Single helically folded aromatic oligoamides that mimic the charge surface of double-stranded B-DNA

Krzysztof Ziach [1], Céline Chollet[1], Vincent Parissi[2], Panchami Prabhakaran [1], Mathieu Marchivie Valentina Corvaglia [1], Partha Pratim Bose [1], Katta Laxmi-Reddy[1], Frédéric Godde[1], Jean-Marie Schmitter[1], Stéphane Chaignepain[1], Philippe Pourquier [4] and Ivan Huc

Numerous essential biomolecular processes require the recognition of DNA surface features by proteins. Molecules mimicking these features could potentially act as decoys and interfere with pharmacologically or therapeutically relevant protein–DNA interactions. Although naturally occurring DNA-mimicking proteins have been described, synthetic tunable molecules that mimic the charge surface of double-stranded DNA are not known. Here, we report the design, synthesis and structural characterization of aromatic oligoamides that fold into single helical conformations and display a double helical array of negatively charged residues in positions that match the phosphate moieties in B-DNA. These molecules were able to inhibit several enzymes possessing non-sequence-selective DNA-binding properties, including topoisomerase 1 and HIV-1 integrase, presumably through specific foldamer–protein interactions, whereas sequence-selective enzymes were not inhibited. Such modular and synthetically accessible DNA mimics provide a versatile platform to design novel inhibitors of protein–DNA interactions.

Nucleic acids (NAs) play a central role in biomolecular processes. One key property from which this role derives is the ability of complementary sequences to hybridize through base-pairing. NA mimics have been designed and synthesized that also exhibit base-pairing abilities while possessing backbones different from natural ribo- (or deoxyribo-) phosphates. These mimics resemble NAs but also feature essential differences that can result in improved behaviour. For example, peptide NAs[1] and locked NAs[2,3] have enhanced abilities to recognize NAs and to interfere with NA–NA recognition while resisting protease and nuclease degradation, and have found applications in therapeutic approaches such as antisense strategies and diagnostic assays[4,5]. Another key property of NAs is their involvement in interactions with proteins through their surface features, that is, their overall shape, the array of phosphate ions and the chemical groups belonging to bases that are exposed in grooves[6,7]. Molecules that reproduce these features may interfere with biological functions requiring NA–protein interactions, including DNA replication and repair, and gene expression. However, no systematic development of synthetic NA mimics displaying the surface features of NAs has been undertaken until now. Yet, the viability of such mimics is validated by the existence of their natural counterparts. Indeed, DNA-mimic proteins have been found to possess helical arrays of negatively charged residues at their surface and to recognize DNA-binding proteins[8–10].

Here, we introduce a modular molecular platform based on single helically folded aromatic oligoamides that mimic the charge surface of double-stranded B-DNA, and show their potential to selectively inhibit several DNA-processing enzymes. These molecules bind to some DNA-binding proteins, presumably at DNA-binding sites, but not to proteins such as transcription factors that recognize DNA sequence features in grooves, features that have not been reproduced in our mimics. Their mode of action mirrors that of other oligomeric molecules that may inhibit DNA–protein interactions, not through the binding of proteins, but instead through the binding of DNA, for example, pyrrole-imidazole oligoamides[11,12], metallohelices[13,14], or DNA itself[15]. They are related to synthetic oligomers that display arrays of side chains similar to peptidic helices[16–18] and also to deoxyribonucleohelicates[19], although these have, to the best of our knowledge, not been tested for enzyme inhibition. They share a polyanionic nature with some heterogeneously sulfated polysaccharides, such as heparin[20] and other polyanions[21,22] that also inhibit DNA-binding enzymes as well as many enzymes unrelated to DNA. Helical aromatic amides thus combine distinct advantages: they possess a defined chemical composition, length and conformation; they can be subjected to selective chemical modifications at defined positions in space; and their structural resemblance to DNA is in principle amenable to rationalization and design.

Results and discussion

Foldamer design. Oligoamides derived from 8-amino-2-quinoline carboxylic acid **Q** fold into single helical conformations in the solid state and in solution[23,24]. Folding is driven by electrostatic repulsions and hydrogen bonds between the amide functions and adjacent residues (Supplementary Fig. 12). Aromatic stacking within the helix brings additional stabilization. Thus, exceptional conformational stability has been observed in a range of solvents, in particular in protic media[25]. For example, a $\mathbf{Q}_8$ strand forms a three-turn helix that resists denaturation at 120 °C in dimethyl sulfoxide[23]. The diameter of the inner aromatic core of $\mathbf{Q}_n$ helices (~9.4 Å between C4 atoms of contiguous **Q** rings) matches the size of a B-DNA base pair

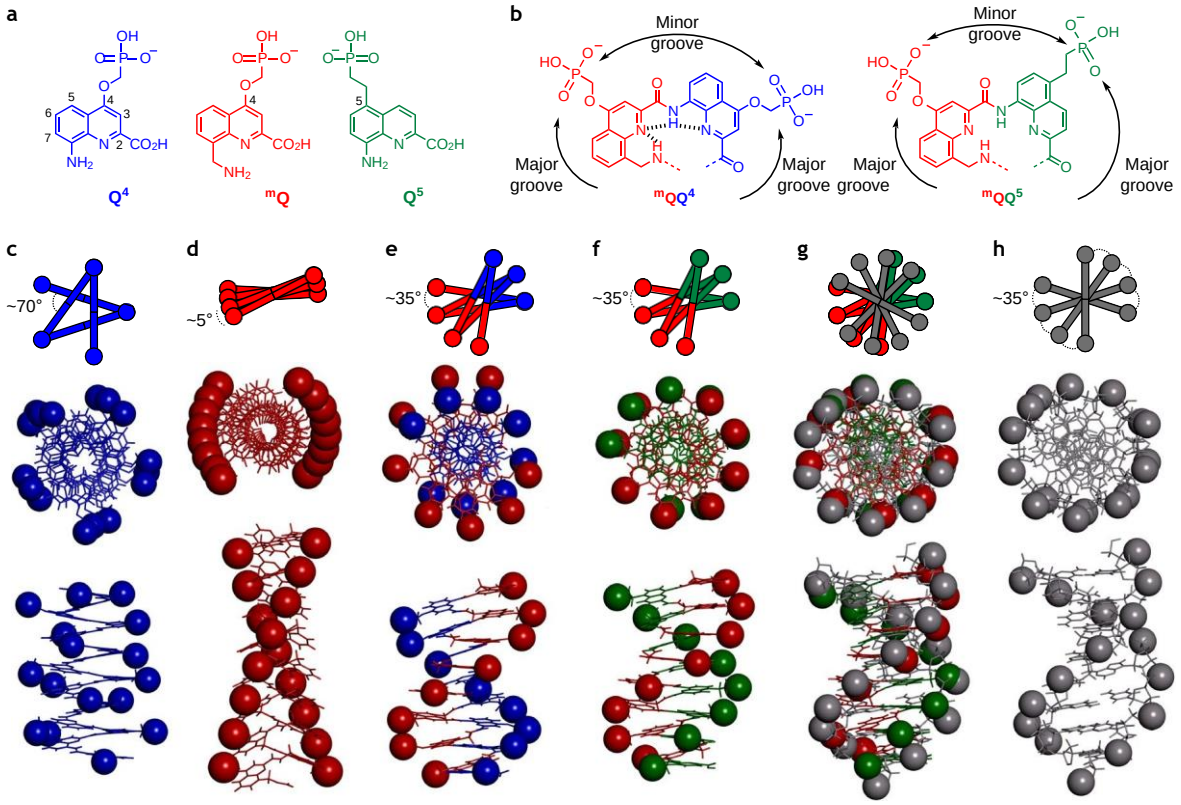

Fig. 1 | DNA mimic design. **a**, Formulae of amino-acid monomers Q^4, ^mQ and Q^5. **b**, Formulae of $^mQQ^4$ and $^mQQ^5$ blocks. **c**–**g**, Schematic representations, top views and side views of molecular models of the structures of $(Q^4)_{16}$ (**c**); $(^mQ)_{16}$ (**d**); $(^mQQ^4)_8$ (**e**); $(^mQQ^5)_8$ (**f**); and an eight-base-pair B-DNA (**h**). The structure in **g** is an overlay of **f** and **h**. The root-mean-square deviation of the phosphorus atom positions of $(^mQQ^5)_8$ and an eight-base-pair B-DNA duplex is 1.83 Å. The schemes show how the phosphorus atoms (circles) distribute when viewed down the helix axis. Angular shifts between adjacent dimeric blocks (**c**–**f**) or base pairs (**h**) are rounded to the nearest multiple of 5. Models are shown at the same scale as the stick representations except that the phosphorus atoms are shown as spheres. The same blue, red and green colours are used in the models and schemes as in the formulae shown in **a**. DNA is shown in grey. The model of $(Q^4)_{16}$ in **c** is based on the crystal structures of related sequences[4,7]. Other models are energy minimized structures.

