Chemistry–A European Journal
Research Article
doi.org/10.1002/chem.202303650
Chemistry Europe
European Chemical Societies Publishing
www.chemeurj.org

Hot Paper

Enhancing the Features of DNA Mimic Foldamers for Structural Investigations

Valentina Corvaglia,[a, b] Jiaojiao Wu,[a] Deepak Deepak,[a] Manuel Loos,[a] and Ivan Huc*[a]

DNA mimic foldamers based on aromatic oligoamide helices bearing anionic phosphonate side chains have been shown to bind to DNA-binding proteins sometimes orders of magnitude better than DNA itself. Here, we introduce new features in the DNA mimic foldamers to facilitate structural investigations of their interactions with proteins. Thirteen new foldamer sequences have been synthesized and characterized using NMR, circular dichroism, molecular modeling, and X-ray crystallography. The results show that foldamer helix handedness can be quantitatively biased by means of a single stereogenic center, that the foldamer structure can be made C_2-symmetrical as in palindromic B-DNA sequences, and that associations between foldamer helices can be promoted utilizing dedicated C-terminal residues that act as sticky ends in B-DNA structures.

Introduction

DNA mimic foldamers are aromatic oligoamides bearing anionic side chains that adopt helically folded conformations in water, reproducing the shape and charge surface distribution of the B-DNA double helix.[1] They have been shown to bind to some DNA-binding proteins better than DNA itself and thus stand as candidates for the competitive inhibition of protein-DNA interactions. Conceptually, DNA mimic foldamers relate to naturally occurring DNA mimic proteins,[2] *i.e.*, proteins that mimic the shape and surface features of DNA and that inspire the design of nonnatural proteins to interfere with DNA-protein interactions.[3] DNA mimic foldamers may also be compared to the so-called decoy oligonucleotides that have been proposed to target DNA-binding proteins such as transcription factors.[4] More remotely, analogies can be drawn with sulfated polysaccharides such as heparin. These sulfated polysaccharides bind to DNA-binding proteins, and are used for their purification by affinity chromatography.[5] An original feature of DNA mimic foldamers is therefore their abiotic nature. Their chemical constitution is distinct from those of peptides, nucleotides, and saccharides. In addition, the way their structure can be modulated is also distinct.

The first proposed DNA mimic foldamers consisted of oligoamide sequences with alternating 8-amino-2-quinolinecarboxylic acid monomer Q^{Pho} and 8-aminomethyl-2-quinolinecarboxylic acid monomer M^{Pho}, both bearing negatively charged phosphonate residues (Figure 1).[1a] Carboxylate residues have been introduced subsequently.[1b] These sequences adopt stable single helical conformations stabilized by hydrogen bonds between amide NH protons and adjacent endocyclic quinoline nitrogen atoms and by the hydrophobic effect associated with

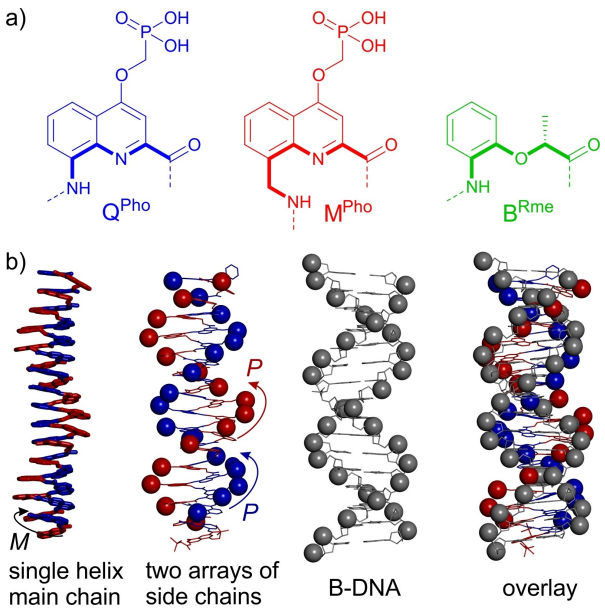

Figure 1. a) Amino acid monomers used to produce oligoamide sequences **1**–**7**. Bonds in bold indicate the inner rim of the helices involving these monomers. b) Crystal structure of the single helix of oligoamide $(M^{Pho}Q^{Pho})_{16}$ with protected ethyl phosphonate (left: aromatic oligoamide main chain only in tube representation; center left: same view with side chains),[1a] and of an ideal computationally-generated 16 base-pair B-DNA duplex (center right) and overlay of the two (right). Colors in b) are the same as in a). The structures are shown at the same scale in tube representation. Phosphorus atoms are shown in space-filling representation. Arrows show the helical sense of the main chain helix (*M*, in black) and of the arrays of side chains (*P* in red and blue).

[a] *Dr. V. Corvaglia, J. Wu, D. Deepak, M. Loos, Prof. Dr. I. Huc*
Department of Pharmacy
Ludwig-Maximilians-Universität München
Butenandtstr. 5–13, 81377 Munich (Germany)
E-mail: ivan.huc@cup.lmu.de

[b] *Dr. V. Corvaglia*
Current address: Institute for Stem-Cell Biology, Regenerative Medicine and Innovative Therapies, IRCCS Casa Sollievo della Sofferenza, San Giovanni Rotondo (Italy) & Center for Nanomedicine and Tissue Engineering (CNTE), ASST Grande Ospedale Metropolitano Niguarda, 20162 Milan (Italy)

Supporting information for this article is available on the WWW under https://doi.org/10.1002/chem.202303650

aromatic stacking. The single helical nature of the main chain can be seen in Figure 1b (left). At the surface of the single helix, the phosphonate side chains of Q^{Pho} and M^{Pho} form a double helical array that matches the positions of phosphates in B-DNA (red and blue spheres in Figure 1b). The single helical nature of DNA mimic foldamers means they cannot be involved in Watson-Crick base pairing interactions, unlike many other DNA analogs such as peptide nucleic acids and locked nucleic acids.[6] It follows that the DNA mimic foldamers do not undergo dissociation, a feature that can be put to an advantage. For example, $(M^{Pho}Q^{Pho})4$ is a stable mimic of four base-pair DNA duplexes, that is, a duplex so short that it would not be stable at submillimolar concentrations.

The structural features of the DNA mimic foldamers enable them to tightly bind to some DNA-binding proteins, particularly nonsequence selective proteins that recognize DNA through its overall shape and charges. In some cases, $(M^{Pho}Q^{Pho})_n$ oligomers outcompeted DNA and bound to the proteins even in the presence of a large excess of DNA. This led, for example, to the strong inhibition of therapeutically relevant enzymes such as human topoisomerase 1 (Top1) and human immunodeficiency virus 1 integrase (HIV-1 IN).[1] In a recent study, DNA mimic foldamers have been shown to affect chromatin composition and perturb cell cycle progression.[7] These discoveries call for further developments, in particular with regard to the elucidation of the interaction modes between the foldamers and DNA-binding proteins. However, $(M^{Pho}Q^{Pho})_n$ foldamers are not ideally suited for this purpose. For instance, they do not possess any stereogenic center and thus exist as racemic mixtures of right-handed (*P*) and left-handed (*M*) helices, potentially producing mixtures of diastereomeric complexes with proteins. In addition, they possess an N terminus and a C terminus which may lead to alternate protein binding modes depending on the N→C orientation of the foldamer with respect to the protein surface. With DNA, this potential problem can be alleviated by using C_2-symmetrical palindromic sequences. Furthermore, growing crystals of protein-DNA complexes often benefit from promoting associations between DNA strands using overhanging bases that mediate inter-duplex base-pairing.[8] The DNA mimic foldamers possess may not have the same ability to self-assemble. Here, we present the implementation of various modifications of the DNA mimic foldamers in order to bias their handedness, make them C_2-symmetrical, *i.e.*, palindromic-like, and promote their association through helix cross-sections. Some of these enhanced features have already helped to solve a crystal structure of a complex between a DNA mimic foldamer and a protein.[9]

