ACS OMEGA

Open Access

http://pubs.acs.org/journal/acsodf

Article

Sequence-Selective Recognition of the d(GGCGCC)$_2$ DNA Palindrome by Oligopeptide Derivatives of Mitoxantrone. Enabling for Simultaneous Targeting of the Two Guanine Bases Upstream from the Central Intercalation Site in Both Grooves and along Both Strands

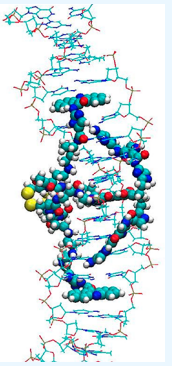
Nohad Gresh,* Alberto Ongaro, Luc Demange, Giuseppe Zagotto, and Giovanni Ribaudo

Cite This: *ACS Omega* 2024, 9, 42309−42328

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The d(GGCGCC)$_2$ palindrome is encountered in several oncogenic and retroviral sequences. In order to target it, we previously designed several oligopeptide derivatives of the mitoxantrone and ametantrone anticancer intercalators. These have two arms with a cationic side-chain in the major groove, each destined to bind along each strand O_6/N_7 of the two successive guanine bases (G1-G2/G1′-G2′) upstream from the central anthraquinone intercalation site. We retained from a previous study (El Hage et al., 2022) a tris-intercalating molecule with two outer 9-aminoacridine (9-AA) intercalators, denoted as **III**. We sought enhancements in both affinity and selectivity by simultaneously targeting the minor groove of the extracyclic $-NH_2$ groups of these bases and G4-G4′ of the intercalation site. We considered derivatives of distamycin, having each pyrrole ring replaced by an imidazole to act as an in-register electron acceptor from the $-NH_2$ group of a target guanine. We substituted the C_6 and C_7 carbons of anthraquinone, or the C_8 and C_9 ones of anthracycline, by an (imidazole-amide)3 chain. Four different derivatives of **III** were designed with different connectors to the anthraquinone/anthracycline and 9-AA. Polarizable molecular dynamics simulations of their complexes with a double-stranded DNA 18-mer with a central d(C GGGC GCCC G)$_2$ palindrome sequence showed in-register minor groove binding to $-NH_2$ of G1-G$_2$/G1′-G2′ to coexist with major groove recognition of O_6/N_7. Up to 12 H-bonds could be stabilized in the minor groove coexisting with four bidentate interactions of the alkyl diammonium moieties in the major groove. Since there is no mutual interference, the binding enthalpies, ΔH, contributed by each groove could add up and enable significant enhancements of the affinity constants. As was the case for their Lys precursor, these derivatives are amenable to chemical syntheses and in vitro and in vivo tests, for which the present results provide an incentive. The construction of derivatives **III-A**−**III-D** is modular. For in vitro experiments, this should enable unraveling the most important structural elements to further optimize both ΔH and $T\Delta S$ and sequence selectivity and how this could translate to in vivo tests.



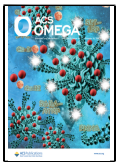
1. INTRODUCTION

The palindromic d(GGCGCC)$_2$ sequence is encountered in numerous noncoding, regulatory DNA sequences. Examples are its occurrence as a recognition site for seven Type II restriction enzymes[1] and as a *Smad*-binding element.[2] It is encountered twice, at a 50 base-pair interval, in the mouse *Id2* gene which is a target for bone morphogenetic proteins belonging to the transforming growth factor family. The two hexamers in this gene bracket two Sp1 hexamer binding sites and one p53 heptamer binding site.[3] The GGCGCC sequence is part of the Long Terminal Repeat of the HIV-1 retrovirus.[4] It is involved in retrovirus integration by forming a double-stranded complex with a complementary DNA sequence from the host that binds the HIV-1 polymerase.[5,6]

A compelling motivation for the present study is its presence in *Alu* repeats.[7] *Alu* sequences are transposon polynucleotides that can detach from the genome and integrate into other DNA sequences but also into extrachromosomal circulating DNA (ecc-DNA). The existence of ecc-DNA was discovered in 1982.[8] It is now identified as a major cause in tumor progression and tumor resistance to chemotherapy.[9,10] The occurrence of an *octa* nucleotide palindrome d(CGGGCGCC)$_2$ rather than the shorter hexanucleotide is worth noting at three *Alu* insertion sites: 250−257 at chr7, 223−230 at chr8, and 290−298 at chr11.[7]

The occurrence of this palindrome in retroviruses, oncogenes, and ecc-DNA is a strong incentive for the search of ligands able to selectively complex it among the 2080 unique hexanucleotide sequences. We designed several oligopeptide derivatives of the antitumor intercalating drugs mitoxantrone

Received: May 30, 2024
Revised: August 2, 2024
Accepted: August 9, 2024
Published: October 3, 2024



(MTX) and ametantrone (AMT). The two G bases on both strands upstream from the central d(CpG)$_2$ intercalation site were targeted in the major groove by an arginine[11] or a lysine[12,13] side-chain. There are two experimental results supporting the design of such derivatives. A Lys derivative of AMT was experimentally shown to result into a 12° larger stabilization (ΔT_m) against thermal melting of this palindrome than a random sequence, while the parent AMT stabilized by equal amounts both the palindrome and a random sequence with a d(CpG)$_2$ intercalation site.[13] This implies that the short side-chain of AMT, $-CH_2-CH_2-NH_2^+-CH_2-CH_2-OH$, cannot reach out to the bases beyond the intercalation site. An AMT Lys-derivative recently synthesized in two of our Laboratories[14] proved active in cellular assays against tumor cells. While it was ten times less active than MTX against tumor cells, it was 50 times less toxic than it on healthy cells. This led us to the search for novel derivatives with augmented affinities.[15] The Lys side-chains on both arms were replaced by an alkyl diammonium entity $-CH_2-CH_2-NH_2^+-CH_2-CH_2-NH_2^+-CH_3$. The first and second ammonium groups were destined to target in-register, by bidentate interactions, O_6/N_7 of G2 and G1 bases, respectively. To further augment the binding affinities, the last methyl group was extended by an amide connected to an imidazole-amide itself linked to an acridine. This led to compound **III** of,[15] a tris-intercalating ligand now destined to target the decanucleotide sequence d(CGGGCGCCCG)$_2$, ie one site out of 524 800 ones.

Are there means to further augment both the affinities and the selectivity of compounds related to **III**? The earliest DNA groove binders are minor groove binding ligands. The first ones discovered are naturally occurring compounds, netropsin and distamycin.[16] Netropsin has an (amide-pyrrole)2 motif capped by guanidinium on one end and amide-amidinium on the other. Distamycin has an (amide-pyrrole)3 motif terminated by an alkylamidinium group. The extracyclic -NH_2 on guanine in the minor groove exerts a repulsive electrostatic interaction on these cationic entities, resulting in a much greater attraction by the AT-rich sequences than by the GC-rich ones for both ligands. Their crescent-shaped shape enables H-bond donation by three successive amide-NH groups to N_3/O_2 to three successive A/T bases on one strand.

The preferential recognition of guanine rather than adenine or thymine could occur following removal of the cationic ends and replacement of the pyrrole by imidazole, which could act, through its deprotonated nitrogen, as an H-bond acceptor from the extracyclic $-NH_2$ of guanine. It was thus shown earlier that such a replacement could revert from an A/T to a G preferential recognition at the site of replacement.[17−24]

Polyamide conjugates of polyamide-pyrrole-imidazole could also be grafted onto an intercalator. This was achieved by Dervan et al. with a protonated aminoacridine[25] and by Perrée-Fauvet al. with a tricationic porphyrin intercalator.[26,27] The ligand designed in[25] has two successive lexitropsin arm connected by an alkylamine, enabling the second arm to fold back and bind in the minor groove to the DNA strand complementary to the one bound by the first arm. By contrast, the ligands designed in the present study have their imidazole-amide arms (denoted as im3 subsequently) expanded over both sides of the central intercalation site rather than folded over one side.

The hairpin polyamide-intercalator conjugate reported in ref 25 had three successive imidazole-amides at its end, and each acted as a H-bond acceptor from one of the three successive guanines in the d(GGGTA) sequence.

This raises the following point. As an alternative to major groove targeting, could not G1-G2/G1′-G2′ targeting by anthraquinone derivatives be enabled in the minor groove by in-register H-bonds between the imidazole nitrogens of grafted polyamide-imidazole arms? If appropriate connectors to the anthraquinone ring were found, should such interactions not coexist with those taking place in the major groove?

Several "threading" intercalators were reported previously. Although this list is not limitative, it includes nogalamycin,[28] binuclear Ru(II) complexes,[29−31] polyamide amidine anthraquinone Pt(II) complexes,[32] imidazoyl naphtalene diimide intercalators,[33] and several bis- and poly naphthalene diimide polyintercalators.[34−36] The latter class of compounds can recognize the major and minor grooves in an alternating fashion. By contrast, simultaneous O_6/N_7 major groove and $-NH_2$ minor groove targeting of four bases, G1-G2/G1′-G2′ appears to have no precedent.

We denote by (im)3 a sequence of three imidazole (im) moieties connected by formamides, Im-NHCO-Im-NHCO-Im-NH_2. This work is organized as follows. A rationale for the design of the derivatives is shortly presented. The first section of results focuses on tris-intercalators, which were designed in succession and have two (im)3 arms in the minor groove. The second section focuses on monointercalators derived from the monointercalating compound **II** from ref 15, which is terminated at the alkyl-diammonium side-chains, thus devoid of the aminoacridine rings of **III**. The possible onset of cooperativity effects and proposals to enhance the binding affinity in the minor groove are then briefly mentioned. The most salient results are then summarized, along with prospects of validation, in the Conclusion and Perspectives section.

2. RATIONALE FOR THE DESIGN OF DERIVATIVES III-A−III-D

With the help of computer graphics and model building, it was first found that a simple $-CH_2-CH_2-$ connector between one C *ortho* to the MTX hydroxyl group and the C_2 carbon of the first imidazole enabled, prior to energy-minimization (EM) and MD, for the sought-for proximity between the deprotonated N of the second and third imidazoles and $-NH_2$ of G1 and G2, but also, on the opposite (primed) strand, between that of the first imidazole to $-NH_2$ of G4′, a 5′ guanine of the intercalation site. The starting H-bond distances were 2.3 Å. These are slightly larger than the known range of optimized H−N(imidazole) distances from the reported quantum chemical calculations. The subsequent rounds of unrestrained MD should nevertheless enable the recovery of appropriate DNA-ligand H-bond distances.

Furthermore, the likely stabilization of three additional H-bonds with the cytosine O_2 atoms was uncovered at this stage. The first was between the C3 base, H-bonded to G4′ in the intercalation site, and the amide NH following im1. The second and third involved two successive C bases on the primed strand: between C6′, H-bonded to G1, and the amide NH following im2, and between C7′, a 3′ cytosine of the acridine intercalation site, and the terminal $-NH_2$ group. The same proximity was found between the corresponding bases of the primed and unprimed strands (respectively) and the second (im)3 arm grafted on the C atom *ortho* to the other MTX hydroxyl. These would amount to a total of six H-bonds contributed by each (im)3 arm.

https://doi.org/10.1021/acsomega.4c05099
ACS Omega 2024, 9, 42309−42328

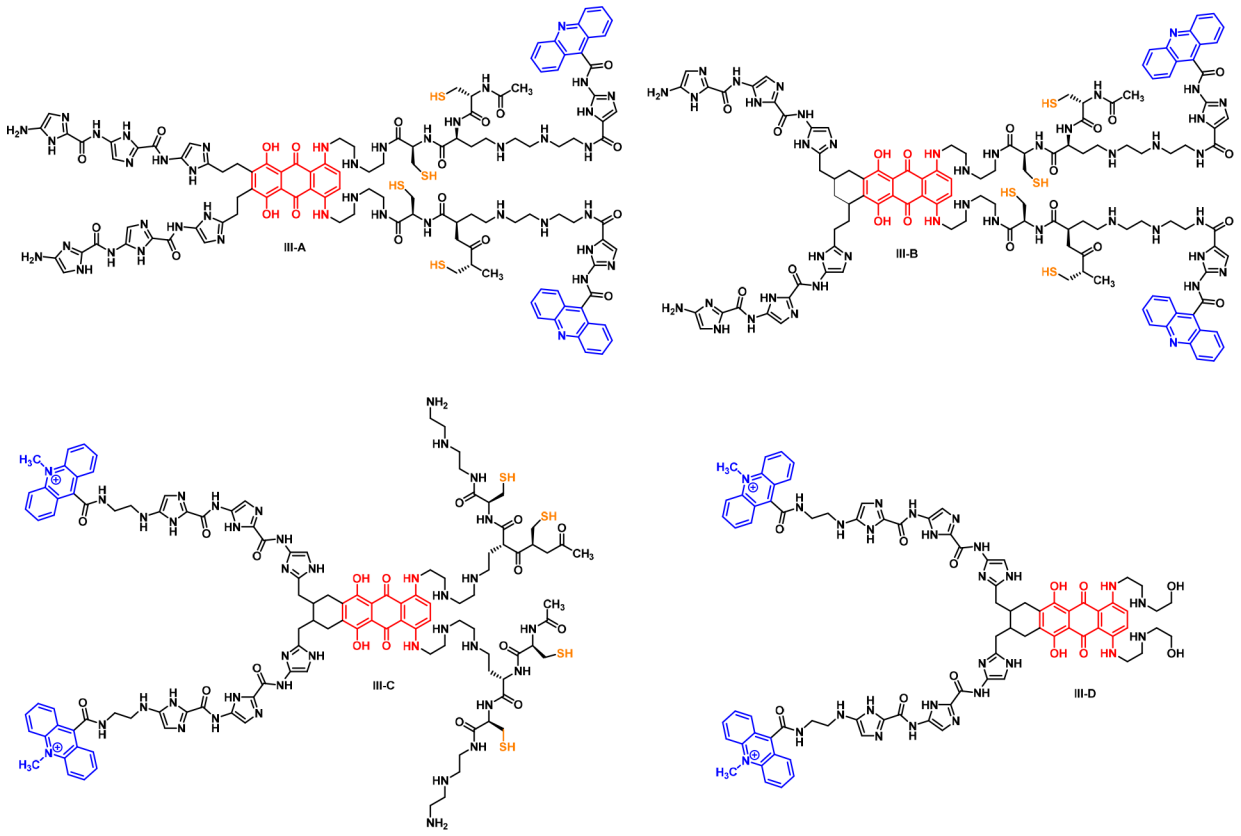


Figure 1. Molecular structures of compounds **III-A**, **III-B**, **III-C**, and **III-D.** On each arm, the aliphatic N atom of the starting MTX arm and the two aliphatic N atoms of the side-chains are to be considered as protonated, i.e., (NH_2+)-, and the S atom of each Cys residue forms a disulfide bridge with the corresponding S on the other arm.