(~9.0 Å between the purine and pyrimidine N1 atoms). Moreover, the oligoamide helix has a pitch of 3.5 Å that matches the base-pair distances in DNA. We thus considered that functionalizing $\mathbf{Q}_n$ oligomers with negatively charged surface residues such as in $\mathbf{Q^4}$ (Fig. 1a) may give rise to a novel generation of DNA mimics. Nevertheless, helix curvature, as imparted by **Q** monomers, is exactly five units per two turns, or 0.8 turns per two units. It follows that negatively charged residues in position 4 of each quinoline monomer would be arranged very differently from the positions of phosphates in DNA: a top view of a model of $(\mathbf{Q^4})_{16}$ shows a five-pointed star (Fig. 1c).

New monomer $^m\mathbf{Q}$ (Fig. 1a) is a slightly elongated version of $\mathbf{Q^4}$ possessing an additional methylene group. In related pyridine-derived monomers, such a benzylic methylene group is compatible with single helix folding, including in water[26,27]. Molecular modelling predicts that two $^m\mathbf{Q}$ units would span almost exactly one helix turn; an $(^m\mathbf{Q})_n$ helix would display two arrays of negative charges nearly parallel to the helix axis (Fig. 1d). This again differs greatly from the positions of phosphates in DNA. However, the outcome is different in $(^m\mathbf{QQ^4})_n$ mixed sequences. An $^m\mathbf{QQ^4}$ block averages the curvature of $^m\mathbf{Q}$ and $\mathbf{Q^4}$ and thus spans 0.9 helix turns. Thus, in an $(^m\mathbf{QQ^4})_n$ helix, the angular shift between adjacent $^m\mathbf{QQ^4}$ blocks is predicted to be a tenth of a turn, which is equal to the angular shift between base pairs in double-stranded B-DNA (ten base pairs span a full double helix turn). Consequently, the negatively charged residues at the surface of an $(^m\mathbf{QQ^4})_n$ single helix are expected to form a double helical array that matches the positions of phosphates in duplex B-DNA (Fig. 1e).

Further considerations allow us to improve the design. When following an $(^m\mathbf{QQ^4})_n$ helix from its N terminus to its C terminus, main-chain methylene bridges are found between $\mathbf{Q^4}$ and $^m\mathbf{Q}$ monomers, but not between $^m\mathbf{Q}$ and $\mathbf{Q^4}$ monomers. Thus, the distances between the side chains of these pairs of monomers are slightly different, allowing us to define a major groove and a minor groove, as in B-DNA (Fig. 1b). Nevertheless, the grooves of the $(^m\mathbf{QQ^4})_n$ helix remain very close in width. To widen the major groove and narrow down the minor groove without changing the main-chain properties, the side chains of the **Q** monomers may be installed at position 5 instead of 4, as in monomer $\mathbf{Q^5}$ (Fig. 1a,b). Molecular modelling predicts an excellent match between the positions of phosphorus atoms in duplex B-DNA and at the surface of $(^m\mathbf{QQ^5})_n$ (Fig. 1f–h).

Thus, structural considerations establish $(^m\mathbf{QQ^5})_n$ and $(^m\mathbf{QQ^4})_n$ as potential mimics of B-DNA charge surface. As with other DNA mimics, the differences that remain may not be deleterious and could, on the contrary, confer advantageous properties. For example, the side chains of $(^m\mathbf{QQ})_n$ are anchored by a single attachment to the helix backbone that leaves some conformational freedom, as opposed to the two ester functions of each phosphate in B-DNA. Also, contiguous $^m\mathbf{QQ}$ blocks are connected by one covalent bond in a single helical arrangement, whereas two covalent bonds (one per nucleotide) connect contiguous base pairs in a DNA duplex. Thus, $(^m\mathbf{QQ})_n$ single helices may bend upon rotation about a single bond of the main chain, allowing us to speculate that their shape, an essential parameter in NA–protein recognition, may be altered more easily than that of duplex B-DNA.

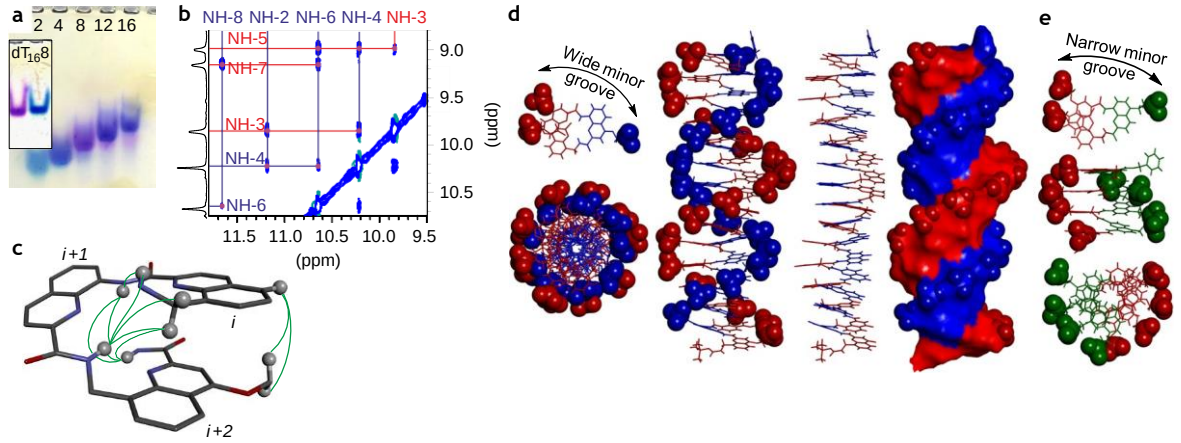

Fig. 2 | Structural characterization. **a**, PAGE migration of $(^{m}QQ^{4})_n$ revealed by 'stains all'. The values of *n* are indicated at the top of the lanes. The inset shows a comparison between the migration of $(^{m}QQ^{4})_8$ and of dT_{16}. **b**, Expansion of the 700 MHz ^1H–^1H NOESY spectrum recorded with 800 ms mixing time of $(^{m}QQ^{4})_4$ (2 mM) in 50 mM ammonium bicarbonate H_2O/D_2O (9:1 vol/vol) at 298 K, showing dipolar couplings between the amide protons that line the inner rim of the helix. Quinoline rings are numbered from 1 to 8 starting from the N terminus. **c**, Perspective view of the structure of a ^{m}Q-Q^4-^{m}Q segment showing NOE correlations (green curved lines) relevant to the assignment of a helical conformation. Hydrogen atoms involved in the correlations are shown as white balls. Other hydrogen atoms have been deleted for clarity. **d**, Crystal structure of fully protected $(^{m}QQ^{4})_{16}$. Stick representations are used except for the phosphonate functions, which are shown as space-filling models. The top left inset illustrates the groove width. Phosphonates have been removed from the middle-right structure, revealing a long-range undulation of the single-helix backbone as a result of ^{m}Q units protruding away from the helix axis, as can also be seen in the view shown at bottom left. The right structure shows the solvent-accessible surface and reveals the grooves. **e**, Crystal structure of $(^{m}QQ^{5})_4$. In the bottom right view, the side chains of the two types of monomer are at the same distance from the helix axis. In **d** and **e**, included solvent molecules and ethyl ester functions at the side chains have been removed for clarity. The same blue, red and green colours are used as in Fig. 1.