Results and Discussion

Biasing Handedness in DNA Mimic Foldamers

Natural B-DNA contains D-deoxyribose and is a right-handed (*P*) double helix. To mimic the structure of B-DNA, the single helical foldamers should thus also display a *P* double helical array of side chains at the surface of the main chain foldamer single helix. These arrays are called exo-helices. It turns out that *P* exo-helicity is achieved when the aromatic oligoamide main chain has *M* helicity, that is with an opposite sense (Figure 1b). This may not be intuitive and can be explained as follows: the repeat units of the foldamer single helix, $M^{Pho}Q^{Pho}$ dimers, span ~0.9 helix turn. Within a sequence, two consecutive $M^{Pho}Q^{Pho}$ dimers thus have an angular shift of ~0.1 turn while stacking on top of each other, which results in a vertical rise of 3.4 Å. These are the same vertical rise and angular shift as between consecutive base pairs in B-DNA duplexes, hence the resemblance between DNA mimic foldamers and B-DNA. However, because $M^{Pho}Q^{Pho}$ dimers span less than a turn, the 0.1 turn angular shift of the next $M^{Pho}Q^{Pho}$ dimer takes place "backwards" along the helix backbone. As a result, the handedness of the side chain exo-helices is opposite to that of the main chain. Exo-helices and main chain helix would have the same handedness if the repeat unit would span more than a turn. One might, for example, achieve B-DNA mimicry with a repeat unit spanning 1.1 turn, which would also result in a 0.1 angular shift, this time "forward", between repeat units.

Quantitative handedness control has been achieved in the context of helical oligoamides of 8-amino-2-quinolinecarboxylic acid Q, *i.e.*, analogs of Q^{Pho} bearing various types of side chains, through appended chiral groups at the N or C termini.[10] Recently, we have shown that a 2-(2-aminophenoxy)propionic acid unit B^{Rme} (Figure 1a) within a Q_n sequence is also efficient at biasing *M* handedness in water.[11] However, it was unknown whether this approach would also be effective in the context of $(M^{Pho}Q^{Pho})_n$ sequences, in particular with regards to the additional flexibility imparted by the methylene groups of M^{Pho} which increase the number of main chain rotatable bonds. We thus designed DNA mimic foldamers **2**–**7** as analogs of reference sequence **1** to assess the effect of B^{Rme} in various sequence contexts (Figure 2a).

In **2** and **3**, B^{Rme} replaces one Q^{Pho} or one M^{Pho}, respectively, within a $(M^{Pho}Q^{Pho})_n$ motif. In constrast, in **4** and **5**, B^{Rme} is inserted after Q^{Pho} or M^{Pho}, respectively. Because an insertion results in a frameshift in the alternation of Q^{Pho} and M^{Pho}, sequences **6** and **7** were also designed where the insertion of B^{Rme} is compensated by the insertion of an additional M^{Pho} later or earlier in the sequence, respectively. For the preparation of these oligomers by solid phase synthesis (SPS), new monomers Fmoc-Q^{Pho}-OH and Fmoc-M^{Pho}-OH were synthesized with their phosphonate side chain protected as diethyl esters (Schemes S1–S2), a protection previously validated for solution phase synthesis.[1a] The di-*tert*-butyl esters of these Fmoc-acid precursors have been introduced earlier,[1b] but the high acid lability of the *tert*-butyl groups make the synthesis and handling of the monomers more delicate. SPS was performed on low loading Wang resin (100–200 mesh) using previously reported protocols (Scheme S3).[12] Fmoc-acid building blocks were activated *in situ* by generating the respective acid chlorides prior to coupling. The oligomers were cleaved from the resin with trifluoroacetic acid (TFA) and purified using semipreparative reverse phase high performance liquid chromatography (RP-HPLC) under acidic conditions (water/acetonitrile containing 0.1% TFA). Removal of the diethylphosphonate protections was then

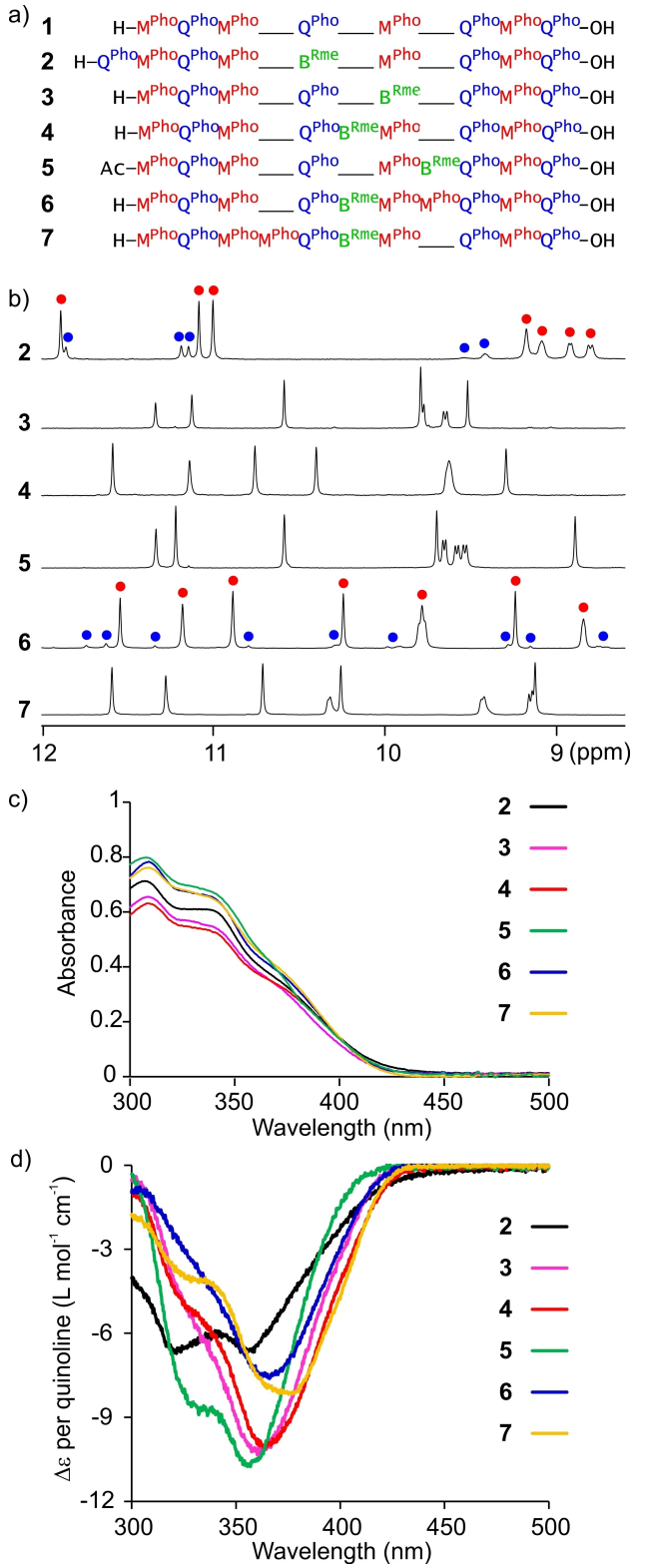

Figure 2. a) Foldamer sequences synthesized to investigate the handedness induction properties of B^{Rme}. Long dashes indicate that there is no monomer and serve to align the identical segments of the various sequences. b) Part of the ^1H NMR (500 MHz) spectra at 333 K of **2**–**7** in 50 mM NH_4HCO_3 (pH 8.5) in H_2O/D_2O (9:1, v/v) showing amide N*H* resonances. For **2** and **6**, major and minor sets of signal are marked with red and blue circles, respectively. c) UV absorption spectra of **2**–**7** (60 μM, 333 K in 50 mM NH_4HCO_3 pH 8.5) in the region of interest for the CD spectra. d) CD spectra of **2**–**7** (60 μM, 333 K in 50 mM NH_4HCO_3 pH 8.5). Molar extinction (Δε) is normalized for the number of quinoline (Q^{Pho} or M^{Pho}) units.

performed using trimethylsilyl bromide (TMSBr), and the oligomers were further purified by semipreparative RP-HPLC using a basic triethylammonium acetate buffer (pH 8.5). In a final step, they were submitted to ion exchange to deliver the side chains as water-soluble ammonium phosphonate salts providing the foldamers in 10–15% isolated yields (after two HPLC purifications and ion exchange) and a purity > 97%.

