $\mathbf{Q}$ is the original structure with N atoms and coordinates $\{\mathbf{Q}_k, k = 1, \ldots, N\}$, $\mathbf{Q}_0$ is the coordinate vector of the geometric center of mass of the original structure $\mathbf{Q}$. $\mathbf{P}$ is a symmetric structure with coordinates $\{\mathbf{P}_k, k = 1, \ldots, N\}$ that belongs to G and maintains the same number of vertices and connectivity as $\mathbf{Q}$. As defined in Eq. (2), $\mathbf{P}$ is chosen such that it minimizes the sum of the square distances between the vertices of $\mathbf{Q}$ and the vertices of $\mathbf{P}$. The CSM is independent of the position, orientation, and size of the original structure. $S(G)$ obtains the value of 0 when the original structure belongs to the G point group, and the maximum value of 100 when $\mathbf{P}$ reduces to the center of mass, $\mathbf{Q}_0$. The CSM becomes a continuous measure of linearity by setting $G = C_n$ with $n \geq 2$ and using the identity permutation in which each atom is permuted with itself. The code that was implemented in this study can be accessed online at http://csm.ouproj.org.il/.

ChemPhysChem

Articles
doi.org/10.1002/cphc.202000630

Chemistry Europe
European Chemical
Societies Publishing

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

[1] a) J. Wang, X. Qu, *Nanoscale* **2013**, *9*, 3589–3600; b) J. J. Xu, W. W. Zhao, S. Song, C. Fan, H. Y. Chen, *Chem. Soc. Rev.* **2014**, *43*, 1601–1611; c) Y. Wu, R. D. Tilley, J. Gooding, *J. Am. Chem. Soc.* **2019**, *141*, 1162–1170.

[2] a) S. M. Borisov, O. S. Wolfbeis, *Chem. Rev.* **2008**, *108*, 423–461; b) R. Radhakrishnan, I. I. Suni, C. S. Bever, B. D. Hammock, *ACS Sustainable Chem. Eng.* **2014**, *2*, 1649–1655; c) B. Jurado-Sanchez, *Biosensors* **2018**, *8*, E66.

[3] a) W. C. Tse, D. L. Boger, *Chem. Biol.* **2004**, *11*, 1607–1617; b) S. Bhaduri, N. Ranjan, D. P. Arya, *Beilstein J. Org. Chem.* **2018**, *14*, 1051–1086.

[4] a) B. Rafique, M. Iqbal, T. Mehmood, M. A. Shaheen, *Sens. Rev.* **2019**, *39*, 34–50; b) K. Vikrant, N. Bhardwaj, S. K. Bhardwaj, K.-H. Kim, A. Deep, *Biomaterials* **2019**, *214*, 119215.

[5] a) K. Wang, M. E. Rodgers, D. Toptygin, V. A. Munsen, L. Brand, *Biochemistry* **1998**, *37*, 41–50; b) B. J. Anderson, C. Larkin, K. Guja, J. F. Schildbach, *Methods Enzymol.* **2008**, *450*, 253–272.

[6] a) M. J. Hannon, *Chem. Soc. Rev.* **2007**, *36*, 280–295; b) P. Samorì, F. Biscarini, *Chem. Soc. Rev.* **2018**, *47*, 4675–4676; c) J. Wang, F. Lv, L. Liu, Y. Ma, S. Wang, *Coord. Chem. Rev.* **2018**, *354* 135–154; d) Q. Wang, Z. Li, D.-D. Tao, Q. Zhang, P. Zhang, D.-P. Guo, Y.-B. Jiang, *Chem. Commun.*, **2016**, *52*, 12929–12939; e) C. König, R. Skanberg, I. Hotz, A. Ynnerman, P. Norman, M. Linares, *Chem. Commun.* **2018**, *54*, 3030–3033.

[7] a) A. Elouahabi, J. M. Ruysschaert, *Mol. Ther.* **2005**, *11*, 336–347; b) C. Tros de Ilarduya, Y. Sun, N. Düzgünes, *Eur. J. Pharm. Sci.* **2010**, *40*, 159–170.