We followed a related MD protocol as in ref 15 (recalled below). This augmented derivative of **III** will be denoted as **III-A**. After an initial phase of restrained MD enforcing the 12 H-bond distances, long-duration unrestrained MD showed their lasting persistence. This led us to consider another means of connecting (im)3: the anthraquinone is augmented with an additional ring, cyclohexene, as with anthracyclines. Each edge C is substituted with a $-CH_2-$ group itself connected to the C_2 carbon of the first (im)3 imidazole. Both C edge atoms have an S configuration. This enabled again for directional H-bonds between both the (im)3 arm and the two DNA strands. This derivative is denoted as **III−B**.

With both compounds at the outcome of long-duration unrestrained MD, the terminal $-NH_2$ group was found at distances from the acridine N atom in the range of 5.2−5.7 Å. This range of distances could correspond to the length of an appropriate connector between (im)3 and a central acridine atom. We then designed a third derivative, **III−C**. The terminal (im)3$-NH_2$ group on each arm is alkylated by the motif $-CH_2-CH_2-NHCO$. The terminal CO is connected to the C_9 carbon of an acridine ring. The acridine N atom is methylated, conferring on it a cationic charge. We removed the connector between the alkyldiammonium group and acridine in the major groove. Tris-intercalation now takes place from the minor groove side, rather than from the major groove. The alkylated aminoacridine N then protrudes into the major groove side, where it undergoes the favorable electrostatic potential exerted by O_6/N_7 of the two 5′ guanines of the intercalation site. Since unrestrained long-duration MD showed lasting interactions between both (im)3 arms and G1-G2/G1′-G2′ in the minor groove, we considered a fourth tris-intercalating compound, **III-D**, now devoid of the two oligopeptide arms in the major groove, limited to the sole $-CH_2-CH_2-NH_2^+-CH_2-CH_2-OH$ arms of the parent MTX.

The structures of **III-A** until **III-D** are represented in Figures 1a-d. For completeness, we provide as Supporting Information S1 the chemical structures of their precursors, denoted as **I, II**, and **III** in ref 15.

We consider next monointercalating compounds, devoid of amino-acridine. The first two are derived from **III-A** and **III-B**, and the third is derived from **III-D**, with the sole MTX arms in the major groove. They will be denoted as **III-A′**, **III-B′**, and **III-D′**. **III-A′**, and **III−B′** can be considered as well as derived from compound **II** of[22] with two (im)3 arms in the minor groove.

It is necessary at the outset to address a reservation expressed by Reviewers. It will be further considered in the Conclusion and Perspectives section. It is mentioned here that a multistep synthetic protocol of **III** and its augmented derivatives, **III-A—III-D**, was recently designed by some of us (Ongaro et al., unpublished). It should enable completion of their syntheses in a time span evaluated to two years. Thereafter the major issue of bioavailability will be faced, since their polycationic nature could possibly oppose membrane crossing. Nevertheless, it is recalled that the Lys-anthraquinone precursor derivative of **III** and **III-A—III-C** did show cellular activity against three tumor cell lines in a 5−15 μM range.[14] It bears four cationic charges: one on each alkylammonium of the "carrier" chain of ametantrone whence it is derived, and one on each Lys side-chain: these did not

https://doi.org/10.1021/acsomega.4c05099
ACS Omega 2024, 9, 42309−42328

thus appear to hamper bioavailability. Derivatives **III-A**−**III-B** bear two additional charges, having a dicationic alkyldiammonium (ADAM) side-chain instead of the monocationic methylammonium Lys end-side chain. Could these two additional charges possibly oppose membrane crossing? Derivative **III-C** has two additional charges, with one 9-amino-acridinium replacing each amino-acridine. As commented on in the Conclusions and Perspectives section, the presence of cationic conjugated groups could facilitate, rather than oppose, membrane crossing. This should also be the case for derivative **III-D**, devoid of the ADAM connector, now with a reduced number of cationic charges, namely, four, including two conjugated ones.

There should, furthermore, be an instructive issue related to in vitro experiments. These are destined to monitor the staged evolutions of the binding affinities and sequence selectivity upon passing from **III** to each of its im3-augmented derivatives in the tris-intercalating series, but also along the series of monointercalators **III-A′**−**III-D′** as well as upon passing from the mono- to the tris-intercalating series. This will also be addressed in the Conclusions and Perspective section.

3. COMPUTATIONAL PROCEDURE

We used the same procedure as in.[15] MD simulations use the AMOEBA polarizable multipole potential[37] coded in the Tinker-HP software[38] in its recent GPU implementation.[39] We use the 2018 DNA parameters.[40] The distributed multipoles of all derivatives compounds were derived by a Stone analysis[41] on their molecular wave functions computed with the B3LYP DFT functional[42,43] and a cc-pVDZ basis set[44,45] with the G09[46] software.

All seven compounds were docked into a tris-intercalated DNA conformation extracted from one of the last poses of a long-duration MD of its complex with compound **III** of.[15] Its sequence is d(CGTA**C** **GGGC** **GCCC** **G**TACG)$_2$. A void is used to designate the three d(CpG)$_2$ intercalation sites. The tris-intercalating ligands were derived from ligand **III** upon connecting both (im)3 arms to the anthraquinone or to the aminoacridine as mentioned above. With the help of computer graphics with the Insight II software (Accelrys Inc., San Diego), torsions around the rotatable bonds of the connectors enabled proximity between the deprotonated N atom of each imidazole on each arm and the −NH$_2$ of a facing guanine base. Thus, owing to its crescent shape, each (im)3 arm would "straddle" the two strands. In the first arm, im$_1$ could bind to G4′, while im2 and im3 could bind to G2 and G1, respectively. In the second arm, im$_2$ could bind to G4, and im2 and im3 to G2′ and G1′, respectively. Manual docking was followed by constrained EM with the six corresponding distance restraints; DNA first held rigid. The DNA-ligand complex was then immersed in a bath of water molecules in a box of dimensions 51, 54, and 90 Å on the *x*, *y*, and *z* dimensions. Neutrality was ensured by placing 28 Na^+ cations in initial random positions for the complexes of **III-A** and **III-B**, and 26 and 30 ones for those of **III-C** and **III-D**, respectively. We accordingly removed for compound **III-D** the connector in the major groove between the alkyldiammonium group and acridine. Periodic boundary conditions (PBC) were applied along with Smooth Particle Mesh Ewald (PME).[47,48] We used cutoff values of 12 and 9 Å for van der Waals and Ewald interactions, respectively.

We then followed the same procedure as in ref 15.

Constrained EM was resumed, now also relaxing DNA, water, and counterions. Molecular Dynamics was subsequently run. Equilibration was started by 50 K stepwise raises in temperature for a duration of 1 ns at constant volume, from 0 to 300 K. Production was then started at 300 K with a Bussi thermostat[49] and at constant pressure 1 atm. Coordinates were saved at every 100 ps.

As motivated in ref 15 four additional 3.2 Å restraints between O$_4$(T) and N$_1$(A) and between N$_3$(T) and N$_6$(A) of the two-end base-pairs, T1-A18′ and T1′-A18 were introduced, which prevent the "fraying" of these base-pairs in the course of MD. This inclusion has precedents in DNA MD simulations including those with polarizable potentials.[40] This was justified by the limited length of the oligonucleotide. It prevents unwanted "danglings" of the end base-pairs.

During the first 40 ns of production stage, the ligand-DNA distance restraints were maintained. This was motivated in ref 15: large amplitude motion of DNA and the solvent were observed during the initial phases of the production runs, which cause an excess kinetic energy transfer to the bound ligand, which could destabilize and disrupt its complex with DNA. Unrestrained MD was then started, retaining, however, the restraints enforcing formamide planarity and the four restraints enforcing H-bonding between the two-end base-pairs.

A concern was raised by one Reviewer regarding the use of six starting intermolecular restraint distances that could bias the MD simulations. We have addressed this point regarding the complexes of **III-A** and **III**−**B**. In their last poses, water and counterions were removed. Using computer graphics, each im3 arm underwent conformational changes around the saturated C−C bonds that connect it to the anthraquinone/anthracycline ring. These enabled us to stretch each arm at a far distance away from the minor groove. In vacuo EM was then performed, keeping DNA rigid, and restoring a single distance restrain for each arm, namely the H-bond connecting the N atom of the first imidazole, im1/im1′, to H$_2$N of G4′/G4 of the central intercalation site. The resulting DNA complex was then immersed in a water bath with counterions, and the same protocol as above was resumed: EM on the entirety of the solvated complex, NVT with stepwise 50 K heating from 0 to 300 K, then restrained MD, now for a 64 ns duration. After this duration, both N(im1/im1′)−H$_2$N(G4′/G4) were removed, and a 140 ns unrestrained MD production was run. The outcome of these MDs is reported accordingly. Retaining one constraint for each arm in the starting stages was deemed necessary, since otherwise the simulation could take an unrealistically long time before the arm is enabled to bind to the minor groove.

To unravel the contribution of polarization in model complexes, energy-decomposition analyses on ternary complexes involving methylammonium, guanine and imidazole extracted from the last pose of unrestrained MD on the DNA complex with **III-A** used the ALMOEDA energy decomposition procedure[50] coded in the QChem package[51] and the ω-B97-XD DFT functional.[52] The search for alternatives to the imidazole ring and/or the formamide backbone (see below) was done by performing energy-minimization on model complexes of the (im)$_3$-like chain with the four base-pairs C3-G4′, G2-C5′, G1-C6′, and G-1-C7′ and the alkyldiammonium group on the major groove side, relaxing the sole chain, upon resorting to the G09 software with the cc-pVTZ(-f) basis set and the PBE0 functional.[53] G1 and G2 were considered with their deoxyribose sugars to prevent intrusion of the chain inside the groove which could occur due to the simplicity of the model. The ΔE_{tot} values at the converged minima were corrected for BSSE.[54]

https://doi.org/10.1021/acsomega.4c05099
ACS Omega 2024, 9, 42309−42328

4. RESULTS AND DISCUSSION

4.1. Complexes of the Trisintercalating Compounds.

Figure 2a gives a representation of the complex of **III-A** with the target oligonucleotide in the last frame of unrestrained MD trajectory of 140 ns. (Figures 2a-c). An overall view of the complex is shown in Figure 2a. The detailed interactions of the separate arms, as found in the representative last pose, are given in Figures 2b and 2c, respectively.

The radial distribution functions of the H-bond distances in complex 2a are shown in Figures 3(a-h). Figures 4, 5 and 6 give a representation of the complexes of **III-B**, **III-C**, and **III-D** with the target oligonucleotide in the last frame of unrestrained 140 ns MD trajectory. The figures labeled 'a' represent the complexes with the entirety of the targeted DNA. Those labeled 'b' and 'c' give close-ups of the complexes of each (im)3 arm with the targeted bases viewed from the minor groove side. They relate to the first and second arm, respectively. We denote by 'N' the deprotonated N atom of the imidazoles. For conciseness, we will denote by 'interaction with the $-NH_2$ group of guanine' the interaction of this nitrogen with the H atom belonging to the G extracyclic $-NH_2$ group, *cis* to N_3 of this base, thus not involved in the Watson−Crick pairing. The H-bonds described below relate, for each complex, to the last pose of its unrestrained MD. The distances have values representative of those occurring during the trajectory. Their time evolutions are shown separately. We focus on the central intercalation site and the two base-pairs above and below it. The bases are numbered for G1 to C6 and from G1′ to C6′ on the unprimed and primed strands, respectively. Thus, the paired bases are G1-C6′, G2-C5′, C3-G4′, G4-C3′, C5-G2′, and C6-G1′, base-pairs C3-G4′ and G4-C3′ being those of the central intercalation site.

*4.1.1. Complex of **III-A** (Figures 2a−c).* An overall view of the complex is given in Figure 2a. The detailed interactions of the separate arms, as found in the representative last pose, are given in Figures 2b and 2c, respectively. An overall view of the complex at the outcome of unrestrained MD with a single restraint on each arm is given in Figure 2d and the detailed interactions of each arm are given in Figures 2e-f.

The radial distribution functions of the H-bond distances described below for the first complex are shown in Figure 3(a-h). Their time evolutions are given in Supporting Information S2. The relevant H-bond distances between each (im)3 arm and G and C bases in the representative last pose are tabulated below in Table 1.

They can be described as follows. Regarding the first arm (Figures 2b, 3a, and 3b): N of im1 accepts a proton from $-NH_2$ of G4′ of the intercalation site on the primed strand. The amide substituting im1 on its C_4 carbon donates its NH proton to the O_2 of C5′, downstream from G4′ on the primed strand. N of im2 accepts a proton from $-NH_2$ of G2 on the unprimed strand, and the amide substituting im_2 on its C4 carbon donates a proton to N_3 of G2. N of im3 accepts a proton from $-NH_2$ of G1 on the unprimed strand and the terminal $-NH_2$ donates a proton to N_3 of G1. Thus, in the minor groove G1 and G2 are both involved in bidentate interactions with im3 and im2, simultaneously to their bidentate interactions with the alkyldiammonium interactions in the major groove (Figures 2b and 3c). A similar pattern is found with the second arm (Figures 2c, 3d, and 3e). In the excerpt above, the "prime" notations for the arm and the bases were simply permuted. Thus, as on the unprimed strand, G1′ and G2′ are both involved in bidentate interactions in the minor groove, and these occur simultaneously with those occurring with their O_6/N_7 atoms in the major groove (Figures 2c and 3f). A total of 12-13 H-bonding interactions is thus found in the minor groove. The interactions occurring in the major groove are not perturbed by those in the minor groove. They could be, in fact, reinforced by some cooperativity effects occurring across the two grooves (see below): the four bidentate interactions of the alkyldiammonium groups with O_6/N_7 of G1-G2/G1′-G2′ are closely similar to those with the parent compound **III**.[15] As with **III**, each $-NH_2^+-$ group of the MTX "carrier" arms interacts with a phosphate group of the central intercalation site (Figure 3g). There are also the "ancillary" interactions taking place in the acridine intercalation sites between N_7 of the 5′ guanines G-1 and G-1′ and either the protonated nitrogen of the imidazole group of the connector (Figure 3h), or the terminal amide connecting to the acridine.