Synthesis and structure elucidation. To test these predictions, we first synthesized $(^{m}\mathbf{QQ^4})_n$ oligomers (Supplementary Figs. 1, 2, 4, 5 and 8). The choice to initially target $\mathbf{Q^4}$ instead of $\mathbf{Q^5}$ rested on the availability of procedures to introduce side chains in position 4 of a 4-hydroxy-8-nitro-2-quinolinecarboxylate precursor under Mitsunobu conditions[28]. Similarly, monomer $^{m}\mathbf{Q}$ was prepared via an 8-cyano-4-hydroxy-2-quinolinecarboxylate intermediate. Reduction of the cyano group yielded the desired 8-aminomethyl function. The side chain length was adjusted to make the overall foldamer helix diameter match that of B-DNA. Phosphonate side chains resembling the phosphates of NAs were introduced as a diethyl ester protected form. Main-chain benzylic amines were protected with an HCl-labile *tert*-butyl-carbamate (Boc). Main-chain acids were protected as benzyl (Bn) esters. A segment doubling approach was followed in which the diethyl-phosphonate forms of BocNH-$(^{m}\mathbf{QQ^4})_x$-CO_2H and H_2N-$(^{m}\mathbf{QQ^4})_y$-CO_2Bn were coupled to yield BocNH-$(^{m}\mathbf{QQ^4})_{(x+y)}$-$CO_2$Bn. A series of $(^{m}\mathbf{QQ^4})_n$ oligomers with n = 2, 4, 8, 12 and 16 were synthesized in this way, then deprotected using trimethylsilyl bromide (that is, up to 64 ester functions for n = 16), purified by reversed-phase liquid chromatography, and submitted to ion exchange to deliver the side chains as water-soluble ammonium phosphonate salts. $(^{m}\mathbf{QQ^4})_n$ oligomers do not possess stereogenic centres and are expected to fold as a racemic mixture of right-handed (*P*) and left-handed (*M*) helices.

Most of the structural and biological investigations were carried out using the $^{m}\mathbf{QQ^4}$ series, but, in a second phase, we also prepared $\mathbf{Q^5}$, $(^{m}\mathbf{QQ^5})_4$ and $(^{m}\mathbf{QQ^5})_8$ (Supplementary Figs. 3 and 6–8). The synthesis of $\mathbf{Q^5}$ required several aromatic substituent manipulations. A 5-bromo-quinoline was eventually produced for which introduction of the side chain by means of a new carbon–carbon bond was most simple. Consequently, the side chains of $\mathbf{Q^5}$ are not exactly the same as those of $\mathbf{Q^4}$ and $^{m}\mathbf{Q}$, as they possess no ether function.

The oligoamides were structurally characterized. PAGE analysis revealed a length-dependent migration of $(^{m}\mathbf{QQ^4})_n$ oligomers (Fig. 2a). Staining with the carbocyanine dye 'stains all' resulted in a blue/purple colour comparable to the staining of NAs. The hexadecamer $(^{m}\mathbf{QQ^4})_8$ migrated similarly to a single-stranded 2′-deoxy-thymidine hexadeca-nucleotide dT_{16}. Mass spectrometry analysis under conditions optimized for DNA also revealed the correct masses of the $(^{m}\mathbf{QQ^4})_n$ oligomers (Supplementary Figs. 32–35). Structural evidence of the helix conformation was obtained in the solid state through crystal structures of the fully protected precursors of $(^{m}\mathbf{QQ^4})_8$, $(^{m}\mathbf{QQ^4})_{16}$ and $(^{m}\mathbf{QQ^5})_4$ (Fig. 2d,e, Supplementary Figs. 13–15 and Supplementary Tables 1–4). The structure of $(^{m}\mathbf{QQ^4})_{16}$ matches the predicted model. The main-chain single helix has a pitch of 3.5 Å and spans about 13 turns. At its surface, the expected double-helical array of side chains only undergoes about one and a half turns in the opposite direction; that is, the main-chain *M* helix has its side chains organized as the *P* B-DNA duplex. An overlay between the structure of $(^{m}\mathbf{QQ^4})_{16}$ and a 16-base-pair B-DNA shows a good match in curvature over 8 base pairs (Supplementary Figs. 16 and 17 and Supplementary Table 5), a length typically involved in protein–DNA contacts. Two crystal structures of $(^{m}\mathbf{QQ^4})_8$ further allowed us to sample structural variations in the solid state (Supplementary Figs. 18–20 and Supplementary Table 6). The side chain positions were found to vary, but, unlike in DNA, they do so independently from each other. The structure of $(^{m}\mathbf{QQ^5})_4$ confirms excellent matches between observed and calculated[29] structures. Its expanded major groove and reduced minor groove match those of B-DNA better than the $^{m}\mathbf{QQ^4}$ series. In $\mathbf{Q^5}$, conformational preferences at the ethylene linkage between phosphonates and the aromatic backbone differ from those of the oxomethylene groups in $^{m}\mathbf{Q}$ and $\mathbf{Q^4}$. In the latter, the phosphorus atom is found in or near the plane of the quinoline ring to which it belongs. In contrast, three of four $\mathbf{Q^5}$ monomers in the structure of $(^{m}\mathbf{QQ^5})_4$ have their phosphorus atoms out of the plane of the corresponding quinoline rings (Supplementary Fig. 19). The deprotected phosphonato foldamers also crystallized from water, but the diffraction intensity was too low to allow for structure elucidation. This difficulty is common for water-soluble aromatic foldamers possessing flexible side chains and might be overcome through the use of shorter and more rigid side chains[30].