[8] L. Strekowski, B. Wilson, *Mutat. Res.* **2007**, *623*, 3–13.

[9] F. Xia, X. Zuo, R. Yang, Y. Xiao, D. Kang, A. Vallée-Bélisle, X. Gong, A. J. Heeger, K. W. Plaxco, *J. Am. Chem. Soc.* **2010**, *132*, 1252–1254.

[10] M. N. Antipina, I. Schulze, M. Heinze, B. Dobner, A. Langner, G. Brezesinski, *ChemPhysChem* **2009**, *10*, 2471–2479.

[11] V. Tudor, C. Cojocaru, A. Rotaru, G. Pricope, M. Pinteala, L. Clima, *Int. J. Mol. Sci.* **2017**, *18*, 1291.

[12] a) U. Lächelt, E. Wagner, *Chem. Rev.* **2015**, *115*, 11043–11078; b) M. Surin, *Polym. Chem.* **2016**, *7*, 4137–4150; c) A. Hall, U. Lächelt, J. Bartek, E. Wagner, S. M. Moghimi, *Mol. Ther.* **2017**, *25*, 1476–1490.

[13] a) J. Rubio-Magnieto, A. Thomas, S. Richeter, A. Mehdi, Ph. Dubois, R. Lazzaroni, S. Clément, M. Surin, *Chem. Commun.* **2013**, *49*, 5483–5485; b) J. Rubio-Magnieto, E. G. Azene, J. Knoops, S. Knippenberg, C. Delcourt, A. Thomas, S. Richeter, A. Mehdi, Ph. Dubois, R. Lazzaroni, D. Beljonne, S. Clément, M. Surin, *Soft Matter* **2015**, *11*, 6460–6471.

[14] M. Fosse-pré, M. E. Trévisan, V. Cyriaque, R. Wattiez, D. Beljonne, S. Richeter, S. Clément, M. Surin, *ACS Appl. Bio Mater.* **2019**, *2*, 2125–2136.

[15] a) E. H. Lee, J. Hsin, M. Sotomayor, G. Comellas, K. Schulten, *Structure* **2009**, *17*, 1295–1306; b) M. Karplus, R. Lavery, *Isr. J. Chem.* **2014**, *54*, 1042–1051.

[16] a) M. F. Lucas, I. Cabeza de Vaca, R. Takahashi, J. Rubio-Martinez, V. Guallar, *Biophys. J.* **2014**, *106*, 421–429; b) C. Maffeo, J. Yoo, J. Comer, D. B. Wells, B. Luan, A. Aksimentiev, *J. Phys. Condens. Matter.* **2014**, *26*, 413101; c) N. Holmgaard List, J. Knoops, J. Rubio-Magnieto, J. Idé, D. Beljonne, P. Norman, M. Surin, M. Linares, *J. Am. Chem. Soc.* **2017**, *139*, 14947–14953; d) Y. Wang, C. S. Lupala, H. Liu, X. Lin, *Curr. Top. Med. Chem.* **2018**, *18*, 2268–2277; e) Y. Guan, R. Shi, X. Li, M. Zhao, Y. Li, *J. Phys. Chem. B* **2007**, *111*, 7336–7344; M. Sayed, B. Krishnamurthy, H. Pal, *Phys. Chem. Chem. Phys.* **2016**, *18*, 24642–24653; f) A. Tafvizi, L. A. Mirny, A. M. van Oijen, *ChemPhysChem* **2011**, *12*, 1481–1489; g) W. Zheng, A. V. Vargiu, M. A. Rohrdanz, P. Carloni, C. Clementi, *J. Chem. Phys.* **2013**, *139*, 145102.

[17] D. A. Kondinskaia, A. Y. Kostritskii, A. M. Nesterenko, A. Y. Antipina, A. A. Gurtovenko, *J. Phys. Chem. B* **2016**, *120*, 6546.

[18] G. Alon, I. Tuvi-Arad, *J. Math. Chem.* **2018**, *56*, 193–212.