Helix handedness bias by B^{Rme} was assessed by nuclear magnetic resonance (NMR) and circular dichroism (CD) spectroscopies (Figure 2b, d). Both CD and NMR confirmed that all sequences are helically folded. NMR spectra show sharp amide and aromatic protons signals distributed over a wide range of chemical shift values (from 12 to 9 ppm for amide protons and from 8.5 to 6.5 ppm for aromatic protons, Figures 2b, S1) in agreement with previous observations.[1,12a] All CD spectra showed an intense negative band near 360 nm, typical of helically folded quinoline oligoamides with predominant *M* helicity (Figure 2d).[11,13] As a consequence of the different sequence composition and the different monomer contacts in the helices, the CD profiles show some differences. The λ_{max} values vary slightly from compound to compound and so does the relative contribution of a second negative band at 320–330 nm. Compound **2** has a distinct aromatic amine chromophore and compound **5** is acetylated at its N terminus whereas other sequences have a benzylic ammonium terminus. These features may perhaps relate to the distinct CD signatures of **2** and **5** which both have bands near 330 and 360 nm of comparable intensities. However, finding a trend and assigning the intensities to specific interactions within the helices is difficult and was not attempted. For example, sequences **6** and **7** differ minimally through the position of one M^{Pho} unit. Yet the contribution of the weaker CD band near 330 nm is barely visible in the spectrum of **6** while it is very clear for **7**. For comparison, the UV-vis absorption spectra of the various sequences are similar (Figure 2c).

Sequences **2**–**7** contain eight to ten residues and should all span over three full helix turns. For such long sequences in water, equilibrium between *M* and *P* diastereomeric conformers is expected to be slow on the NMR timescale. This is illustrated in the spectrum of achiral **1** (Figure S1) in which the signals of $C\underline{H}_2$ protons form AB systems reflecting their anisochronicity. Under a fast exchange regime, these signals would average. For chiral oligomers **2**–**7**, *P* and *M* conformational diastereomers should thus appear as distinct sets of NMR signals. The single set of resonances observed for **3**, **4**, **5**, and **7** therefore demonstrates that handedness bias towards the *M* helix is quantitative – as far as NMR can detect – in these cases (Figure 2b). In contrast, the second set of signals amounting to 30% in the spectrum of **2** and that, less intense (4%), in the spectrum of **6**, indicate the presence of some *P*-helical conformer. The incomplete handedness bias in **2** and **6** is also reflected in less intense CD bands for these compounds (Figure 2d). However, CD intensity is unreliable to quantitatively assess handedness bias because intensity also varies as a function of sequences even when handedness bias is quantitative. An illustration is the weaker CD band of **7** (Figure 2d). Its normalized CD intensity is 25% smaller than that of **5** whereas

these two compounds absorb similarly (Figure 2c). One should point that the Δε normalized "per quinoline ring" reported in Figure 2d does not consider possible differences in the contributions of Q^{Pho} and M^{Pho} monomers to the CD.

In summary, B^{Rme} efficiently biases helix handedness in DNA mimic foldamers but only in the context of certain sequence patterns. Of note, the partial handedness bias observed for **2** and **6** concerns the only two sequences where the B^{Rme} monomer is part of three consecutive more flexible, methylene-containing, monomers ($M^{Pho}B^{Rme}M^{Pho}$ or $B^{Rme}M^{Pho}M^{Pho}$).

Structural Consequences of a B^{Rme} Monomer in DNA Mimic Foldamer Sequences

B^{Rme} is a δ-amino acid and may thus be considered to be a structural analog of Q^{Pho} rather than ε-amino acid M^{Pho}. The Q^{Pho}/B^{Rme} mutation, as in **2**, might appear to be the most desirable way to introduce B^{Rme} in a sequence while preserving DNA mimicry. Yet, we have seen that handedness bias is insufficient in this case. To evaluate how the replacement of Q^{Pho} or M^{Pho} by B^{Rme} or the insertion of B^{Rme} alter the shape of DNA mimic foldamers, energy minimized molecular models were produced (Figures 3, S2–S4).

For a better visualization of the double helical arrangement of negative charges and of the major and minor grooves, model oligomers were elongated up to 48 units. Segments corresponding to **2**–**7** were introduced in the middle of these sequences and final objects were compared to a reference $(M^{Pho}Q^{Pho})_{24}$ DNA mimic foldamer with no stereogenic center. Three representative examples are shown in Figure 3, including the replacement of M^{Pho} by B^{Rme}, the insertion of B^{Rme}, and the double insertion of B^{Rme} and M^{Pho}. As predicted, the energy-minimized model of the sequence corresponding to an extended-**2** shows the smooth insertion of the B^{Rme} unit replacing one Q^{Pho} (Figure S2). Interestingly, molecular modeling also predicts a smooth insertion when B^{Rme} replaces M^{Pho} (extended-**3** in Figure 3a, b). At the scale of these large objects, replacing Q^{Pho} or M^{Pho} does not cause significant structural differences. In both extended-**4** (Figure 3c, d) and extended-**5** (Figure S3), the inserted B^{Rme} causes a swap of the major and minor grooves of the DNA mimic. We have described in earlier publications that, in the case of $(M^{Pho}Q^{Pho})_n$ sequences, and unlike B-DNA, the two grooves have similar widths.[1,14] Because of these matching widths, the swap then has limited structural consequences for what concerns the overall shape. It remains that M^{Pho} and Q^{Pho} line opposite sides of each groove and that this role is reverted if there is a groove swap. Furthermore, other DNA mimics have been designed where the groove widths differ more[1] and for which a swap of grooves may constitute an important structural change. Nevertheless, one cannot exclude potential benefits of such a swap to recognize some DNA-binding proteins. In extended-**6** (Figure 3e, f) and extended-**7** (Figure S4), the grooves do not swap so groove swapping can be avoided using such sequences. Initial attempts to crystallize sequences **2**–**7** to consolidate the results of