4.1.1.1. Results from Unrestrained MD Started from Restrained EM/MD with a Single Starting Distance Restrain on Each Arm. An overall view of the complex is given in Figure 2d and the detailed interactions of each arm are given in Figures 2e-f. The H-bond distances from the last pose are listed below in Table 2. The time evolutions of the three H-bonds between the N atoms of im1, im2, and im3 with the amino $-NH_2$ of G4′, G2, and G1 (respectively) are also given in Supporting Information 2.

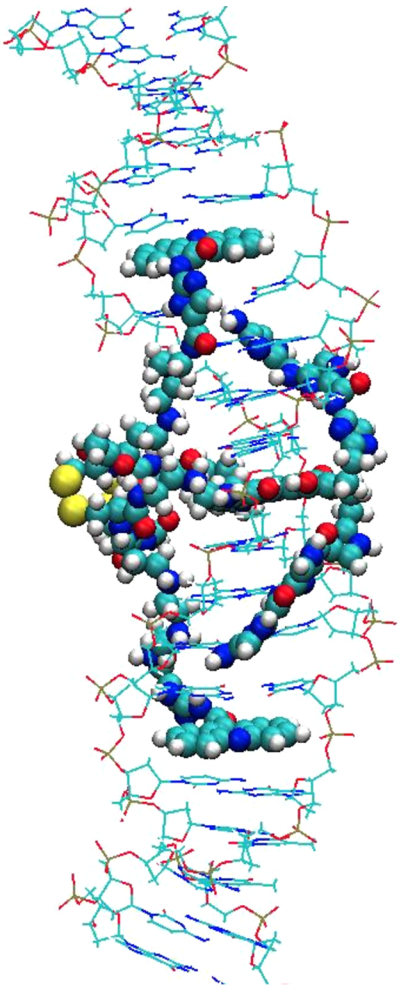
Along each arm, the interactions of each of the three imidazole N atoms with the guanine NH_2 atoms are fully recovered. A different pattern of interactions of the amide NH protons is however observed. For the first arm, the first NH proton interacts with O_2 of C3 in the unprimed strand, the base which is H-bonded to G4′, instead of O_2 of C5′ on the primed strand downstream from G4′ as was the case from the previous simulation. Therefore, the im1 N and its following HN atom bind to the central base-pair G4′-C3. While the second im2 N binds to NH_2 of G2, its following HN binds to the O_2 of C6′, downstream along the primed strand, rather than to N_3 of G2 as previously. The H-bond pattern involving im1 and im2 is, in fact, the same as that found for ligand **III-B** as detailed below. Regarding the third imidazole, a similar pattern is found to that from the previous simulation, but the H-bond distance between the terminal NH_2 group and N3 of G1 has lengthened from 2.5 to 3.0 Å. A closely related pattern is found in the primed strand for the first two imidazoles. N of im1′ accepts a proton from NH_2 of G4, while the following amide NH is H-bonded to the amide O_2 of both C3′, the complementary base to G4, and C5, downstream for G4. N of im2′ accepts a proton from G2 and the following NH group donates a proton to the O_2 of C6. N of im3′ accepts a proton from G1′ and the terminal NH group donates a proton to the O_2 of C7, the upstream base of the 9-AA intercalation site. The pattern of H-bonding interactions of the second arm appears to be the same as that found for ligand **III-B** along the primed strand, as detailed below.

*4.1.2. Complex of **III-B** (Figures 4a−c).* An overall view of the complex is given in Figure 4a. The detailed interactions of the separate arms, as found in the representative last pose, are given in Figures 4b and 4c, respectively.

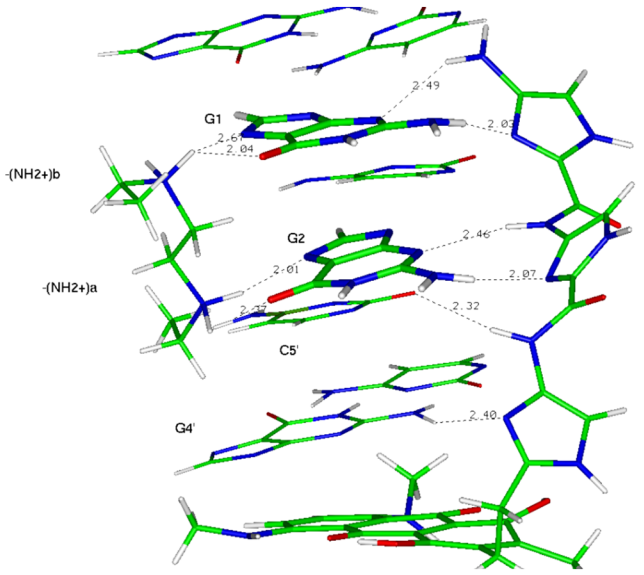
The relevant H-bond distances in Å units between each (im)3 arm and G and C bases, as found in the last pose, are tabulated in Table 3 below.

On the first (im)3 arm (Figure 4b), N of im1 accepts a proton from the $-NH_2$ of G4′ of the central intercalation site. The amide group substituting im_1 donates its proton to the O_2

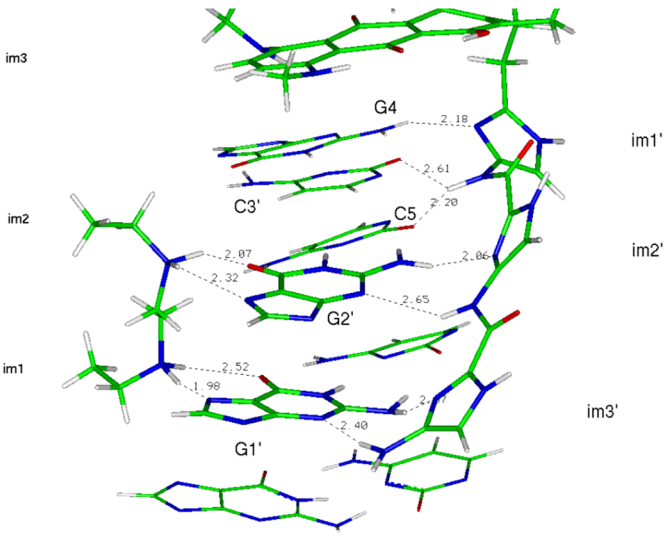
https://doi.org/10.1021/acsomega.4c05099
ACS Omega 2024, 9, 42309−42328



2a



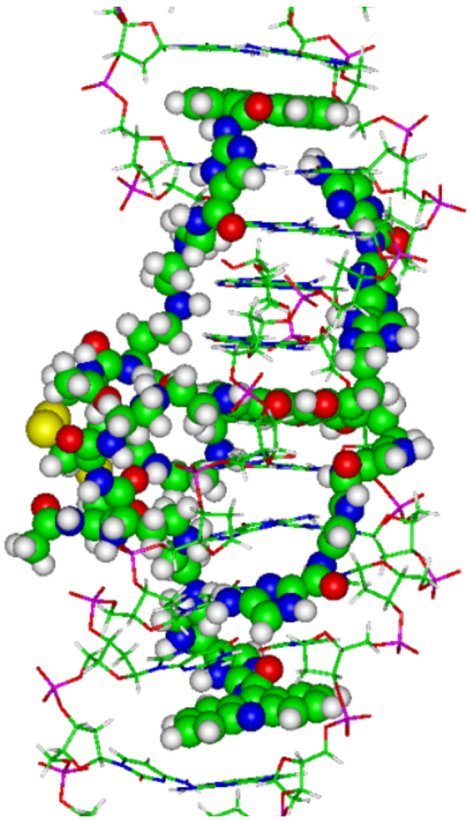
2b



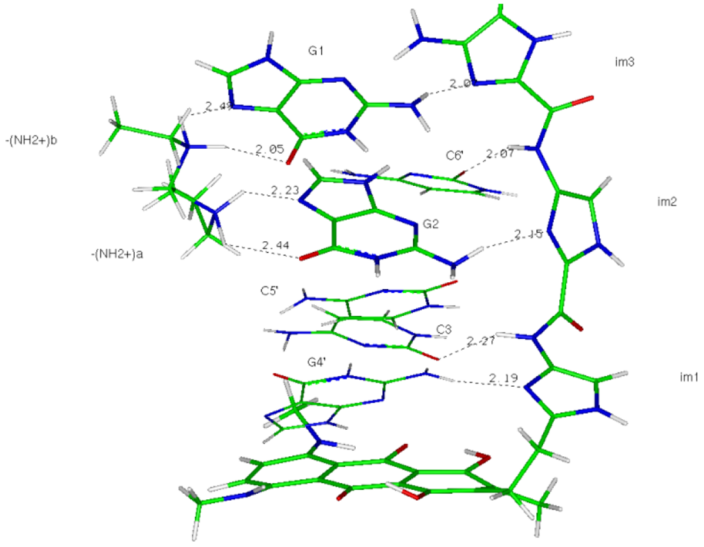
2c

Figure 2. continued

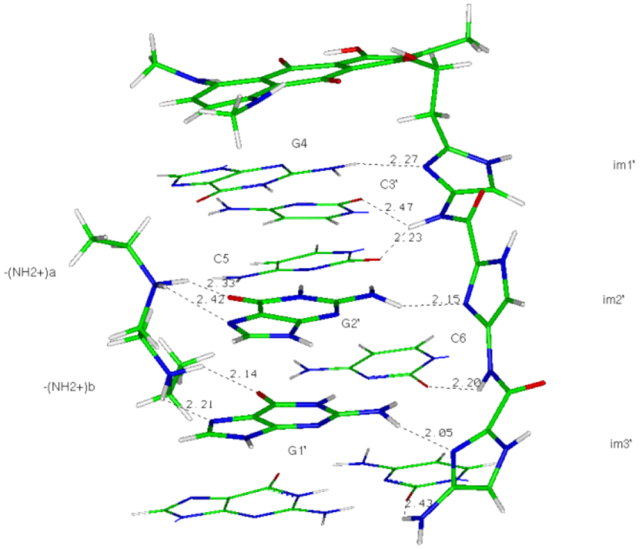
https://doi.org/10.1021/acsomega.4c05099
ACS Omega 2024, 9, 42309−42328



2d



2e



2f

Figure 2. (a–f) Complex of **III-A** with the decanucleotide target. (a) The entire complex; and from the minor groove side: (b, c) interactions involving the first and the second (im)3 arm, respectively; and from unrestrained MD started from restrained MD with one constraint per (im)3 arm; (d) the entire complex; (e, f) interactions involving the first and the second (im)3 arm, respectively.

https://doi.org/10.1021/acsomega.4c05099
ACS Omega 2024, 9, 42309−42328

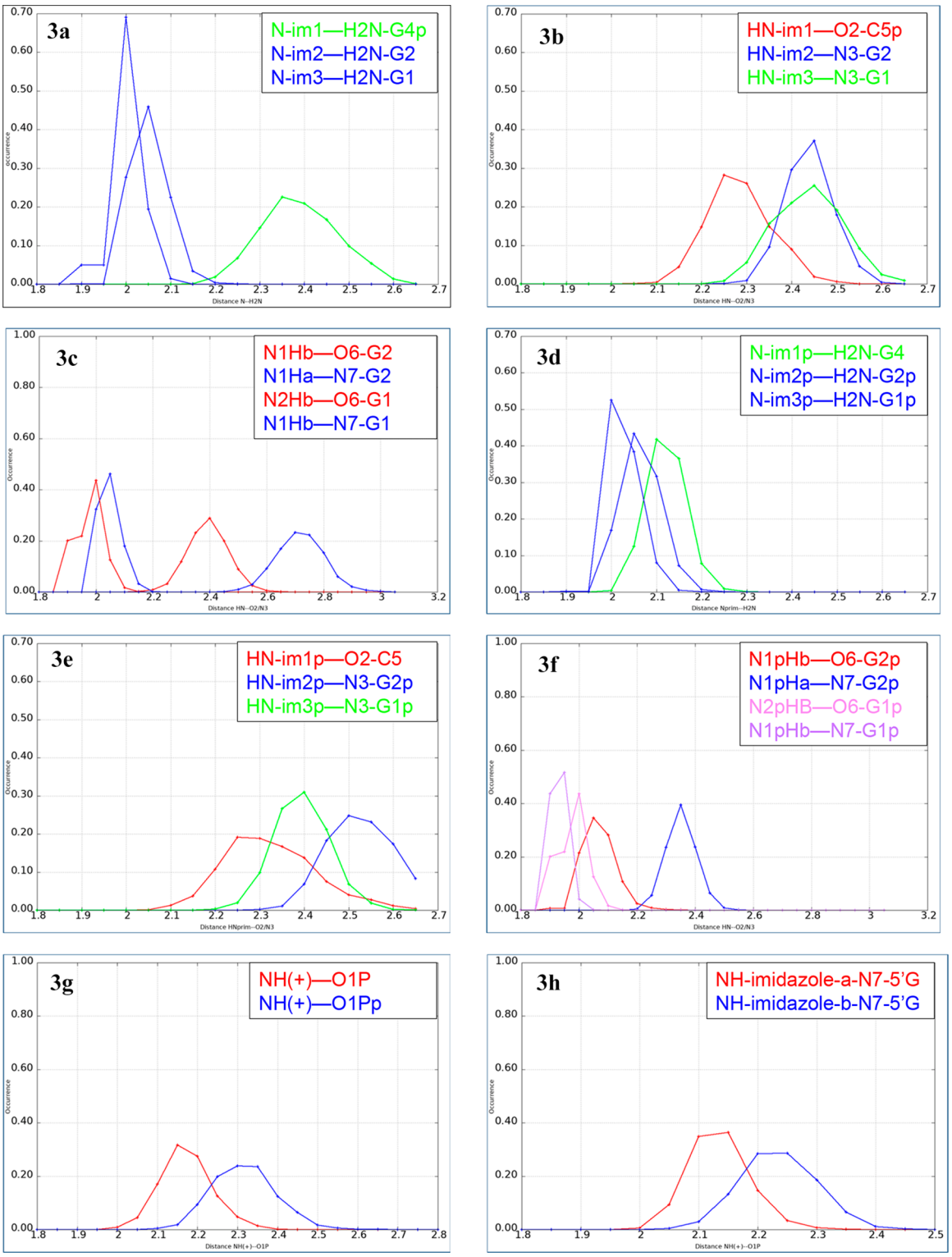


Figure 3. (a–h) Radial distribution functions (rdf) of the H-bonds distances stabilizing the **III-A** complex in the minor and major grooves. (a) Rdfs between N-im1, N-im2, and N-im3 with H2N of G4′, G2, and G1, respectively; (b) rdfs between the HN amides substituting im_1, im2, and im3 and O2(C5′), N3(G2), and N3(G1), respectively; (c) rdfs in the major groove between protons Ha and Hb of the first ammonium group and O6 and N7 of G2, respectively and between the protons Ha and Hb of the second ammonium group and O6 and N7 of G1, respectively; (d) rdfs between N-im1′, N-im2′, and N-im3′ with H_2N of G4, G2′ and G1′, respectively; (e) rdfs between the HN amides substituting im1′, im2′, and im3′ and O_2(C5), N_3(G2′,) and N_3(G1), respectively; (f) rdfs in the major groove between protons H_a and H_b of the first ammonium group and O_6 and N_7 of G2′, respectively, and protons H_a and H_b of the second ammonium group and O_6 and N_7 of G1′, respectively; (g) rdfs between the ammonium group of the carrier chain of MTX and O_1 of the phosphate group of the central intercalation site on both strands; (h) rdfs between the NH group of the imidazole connecting to 9-AA and N_7 of the 5′ G of the corresponding 9-AA intercalation site.