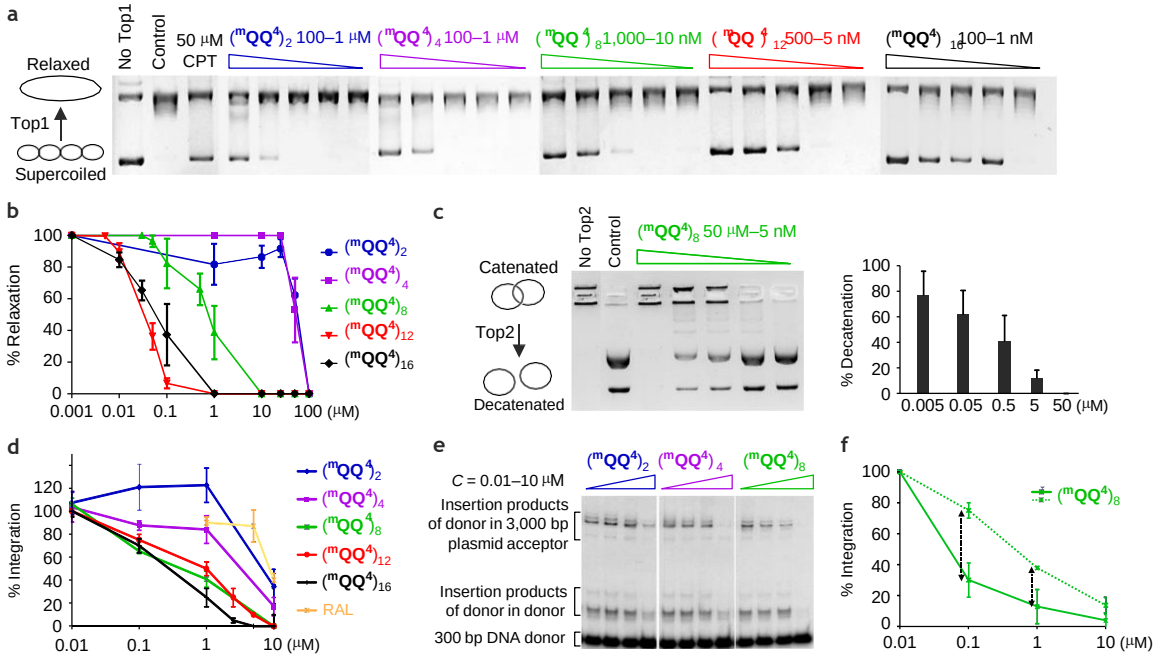

Fig. 3 | Enzyme inhibition. **a,** Inhibition of Top1-mediated relaxation of supercoiled circular DNA by $(^{m}QQ^{4})_n$ oligomers assessed by gel electrophoresis. The control lane shows full Top1 activity. Camptothecin (CPT) is a classical Top1 inhibitor. Note that the $(^{m}QQ^{4})_n$ concentration is in the nM range for $n > 4$ and the μM range for $n \leq 4$. **b,** Quantitation of Top1 inhibition expressed as a percentage of relaxation, compared to the control. **c,** Inhibition of Top2-mediated decatenation of catenated DNA by $(^{m}QQ^{4})_8$. The control lane shows full Top2 activity. Quantitation is expressed as a percentage of decatenation as compared to the control. **d,e,** Inhibition of in vitro HIV-IN integration (200 nM of IN) of a radio-labelled 246-base-pair (bp) viral DNA donor into a 2,700 bp pZeo plasmid acceptor assessed by agarose gel electrophoresis (Supplementary Fig. 37). Raltegravir (RAL) is a strong HIV-IN inhibitor[37]. **f,** Inhibition of wild-type HIV-IN (solid line) and K236A HIV-IN (dashed line) by $(^{m}QQ^{4})_8$. Double-headed arrows indicate selectivity. Quantitation in **d** and **f** is expressed as a percentage of integration related to a control in the absence of inhibitors. In **a,c,e** and **f,** quantitation was obtained by analysis of the gels using ImageJ software. Results are the means of at least triplicate experiments ± standard deviation.

NMR provided unequivocal evidence for helical folding in aqueous solution for deprotected oligomers and in chloroform for fully protected precursors. Ring current effects associated with aromatic stacking within the helix result in upfield shifts of aromatic and amide proton resonances (Supplementary Figs. 21 and 22). Main-chain benzylic CH_2 protons were found to be anisochronous ($\Delta\delta$ up to 1.2 ppm), consistent with a chiral helical structure undergoing slow $P \leftrightarrow M$ conformational exchange on the NMR timescale (Supplementary Fig. 23). These CH_2 protons signals are upfield shifted from expected resonances ($\Delta\delta$ up to 2.5 ppm) except those of the ^{m}**Q** unit closest to the C terminus, in agreement with the crystal structures that show that the benzylic CH_2 protons closest to the C terminus are the only ones not exposed to aromatic ring current effects. A complete assignment of ^{1}H and ^{13}C resonances of $(^{m}QQ^{4})_4$ in H_2O/D_2O was carried out (Supplementary Figs. 24–27 and Supplementary Table 7). Nuclear Overhauser effect spectroscopy (NOESY) allowed us to determine dipolar interactions (that is, close proximity, <5 Å) between protons that are fully consistent with the X-ray structures (Fig. 2b,c and Supplementary Figs. 28–31). The clear demonstration of helical folding in water corroborates earlier studies[27,31] showing the robustness of these structures despite the proximity of multiple identical charges. Hydrophobic effects associated with aromatic stacking overcome electrostatic repulsions between side chains, unlike what happens in peptidic α-helices[32].

DNA binding enzyme inhibition. We assessed the effect of the DNA mimics on the activity of several DNA-binding enzymes. Sequence-selective enzymes such as type II restriction enzymes XhoI and NdeI were not inhibited. Some non (or poorly) sequence-selective enzymes were not inhibited either (Supplementary Fig. 40): deoxyribonuclase 1, which cleaves both single- and double-stranded DNA; S1 nuclease, which cleaves single-stranded DNA; and benzonase, which degrades all forms of DNA and RNA (Supplementary Figs. 44–46). This set of experiments also demonstrated that $(^{m}QQ^{4})_8$ resists degradation by these enzymes. Furthermore, $(^{m}QQ^{4})_8$ is not degraded by, and does not alter the function of, a variety of proteases (trypsin, pepsin, protease XIII). Moderate inhibition of some DNA-binding enzymes was detected. Topoisomerase 2 (Top2), an enzyme that simultaneously cleaves both strands of double-stranded DNA, was inhibited by $(^{m}QQ^{4})_8$ with an inhibitory concentration (IC_{50}) near 1 μM (Fig. 3c). Flap endonuclease 1, a key DNA repair enzyme that removes overhanging single strands from double-stranded DNA, was also inhibited by $(^{m}QQ^{4})_8$ in the μM range (Supplementary Fig. 41), with lower efficiency than known inhibitors[33].

Remarkable inhibition of two therapeutically important enzymes was discovered. Topoisomerase 1 (Top1)[34], which cleaves one strand of DNA duplexes to allow relaxation of supercoiling, was found to be strongly inhibited (Fig. 3a,b) by $(^{m}QQ^{4})_n$. IC_{50} values decreased upon increasing n, reaching the low nanomolar range for $n \geq 12$. By comparison, classical Top1 inhibitor camptothecin (CPT) completely inhibited Top1 activity at a concentration of 50 μM in our settings. Human immunodeficiency virus 1 integrase (HIV-IN)[35] was also found to be inhibited (Fig. 3d,e). This enzyme is a site-specific nuclease that shares some structural and mechanistic features with eukaryotic Top1 (ref. [36]). It catalyses the insertion into the host genome of viral DNA produced by reverse transcription of the RNA of incoming viruses. HIV-IN possesses DNA-binding and cleavage activities sequence-specific to viral DNA-ends and sequence-non-specific to target DNA. Inhibition by DNA mimics occurred at concentrations below 1 μM, matching the best IN-specific inhibitors such as raltegravir (RAL)[37]. We also note that