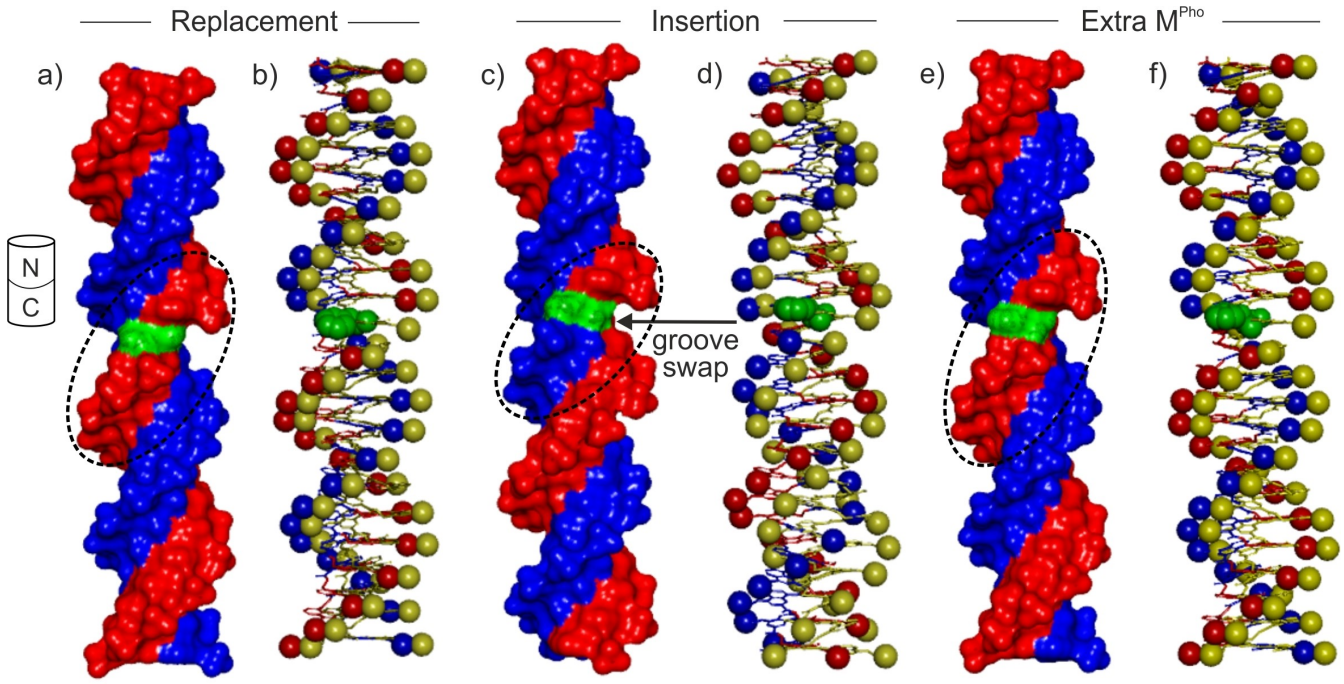

Figure 3. Energy minimized molecular models (Maestro, MMFFs force field, water as implicit solvent)[15] of: extended-**3** (a) and its overlay with $(M^{Pho}Q^{Pho})_{24}$ (b); extended-**4** (c) and its overlay with $(M^{Pho}Q^{Pho})_{24}$ (d); and extended-**6** (e) and its overlay with $(M^{Pho}Q^{Pho})_{24}$ (f). In a), c), and e), the molecule isosurface is shown with monomers color-coded as in Figure 1. In b), d) and f), molecules are shown in tube representation except the phosphorous atoms which are shown in space-filling representation. $(M^{Pho}Q^{Pho})_{24}$ is in yellow and the other molecule is color-coded as in Figure 1. The area near the B^{Rme} monomer is encircled by a dashed line in a), c), and e). The arrow in c) indicates a minor/major groove swapping on each side of the B^{Rme} monomer: monomers on either side of the B^{Rme} monomer have different colors. In contrast, this does not occur in a) and e).

molecular modeling by experimental observations did not yield diffracting crystals. However, a crystal structure was eventually obtained upon adding further features to the sequence, as described in the last section below.

Palindromic-Like C_2-Symmetrical DNA Mimic Foldamers

The next DNA mimic foldamer feature we envisaged is C_2 symmetry, by analogy with palindromic DNA sequences,[16] *i.e.*, DNA sequences that are self-complementary and can thus form C_2-symmetrical duplexes. Note that palindromes also exist in peptides but with a definition that does not entail symmetry.[17] Interest in C_2-symmetrical DNA mimic foldamers stemmed from the prospect of targeting proteins that recognize palindromic DNA, such as some transcription factors[18] and restriction enzymes,[19] which are often themselves C_2-symmetrical homodimers. Furthermore, the geometry of complexes between palindromic DNA and a given protein does not depend on the 5′-3′ DNA orientation. This degeneracy is desirable for crystal growth where the presence of distinct complexes can be an impediment. Thus, palindromic DNA sequences have been used to grow crystals of DNA-protein complexes, even with proteins that do not typically target such sites.[20] We therefore anticipated that this feature would be useful for DNA mimic foldamers as well.

C_2 symmetry was introduced in the DNA mimic foldamers by means of central diamine or diacid linkers that locally revert the C→N strand polarity. Four different linkers were considered, leading to the design of sequences **8**–**11** (Figure 4). Many others may be conceived depending on the desired outcome. The aliphatic linker of **8** and the aromatic linker of **9** are expected to contribute similarly to helix curvature. The rigid 2,6-pyridinedicarboxylic acid linker of **9** is a classical helicogenic monomer in aromatic oligoamide foldamers.[21] The linker of **8** is more flexible, even though its central oxygen atom may form hydrogen bonds with neighbor amide protons, stabilizing a helical structure as does the endocyclic nitrogen atom of **9**. The diamine linkers of **10** and **11** have been explored previously in the context of organic solvent-soluble Q_n oligomers.[22] Molecular models of **8**–**11** highlight their predicted symmetrical structures (Figure 4). The linkers provide various alternatives to put the grooves of the two halves of the molecules more or less in register. They may allow to deviate from ideal DNA mimicry, depending on what is desired. In this respect, the more flexible aliphatic linkers may provide a range of acceptable local helix curvature. They may also play the role of a hinge and permit a local kink in the foldamer structure to target proteins known to kink DNA.[23]

To prepare palindromic-like DNA mimic foldamers **8**–**11**, we developed a solution-phase fragment condensation approach using chiral sequence **4** as a building block (Figure 5a). The diethylphosphonate-protected precursor of **4** prepared by SPS could be cleaved from the resin before or after acetylation of the N terminus to provide fragments A and B, respectively, which were purified by RP-HPLC. Fragment A was reacted with diacid linkers L1 activated as *N*-hydroxysuccinimide (NHS) esters