[19] M. Pinsky, C. Dryzun, D. Casanova, P. Alemany, D. Avnir, *J. Comput. Chem.* **2008**, *29(16)*, 2712–2721.

[20] H. Zabrodsky, S. Peleg, D. Avnir, *J. Am. Chem. Soc.* **1992**, *114*, 7843–7851.

[21] H. Zabrodsky, D. Avnir, *J. Am. Chem. Soc.* **1995**, *117*, 462–473.

[22] M. Pinsky, D. Avnir, *Inorg. Chem.* **1998**, *37*, 5575–5582.

[23] P. Alemany, D. Casanova, S. Alvarez, C. Dryzun, D. Avnir, *Continuous symmetry measures: a new tool in quantum chemistry. In Reviews in Computational Chemistry*; A. L. Parrill, K. B. Lipkowitz, Eds.; Wiley, **2017**; Vol. 30, pp 289–352.

[24] I. Tuvi-Arad, T. Rozgonyi, A. Stirling, *J. Phys. Chem. A* **2013**, *117(48)*, 12726–12733.

[25] S. Keinan, D. Avnir, *J. Am. Chem. Soc.* **2000**, *122*, 4378–4384.

[26] A. F. Zahrt, S. E. Denmark, *Tetrahedron.* **2019**, *75*, 1841–1851.

[27] I. Ergaz, R. A. Toscano, G. Delgado, A. Steinberg, R. Glaser, *Cryst. Growth Des.* **2008**, *8*, 1399–1405.

[28] E. F. Sheka, B. S. Razbirin, D. K. Nelson, *J. Phys. Chem. A* **2011**, *115*, 3480–3490.

[29] D. Yogev-Einot, D. Avnir, *Tetrahedron: Asymmetry* **2006**, *17*, 2723–2725.

[30] A. W. Kaspi-Kaneti, I. Tuvi-Arad, *Organometallics.* **2018**, *37*, 3314–3321.

[31] I. Tuvi-Arad, A. Stirling, *Isr. J. Chem.* **2016**, *56*, 1067–1075.

[32] A. Carreras, E. Bernuz, X. Marugan, M. Llunell, P. Alemany, *Chem. Eur. J.* **2019**, *25*, 673–691.

[33] M. Bonjack-Shterengartz, D. Avnir, *PLoS One* **2017**, *12*, e0180030.

[34] D. A. Case, T. E. Cheatham III, T. Darden, H. Gohlke, R. Luo, K. Merz Jr, A. Onufriev, C. Simmerling, B. Wang, R. J. Woods, *J. Comput. Chem.* **2005**, *26*, 1668–1688.

[35] I. Ivani, P. D. Dans, A. Noy, A. Peŕez, I. Faustino, A. Hospital, J. Walther, P. Andrio, R. Goni, A. Balaceanu, G. Portella, F. Battistini, J. L. Gelpi, C. Gonzales, M. Vendruscolo, C. A. Laughton, S. A. Harris, D. A. Case, M. Orozco, *Nat. Methods* **2016**, *13*, 55–58.

[36] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, *25*, 1157–1174.

[37] J. Knoops, J. Rubio-Magnieto, S. Richeter, S. Clément, M. Surin in *Conjugated Polymers and Oligomers: Structural and Soft Matter Aspects* (Ed.: M. Knaapila), World Scientific Publishing, **2018**, Vol. 9, Chapter 5, pp 139–157.

[38] a) F. Mocci, A. Laaksonen, *Soft Matter* **2012**, *8*, 9268–9284; b) M. Pasi, J. H. Maddocks, R. Lavery, *Nucleic Acids Res.* **2015**, *43*, 2412–2423.

[39] G. Hawkins, C. Cramer, D. Truhlar, *Chem. Phys. Lett.* **1995**, *246*, 122–129.

[40] *The Pymol Molecular Graphics System, version 2.0*; Schrödinger LLC; www.pymol.org.

[41] R Core Team. R; *R Foundation for Statistical Computing*: Vienna, Austria, 2013; www.R-project.org.