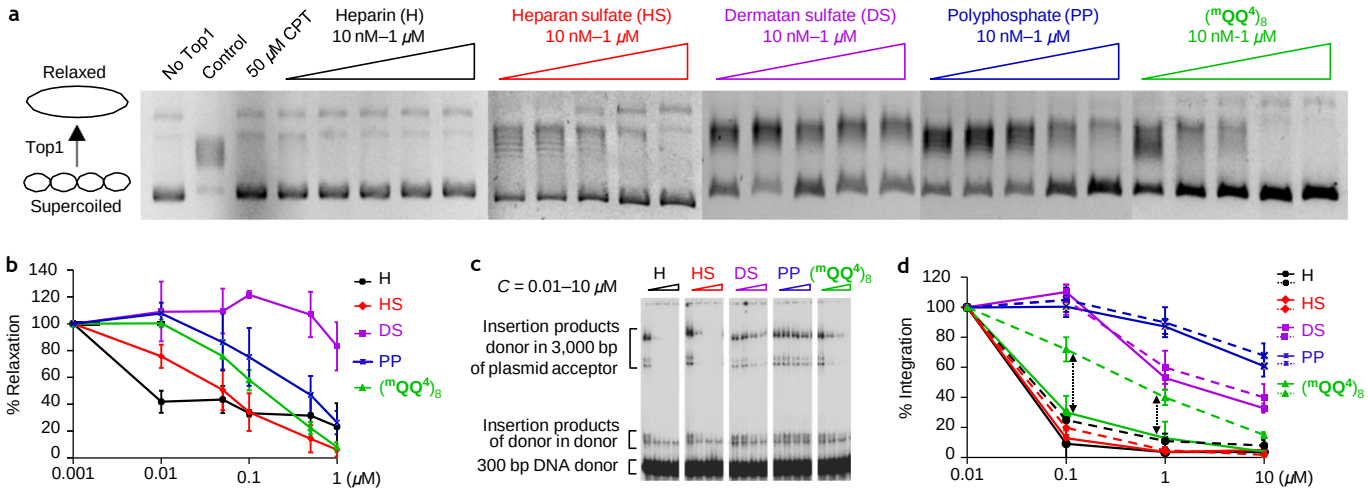

Fig. 4 | Inhibition of Top1 and HIV-IN by polyanions. **a**, Inhibition of Top1-mediated relaxation of supercoiled circular DNA by various polyanions and $(^{m}QQ^{4})_8$ assessed by gel electrophoresis. The control lane shows full Top1 activity. Camptothecin (CPT) is a classical Top1 inhibitor. **b**, Quantitation of Top1 inhibition expressed as a percentage of relaxation as compared to the control. **c**, Inhibition of in vitro HIV-IN integration (200 nM of IN) of a radio-labelled 246 bp viral DNA donor into a 2,700 bp pZeo plasmid acceptor assessed by agarose gel electrophoresis using the same polyanions. **d**, Quantitation of the inhibition of wild-type HIV-IN (solid lines) and K236A HIV-IN (dashed lines) by polyanions and $(^{m}QQ^{4})_8$, expressed as a percentage of integration relative to a control in the absence of inhibitors. Double-headed arrows indicate the selectivity of $(^{m}QQ^{4})_8$. In **a** and **c**, quantitation was obtained by analysis of the gels using ImageJ software. Results are the means of triplicate experiments ± standard deviation.

low concentrations of the short $(^{m}QQ^{4})_2$ slightly enhance HIV-IN activity (Fig. 3d,e). This effect is similar to the effect of short oligodeoxynucleotides, which are believed to stabilize the conformation of HIV-IN before it binds to its substrate[38]. The inhibitory activity of $(^{m}QQ^{5})_8$ was found to be similar (for Top1) or only slightly better (for HIV-IN) than that of $(^{m}QQ^{4})_8$ (Supplementary Figs. 38 and 42).

Comparison with the inhibitory effects of other polyanions. These results brought up the question of the structural requirements of the DNA mimics for inhibition. Inhibition actually takes place in the presence of an excess of DNA substrate. In Top1 inhibition assays, the amount of DNA base pairs (pcDNA3.1 plasmid) is two to three orders of magnitude larger than the amount of base pair equivalents of the largest foldamers at inhibitory concentrations. If competitive binding of the foldamer to the DNA-binding site of the protein is to be hypothesized—a mechanism that would be completely different from that of CPT for Top1 (ref. [39]) and RAL for HIV-IN[35,37]—then the mimics must bind to the proteins considerably better than DNA itself. A first series of controls consisted in assessing whether DNA could inhibit these enzymes: a random sequence of an eight-base-pair double-stranded DNA (to be compared to $(^{m}QQ^{4})_8$) had no significant effect on the activity of Top1 and Top2 (Supplementary Fig. 43), and little effect on HIV-IN activity (Supplementary Fig. 38). Thus, structural components of the foldamers that make them distinct from DNA eventually give rise to enhanced inhibition. The nature, number and positions of the anions are the only features that were purposely designed in the foldamer-based DNA mimics and are certainly critical. One may note that the phosphonate side chains have the potential for being dianionic, whereas DNA phosphodiesters are monoanionic. The second p*K*a of a phosphonic acid is typically around 6.5, which would result in a fraction of dianions at pH 7.0 (Supplementary Figs. 11 and 36). However, we measured the second p*K*a of the phosphonic acids of $(^{m}QQ^{4})_8$ to be shifted to a range of higher values centred above 8.5 (Supplementary Fig. 36). This shift was assigned to the negative charge density at the surface of the foldamer helix. Phosphonic acids in $(^{m}QQ^{4})_8$ are thus mostly monoanionic at neutral pH, but nevertheless retain the ability to become dianionic when in contact with the basic residues of a DNA-binding protein.

Some polyanions, including polyphosphate[22] and sulfated polysaccharides[40] such as heparin, heparan sulfate or dermatan sulfate, have been known to inhibit DNA-binding enzymes and enzymes not related to DNA. We compared the inhibitory activity of these polymers on Top1 and HIV-IN to that of $(^{m}QQ^{4})_8$ at a comparable charge concentration and found different patterns (Fig. 4). Polyphosphate inhibits Top1 but not HIV-IN. Dermatan sulfate weakly inhibits HIV-IN, but not Top1. Heparin and heparan sulfate strongly inhibit Top1 (ref. [40]) and HIV-IN. These qualitative comparisons are valid, but they do not allow for a direct ranking of inhibitory activities, because the polymers differ considerably in average length (usually much longer than the foldamers), polydispersity and chemical variability (Supplementary Table 8). These different substances nevertheless share an ability to fit with one or another protein surface. Thus, heparin is a very broad-range protein ligand, so much so that heparin-binding chromatography is a well-known technique in DNA-binding protein purification[41]. Among these substances, the foldamers stand out as the only molecules that are structurally and conformationally defined and amenable to targeted modifications. A strong manifestation of the importance of the foldamer structure in its binding and inhibitory activity was found while studying the K236A HIV-IN mutant. This single mutation reduces both DNA binding to HIV-IN (see the section on Foldamer to protein and to DNA–ligand binding assessment below) and HIV-IN activity. We found that it also significantly alters the inhibitory activity of $(^{m}QQ^{4})_8$ (Fig. 3f and Supplementary Fig. 48). In contrast, the mutation has no significant effect on the inhibitory activity of the other polyanions, be they strong or weak inhibitors, pointing at their lack of structural specificity (Fig. 4d).

Foldamer to protein and to DNA-ligand binding assessment. Several experiments were carried out to detect direct binding between the foldamers and proteins that could be expressed in sufficiently high quantity and purity. The requirement for purity excluded most of the commercial enzymes (tested above) that are guaranteed to have a certain activity level but are not sold in a pure form and often in very small amounts (that is, XhoI, NdeI, deoxyribonuclease 1, S1 nuclease, benzonase, Flap endonuclease