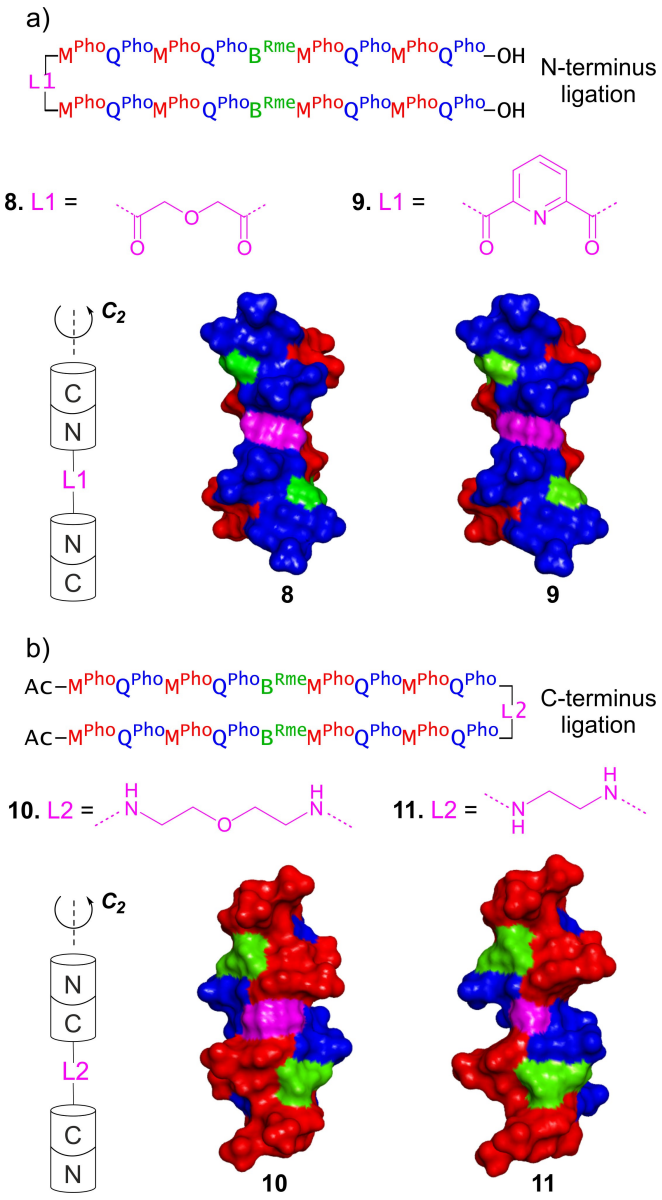

Figure 4. Palindromic-like DNA mimic foldamers. a) Formulas and energy-minimized molecular models of **8** and **9** with central diacid linkers. b) Formulas and energy-minimized molecular models of **10** and **11** with central diamine linkers. Models were obtained with Maestro (MMFFs force field, water as implicit solvent).[15] They are shown at the same scale down their C_2 symmetry axis (through the linker in purple) as solvent accessible isosurfaces color-coded according to the type of monomer as in Figure 1.

to yield the precursors of **8** and **9** which were further purified by RP-HPLC (Scheme S4). Conversely, acetylated fragment B was coupled to the diamine linkers L2 using benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) activation to yield the protected precursors of **10** and **11** which were also purified by RP-HPLC (Scheme S5). Cleavage of the diethylphosphonate esters with TMSBr yielded the final products which were again purified by RP-HPLC and subjected to ion exchange to generate ammonium salts.

The crude HPLC traces shown in Figure 5b highlight the high conversion yields of the double couplings on the linkers

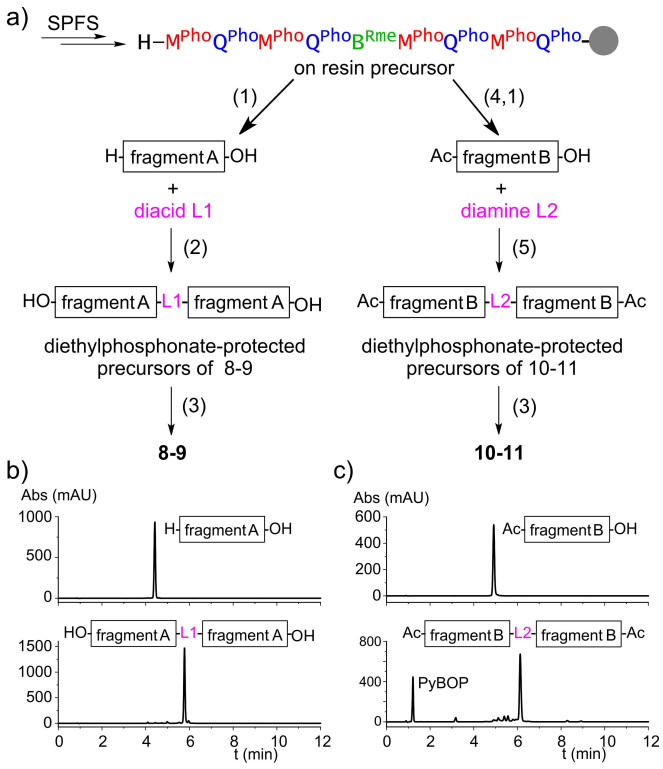

Figure 5. a) Scheme for the solution phase synthesis of **8**–**11** from a common precursor produced by SPS. (1) TFA-mediated resin cleavage; (2) double condensation of the N terminus of fragment A on diacid linkers activated as bis-NHS esters; (3) deprotection of diethyl-phosphonate esters with TMSBr; (4) acetylation of N terminus; (5) double condensation of diamine linkers on the C terminus of fragment B activated with PyBOP. b) Representative examples of the RP-HPLC traces of crude products after step (2) (left) or (5) (right). The chromatograms of the pure fragment precursors are shown above for comparison. Conditions: linear gradient from 30% to 80% solvent B in 15 min, and C8 column; A: H_2O + 0.1% TFA and B: acetonitrile + 0.1% TFA.

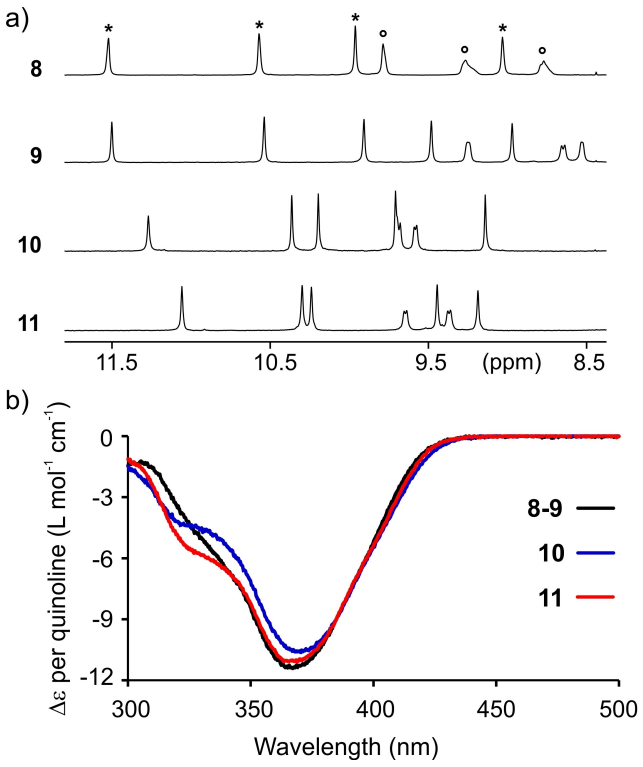

Figure 6. a) Part of the ^{1}H NMR (500 MHz) spectra at 333 K of **8**–**11** in 50 mM NH_4HCO_3 (pH 8.5) in H_2O/D_2O (9:1, vol/vol) showing amide resonances. Signals assigned to aromatic N*H*s and aliphatic N*H*s are indicated by stars and circles, respectively. The signals of other aliphatic N*H*s are found at higher fields and overlap with aromatic protons. b) CD spectra of **8**–**11** recorded at 333 K in 50 mM aqueous NH_4HCO_3 (pH 8.5). The molar extinction coefficient (Δε) is normalized for the number of quinoline units (Q^{Pho} and M^{Pho}). The linkers and B^{Rme} units do not absorb in this region.

(see also the Supporting Information). With both diamine and diacid linkers, intermediates where the linker had reacted only once were consumed or kept to negligible amounts. In turn, the high yields greatly facilitated HPLC purification and the final sequences were obtained in 7–11% isolated yield from initial resin loading with >97% purity. Even though these reactions are not the first fragment couplings between foldamer segments prepared on solid phase,[24] the conversion yields reflect particularly well-behaved couplings without requiring excess reagent. This chemistry is currently being successfully extended to foldamer segments twice as long as **8**–**11**. Progress will be reported in due course.