https://doi.org/10.1021/acsomega.4c05099
ACS Omega 2024, 9, 42309−42328

Table 1. List of H-bond Distances (Å) between the im3 Arms of III-A and the Bases in the Minor Groove (Figures 2b,c)

N-im1−−H2N(G4′)	HN-im1−O2(C5′)	N-im2−H2N(G2)	HN-im2−N3(G2)	N-im3−H2N(G1)	HN-im3−N(G1)
2.40	2.32	2.07	2.46	2.03	2.49
N-im1−H2N(G4)	HN-im1′−O2(C5)	N-im2′−H2N(G2′)	HN-im2′−N3(G2′)	N-im3′−H2N(G1′)	HN-im3′−N(G1′)
2.18	2.20	2.06	2.65	2.07	2.40

Table 2. List of H-Bond Distances (Å) between the im3 Arms of III-A and the Bases in the Minor Groove (Figures 2e,f)

N-im−H2N(G4′)	HN-im1−O2(C3)	N-im−H2N(G2)		HN-im2−O2(C6′))	N-im3−H2N(G1)	HN-im3−N(G1)
2.19	2.27	2.15		2.07	2.07	3.02
N-im1′−H2N(G4)	HN-im1′−O2(C3′)	HN-im1′−O2(C5)	N-im2′−H2N(G2′)	HN-im2′−O2(C6)	N-im3′−H2N(G1′)	HN-im3′− O2(C)
2.27	2.46	2.24	2.15	2.19	2.04	2.43

of C3. G4′ and C3 belonging to one of the two base-pairs of the central intercalation site are thus each H-bonded to the first imidazole-amide entity of the first (im)3 arm. N of im2 then accepts a proton from $-NH_2$ of G2 in the unprimed strand while the following amide donates its proton to O_2 of C6′ in the primed strand. N of im3 accepts a proton from $-NH_2$ of G1 and the terminal $-NH_2$ donates a proton to O_2 of the 3′ cytosine of the acridine intercalation site on the primed strand. The same pattern is found between the second (im)3 arm and the corresponding bases (Figure 4c). As with **III-A**, the interactions in the major groove remain very close to those with parent compound **III.**

With **III-B**, both (im)3 arms thus bind simultaneously, through im1 and its substituting amide, to both bases of the two G-C pairs of the intercalation site, C3-G4′ and C3′-G4. On the other hand, both im2−NHCO- and im3-NH_2 fragments bind in "transverse" modes, namely H-bond acceptance by the im N atom from G2/G2′ and from G1/G1′ respectively, and H-bond donation by NHCO and − NH_2 to the cytosine bases on the opposite strand, one step downstream along this strand. By contrast, with **III-A**, im2 with its following substituting amide and im3 and the following $-NH_2$ bound in bidentate fashion to the G2-C5′/G2-C5 and G1-C6′/G1′-C6 base-pairs, respectively, rather than in a transverse mode.

The radial distribution functions of the H-bond distances are shown in Supporting Information S3.

4.1.2.1. Results from Unrestrained MD Started from Restrained EM/MD with a Single Starting Distance Restrain on Each Arm. An overall view of the complex is given in Figure 4d and the detailed interactions of each arm are given in Figures 4e-f.

The H-bond distances from the last pose are listed below in Table 4.

The pattern of interactions of the first arm is consistent with that from the previous simulation. One difference relates to HN of im1 which H-bonds with elongated distances to both O_2 of C3 and O_2 of C5′, rather than with solely O_2 of C3. The H-bond between terminal $-NH_2$ of im3 and C7 of the 9-AA intercalation site is lengthened. Similarly, the pattern of interactions of the second arm is consistent with the one from the previous simulation. The sole difference concerns the NH group of im1′, which donates its proton to the O_2 of C5 instead of the O_2 of C3′.

For both **III-A** and **III-B**, unrestrained MD starting from restrained MD with a single "triggering" restrain per im3 arm have thus recovered the H-bond interactions of each im N to their target guanine NH_2 groups. For **III-A**, the pattern of the amide NH bond donation differed from the initial simulations. They became similar, for im1/im1′ and im2/im2′, to the pattern found with **III-B**. For **III-B**, the second unrestrained MD simulations recovered the same pattern as the original simulations, with one exception regarding the im1′ arm. Thus, whether started with six or with one restraint per arm, at least 11 intermolecular H-bonds are recovered in the minor groove for both ligands. It is recalled that numerous NMR and X-ray structural studies, as well as past MD simulations, support the binding of lexitropsin in the core of the minor groove and fulfillment of the H-bonds. A structure put forth by Fechter et al. on a DNA complex of a lexitropsin connected to a protonated aminoacridine,[25] displayed an intermolecular complex having all H-bonding possibilities fulfilled.

*4.1.3. Complex of **III-C** (Figures 5a−c).* The binding pattern appears to be similar to that of **III-B**. There is one difference. In both strands, it is the second amide past im_3, connecting to the amino-acridinium ring, that now donates a proton to the O_2 of the 3′cytosine of the 9-AA intercalation site. The H-bonding distances are tabulated below in Table 5.

The radial distribution functions of the H-bond distances are shown in Supporting Information S4. The bidentate interactions with the O_6/N_7 of G1-G2/G1′-G2′ in the major groove are not significantly modified compared to **III.**

*4.1.4. Complex of **III-D**.* The binding pattern in the minor groove is similar to those of compounds **III-B** and **III-C**.

The relevant H-bond distances are tabulated below in Table 6. The essential difference with respect to the complex of III-C is the fact that it is both the HN connected to the last im_3 arm and the following amide NH that donate their proton to the C base of the 9-AA intercalation site.

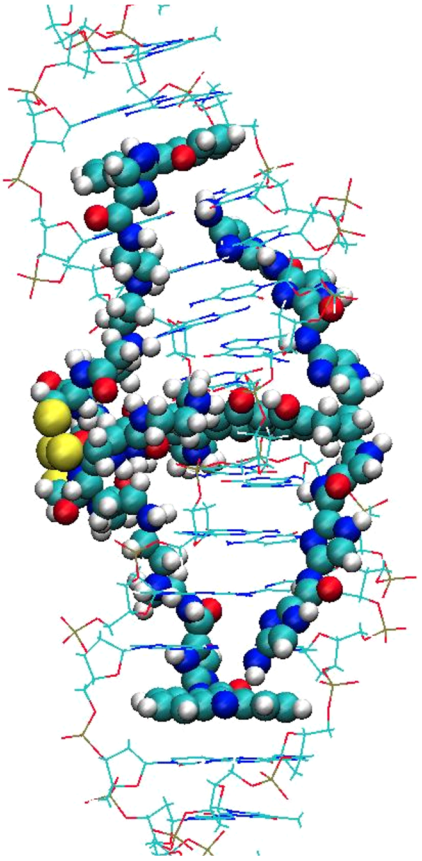
The radial distribution functions of the H-bond distances are shown in Supporting Informationg S5. The sole interactions in the major groove are those of the $-(NH_2)^+$- groups of the MTX side-chains with O_6 of the 5′guanine of the central intercalation site on the unprimed strand (d = 1.97 Å) and with N_7 of the corresponding guanine on the primed strand (d = 2.39 Å).

4.2. Complexes of the Monointercalating Compounds. Supporting Information S6a−c give representations of the complexes of **III-A′**, **III-B′**, and **III-D′** after 140 ns of unrestrained MD.

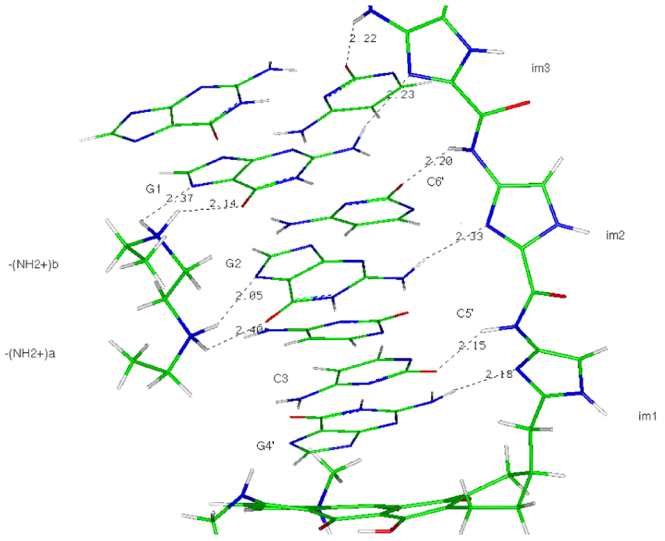
*4.2.1. Complex of **III-A′**.* The interactions of **III-A′** with the targeted palindrome bear a strikingly similar pattern to those occurring with its higher homologue **III-A** whence it is derived. The distances in the last frame have changed to only a limited extent. They are listed below in Table 7 for completeness.

The interactions occurring in the major groove of DNA with the O_6/N_7 of G1/G2 and G1′/G2′′ are maintained throughout

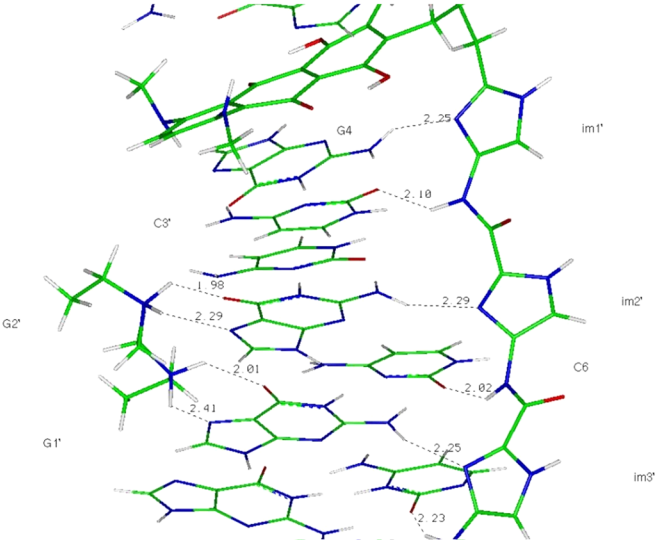
https://doi.org/10.1021/acsomega.4c05099
ACS Omega 2024, 9, 42309−42328



4a



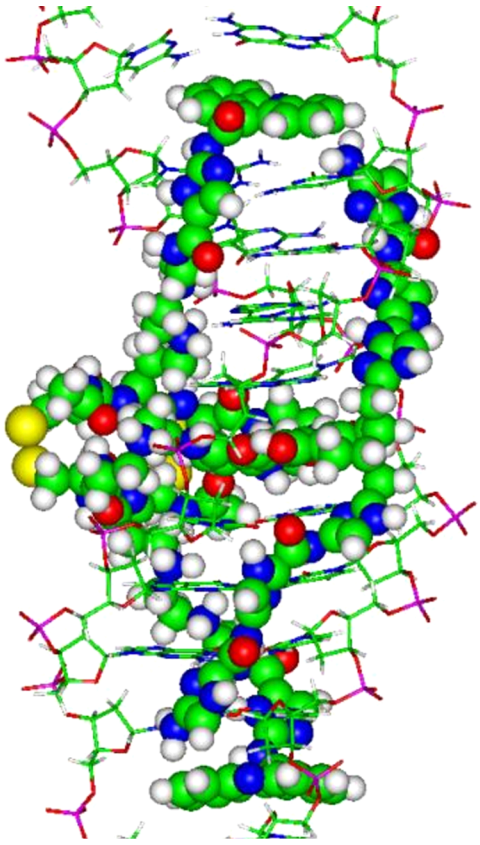
4b



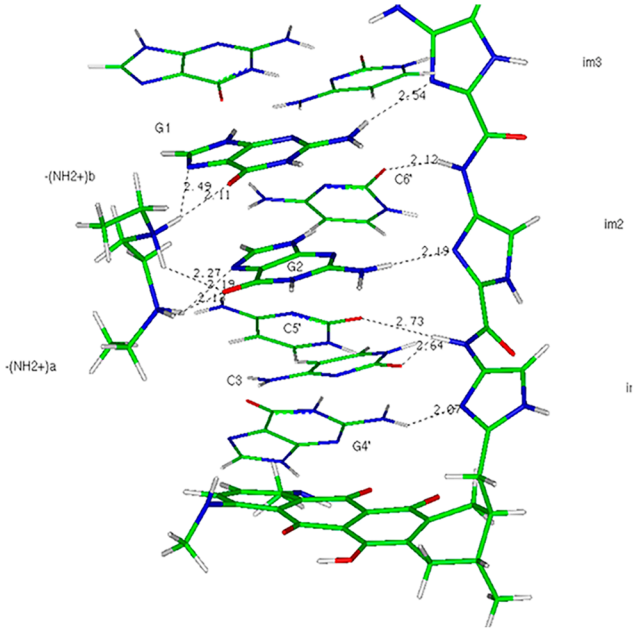
4c

Figure 4. continued

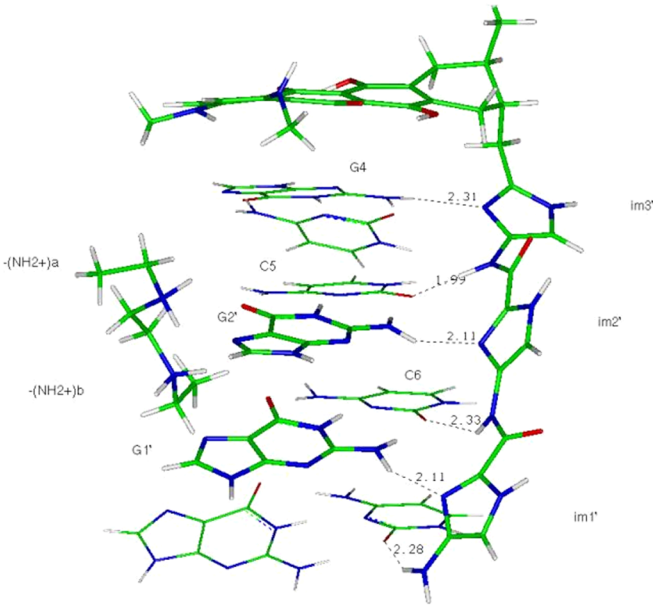
https://doi.org/10.1021/acsomega.4c05099
ACS Omega 2024, 9, 42309−42328