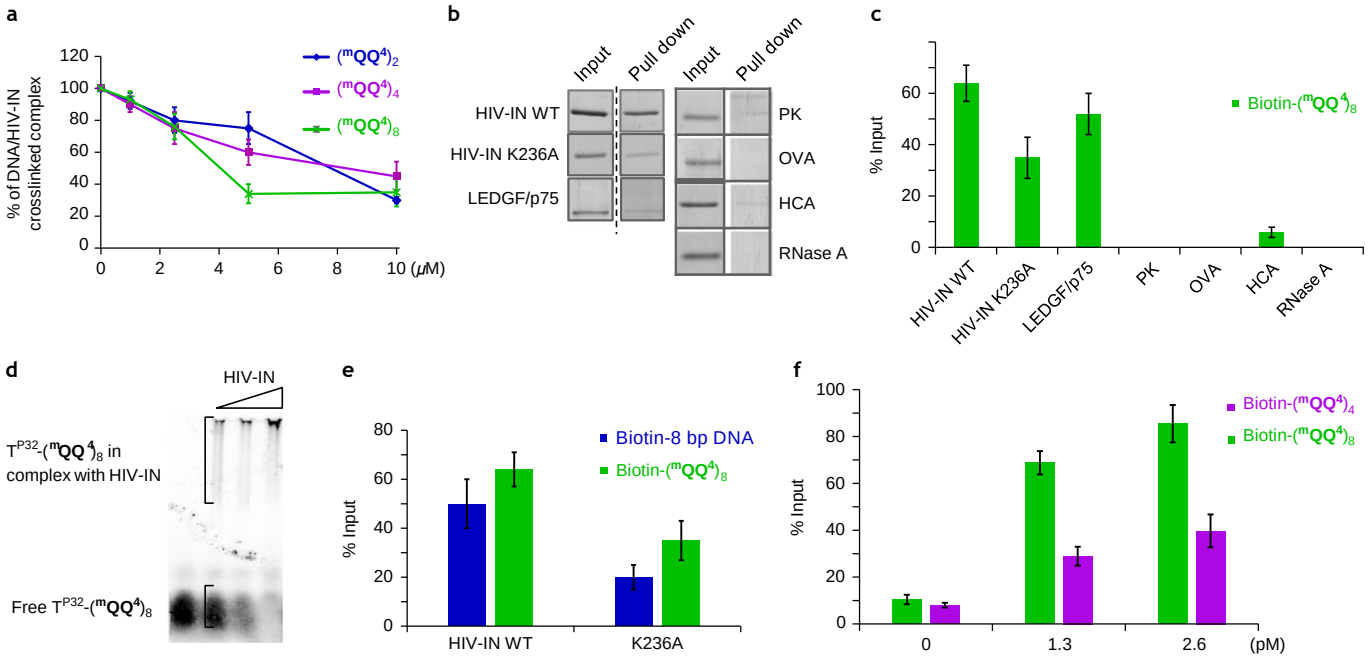

Fig. 5 | Foldamer–protein binding. **a,** Quantitative assessment of DNA/HIV-IN complex captured by photo-induced crosslinking in the presence of increasing amounts of $(^{m}QQ^{4})_{n}$ foldamers. **b,** Excerpts of non-denaturing polyacrylamide gels showing reference protein bands (input) and bands of protein pulled-down by biotin-$(^{m}QQ^{4})_{8}$ on streptavidin beads. **c,** Quantitation of the experiment shown in **b**. The weak extraction of HCA is also present when using streptavidin beads without foldamer as a negative control, indicating nonspecific interactions with streptavidin. **d,** Non-denaturing polyacrylamide gel shift assays revealing the formation of complexes between HIV-IN and T^{P32}-$(^{m}QQ^{4})_{8}$. **e,** Quantitation of pulldown experiments comparing biotin-$(^{m}QQ^{4})_{8}$ and a biotin–8 bp DNA. **f,** Quantitation of HIV-IN pulldown experiments comparing biotin-$(^{m}QQ^{4})_{8}$ and biotin-$(^{m}QQ^{4})_{4}$. Quantitation was obtained by analysis of the gels using ImageJ software. In all panels, results are the means of at least triplicate experiments ± standard deviation.

Top1 and Top2). A thymine-mononucleotide was attached at the N terminus of $(^{m}\mathbf{QQ^{4}})_{8}$ (Supplementary Fig. 9) and radiolabelled with ^{32}P phosphate to produce T^{32P}-$(^{m}\mathbf{QQ^{4}})_{8}$. Gel shift assays showed that the presence of HIV-IN shifted the band of the radiolabelled foldamer on non-denaturing gels, demonstrating the direct association between the two (Fig. 5d). Quantitative analysis in filter binding assays showed that $(^{m}\mathbf{QQ^{4}})_{8}$ binds to HIV-IN more strongly than both specific and non-specific duplex DNA (Supplementary Fig. 39). In a subsequent assay, the amount of DNA/HIV-IN complex was quantified after UV-induced crosslinking and shown to decrease in the presence of foldamers, consistent with the hypothesis of competitive binding of the latter (Fig. 5a and Supplementary Fig. 47). Next, biotinylated-$(^{m}\mathbf{QQ^{4}})_{8}$ and -$(^{m}\mathbf{QQ^{4}})_{4}$ foldamers were produced (Supplementary Fig. 10). It was checked that biotinylation has no effect on catalytic activity in the case of HIV-IN (Supplementary Fig. 49). The biotinylated foldamers were used for pulldown extraction experiments of recombinant proteins of high purity, including DNA-binding proteins (HIV-IN and LEDGF/p75), RNA-binding protein ribonuclease A (RNase A), a protein-binding protein (proteinase K, PK), a broad-range binder (ovalbumin) and a negative control (human carbonic anhydrase, HCA). These experiments (Fig. 5b,c and Supplementary Figs. 50–53) confirmed that $(^{m}\mathbf{QQ^{4}})_{8}$ binds to HIV-IN. The low concentrations used in the pulldown assay allow for an estimate of the $(^{m}\mathbf{QQ^{4}})_{8}$/HIV-IN complex K_{d} in the low nanomolar range. This value may not be directly correlated to inhibitory concentrations, which are determined in the presence of a large excess of DNA substrate that competes for binding to HIV-IN. $(^{m}\mathbf{QQ^{4}})_{8}$ also binds to LEDGF/p75, a growth factor known to bind to DNA and recruited by HIV-IN for viral DNA integration[42,43]. It does not bind to RNase A, to ovalbumin, to HCA, nor to PK. Further experiments confirmed that $(^{m}\mathbf{QQ^{4}})_{8}$ binds to HIV-IN better than to its K236A mutant, in agreement with its weaker inhibition of the latter in the context of a competitive mode of action, but in both cases, binding is superior to the binding of an 8 bp DNA duplex (Fig. 5e). As another hint of a correlation between binding and inhibition, $(^{m}\mathbf{QQ^{4}})_{8}$ binds to HIV-IN better than $(^{m}\mathbf{QQ^{4}})_{4}$ (Fig. 5f).

Mimicry of the charge surface of DNA thus endows the foldamers with binding and inhibition properties towards some DNA-binding enzymes among those that recognize DNA mainly through its array of phosphates. Yet NAs possess other surface features that are essential to their recognition by proteins, including their shape, ability to bend and the base-pair functionalities exposed in the grooves. Although none of these features were deliberately included in the design of the foldamers, we assessed whether the foldamers could be recognized by small-molecule ligands of double-stranded DNA. Thus, biotin-$(^{m}\mathbf{QQ^{4}})_{8}$ was compared to two biotinylated DNA hairpin sequences using surface plasmon resonance (SPR) (Supplementary Methods and Supplementary Fig. 54). Binding of the polycation spermine[44] to all three substrates was negligible under our conditions, that is, elution of 0.25, 0.5 and 1 μM solutions of the ligands. Intercalating agent ethidium bromide[44] showed weak binding to DNA and no binding to $(^{m}\mathbf{QQ^{4}})_{8}$. Actinomycin D[44,45] only bound its DNA sequence target containing a GC site (DNA hairpin 2). Finally, netropsin[46] and Hoechst 33258 (ref. [47]), two minor groove ligands, bound strongly to a privileged AATT-containing DNA target (DNA hairpin 1) and also bound to the other AT-rich DNA hairpin and to $(^{m}\mathbf{QQ^{4}})_{8}$. These two ligands possess multiple hydrogen bond donors that exploit the high electronegative potential of AT tracks. Although their binding mode to $(^{m}\mathbf{QQ^{4}})_{8}$ is still unknown, the presence of amide carbonyl group hydrogen-bond acceptors in the grooves of the foldamer could contribute. Although preliminary, this result hints at potential developments of the foldamer design through the introduction of specific groove functionalities.