The folding and symmetry of **8**–**11** are reflected in their ^{1}H NMR and CD spectra (Figure 6). All NMR spectra show five degenerate aromatic N*H* resonances indicating an overall symmetrical structure (Figure 6a). Not all resonances N*H* can be distinguished in the 1D ^{1}H NMR spectra because some overlap with aromatic C*H* signals. However, 2D ^{15}N–^{1}H heteronuclear single quantum coherence (HSQC) spectra show the expected number of resonances for symmetrical structures, that is, half the total number of NH functions (Figure S5). The CD spectra confirm the *M* helicity (Figure 6b). The single sets of ^{1}H NMR signals show that handedness bias is quantitative. The combination of *M* helicity and overall symmetry can be fulfilled only with an average C_2 symmetry axis; a center of inversion or a symmetry plane can be excluded. The similar ^{1}H NMR spectra of **8** and **9** as well as their perfectly overlapping CD spectra suggest that the two diacid linkers produce a similar arrangement of the two halves of the molecule. In contrast, the slight differences in the CD spectra of **10** and **11** can be interpreted as resulting from different aryl-aryl contacts due to the different lengths of the diamine linkers.

Adding Sticky Ends to DNA Mimic Foldamers

In B-DNA, sticky ends, as opposed to blunt ends, refer to overhanging unpaired nucleotides at the 3′ or 5′ of a strand within a duplex. They occur naturally from the staggered cut of B-DNA by restriction enzymes. Even when such overhangs contain as few as two bases, they may promote base pairing at high (millimolar) concentrations. This property can be advantageous when crystallizing DNA alone or in complex with proteins. The overhanging bases promote inter-duplex base-pairing and may favor the formation of continuous stacks of B-DNA duplexes that bring cohesion to the crystal lattices.[8] We

devised that inter-helix associations may also be implemented in DNA mimic foldamers utilizing different types of sticky ends. For example, some aromatic oligoamide sequences have been shown to promote dimerization via anti-parallel double helix formation.[25] Other sequences mediate the head-to-head stacking of the C-terminal cross-section of aromatic single helices.[26] Sequence **12** (Figure 7a) was thus conceived to contain an $(M^{Pho}Q^{Pho})_6$ DNA mimic segment at its N terminus, handedness control by a B^{Rme} unit, and a $Q^{Ala}Q^{Ala}Q^{Asp}$ C-terminal segment known to promote dimerization by stacking of single helices and that amount here to a sticky end. The lack of a charged side chain in the penultimate residue at the C terminus has been shown to be important for dimerization to occur.[26] The ^{1}H NMR spectrum of **12** was recorded at different concentrations (Figure S6). As with previous oligomers,[26] two sets of signals were observed whose proportions vary with concentration, in agreement with a monomer-dimer equilibrium. Sequences such as **1** show no such effect.

Evidence of head-to-head dimerization via stacking the C-terminal cross-section of two helices was obtained from the solid state structure of **13**. This sequence is an analog of **12** where a Q^{Sem} replaces a Q^{Ala}. The selenium-containing monomer was intended to facilitate X-ray structure elucidation using anomalous scattering, but this proved unnecessary for **13**. Crystals of **13** were obtained using the hanging drop method and diffracted up to 3 Å (see Supporting Information). The structure was solved by molecular replacement using an energy-minimized molecular model. The asymmetric unit contained two molecules (Figure S7). It confirmed the handedness bias imparted by the B^{Rme} monomer as well as dimerization mediated by the sticky ends at the C terminus (Figure 7c box). As with earlier structures, helix-helix interactions bring the two C-terminal carboxylates in close proximity, from which we infer that one of the two must be in its carboxylic acid form and the interaction be mediated by a hydrogen bond.

The crystal structure of **13** is the first of a DNA mimic foldamer bearing phosphonate side chains. Previously, structures of the diethylphosphonate precursors have been obtained after growing crystals from organic solvents.[1a] A structure of a carboxylate analog had also been obtained using a crystal grown from an aqueous solution.[1b] However, all prior attempts to solve the structure of the parent series with phosphonate side chains had been unsuccessful until now. At best, crystals that did not diffract were obtained. Whether the better result with **13** reflects a contribution from the sticky end, for example, by stiffening the lattice, cannot be ascertained from only one example. Obviously, the sticky end was not an impediment. The crystal lattice features remarkably large pores with up to 77% solvent content (Figure 7c, S8) that explain the low diffraction intensity and low resolution of the collected data. One may speculate that the strong negative charge density of DNA mimic foldamers with phosphonate side chains – each side chain may be up to two negative charges – may be responsible for the large pores of the structure of **13** and for the general difficulty in obtaining diffracting crystals. In nucleic acid solid-state structures, negative charge density is often partially screened by divalent metal cations. Such cations were also

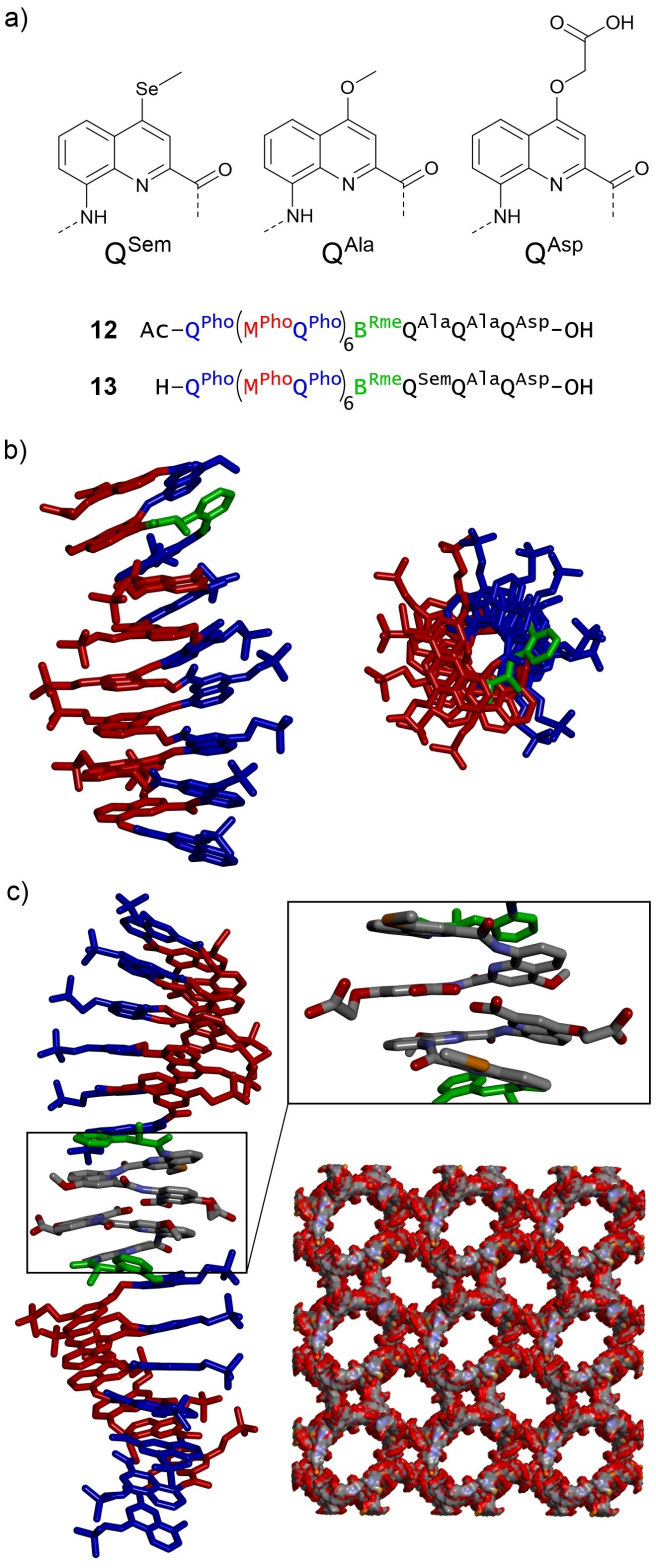

Figure 7. a) Formula of Q^{Sem}, Q^{Ala} and Q^{Asp} monomers and of sequences **12** and **13**. b) Solid state structure of **13** (PDB# 8QHM) showing the side view and top view of one of the two molecules found in the asymmetric unit. c) Crystal packing. Two helices stacked head-to-head via their C-terminal cross-sections are shown on the right. The box enlarges the contact area. On the left, a view of the crystal lattice down the *c* crystallographic axis highlights the very large pore of the structure.

involved in the structure of the carboxylate functionalized DNA mimic foldamers.[1b] However, this effect has until now not been helpful to crystallize phosphonate-functionalized aromatic helices.