4d



4e



4f

Figure 4. Complex of **III-B** with the decanucleotide target: (a) the entire complex; and from the minor groove side: (b) and (c): interactions involving the first and the second (im)3 arm respectively; and from unrestrained MD started from restrained MD with one constraint per (im)3 arm: (d) the entire complex; (e) and (f): interactions involving the first and the second (im)3 arm respectively.

https://doi.org/10.1021/acsomega.4c05099
ACS Omega 2024, 9, 42309−42328

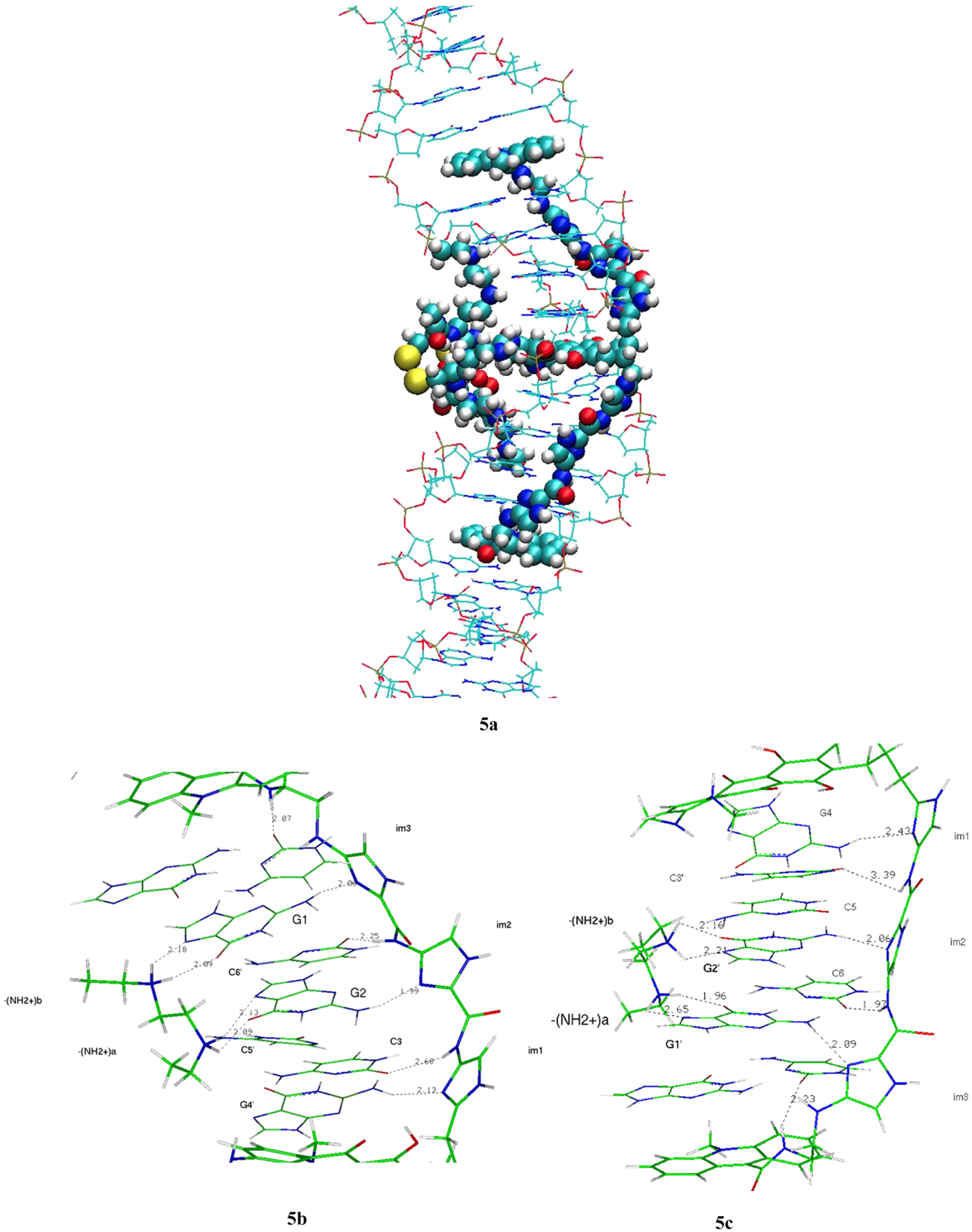


Figure 5. (a–c) Complex of **III-C** with the decanucleotide target. (a) The entire complex; and from the minor groove side; (b, c) interactions involving the first and the second (im)3 arm, respectively.

https://doi.org/10.1021/acsomega.4c05099
ACS Omega 2024, 9, 42309–42328

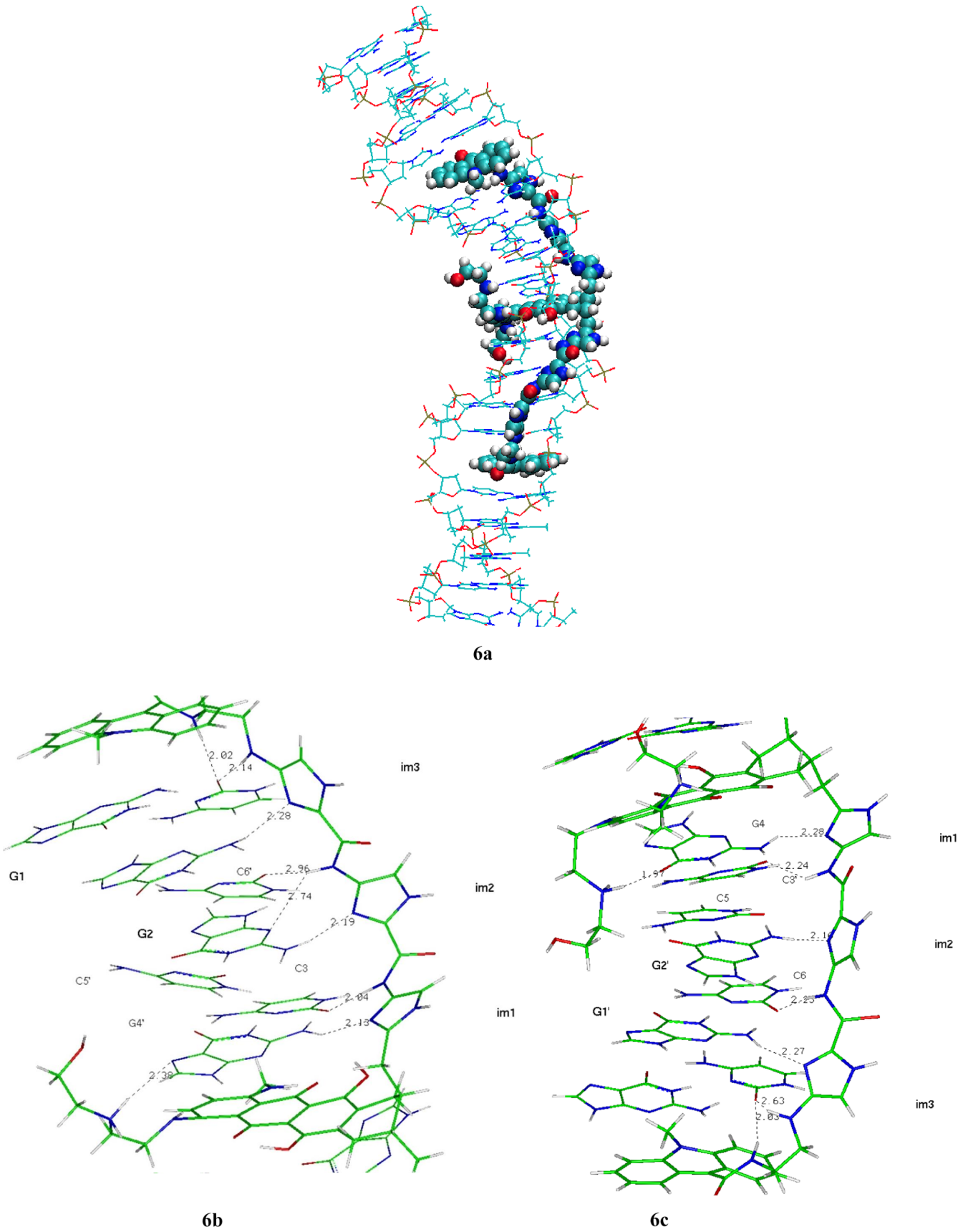


Figure 6. (a–c) Complex of **III-D** with the decanucleotide target. (a) The entire complex; and from the minor groove side; (b, c) interactions involving the first and the second (im)3 arm, respectively.

https://doi.org/10.1021/acsomega.4c05099
ACS Omega 2024, 9, 42309–42328

Table 3. List of H-Bond Distances (Å) between the im3 Arms of III-B and the Bases in the Minor Groove (Figures 5b,c)

N-im1−H2N(G4′)	HN-im1−O2(C3)	N-im2−H2N(G2)	HN-im2−O2(C6′)	N-im3−H2N(G1)	HN-im3−O2(C)
2.18	2.15	2.33	2.20	2.23	2.22
N-im1′−H2N(G4)	HN-im1′−O2(C3′)	N-im2′−H2N(G2′)	HN-im2′−O2(C6)	N-im3′−H2N(G1′)	HN-im3′−O2(C)
2.25	2.10	2.29	2.02	2.25	2.23

Table 4. List of H-Bond Distances (Å) between the im3 Arms of III-B and the Bases in the Minor Groove (Figures 5e,f)

N-im1−H2N(G4′)	HN-im1−O2(C3)	HN-im−O2(C5′)	N-im2−H2N(G2)	HN-im2−O2(C6′)	N-im3−H2N(G1)	HN-im3−O2(C)
2.07	2.64	2.73	2.19	2.12	2.54	3.33
N-im1′−H2N(G4)	HN-im1′−O2(C5)		N-im2′−H2N(G2′)	HN-im2′−O2(C6)	N-im3′−H2N(G1′)	HN-im3′−O2(C)
2.31	1.99		2.11	2.33	2.11	2.28

the trajectory, coexisting with those occurring in the minor groove.

4.2.2. Complex of III−B′. The interactions of **III-B′** with the palindromic sequence have the same pattern as those of **III-B**. The H-bond distances are listed in Table 8 below.

The interactions in the major groove with O_6/N_7 of G1-G2/G1′-G2′ remain close to those with the parent compound **III-B**.

4.2.3. Complex of III-D′. The DNA-**III-D′** interactions in the minor groove, listed in Table 9 below, are very consistent with those found in the complexes of **III-A** and **III-A′**.

In the major groove, the interactions of the two mitoxantrone end-side chains, $-CH_2-(NH_2^+)-CH_2-CH_2-$OH, differ, as a consequence of their larger flexibility compared to those of the oligopeptide entities of compounds **III-A**—**III-B′**. One $-(NH_2^+)-$ group can bridge N_7 of G4' and O_1 of the phosphate on the primed strand of the central intercalation site. The other $-(NH_2^+)-$ group interacts with its facing phosphate indirectly through one water molecule. It is the preceding $-(CH_2)-$ group, on which the cationic charge is partly delocalized, that interacts with G4 (d = 2.51 Å), while the $-(CH_2)-$ group following the $-(NH_2^+)-$ group has an elongated interaction with O_1 of the facing phosphate (d = 2.87 Å). It is worth noting that such interactions, involving these two bracketing $(-CH_2)-$ fragments with N_7 of the 5′ guanine and O_1 of the central phosphate, respectively, were characterized in an earlier energy-minimization study using the SIBFA potential.[55] The terminal −OH groups, on the other hand, are exposed to the water phase.

4.2.3.1. Possible Onset of Local Cooperativity in Trimolecular Complexes between Alkylammonium, Guanine, and Imidazole. A central result of this article is the unraveled possibility for four guanines of the targeted palindrome to be simultaneously bound in both DNA grooves and along both strands. Each G is in the center of a local trimolecular complex and acts simultaneously as an H-bond acceptor from one partner and as an H-bond donor to the second. Earlier simulations showed such arrangements to enable for enhanced cooperativity,[56] We have considered the trimolecular complex of G1, methylammonium, and the third imidazole of the first im3 arm, connected to its substituting $-NH_2$ group. Using the ALMOEDA energy decomposition procedure, we compared the values of the polarization (E_{pol}) and charge-transfer (E_{ct}) in the trimolecular complex to their sums in the three separate bimolecular complexes. This was done for the complex of **III-A** on five snapshots along the unrestrained MD trajectory as well as on the final one. The intermediary snapshots were recorded at 56, 72, 96, 104, and 120 ns. The values of E_{pol} varied in a limited range, of −12.7 to −13.5 kcal/mol. For each snapshot, their magnitudes are increased compared to the summed E_{pol} value in the three separate bimolecular complexes by amounts in the range 1.4−1.7 kcal/mol. The cooperativity of E_{ct} was much smaller in this case. Its values in the range −7.0 to −7.9 kcal/mol were only 0.3 kcal/mol larger in magnitude than those in the three separate bimolecular complexes. The cooperativity of E_{pol} could be considered as relatively small considering the much larger polarizabilities of guanine and imidazole than water. It could be explained by the fact that owing to the size of guanine, the intermolecular separation of the interacting methylammonium and imidazole is much larger than in the case of linear or cyclic water, where the distances of next-to-neighbor waters are much smaller.