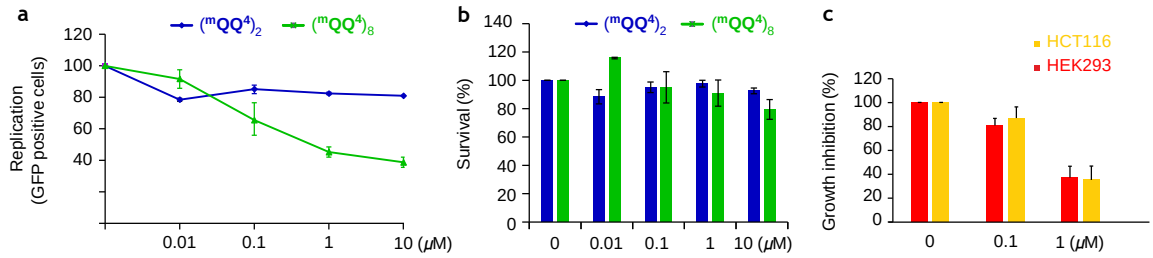

Fig. 6 | Cell-based assays. **a**, In-cell inhibition of HIV-1-derived pseudoviruses by $(^{m}QQ^{4})_8$ (n = 2, 8). **b**, Assessment of cytotoxicity of $(^{m}QQ^{4})_8$ (n = 2, 8) on human HEK293 embryonic kidney cells. **c**, Effect of $(^{m}QQ^{4})_8$ on human HEK293 embryonic kidney and HCT116 colon cancer cell growth in the presence of PULSin transfection reagent. Cells were exposed to the foldamer for 4 h, the medium was then replaced by fresh complete medium, and cell growth was measured 72 h later. Results are expressed as a percentage of growth as compared to cells treated with PULSin alone. In all panels, results are the means of at least triplicate experiments ± standard deviation.

Cell-based assays. Following the discovery of HIV-IN and Top1 inhibition by anionic foldamers, we carried out cell-based studies. We found no or weak cellular toxicity after a 4–8 h treatment of human embryonic kidney cells (HEK293) with $(^{m}\mathbf{QQ^{4}})_n$ (n = 2, 8) at concentrations up to 10 μM (Fig. 6b). Low cytotoxicity was assigned to poor cell penetration, as expected for a polyanion with poor lipophilicity. However, cytotoxicity was greatly enhanced in the presence of a cationic amphiphile—transfection reagent PULSin—in both HEK293 and human colon cancer cells (HCT116) (Fig. 6c). It remains to be determined which effects lead to this cellular toxicity, as inhibition of not only Top1 but also of other DNA-binding processive enzymes may be involved. We also discovered in cell inhibition of the early steps of HIV-1 replication (Fig. 6a) by $(^{m}\mathbf{QQ^{4}})_8$ in the absence of any vector to facilitate cell penetration, meaning that the foldamer does not penetrate in sufficient amounts to be toxic, but enough to inhibit HIV-1. Again, HIV-1 inhibition may have different origins, and the effects at play have yet to be determined. For instance, the direct or indirect (via a regulatory protein) inhibition of (1) reverse transcription catalysed by an RNA-dependent DNA-polymerase, (2) nuclear import of viral DNA bound to HIV-IN and (3) integration of the viral DNA are not discriminated in this assay. Even though further analysis of the cellular behaviour of the foldamers remains to be performed, these results bode well for future developments and optimization of these compounds.

Conclusion

Foldamer-based DNA mimics of an unprecedented kind have been developed. They inhibit some DNA-binding enzymes and display a better binding affinity than DNA itself. Inhibition is particularly strong for two therapeutically relevant enzymes whose mechanisms are related[36,] Top1 and HIV-IN. Although it was the foldamer resemblance to DNA that drove their design, it is actually their differences from DNA that confer them with their most remarkable properties. The foldamers also differ from anionic polymers known to inhibit NA-binding enzymes in terms of their well-defined length, chemical structure and conformation, and also in terms of their amenability to undergo designed modifications at specific sites, two features necessary for selective targeting. Evidence for specific foldamer–protein interactions came from the decrease in both the binding and inhibitory properties of $(^{m}\mathbf{QQ^{4}})_8$ towards HIV-IN due to mutation of a single amino acid known to alter DNA binding by the enzyme. The exact structural requirements for the DNA mimics to achieve strong and selective inhibition of DNA-binding enzymes remains to be investigated. This will entail structural studies of foldamer–protein complexes and the production of foldamer variants by solid-phase synthesis. Given the high modularity of helical aromatic amide foldamers, multiple variations can be envisaged, including the introduction of chemical functions to reproduce nucleobase and shape features as they appear in the grooves of DNA so as to also target sequence-selective DNA enzymes.

Reporting Summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Data availability. The crystallographic data and experimental details of the structural refinement for the X-ray crystal structure of fully protected $(^{m}\mathbf{QQ^{4}})_8$ (benzyl ester), $(^{m}\mathbf{QQ^{4}})_8$ (2-trimethylsilyl-ethyl ester), $(^{m}\mathbf{QQ^{4}})_{16}$ and $(^{m}\mathbf{QQ^{5}})_4$ have been deposited at the Cambridge Crystallographic Data Centre under deposition numbers CCDC 1059495, CCDC 1059497, CCDC 1059493 and CCDC 1059496, respectively. These data can be obtained free of charge from the Cambridge Crystallographic Data Centre (http://www.ccdc.cam.ac.uk/data_request/cif). Other data that support the findings of this study (spectroscopic or mass spectrometric data) are available from the corresponding author upon reasonable request.

Received: 7 August 2016; Accepted: 31 January 2018;
Published online: 2 April 2018

References

1. Nielsen, P. E., Egholm, M., Berg, R. H. & Buchardt, O. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* **254**, 1497–1500 (1991).
2. Koshkin, A. A. et al. LNA (locked nucleic acids): synthesis of the adenine, cytosine, guanine, 5-methylcytosine, thymine and uracil bicyclonucleoside monomers, oligomerisation, and unprecedented nucleic acid recognition. *Tetrahedron* **54**, 3607–3630 (1998).
3. Obika, S. et al. Synthesis of 2′-O,4′-C-methyleneuridine and -cytidine. Novel bicyclic nucleosides having a fixed C3′-endo sugar puckering. *Tetrahedron Lett.* **38**, 8735–8738 (1997).
4. Veedu, R. K. & Wengel, J. Locked nucleic acids: promising nucleic acid analogs for therapeutic applications. *Chem. Biodiv.* **7**, 536–542 (2010).
5. Nielsen, P. E. *Nucleic Acid Backbone Structure Variations: Peptide Nucleic Acids* (Wiley, Chichester, 2014).
6. Rohs, R. et al. Origins of specificity in protein–DNA recognition. *Annu. Rev. Biochem.* **79**, 233–269 (2010).
7. Luscombe, N. M., Austin, S. E., Berman, H. M. & Thornton, J. M. An overview of the structures of protein–DNA complexes. *Genome Biol.* **1**, reviews001.1 (2000).
8. Wang, H.-C., Ho, C.-H., Hsu, K.-C., Yang, J.-M. & Wang, A. H.-J. DNA mimic proteins: functions, structures, and bioinformatic analysis. *Biochemistry* **53**, 2865–2874 (2014).
9. Dryden, D. T. F. DNA mimicry by proteins and the control of enzymatic activity on DNA. *Trends Biotechnol.* **24**, 378–382 (2006).
10. Yüksel, D., Bianco, P. R. & Kumar, K. De novo design of protein mimics of B-DNA. *Mol. BioSyst.* **12**, 169–177 (2016).
11. Chenoweth, D. M., Poposki, J. A., Marques, M. A. & Dervan, P. B. Programmable oligomers targeting 5′-GGGG-3′ in the minor groove of DNA and NF-κB binding inhibition. *Bioorg. Med. Chem.* **15**, 759–770 (2007).
12. Bremer, R. E., Baird, E. E. & Dervan, P. B. Inhibition of major-groove-binding-proteins by pyrrole-imidazole polyamides with an Arg-Pro-Arg positive patch. *Chem. Biol.* **5**, 119–133 (1998).
13. Ducani, C., Leczkowska, A., Hodges, N. J. & Hannon, H. J. Noncovalent DNA-binding metallo-supramolecular cylinders prevent DNA transactions in vitro. *Angew. Chem. Int. Ed.* **49**, 8942–8945 (2010).