Conclusions

In conclusion, we have successfully implemented new features in DNA mimic foldamers that are typical of nucleic acid double helices. Specifically, handedness control, palindromic-like C_2 symmetry, and chain extensions that promote helix-helix associations akin to sticky ends in nucleic acids can all be implemented in the DNA mimics. These features are useful to elucidate the structures of DNA-protein complexes and we anticipate that they will be similarly helpful to investigate the structures of complexes between DNA mimic foldamers and DNA-binding proteins. Steps in this direction have already been made.[9]

Supporting Information

The data that support the findings of this study are available in the supplementary material of this article. The authors have cited additional references within the Supporting Information.[27–33]

Acknowledgements

We acknowledge financial support from the Deutsche Forschungsgemeinschaft (DFG) via projects HU1766/2-1 and 325871075 (CRC1309-C7), from the European Research Council (ERC) under the European Union's Horizon Europe Framework Programme (Grant Agreement No. ERC-2021-ADG-320892), from the German Academic Exchange Service (DAAD, predoctoral fellowship to D. D.) and from the China Scholarship Council (CSC, predoctoral fellowship to J. W.). We thank D. Bindl for providing the B^{Rme} monomer and L. Allmendinger for assistance with NMR measurements. Synchrotron data were collected at beamline ID23-1 at ESRF Grenoble, France. We thank A. McCarthy for his assistance in using the beamline and J. Sigl for his help with the generation of the geometric restraint library. Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interests

The authors declare no conflict of interest.

Data Availability Statement

The crystallographic data that support the findings of this study are openly available in Protein Data Bank at https://doi.org/10.2210/pdb8QHM/pdb. Other data may obtained from the authors upon reasonable request.

Keywords: Aromatic oligoamide · DNA mimicry · Foldamers · Helix handeness control · Structure elucidation

[1] a) K. Ziach, C. Chollet, V. Parissi, P. Prabhakaran, M. Marchivie, V. Corvaglia, P. P. Bose, K. Laxmi-Reddy, F. Godde, J. M. Schmitter, S. Chaignepain, P. Pourquier, I. Huc, *Nat. Chem.* **2018**, *10*, 511–518; b) V. Corvaglia, D. Carbajo, P. Prabhakaran, K. Ziach, P. K. Mandal, V. D. Santos, C. Legeay, R. Vogel, V. Parissi, P. Pourquier, I. Huc, *Nucleic Acids Res.* **2019**, *47*, 5511–5521.

[2] a) S. S. Hegde, M. W. Vetting, S. L. Roderick, L. A. Mitchenall, A. Maxwell, H. E. Takiff, J. S. Blanchard, *Science* **2005**, *308*, 1480–1483; b) P. A. Tsonis, B. Dwivedi, *Biochim. Biophys. Acta Mol. Cell Res.* **2008**, *1783*, 177–187; c) H.-C. Wang, C.-C. Chou, K.-C. Hsu, C.-H. Lee, A. H. J. Wang, *IUBMB Life* **2019**, *71*, 539–548; d) H.-C. Wang, C.-H. Ho, K.-C. Hsu, J.-M. Yang, A. H. J. Wang, *Biochemistry* **2014**, *53*, 2865–2874; e) H.-C. Wang, H.-C. Wang, T.-P. Ko, Y.-M. Lee, J.-H. Leu, C.-H. Ho, W.-P. Huang, C.-F. Lo, A. H. J. Wang, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 20758–20763; f) M. D. Walkinshaw, P. Taylor, S. S. Sturrock, C. Atanasiu, T. Berge, R. M. Henderson, J. M. Edwardson, D. T. F. Dryden, *Mol. Cell* **2002**, *9*, 187–194; g) F. Ye, I. Kotta-Loizou, M. Jovanovic, X. Liu, D. T. F. Dryden, M. Buck, X. Zhang, *eLife* **2020**, *9*, e52125.

[3] a) D. Yüksel, P. R. Bianco, K. Kumar, *Mol. BioSyst.* **2016**, *12*, 169–177; b) F. Haque, C. Freniere, Q. Ye, N. Mani, E. M. Wilson-Kubalek, P.-I. Ku, R. A. Milligan, R. Subramanian, *Nat. Cell Biol.* **2022**, *24*, 1088–1098; c) D. T. F. Dryden, *Trends Biotechnol.* **2006**, *24*, 378–382; d) H.-C. Wang, C.-H. Ho, C.-C. Chou, T.-P. Ko, M.-F. Huang, K.-C. Hsu, A. H. J. Wang, *Nucleic Acids Res.* **2016**, *44*, 4440–4449.

[4] a) R. Crinelli, M. Bianchi, L. Gentilini, M. Magnani, *Nucleic Acids Res.* **2002**, *30*, 2435–2443; b) B. Johari, M. Moradi, in *Methods in Molecular Biology, Vol. 2521* (Ed.: W. Walther), Humana, New York, NY, **2022**, pp. 207–230; c) B. Johari, J. Zargan, *Cell Biol. Int.* **2017**, *41*, 1335–1344; d) G. Casas, F. Perche, P. Midoux, C. Pichon, J. M. Malinge, *Mol. Ther. Nucleic Acids* **2022**, *29*, 162–175.

[5] a) M. C. Z. Meneghetti, A. J. Hughes, T. R. Rudd, H. B. Nader, A. K. Powell, E. A. Yates, M. A. Lima, *J. R. Soc. Interface* **2015**, *12*, 20150589; b) A. K. Powell, E. A. Yates, D. G. Fernig, J. E. Turnbull, *Glycobiology* **2004**, *14*, 17R-30R.

[6] a) A. Porcheddu, G. Giacomelli, *Curr. Med. Chem.* **2005**, *12*, 2561–2599; b) A. Gupta, A. Mishra, N. Puri, *J. Biotechnol.* **2017**, *259*, 148–159; c) D. A. Braasch, D. R. Corey, *Chem. Biol.* **2001**, *8*, 1–7; d) B. Vester, J. Wengel, *Biochemistry* **2004**, *43*, 13233–13241.

[7] V. Kleene, V. Corvaglia, E. Chacin, I. Forne, D. B. Konrad, P. Khosravani, C. Douat, C. F. Kurat, I. Huc, A. Imhof, *Nucleic Acids Res.* **2023**, *51*, 9629–9642.

[8] a) P. A. Rice, S.-w. Yang, K. Mizuuchi, H. A. Nash, *Cell* **1996**, *87*, 1295–1306; b) S. C. Schultz, G. C. Shields, T. A. Steitz, *Science* **1991**, *253*, 1001–1007; c) S. C. Schultz, G. C. Shields, T. A. Steitz, *J. Mol. Biol.* **1990**, *213*, 159–166; d) P. H. Winegar, O. G. Hayes, J. R. McMillan, C. A. Figg, P. J. Focia, C. A. Mirkin, *Chem* **2020**, *6*, 1007–1017; e) A. Shamim, N. Parveen, V. K. Subramani, K. K. Kim, *Crystals* **2020**, *10*, 1093.