For the sake of completeness, Table 10 lists the values of ΔE_{int} and its contributions, averaged from the six considered snapshots. The summed values of the second-order contributions, E_{pol} and E_{ct}, namely E_2, of −20.6 kcal/mol, come close to those of the summed first-order ones, E_C (Coulomb) and E_X (short-range repulsion), E_1, of −24 kcal/mol. Adding the dispersion contribution, E_{disp}, to E_2, results in a magnitude of −30.0 kcal/mol, now larger than that of E_1.

4.2.3.2. Toward Increasing the Affinities in the Minor Groove. Several proposals to increase the binding of minor-groove polyamide binders were put forth.[57−60] These mostly considered replacements of the pyrrole and imidazole rings by alternative five-membered rings, such as hyroxypyrrole, pyrazole, thiazole, etc. We focused here on rings prone to acting as H-bond acceptors from the extracyclic guanine $-NH_2$ group. We evaluated the extent to which substitution of imidazole with electron-donating groups and/or its replacement by other conjugated five-membered rings could enable for improved H-bonding interaction with guanine. We also sought for changes in the formamide backbone. In order to increase backbone polarizability and the acidity of the NH hydrogen, we considered replacement of the formamides by thioformamides or sulfonamides. The choice of sulfonamides was motivated by the ability of this group to give rise to ordered structures.[61] Such backbone modifications do not seem to have been reported before in the context of minor groove binders. We report the results found with six selected model (im)3 derivatives: methylated imidazole; NH_2-substituted triazole; $NHCH_3$-substituted triazole and triazotropsin; and $NHCH_3$-substituted imidazole with thioamide or sulfonamide backbones. The model calculations were done on a trimolecular complex extracted from one of the last unrestrained MD poses of the **III-A** complex. It involved three G-C base pairs: C3-G4′, G2-C5′, and G1-C6′; the methylammonium moiety bound to G2 in the major groove;

https://doi.org/10.1021/acsomega.4c05099
ACS Omega 2024, 9, 42309−42328

Table 5. List of H-Bond Distances (Å) between the im3 Arms of III-C and the Bases in the Minor Groove

N-im1−H2N(G4′)	HN-im1−O2(C3)	N-im2−H2N(G2)	HN-im2−O2(C6′))	N-im3−H2N(G1)	HN-amide−O2(C)
2.13	2.68	1.99	2.25	2.06	2.07
N-im1′−H2N(G4)	HN-im1′−O2(C3′)	N-im2′−H2N(G2′)	HN-im2′−O2(C6)	N-im3′−H2N(G1′)	HN-amide−O2(C)
2.43	3.39	2.06	1.97	2.09	2.23

Table 6. List of H-Bond Distances (Å) between the im3 Arms of III-D and the Bases in the Minor Groove

N-im1−H2N(G4′)	HN-im1−O2(C3)	N-im2−H2N(G2)	HN-im2−O2(C6′)	HN-im2−N3(G2)	N-im3−H2N(G1)	HN-im3−O2(C)	HN amide−O2(C)
2.13	2.04	2.19	2.96	2.74	2.28	2.14	2.02
N-im1′−H2N(G4)	HN-im1′−O2(C3′)	N-im2′−H2N(G2′)	HN-im2′−O2(C6)		N-im3′−H2N(G1′)	HN-im3′−O2(C)	HN amide O2(C)
2.28	2.24	2.16	2.25		2.27	2.63	2.03

and a sequence made of the (im)3 motif or the above-mentioned variants. Considering the preliminary character of these calculations, energy-minimization was done on the sole position of the (im)3 motif. We have resorted to the G09 software with the cc-pVdZ(-f) basis set and the PBE0 functional. The results are reported in Table 11. At this stage, replacing imidazole by triazole, thioxazole, or triazotropsin, even following substitution with electron-donating groups, did not result in improved ΔE_{tot} values and were less effective than the simple methylation of the imidazole ring. Replacement of the formamide backbone by a thioamide one, along with imidazole substitution with an electron-donating -$NHCH_3$ group, did result into an improved ΔE_{tot} value, not overcoming however the one resulting from methylation. On the other hand, replacement of the amide by a sulfonamide enabled a significant ΔE_{tot} gain, by −10.2 and −7.1 kcal/mol with respect to imidazole and methylimidazole. The complex is represented in Figure 7. Even prior to MD simulations on the DNA complexes of $(im)_3$/sulfonamide-augmented ligands and prospective free energy calculations, these results indicate that some alternatives to the formamide backbone deserve further investigating.

CONCLUSIONS AND PERSPECTIVES

The palindromic sequence $d(GGCGCC)_2$ is encountered in oncogenes, regulatory DNA sequences, and retroviruses. It occurs three times in *Alu* repeat sequences present in extrachromosomal DNA (ecc-DNA),[7] the involvement of which in tumor progression and resistance was recently underlined.[9,10] Ecc-DNA thus would represent an emerging target for molecules that could selectively bind to this palindrome. Such molecules do not need to cross the nuclear membrane to reach chromosomal DNA in order to exert their effect.

Our essential purpose is to optimize both binding affinity and selectivity for sequences encompassing this palindrome, by designing novel, augmented derivatives of a tris-intercalating oligopeptide derivative of mitoxantrone, denoted **III** in.[15] This could be enabled by minor groove targeting of the extracyclic $-NH_2$ groups of the G1-G2/G1′-G2′ bases simultaneous with that of their O_6/N_7 atoms in the major groove. For that purpose, two three-imidazole amide chains, denoted as (im)3, were grafted, each on one outermost C atom of the central anthraquinone or anthracycline intercalator, as with compounds **III-A** and **III-B**, respectively, as well as on the C_9 atom of each of the two amino-acridine intercalators, as with **III-C**. The (im)3 motif matches the crescent-shaped backbone of distamycin, with pyrrole replaced by imidazole. This enables the deprotonated imidazole nitrogen to act as an H-bond acceptor from the guanine extra-cyclic NH_2 atom, whence exclusive recognition of this base in the minor groove. Following the same protocol as in,[15] long-duration, unrestrained polarizable MD simulations were performed on the complexes of all three novel derivatives with an oligonucleotide encompassing a central decameric palindrome $d(CGGGCGCCCG)_2$. For all three complexes and on each arm, the persistence of directional H-bonds between each imidazole N and the $-NH_2$ group of a facing G base was shown. With the first (im)3 arm, these were with G4′ of the intercalation site on the primed strand, and G2 and G1 on the unprimed strand. With the second (im)3 arm, these were with G4, G2′, and G1′. In addition, the amide group following each imidazole donated its proton to an electron-rich site in the groove, namely either N_3 of the targeted G base which is thus bidentate bound, or to O_2 of a C base on the strand opposite to the targeted G base. With all three complexes, up to 12 H-bonds were stabilized. As a consequence, the two target G bases on each strand upstream from the intercalation site, G1/G2 and G1′/G2′, can be recognized simultaneously in the major groove through O_6 and N_7 and in the minor groove through $-NH_2$. This should significantly increase both the binding affinity and the sequence selectivity in favor of the palindromic sequence. While alternating major and minor groove recognition by threading intercalators was already reported[28,62,31−33,63,64,36] simultaneous in-register major and minor groove recognition by one single molecule to a given DNA base was to the best of our knowledge never reported before, let alone that of four bases.

The last tris-intercalator derivative, **III-D**, is devoid of the peptide chains in the major groove and limited to the sole side-chains of the parent MTX molecule. It could qualify as a tris-intercalating MTX derivative destined to recognize the palindromic sequence solely from the minor groove. The pattern of persistent H-bond interactions with the (im)3 arms was similar to the one found with derivative **III-C** whence it was derived. A possible asset of this derivative with respect to **III-A**—**III-C** is that from a kinetic standpoint. The MTX arms have a much lesser bulkiness than those of **III-A**—**III-C**. Thus its approach from the minor groove side and subsequent DNA unfolding to generate the tris-intercalation site can be anticipated to incur much smaller rearrangements/transient base openings of the DNA duplex.

We have sought further possible increases of the binding affinities of the (im)3 motif upon substitution of the imidazole with electron-donating groups or its replacement altogether by other five-membered rings. At the preliminary stage that we reported, none provided any ΔE_{tot} improvement. On the other hand, a promising improvement appeared to result from the replacement of the amide backbone by a sulfonamide one,

https://doi.org/10.1021/acsomega.4c05099
ACS Omega 2024, 9, 42309−42328

Table 7. List of H-Bond Distances (Å) between the im3 Arms of III-A′ and the Bases in the Minor Groove

N-im1−H2N(G4′)	HN-im1−O2(C5′)	N-im2−H2N(G2)	HN-im2−N3(G2)	N-im3−H2N(G1)	HN-im3−N(G1)
2.08	2.24	2.63	1.97	2.07	2.63
N-im1′−H2N(G4)	HN-im1′−O2(C5)	N-im2′−H2N(G2′)	HN-im2′−N3(G2′)	N-im3′−H2N(G1′)	HN-im3′−N(G1′)
2.46	2.51	2.07	2.66	2.04	2.54

Table 8. List of H-Bond Distances (Å) between the im3 Arms of III-B′ and the Bases in the Minor Groove

N-im1−H2N(G4′)	HN-im1−O2(C3)	N-im2−H2N(G2)	HN-im2−O2(C5′))	N-im3−H2N(G1)	HN-im3−O2(C)
2.14	2.13	2.33	2.01	2.16	2.06
N-im1′−H2N(G4)	HN-im1′−O2(C3′)	N-im2′−H2N(G2′)	HN-im2′−O2(C6)	N-im3′−H2N(G1′)	HN-im3′−O2(C)
2.25	2.18	2.10	2.17	2.16	2.37

Table 9. List of H-Bond Distances (Å) between the im3 Arms of III-D′ and the Bases in the Minor Groove

N-im1−H2N(G4′)	HN-im1−O2(C5′)	N-im2−H2N(G2)	HN-im2−N3(G2)	N-im3−H2N(G1)	HN-im3−N(G1)
2.31	2.24	2.05	2.30	2.02	2.45
N-im1′−H2N(G4)	HN-im1′−O2(C5)	N-im2′−H2N(G2′)	HN-im2′−N3(G2′)	N-im3′−H2N(G1′)	HN-im3′−N(G1′)
2.41	2.18	2.04	2.62	2.16	2.25

Table 10. Averaged Values (kcal/mol) of ΔE_{int} and Its Contributions

ΔE_{int}	E_C	E_X	E_{pol}	E_{ct}	E_{disp}
−54.0	−48.6	24.6	−13.0	−7.6	−9.4

Table 11. Intermolecular Interaction Energies (kcal/mol) of the $(im)_3$ Amide and Six Other Derivatives with a Model of Central Target Sequence $d(GGCG)_2$

Derivative	ΔE
$[imidazole\text{-}formamide]_3$	−167.2
$[(imidazole\text{-}CH3)\text{-}formamide]_3$	−170.3
$[(triazole\text{-}NH2)\text{-}formamide]_3$	−167.4
$[(thioxazole\text{-}NHCH3)\text{-}formamide]_3$	−165.9
$[(thiazotropsin\text{-}NHCH3)\text{-}formamide]_3$	−165.9
$[(imidazole\text{-}NHCH3)\text{-}thioamide]_3$	−169.7
$[(imidazole\text{-}NHCH3)\text{-}sulfonamide]_3$	**−177.4**

owing to the increased polarizability of this group and acidic character of its protons as compared to formamide. Such a replacement does not seem to have been reported before and would be worth considering both experimentally and in the context of prospective free energy calculations.

The three occurrences of the target palindrome in the *Alu* repeat are as octanucleotides rather than as hexa- or decanucleotides. This is also the case for the host template sequence, which together with the HIV-1 RNA, forms a hybrid DNA-RNA double helix to which the HIV-1 polymerase binds (5). This implies that along with the syntheses of mono- and tris-intercalators, that of bis-intercalators with one rather than two aminoacridine ring, should also be considered. These could target selectively one site out of 32 896. This might be more advantageous than the binding of a tris-intercalator to an eight-base sequence, which could leave one 9-AA ring either unbound or bound to a lesser-affinity site.

Summarizing, three related classes of MTX oligopeptide derivatives were developed in this and in the previous paper. They were all designed to selectively target hexa- and decanucleotide palindromes. In the first class, recognition of G1-G2/G1′-G2′ occurs exclusively in the major groove. In the second, it occurs simultaneously in both grooves. The (im)3 chain can be grafted either to the anthraquinone (**III-A** and **III-B**) or to the acridine (**III-C**) ring. In the third class, targeting occurs exclusively in the minor groove, leaving only the two original $(CH_2)_2-(NH_2^+)-(CH_2)-CH_2OH$ MTX chains in the major groove.

On account of the modular nature of these constructs, it should be instructive to compare the affinities, sequence-selectivity, and cellular activities within and between the three series of derivatives. Regarding the actual affinities, there could be two favorable assets from experimental precedents. The first relates to two bis-intercalator derivatives of an anticancer drug, Adriamycin, endowed with picomolar (10^{-12}) affinities.[65−67] These affinities are reinforced by the presence of two cationic ammonium groups interacting electrostatically in the minor groove with the O_2 atoms of a cytosine of the intercalation site and of a thymine upstream. By comparison, derivatives **III-A**, **III-B**, and **III-C** total three instead of two intercalating groups and four ammonium groups, which can each bind bidentate to a well-defined target G base. It will be informative to compare the DNA-binding affinities of these derivatives to the bis-anthracycline ones. A compelling indication for intercalation acting as a major driving force is provided by earlier designed tris-intercalating acridine derivatives[68] with four ammonium groups and three amino-acridine rings. Such derivatives have affinities amounting to 10^{-14}. They were, however, hampered by apparent lack of DNA sequence selectivity of the ammonium groups, considered to bind to the phosphates.