14. Brabec, V. et al. Metallohelices with activity against cisplatin-resistant cancer cells; does the mechanism involve DNA binding? *Chem. Sci.* **4**, 4407–4416 (2013).
15. Maher, L. J. III, Wold, B. & Dervan, P. B. Inhibition of DNA binding proteins by oligonucleotide-directed triple helix formation. *Science* **245**, 725–730 (1989).
16. Azzarito, V., Long, K., Murphy, N. S. & Wilson, A. J. Inhibition of α -helix-mediated protein–protein interactions using designed molecules. *Nat. Chem.* **5**, 161–173 (2013).
17. Jayatunga, M. K. P., Thompson, S. & Hamilton, A. D. α-Helix mimetics: outwards and upwards. *Bioorg. Med. Chem. Lett.* **24**, 717–724 (2014).
18. Johnson, L. M. & Gellmann, S. H. α-Helix mimicry with α/β-peptides. *Methods Enzymol.* **523**, 407–429 (2013).
19. Koert, U., Harding, M. M. & Lehn, J.-M. DNH deoxyribonucleohelicates: self assembly of oligonucleosidic double-helical metal complexes. *Nature* **346**, 339–342 (1990).
20. Conrad, H. E. *Heparin-Binding Proteins* (Academic, San Diego, 1998).
21. Monien, B. H. & Desai, U. R. Antithrombin activation by nonsulfated, non-polysaccharide organic polymer. *J. Med. Chem.* **48**, 1269–1273 (2005).
22. Rodriguez, R. J. Polyphosphate present in DNA preparations from filamentous fungal species of *Collectrichum* inhibits restriction endonucleases and other enzymes. *Anal. Biochem.* **209**, 291–297 (1993).
23. Jiang, H., Léger, J.-M. & Huc, I. Aromatic delta-peptides. *J. Am. Chem. Soc.* **125**, 3448–3449 (2003).
24. Dolain, C. et al. Solution structure of quinoline- and pyridine-derived oligoamide foldamers. *Chem. Eur. J.* **11**, 6135–6144 (2005).
25. Qi, T. et al. Solvent dependence of helix stability in aromatic oligoamide foldamers. *Chem. Commun.* **48**, 6337–6339 (2012).
26. Sánchez-García, D. et al. Nanosized hybrid oligoamide foldamers: aromatic templates for the folding of multiple aliphatic units. *J. Am. Chem. Soc.* **131**, 8642–8648 (2009).
27. Baptiste, B., Douat-Casassus, C., Laxmi-Reddy, K., Godde, F. & Huc, I. Solid phase synthesis of aromatic oligoamides: application to helical water-soluble foldamers. *J. Org. Chem.* **75**, 7175–7185 (2010).
28. Qi, T., Deschrijver, T. & Huc, I. Large-scale and chromatography-free synthesis of an octameric quinoline-based aromatic amide helical foldamer. *Nat. Protoc.* **8**, 693–708 (2013).
29. Liu, Z., Abramyan, A. M. & Pophristic, V. Helical arylamide foldamer: structure prediction by molecular dynamics simulations. *New J. Chem.* **39**, 3229–3240 (2015).
30. Hu, X. et al. Optimizing side chains for crystal growth from water: a case study of aromatic amide foldamers. *Chem. Sci.* **8**, 3741–3749 (2017).
31. Hu, X., Dawson, S. J., Nagaoka, A. & Huc, I. Solid-phase synthesis of water-soluble helically folded hybrid α-amino acid/quinoline oligoamides. *J. Org. Chem.* **81**, 1137–1150 (2016).
32. Lu, H. et al. Ionic polypeptides with unusual helical stability. *Nat. Commun.* **2**, 206 (2011).
33. Tumey, L. N. et al. The identification and optimization of a N-hydroxy urea series of flap endonuclease 1 inhibitors. *Bioorg. Med. Chem. Lett.* **15**, 277–281 (2005).
34. Redinbo, M. R., Stewart, L., Kuhn, P., Champoux, J. J. & Hol, W. G. J. Crystal structure of human topoisomerase I in covalent and noncolvalent complexes with DNA. *Science* **279**, 1504–1513 (1998).
35. Hare, S., Gupta, S. S., Valkov, E., Engelman, A. & Cherepanov, P. Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature* **464**, 232–236 (2010).
36. Cheng, C., Kussie, P., Pavletich, N. & Shuman, S. Conservation of structure and mechanism between eukaryotic topoisomerase I and site-specific recombinases. *Cell* **92**, 841–850 (1998).
37. Hazuda, D. J. et al. Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells. *Science* **287**, 646–650 (2000).
38. Lesbats, P. et al. In vitro initial attachment of HIV-1 integrase to viral ends: control of the DNA specific interaction by the oligomerization state. *Nucleic Acids Res.* **36**, 7043–7058 (2008).
39. Hsiang, Y.H., Lihou, M. G. & Liu, L. F. Arrest of replication forks by drug-stabilized topoisomerase I–DNA cleavable complexes as a mechanism of cell killing by camptothecin. *Cancer Res.* **49**, 5077–5082 (1989).
40. Ishii, K. et al. Mechanism of inhibition of mammalian DNA topoisomerase I by heparin. *Biochem. J.* **241**, 111–119 (1987).
41. Xiong., S., Zhang, L. & He, Q. Y. Fractionation of proteins by heparin chromatography. *Methods Biol. Mol.* **424**, 213–221 (2008).
42. Maertens, G. et al. EDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. *J. Biol. Chem.* **278**, 33528–33539 (2003).
43. Demeulemeester, J., De Rijck, J., Gijsbers, R. & Debyser, Z. Retroviral integration: site matters: mechanisms and consequences of retroviral integration site selection. *Bioessays* **37**, 1202–1214 (2015).
44. Ciolkowski, M. L., Fang, M. M. & Lund, M. E. A surface plasmon resonance method for detecting multiple modes of DNA–ligand interactions. *J. Pharm. Biomed. Anal.* **22**, 1037–1045 (2000).
45. Lo, Y.S., Tseng, W. H., Chuang, C. Y. & Hou, M. H. The structural basis of actinomycin D-binding induces nucleotide flipping out, a sharp bend and a left-handed twist in CGG triplet repeats. *Nucleic Acids Res.* **41**, 4284–4294 (2013).
46. Nguyen, B., Tanious, F. A. & Wilson, W. D. Biosensor-surface plasmon resonance: Quantitative analysis of small molecule–nucleic acid interactions. *Methods* **42**, 150–161 (2007).
47. Bailly, C. et al. Sequence-specific minor groove binding by bis-benzimidazoles: water molecules in ligand recognition. *Nucleic Acids Res.* **31**, 1514–1524 (2003).

Acknowledgements

This work was supported by the Agence Nationale de la Recherche (project no. ANR-11-BS07-013-01 and project RETROSelect, jcjc2011 program), by the French National Research Agency against AIDS (ANRS, AO2016), by SIDACTION (AO2016, VIH20160721002), by the European Union under the Seventh Framework Programme (grant agreements nos. ERC-2012-AdG-320892 and PEOPLE-2011-IEF-300948) and by the Ligue contre le Cancer (Comité Languedoc Roussillon). The authors thank Mr B. Kauffmann for assistance with crystallographic measurements and resolution, Mr J.-L. Ferrer for beam time and help during data collection on FIP BM30A at the ESRF, and Mr C. Di Primo and Ms L. Minder for assistance with SPR measurements. This work benefited from the facilities and expertise of the Biophysical and Structural Chemistry platform at IECB, CNRS UMS3033, INSERM US001, Bordeaux University, France.

Author contributions

K.Z. and C.C. contributed equally to this work. K.Z., P.Pr., V.C. and P.P.B. synthesized all new compounds. K.L.-R. synthesized earlier lipophilic versions of the DNA mimics that were critical to the design. K.Z. carried out NMR structural studies. C.C., V.P., S.C. and P.Po. carried out biological assays. M.M. resolved the crystal structures. I.H., V.P., F.G., J.-M.S., S.C. and P.Po. designed the study. I.H. wrote the manuscript. All authors discussed the results and commented on the manuscript.