[9] D. Deepak, V. Corvaglia, J. Wu, L. Allmendinger, I. Huc, *Chemarxiv.*, DOI: 10.26434/chemrxiv-2023-9wnb8. This content is a preprint and has not been peer-reviewed.

[10] a) A. M. Kendhale, L. Poniman, Z. Dong, K. Laxmi-Reddy, B. Kauffmann, Y. Ferrand, I. Huc, *J. Org. Chem.* **2011**, *76*, 195–200; b) S. J. Dawson, Á. Mészáros, L. Pethő, C. Colombo, M. Csékei, A. Kotschy, I. Huc, *Eur. J. Org. Chem.* **2014**, *2014*, 4265–4275; c) L. Zheng, D. Zheng, Y. Wang, C. Yu, K. Zhang, H. Jiang, *Org. Biomol. Chem.* **2019**, *17*, 9573–9577; d) L. Yang, C. Ma, B. Kauffmann, D. Li, Q. Gan, *Org. Biomol. Chem.* **2020**, *18*, 6643–6650; e) Z. Liu, X. Hu, A. M. Abramyan, Á. Mészáros, M. Csékei, A. Kotschy, I. Huc, V. Pophristic, *Chem. Eur. J.* **2017**, *23*, 3605–3615.

[11] D. Bindl, E. Heinemann, P. K. Mandal, I. Huc, *Chem. Commun.* **2021**, *57*, 5662–5665.

[12] a) B. Baptiste, C. Douat-Casassus, K. Laxmi-Reddy, F. Godde, I. Huc, *J. Org. Chem.* **2010**, *75*, 7175–7185; b) V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat, I. Huc, *Chem. Eur. J.* **2023**, *29*, e202300898.

[13] C. Dolain, H. Jiang, J.-M. Léger, P. Guionneau, I. Huc, *J. Am. Chem. Soc.* **2005**, *127*, 12943–12951.

[14] The major and minor grooves of the DNA mimic foldamers can be distinguished from the fact that $Q^{Pho}M^{Pho}$ contains a methylene bridge whereas $M^{Pho}Q^{Pho}$ does not. The distance, at the surface of the mimic, from the side chain of Q^{Pho} to the side chain of the M^{Pho} forward in the sequence is a little longer than the distance from the side chain of M^{Pho} to the side chain of Q^{Pho} forward in the sequence. However, groove widths only differ by the contribution of this methylene group.
[15] *Maestro*, *Schrödinger*, LLC, New York, NY, **2021**.
[16] M. K. Ganapathiraju, S. Subramanian, S. Chaparala, K. B. Karunakaran, *Hum. Genome Var.* **2020**, *7*, 40.
[17] M. Giel-Pietraszuk, M. Hoffmann, S. Dolecka, J. Rychlewski, J. Barciszewski, *J. Protein Chem.* **2003**, *22*, 109–113.
[18] a) Y. Fujii, T. Shimizu, T. Toda, M. Yanagida, T. Hakoshima, *Nat. Struct. Biol.* **2000**, *7*, 889–893; b) W. Keller, P. König, T. J. Richmond, *J. Mol. Biol.* **1995**, *254*, 657–667.
[19] a) N. C. Horton, J. J. Perona, *Biochemistry* **2004**, *43*, 6841–6857; b) M. Newman, T. Strzelecka, L. F. Dorner, I. Schildkraut, A. K. Aggarwal, *Nature* **1994**, *368*, 660–664.
[20] a) C. Park, J. L. Campbell, W. A. Goddard, 3rd, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 4892–4896; b) C. Y. Chen, T. P. Ko, T. W. Lin, C. C. Chou, C. J. Chen, A. H. Wang, *Nucleic Acids Res.* **2005**, *33*, 430–438; c) Z. Zhang, Y. Gong, L. Guo, T. Jiang, L. Huang, *Mol. Microbiol.* **2010**, *76*, 749–759.
[21] I. Huc, *Eur. J. Org. Chem.* **2004**, *2004*, 17–29.
[22] N. Delsuc, L. Poniman, J.-M. Léger, I. Huc, *Tetrahedron* **2012**, *68*, 4464–4469.
[23] a) H. Robinson, Y. G. Gao, B. S. McCrary, S. P. Edmondson, J. W. Shriver, A. H. Wang, *Nature* **1998**, *392*, 202–205; b) K. K. Swinger, K. M. Lemberg, Y. Zhang, P. A. Rice, *EMBO J.* **2003**, *22*, 3749–3760; c) R. Sánchez-Giraldo, F. J. Acosta-Reyes, C. S. Malarkey, N. Saperas, M. E. Churchill, J. L. Campos, *Acta Crystallogr. Sect. D* **2015**, *71*, 1423–1432.
[24] a) J. Fremaux, L. Fischer, T. Arbogast, B. Kauffmann, G. Guichard, *Angew. Chem. Int. Ed.* **2011**, *50*, 11382–11385; *Angew. Chem.* **2011**, *123*, 11584–11587; b) R. David, R. Günther, L. Baumann, T. Lühmann, D. Seebach, H.-J. Hofmann, A. G. Beck-Sickinger, *J. Am. Chem. Soc.* **2008**, *130*, 15311–15317.
[25] a) J. Shang, Q. Gan, S. J. Dawson, F. Rosu, H. Jiang, Y. Ferrand, I. Huc, *Org. Lett.* **2014**, *16*, 4992–4995; b) B. Teng, P. K. Mandal, L. Allmendinger, C. Douat, Y. Ferrand, I. Huc, *Chem. Sci.* **2023**, *14*, 11251–11260; c) V. Koehler, G. Bruschera, E. Merlet, P. K. Mandal, E. Morvan, F. Rosu, C. Douat, L. Fischer, I. Huc, Y. Ferrand, *Angew. Chem. Int. Ed.* **2023**, *62*, e202311639; *Angew. Chem.* **2023**, *135*, e202311639.
[26] D. Bindl, P. K. Mandal, L. Allmendinger, I. Huc, *Angew. Chem. Int. Ed.* **2022**, *61*, e202116509; *Angew. Chem.* **2022**, *134*, e202116509.
[27] J. Buratto, C. Colombo, M. Stupfel, S. J. Dawson, C. Dolain, B. Langlois d'Estaintot, L. Fischer, T. Granier, M. Laguerre, B. Gallois, I. Huc, *Angew. Chem. Int. Ed.* **2014**, *53*, 883–887; *Angew. Chem.* **2014**, *126*, 902–906.
[28] X. Hu, S. J. Dawson, Y. Nagaoka, A. Tanatani, I. Huc, *J. Org. Chem.* **2016**, *81*, 1137–1150.
[29] D. Nurizzo, T. Mairs, M. Guijarro, V. Rey, J. Meyer, P. Fajardo, J. Chavanne, J.-C. Biasci, S. McSweeney, E. Mitchell, *J. Synchrotron Radiat.* **2006**, *13*, 227–238.
[30] C. Vonrhein, C. Flensburg, P. Keller, A. Sharff, O. Smart, W. Paciorek, T. Womack, G. Bricogne, *Acta Crystallogr. Sect. D* **2011**, *67*, 293–302.
[31] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr.* **2007**, *40*, 658–674.
[32] D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkoczi, V. B. Chen, T. I. Croll, B. Hintze, L.-W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams, P. D. Adams, *Acta Crystallogr. Sect. D* **2019**, *75*, 861–877.
[33] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. Sect. D* **2010**, *66*, 486–501.

Manuscript received: November 3, 2023
Accepted manuscript online: January 9, 2024
Version of record online: January 30, 2024