Along these lines, worth mentioning are thermal denaturation studies on DNA complexes of a tris-intercalating derivative of a tricationic derivative of porphyrin, bis-acridyl-bis-arginyl-porphyrin, BABAP.[69] BABAP has two Arg arms, each substituting for an O atom of a diphenyl substituent of the porphyrin. As with the present molecules, each arm targets O_6/N_7 of the two G bases upstream from the central intercalation site. Each is extended with a Gly peptide connected by a C_6 chain to another intercalator, 9-aminoacridine. BABAP stabilized the target d(TC GGGC GCCC GA)$_2$ palindrome by 30 °C. Such a ΔT_M is similar to the one resulting from the binding to DNA of a bis-intercalating dimer of daunorubicin, which has a DNA-binding affinity amounting to 10^{-12}.[67] Comparable ΔT_M magnitudes could be anticipated from the tris-intercalators designed in the present work.

https://doi.org/10.1021/acsomega.4c05099
ACS Omega 2024, 9, 42309−42328

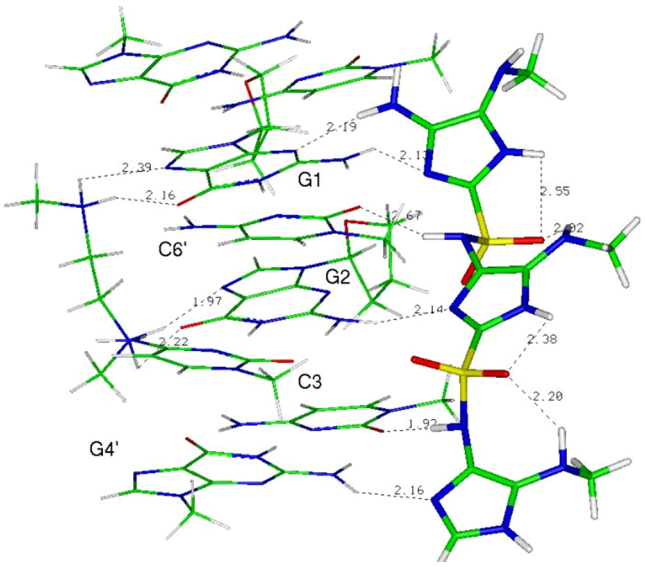


Figure 7. Complex of a model four-base pair with an (im)3 chain having a sulfonamide backbone.

The second asset relates to results on minor groove, hairpin polyamide derivatives. Several compounds from such class have affinities in the 10^{-9} to 10^{-10} range, with one amounting to 10^{-11}.[57] This is ascribed to a large part to the number of directional ligand-DNA H-bonds, 12 to 13. In this respect, all (im)3−augmented derivatives of the present study were found to be stabilized by the same number of H-bonds. DNA-binding affinities in a submicromolar range could thus be contributed by the two expanded (im)3 arms.

Worth recalling in this connection are the earlier results by Fesik et al.[70] bearing on the complexation of the FKBP protein. A ligand resulting from the linking together of two micromolar ligands, each bound to a distinct FKBP cavity, was endowed with nanomolar affinities. On the basis of the above-mentioned DNA-ligand experimental results, tris-intercalation/in-register major groove binding on one side and minor groove binding with up to 12 directional H-bonds on the other side could be expected to enable each for submicromolar affinities. It will be revealing to evaluate the affinities of **III-A**, **III-B**, and **III-C** where the two modes of binding are brought to act in synergy. This should provide the incentive for their syntheses and the evaluation of their affinities and selectivity, parallel to theoretical computations of their Binding Free Energies for target and nontarget sequences. Following their preliminary validations on experimentally reported results on mono- and bis-intercalation of anthracyclines,[67] we plan to apply newly developed approaches, such as lambda dynamics (Lagardere et al., to be published) along the present series. The modular nature of the present approach is to be mentioned. Thus, **III-B** has a connector to the im3 arms shorter than **III-A**, a single methylene instead of two ones. This could reduce to some extent its solvation entropy in solution and possibly favor **III-B** over **III-A** in the energy balances. **III-B** also has a more expanded ring than **III-A**, an anthracycline instead of an anthraquinone, enabling it to reach out and interact closer to the 9-AA ring than **III-A**. This could provide additional stability to the complex of **III-B**. **III-C** has each im3 arm connected to both the central intercalator and to one 9-AA intercalator. The im3 arms have, on account of their conjugated nature, less flexibility than the ADAM connectors of **III-A** and **III-B**. Uncomplexed **III-C** would thus have a further reduced desolvation entropy than **III-A** and **III-B**, favoring it in the energy balances. Past such starting considerations, only Free Binding Energy calculations should quantify these outlined entropy trends and the corresponding weights and trends of the complexation/desolvation enthalpies.

As mentioned above, a multistep synthetic protocol for all derivatives has been developed by some of us. It should lead to completion of syntheses in a foreseeable time span of two years. This should first enable, as a proof of principle, in vitro evaluations of their sequence selectivity and DNA binding affinities and how these are modulated upon evolving the series.

Such evaluations should foster extension to cellular tests and the search, if needed, for appropriate transporters since concerns could indeed be raised regarding their bioavailability owing to their high molecular weights and their polycationic nature. Amino-acridines are reported to favor cellular uptake (25; 71 and refs therein). This property is enhanced when upon methylation of the N_9 nitrogen the amino-acridine bears a positive charge. This is the case for compounds **III-C** and **III-D**. It was reported that polycationic conjugated molecules have a propensity to cross membranes. This is the case with "cargo" molecules designed as CPP (Cell Penetrating Peptides),[72−74] which enable transport across membranes of drug molecules, peptides, proteins, and even polyanionic molecules, such as nucleic acids (reviewed in refs 75 and 76). Conjugation of the designed compounds with CPP could, if needed, provide an asset to enhance the cellular transport. Of note is also a report that nonconjugated polycationic molecules could themselves enable facilitated transport of acridines. This is the case for spermine, spermidine and related molecules with up to four cationic charges grafted, as for compounds **III-A**−**III-D**, by an amide linkage to the C carbon of the central ring of 9-AA properties.[77] This raises the possibility that the dicationic ADAM connectors used in **III**, **III-A**, and **III-B** destined to target the G bases might facilitate rather than hamper cellular transport.

■ ASSOCIATED CONTENT

Data Availability Statement

Four AMOEBA MD input files are provided for ligands **IIIA**−**IIID**, with extensions *xyz (coordinates), *key and *prm (parameters), *dyn (for MD restarts), *sl (a slurm file to run MD on GPU processors), and regrouped for each ligand as one compressed file (lig-*-MD.doc). They are provided with the paper as Supporting Information. The calculations were performed with the publicly available free open-source software Tinker-HP (https://github.com/TinkerTools/tinker-hp).

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.4c05099.

Figure S1, molecular structures of the precursor ligands I, II, and III (ref 15); Figure S2, time evolutions of the H-bond distances stabilizing the **III-A** complex; normalized radial distribution functions in the **III-B** complex (Figure S3), the **III-C** complex (Figure S4), and the **III-D** complex (Figure S5); Figure S6a−c, representation of the complexes of **III-A′**, **III-B′**, and **III-D′** at the outcome of 140 ns MD (PDF)

Data for MD simulations (PDF)

https://doi.org/10.1021/acsomega.4c05099
ACS Omega 2024, 9, 42309−42328

AUTHOR INFORMATION

Corresponding Author

Nohad Gresh − *Laboratoire de Chimie Théorique, UMR 7616 CNRS Sorbonne Universités, Paris 75005, France;* orcid.org/0000-0001-7174-2907; Email: gresh@lct.jussieu.fr

Authors

Alberto Ongaro − *Department of Pharmaceutical and Pharmacological Sciences, University of Padova, Padova 35131, Italy*

Luc Demange − *UMR 8038 CNRS CiTCoM, Team PNAS, Faculté de Pharmacie, Université Paris-Cité, Paris 75006, France*

Giuseppe Zagotto − *Department of Pharmaceutical and Pharmacological Sciences, University of Padova, Padova 35131, Italy;* orcid.org/0000-0001-7053-4196

Giovanni Ribaudo − *DMMT, University of Brescia, Brescia 25121, Italy;* orcid.org/0000-0003-3679-5530

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.4c05099

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We wish to thank Dr. Krystel El Hage (Qubit Pharmaceuticals, France) for her help in the early stages of this work. We wish to thank the Grand Equipement de Calcul Intensif (GENCI): Institut du Développement et des Ressources en Informatique (IDRIS), Centre Informatique de l'Enseignement Supérieur (CINES), France, project x-2009-07509, and the Centre Régional Informatique et d'Applications Numériques de Normandie (CRIANN), project 19980853.

REFERENCES

(1) Gowers, D. M.; Bellamy, S. R. W.; Halford, S. E. One recognition sequence, seven restriction enzymes, five reaction mechanisms. *Nucleic Acids Res.* **2004**, *32*, 3469−3479.

(2) Korchynskyi, O.; ten Dijke, P. Identification and functional characterization of distinct critically important bone morphogenetic protein-specific response elements in the Id1 promoter. *J. Biol. Chem.* **2002**, *277*, 4883−4891.

(3) Nakahiro, T.; Kurooka, H.; Mori, K.; Sano, K.; Yokota, Y. Identification of BMP-responsive elements in the mouse Id2 gene. *Biochemical and biophysical research communications* **2010**, *399*, 416−421.

(4) Wain-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Alizon, M. Nucleotide sequence of the AIDS virus. *LAV. Cell* **1985**, *40*, 9−17.

(5) Jacobo-Molina, A.; Ding, J.; Nanni, R. G.; Clark, A. D.; Lu, X.; Tantillo, C.; Williams, R. L.; Kamer, G.; Ferris, A. L.; Clark, P.; et al. Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 A resolution shows bent DNA. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 6320−6324.

(6) Ding, J.; Das, K.; Hsiou, Y.; Sarafianos, S. G.; Clark, A. D.; Jacobo-Molina, A.; Tantillo, C.; Hughes, S. H.; Arnold, E. Structure and functional implications of the polymerase active site region in a complex of HIV-1 RT with a double-stranded DNA template-primer and an antibody fab fragment at 2.8 Å resolution11Edited by J. Karn. *J. Mol. Biol.* **1998**, *284*, 1095−1111.

(7) Wildschutte, J. H.; Baron, A.; Diroff, N. M.; Kidd, J. M. Discovery and characterization of Alu repeat sequences via precise local read assembly. *Nucleic Acids Res.* **2015**, *43*, 10292−10307.

(8) Krolewski, J. J.; Bertelsen, A. H.; Humayun, M. Z.; Rush, M. G. Members of the Alu family of interspersed, repetitive DNA sequences are in the small circular DNA population of monkey cells grown in culture. *Journal of molecular biology* **1982**, *154*, 399−415.

(9) Wu, S.; Bafna, V.; Mischel, P. S. Extrachromosomal DNA (ecDNA) in cancer pathogenesis. *Current Opinion in Genetics & Development* **2021**, *66*, 78−82.

(10) Kim, H.; Nguyen, N.-P.; Turner, K.; Wu, S.; Gujar, A. D.; Luebeck, J.; Liu, J.; Deshpande, V.; Rajkumar, U.; Namburi, S.; et al. Extrachromosomal DNA is associated with oncogene amplification and poor outcome across multiple cancers. *Nat. Genet.* **2020**, *52*, 891−897.

(11) Gresh, N.; Kahn, P. H. Theoretical design of a bistetrapeptide derivative of mitoxantrone targeted towards the double-stranded hexanucleotide sequence d(GGCGCC)2. *J. Biomol. Struct. Dyn.* **1991**, *8*, 827−846.

(12) Gresh, N.; Kahn, P. H. Theoretical Design of Novel, 4 Base Pair Selective Derivatives of Mitoxantrone. *J. Biomol. Struct. Dyn.* **1990**, *7*, 1141−1160.

(13) Gianoncelli, A.; Basili, S.; Scalabrin, M.; Sosic, A.; Moro, S.; Zagotto, G.; Palumbo, M.; Gresh, N.; Gatto, B. Rational design, synthesis, and DNA binding properties of novel sequence-selective peptidyl congeners of ametantrone. *ChemMedChem.* **2010**, *5*, 1080−1091.

(14) Ongaro, A.; Ribaudo, G.; Braud, E.; Ethève-Quelquejeu, M.; De Franco, M.; Garbay, C.; Demange, L.; Gresh, N.; Zagotto, G. Design and synthesis of a peptide derivative of ametantrone targeting the major groove of the d(GGCGCC)2 palindromic sequence. *New J. Chem.* **2020**, *44*, 3624−3631.

(15) El Hage, K.; Ribaudo, G.; Lagardère, L.; Ongaro, A.; Kahn, P. H.; Demange, L.; Piquemal, J.-P.; Zagotto, G.; Gresh, N. Targeting the Major Groove of the Palindromic d(GGCGCC)2 Sequence by Oligopeptide Derivatives of Anthraquinone Intercalators. *J. Chem. Inf. Model.* **2022**, *62*, 6649−6666.

(16) Zimmer, C.; Puschendorf, B.; Grunicke, H.; Chandra, P.; Venner, H. Influence of Netropsin and Distamycin A on the Secondary Structure and Template Activity of DNA. *Eur. J. Biochem.* **1971**, *21*, 269−278.

(17) Lown, J. W.; Krowicki, K.; Bhat, U. G.; Skorobogaty, A.; Ward, B.; Dabrowiak, J. C. Molecular recognition between oligopeptides and nucleic acids: novel imidazole-containing oligopeptides related to netropsin that exhibit altered DNA sequence specificity. *Biochemistry* **1986**, *25*, 7408−7416.

(18) Zakrzewska, K.; Lavery, R.; Pullman, B. A theoretical study of the sequence specificity in binding of lexitropsins to B-DNA. *J. Biomol. Struct. Dyn.* **1987**, *4*, 833−843.

(19) Zakrzewska, K.; Randrianarivelo, M.; Pullman, B. Drug recognition of DNA. Proposal for GC minor groove specific ligands: vinylexins. *J. Biomol. Struct. Dyn.* **1988**, *6*, 331−344.

(20) Dervan, P. B. Molecular recognition of DNA by small molecules. *Bioorganic & medicinal chemistry* **2001**, *9*, 2215−2235.

(21) Kopka, M. L.; Goodsell, D. S.; Han, G. W.; Chiu, T. K.; Lown, J.; Dickerson, R. E. Defining GC-specificity in the minor groove: side-by-side binding of the di-imidazole lexitropsin to CATGGCCATG. *Structure* **1997**, *5*, 1033−1046.

(22) Chenoweth, D. M.; Harki, D. A.; Phillips, J. W.; Dose, C.; Dervan, P. B. Cyclic Pyrrole−Imidazole Polyamides Targeted to the Androgen Response Element. *J. Am. Chem. Soc.* **2009**, *131*, 7182−7188.

(23) Muzikar, K. A.; Meier, J. L.; Gubler, D. A.; Raskatov, J. A.; Dervan, P. B. Expanding the repertoire of natural product-inspired ring pairs for molecular recognition of DNA. *Org. Lett.* **2011**, *13*, 5612−5615.

(24) Kurmis, A. A.; Yang, F.; Welch, T. R.; Nickols, N. G.; Dervan, P. B. A Pyrrole-Imidazole Polyamide Is Active against Enzalutamide-Resistant Prostate Cancer. *Cancer Res.* **2017**, *77*, 2207−2212.

(25) Fechter, E. J.; Dervan, P. B. Allosteric Inhibition of Protein−DNA Complexes by Polyamide−Intercalator Conjugates. *J. Am. Chem. Soc.* **2003**, *125*, 8476−8485.

(26) Anneheim-Herbelin, G.; Perree-Fauvet, M.; Gaudemer, A.; Helissey, P.; Giorgi-Renault, S.; Gresh, N. Porphyrin-Netropsin: A potential ligand of DNA. *Tetrahedron Lett.* **1993**, *34*, 7263.
(27) Perrée-Fauvet, M.; Gresh, N. Structure and energetics in the complexes of a double-stranded B-DNA dodecamer with netropsin derivatives of a tricationic water-soluble porphyrin: a theoretical investigation. *J. Biomol. Struct. Dyn.* **1994**, *11*, 1203−1224.
(28) Williams, L. D.; Egli, M.; Qi, G.; Bash, P.; van der Marel, G. A.; van Boom, J. H.; Rich, A.; Frederick, C. A. Structure of nogalamycin bound to a DNA hexamer. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87*, 2225−2229.
(29) Yarden, Y.; Kuang, W. J.; Yang-Feng, T.; Coussens, L.; Munemitsu, S.; Dull, T. J.; Chen, E.; Schlessinger, J.; Francke, U.; Ullrich, A. Human proto-oncogene c-kit: a new cell surface receptor tyrosine kinase for an unidentified ligand. *EMBO journal* **1987**, *6*, 3341−3351.
(30) Johansson, J. R.; Wang, Y.; Eng, M. P.; Kann, N.; Lincoln, P.; Andersson, J. Bridging ligand length controls AT selectivity and enantioselectivity of binuclear ruthenium threading intercalators. *Chemistry−A European Journal* **2013**, *19*, 6246−6256.
(31) Almaqwashi, A. A.; Andersson, J.; Lincoln, P.; Rouzina, I.; Westerlund, F.; Williams, M. C. DNA intercalation optimized by two-step molecular lock mechanism. *Sci. Rep.* **2016**, *6*, 37993.
(32) Lo, A. T.; Bryce, N. S.; Klein, A. V.; Todd, M. H.; Hambley, T. W. Novel polyamide amidine anthraquinone platinum (II) complexes: cytotoxicity, cellular accumulation, and fluorescence distributions in 2D and 3D cell culture models. *JBIC Journal of Biological Inorganic Chemistry* **2021**, *26*, 217−233.
(33) Suseela, Y.; Das, S.; Pati, S. K.; Govindaraju, T. Imidazolyl-Naphthalenediimide-Based Threading Intercalators of DNA. *ChemBioChem.* **2016**, *17*, 2162−2171.
(34) Guelev, V.; Lee, J.; Ward, J.; Sorey, S.; Hoffman, D. W; Iverson, B. L Peptide bis-intercalator binds DNA via threading mode with sequence specific contacts in the major groove. *Chemistry & Biology* **2001**, *8*, 415−425.
(35) Murr, M. M.; Harting, M. T.; Guelev, V.; Ren, J.; Chaires, J. B.; Iverson, B. L. An octakis-intercalating molecule. *Bioorganic & medicinal chemistry* **2001**, *9*, 1141−1148.
(36) Rhoden Smith, A.; Iverson, B. L. Threading polyintercalators with extremely slow dissociation rates and extended DNA binding sites. *J. Am. Chem. Soc.* **2013**, *135*, 12783−12789.
(37) Ren, P.; Wu, C.; Ponder, J. W. Polarizable Atomic Multipole-Based Molecular Mechanics for Organic Molecules. *J. Chem. Theory Comput.* **2011**, *7*, 3143−3161.
(38) Lagardère, L.; Jolly, L.-H.; Lipparini, F.; Aviat, F.; Stamm, B.; Jing, Z. F.; Harger, M.; Torabifard, H.; Cisneros, G. A.; Schnieders, M. J.; et al. Tinker-HP: a massively parallel molecular dynamics package for multiscale simulations of large complex systems with advanced point dipole polarizable force fields. *Chemical Science* **2018**, *9*, 956−972.
(39) Adjoua, O.; Lagardère, L.; Jolly, L.-H.; Durocher, A.; Very, T.; Dupays, I.; Wang, Z.; Inizan, T. J.; Célerse, F.; Ren, P.; et al. Tinker-HP: Accelerating Molecular Dynamics Simulations of Large Complex Systems with Advanced Point Dipole Polarizable Force Fields Using GPUs and Multi-GPU Systems. *J. Chem. Theory Comput.* **2021**, *17*, 2034−2053.
(40) Zhang, C.; Lu, C.; Jing, Z.; Wu, C.; Piquemal, J. P.; Ponder, J. W.; Ren, P. AMOEBA Polarizable Atomic Multipole Force Field for Nucleic Acids. *J. Chem. Theory Comput* **2018**, *14*, 2084−2108.
(41) Stone, A. J. Distributed Multipole Analysis: Stability for Large Basis Sets. *J. Chem. Theory Comput.* **2005**, *1*, 1128−1132.
(42) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785−789.
(43) Becke, A. D. A new mixing of Hartree−Fock and local density-functional theories. *J. Chem. Phys.* **1993**, *98*, 1372−1377.
(44) Dunning, T. Gaussian-Basis Sets for Use in Correlated Molecular Calculations.1. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(45) Feller, D. The role of databases in support of computational chemistry calculations. *J. Comput. Chem.* **1996**, *17*, 1571−1586.
(46) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H. et al. *Guassian 16*; Gaussian, Inc.: Wallingford, CT, 2016.
(47) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(48) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(49) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.
(50) Mao, Y.; Ge, Q.; Horn, P. R.; Head-Gordon, M. On the Computational Characterization of Charge-Transfer Effects in Noncovalently Bound Molecular Complexes. *J. Chem. Theory Comput.* **2018**, *14*, 2401−2417.
(51) Epifanovsky, E.; Gilbert, A. T. B.; Feng, X.; Lee, J.; Mao, Y.; Mardirossian, N.; Pokhilko, P.; White, A. F.; Coons, M. P.; Dempwolff, A. L.; et al. Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package. *J. Chem. Phys.* **2021**, *155*, 084801.
(52) Mardirossian, N.; Head-Gordon, M. ωB97X-V: A 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, designed by a survival-of-the-fittest strategy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904−9924.
(53) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158−6170.
(54) Boys, S. F.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* **1970**, *19*, 553−566.
(55) Chen, K. X.; Gresh, N.; Pullman, B. A theoretical investigation on the sequence selective binding of mitoxantrone to double-stranded tetranucleotides. *Nucleic Acids Res.* **1986**, *14*, 3799−3812.
(56) Masella, M.; Gresh, N.; Flament, J.-P. A theoretical study of nonadditive effects in four water tetramers. *Journal of the Chemical Society, Faraday Transactions* **1998**, *94*, 2745−2753.
(57) Dervan, P. B.; Doss, R. M.; Marques, M. A. Programmable DNA binding oligomers for control of transcription. *Current medicinal chemistry. Anti-cancer agents* **2005**, *5*, 373−387.
(58) Wittayanarakul, K.; Anthony, N. G.; Treesuwan, W.; Hannongbua, S.; Alniss, H.; Khalaf, A. I.; Suckling, C. J.; Parkinson, J. A.; Mackay, S. P. Ranking ligand affinity for the DNA minor groove by experiment and simulation. *ACS medicinal chemistry letters* **2010**, *1*, 376−380.
(59) Alniss, H. Y. Thermodynamics of DNA Minor Groove Binders. *J. Med. Chem.* **2019**, *62*, 385−402.
(60) Alniss, H. Y.; Witzel, I.-I.; Semreen, M. H.; Panda, P. K.; Mishra, Y. K.; Ahuja, R.; Parkinson, J. A. Investigation of the Factors That Dictate the Preferred Orientation of Lexitropsins in the Minor Groove of DNA. *J. Med. Chem.* **2019**, *62*, 10423−10440.
(61) Ramesh, V. V. E.; Kale, S. S.; Kotmale, A. S.; Gawade, R. L.; Puranik, V. G.; Rajamohanan, P. R.; Sanjayan, G. J. Carboxamide versus Sulfonamide in Peptide Backbone Folding: A Case Study with a Hetero Foldamer. *Org. Lett.* **2013**, *15*, 1504−1507.
(62) Feng, B.; Sundin, E.; Lincoln, P.; Mårtensson, A. K. DNA threading intercalation of enantiopure [Ru (phen) 2 bidppz] 2+ induced by hydrophobic catalysis. *Phys. Chem. Chem. Phys.* **2021**, *23*, 2238−2244.
(63) Guelev, V.; Lee, J.; Ward, J.; Sorey, S.; Hoffman, D. W.; Iverson, B. L. Peptide bis-intercalator binds DNA via threading mode with sequence specific contacts in the major groove. *Chemistry & Biology* **2001**, *8*, 415−425.
(64) Murr, M. M.; Harting, M. T.; Guelev, V.; Ren, J.; Chaires, J. B.; Iverson, B. L. An octakis-intercalating molecule. *Bioorg. Med. Chem.* **2001**, *9*, 1141−1148.

https://doi.org/10.1021/acsomega.4c05099
ACS Omega 2024, 9, 42309−42328

(65) Portugal, J.; Cashman, D. J.; Trent, J. O.; Ferrer-Miralles, N.; Przewloka, T.; Fokt, I.; Priebe, W.; Chaires, J. B. A New bisintercalating anthracycline with picomolar DNA binding affinity. *J. Med. Chem.* **2005**, *48*, 8209−8219.
(66) Chaires, J. B. A thermodynamic signature for drug−DNA binding mode. *Arch. Biochem. Biophys.* **2006**, *453*, 26−31.
(67) Chaires, J. B. Calorimetry and thermodynamics in drug design. *Annual Review of Biophysics* **2008**, *37*, 135−151.
(68) Laugâa, P.; Markovits, J.; Delbarre, A.; Le Pecq, J. B.; Roques, B. P. DNA tris-intercalation: first acridine trimer with DNA affinity in the range of DNA regulatory proteins. Kinetic studies. *Biochemistry* **1985**, *24*, 5567−5575.
(69) Far, S.; Kossanyi, A.; Verchère-Béaur, C.; Gresh, N.; Taillandier, E.; Perrée-Fauvet, M. Bis- and tris-DNA intercalating porphyrins designed to target the major groove: synthesis of acridylbis-arginyl-porphyrins, Molecular modelling of their DNA complexes, and experimental tests. *Eur. J. Org. Chem.* **2004**, *2004*, 1781−1797.
(70) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. Discovering high-affinity ligands for proteins: SAR by NMR. *Science (New York, N.Y.)* **1996**, *274*, 1531−1534.
(71) Spicer, J. A.; Gamage, S. A.; Atwell, G. J.; Finlay, G. J.; Baguley, B. C.; Denny, W. A. Structure-Activity relationships for acridine-Substituted analogues of the mixed topoisomeraseI/II inhibitor N-[(2-(dimethylamino)ethyl]acridine-4-carboxamide. *J. Med. Chem.* **1997**, *40*, 1919−1929.
(72) Di Pisa, M.; Chassaing, G.; Swiecicki, J.-M. Translocation mechanism(s) of cell penetrating peptides: biophysical studies using artificial membrane bilayers. *Biochemistry* **2015**, *54*, 194−207.
(73) Swiecicki, J.-M.; Di Pisa, M.; Burlina, F.; Lecorche, P.; Mansuy, C.; Chassaing, G.; Lavielle, S. Accumulation of cell-penetrating peptides in large unilamellar vesicles: a straightforward screening assay for investigating the internalization mechanism. *Biopolymers (Peptide Science)* **2015**, *104*, 533−543.
(74) Yousef, M.; Szabo, I.; Muranyi, J.; Illien, F.; Soltesz, D.; Bato, C.; Toth, G.; Batta, G.; Nagy, P.; Sagan, S.; Banoczi, Z. Cell-penetrating dabcyl-conting tetraarginines with backbone aromatics as uptake enhancers. *Pharmaceutics* **2023**, *15*, 141.
(75) Koren, E.; Torchilin, V. P. Cell-penetrating peptides: breaking through the other line. *Trends in Molecular Medicine* **2012**, *18*, 385−393.
(76) Srimanee, A.; Arvanitidou, M.; Kim, K.; Hallbrink, M.; Langel, U. Cell-penetrating peptides for siRNA delivery to glioblastomas. *Peptides* **2018**, *104*, 62−69.
(77) Delcros, J.-G.; Tomasi, S.; Carrington, S.; Martin, B.; Renault, J.; Blagbrough, I. S.; Uriac, P. Effect of spermine conjugation on the cytotoxicity and cellular transport of acridine. *J. Med. Chem.* **2002**, *45*, 5098−5111.

https://doi.org/10.1021/acsomega.4c05099
ACS Omega 2024, 9, 42309−42328