Dissertation zur Erlangung des Doktorgrades der Fakultät für Chemie und Pharmazie der Ludwig-Maximilians-Universität München

DNA Mimic Foldamers to Interfere in DNA-Protein Interactions

von Manuel Julian Loos

aus Starnberg

2025

Erklärung

Diese Dissertation wurde nach den Richtlinien der Promotionsordnung der Fakultät für Chemie und Pharmazie nach §7 betreut und nach §8 geschrieben.

Eidesstattliche Versicherung

Hiermit versichere ich, dass diese Dissertation eigenständig und ohne unerlaubte Hilfe angefertigt wurde.

München, den 17.12.2025

_____

Manuel Loos

Dissertation eingereicht am:	17.12.2025
1. Gutachter:	Prof. Dr. Ivan Huc
2. Gutachter:	Prof. Dr. Thomas Carell
Mündliche Prüfung am:	13.02.2026

Acknowledgements

Before diving into the scientific depths, I would like to acknowledge the many people who deserve appreciation. Without many of them, this work would not have come to fruition.

Prof. Dr. Ivan Huc – I want to thank you for giving me the opportunity to pursue my doctoral work in your research group, and for your excellent supervision and continuous support throughout all stages of the development of my thesis.

Prof. Dr. Thomas Carell – Thank you for taking on the role of the second reviewer and your scientific input about my PhD project and your guidance for my future endeavors.

The examination committee – I want to thank Prof. Dr. Andrea Rentmeister, Prof. Dr. Daniel Merk, Dr. Sabine Schneider and Dr. Pavel Kielkowski for completing my committee and taking the time to evaluate my work.

Huc research group – My gratitude also goes to the entire research group and all my colleagues, who created a pleasant working atmosphere for me over the past years. The daily exchanges have made an important contribution to this work. I would like to thank Dr. Valentina Corvaglia, who supported me particularly well at the start and during the transition towards my doctoral studies. Further I would like to thank Dr. Céline Douat for her guidance in the world of chromatography, and for always helping with maintaining and troubleshooting the instruments. I have thoroughly enjoyed the Praktikum supervision sessions with you. I want to thank Dr. Lars Allmendinger, I enjoyed the time teaching in the practical courses organized by you and want to thank you for all the NMR measurements and guidance you provided. Further, I want to thank Ivan Alonso for his constant support at any HPLC line.

My gratitude goes to all the people that I have collaborated with, both internally in this research group as well as all the people externally. My projects heavily relied on your achievements, and I wish you the best of luck in your future scientific endeavors.

I would like to thank my interns Jonathan Viehöver, Johanna Hofer, Leon Stanislawski, Julian Heeg, and Marla Dorfner for their enthusiasm and their great commitment during their bachelor's theses or research internships.

Further, I want to thank the people that started studying with me together at this university more than 10 years ago and all other ones that I met during these times. I got to know many new people over this time and there are many that I now can call friends.

Lunch at 1 & boulder gang – Lisa, Niko, Giorgio, Ilaria, Paul – For all the fun discussion at lunch, all the routes we have closed at the boulder gym. Most of you will continue your journey here for a while longer. I wish you the same energy and spirit that you have always brought into our shared moments. Never hesitate to reach out to me, if you think I might be able to help you in one way or another.

Mensa Bois – Daniel, Niklas, Sebastian – Auf die Zeiten, in der die Mahlzeit noch 3,50 gekostet hat und die Nachmittage, die man trotz Haxe im Magen durchgearbeitet hat. Ich wünsche euch das Beste, sowohl wissenschaftlich als auch privat. Sebastian, ich hoffe wir werden zusammen noch viele Berge erklimmen und Kletterrouten durchsteigen.

The Boys –Nacho, Florian, Mathieu, Johannes – You were the cornerstone that kept me going. Nacho, our iteration of "the final compound" makes me smile every time when I see you finish your synthetic route. Mathieu, we got scrambled together in the lab and were separated again. I have learned so much from you. I have already started to miss you being around in the lab and our wooden drying rack. Florian, we started together and finished together. I am sure that we will both find our ways, hopefully our paths might cross again in the future. Johannes, we shared the past 10 years of our studies and our PhD journey. Without you, I would probably not have completed this path. Thank you for keeping me grounded and for being there every time I needed you.

Vielen Dank an meine Freunde, ob aus alten oder neuen Zeiten. Die Schwabinger Gang, Moritz, und alle anderen.

Consti, Lukas, Bene. Auf die Mittwoche mit hervorragender Kulinarik zwischen WT, 206 und 213. Ihr seid von Kommilitonen in den ersten Semestern zu meinen engsten Freunden geworden. Ich kann mich nicht glücklicher schätzen euch kennengelernt zu haben.

Während eines großen Teils meiner Promotionszeit wurde ich von Menschen begleitet, die mir sehr nahestanden. Auch wenn sich manche Wege am Ende getrennt haben, möchte ich sagen, dass Zeit und diese Menschen immer einen besonderen Platz in meinem Herzen haben werden.

Liebe Mama, Lieber Papa. Ohne euch wäre ich nicht der Mensch der ich heute bin. Ich will euch dafür danken, dass ihr mich immer unterstützt habt, egal in welche Situation ich mich begeben habe.

Table of Contents

1 Scientific contributions

1.1 First-author publications

Interfacing DNA and DNA mimic foldamers (accepted)
Authors: **Manuel Loos**, Felix Xu, Pradeep K. Mandal, Tulika Chakrabortty, Céline Douat, David B. Konrad, Melis Cabbar, Johannes Singer, Valentina Corvaglia, Thomas Carell, and Ivan Huc, *Angew. Chem., Int. Ed.*, 2025, 64, e202505273, DOI: 10.1002/anie.202505273.

Structural dynamics of DNA mimic foldamers (accepted)
Authors: **Manuel Loos**, Lion Thurecht, Jiaojiao Wu, Valentina Corvaglia, Zhiwei Liu, Vojislava Pophristic, and Martin Zacharias, Ivan Huc, *Chemical Science*.
DOI: 10.1039/D5SC08567E.

1.2 Co-author publications

Enhancing the Features of DNA Mimic Foldamers for Structural Investigations (accepted)
Authors: Valentina Corvaglia, Jiaojiao Wu, Deepak Deepak, **Manuel Loos** and Ivan Huc, *Chem. - Eur. J.*, 2024, **30**, e202303650. DOI: 10.1002/chem.202303650.

1.3 To be submitted

Long DNA mimic foldamers enable binding of multi-protein complexes (to be submitted)
Authors: **Manuel Loos**, Franziska Därr, Céline Douat, Lisa Gourdon-Grünewaldt, Alberto López-Francos López-Romero, Karl-Peter Hopfner and Ivan Huc.

1.4 Conference & symposia contributions

Oral contribution: International symposium on foldamers London. *Interfacing B-DNA and DNA mimic foldamers*. **Manuel Loos,** Felix Xu, Pradeep Mandal, Tulika Chakrabortty, Céline Douat, David Konrad, Melis Cabbar, Johannes Singer, Valentina Corvaglia, Thomas Carell, and Ivan Huc.

Poster: SFB1309 Summer symposium 2025. *Long DNA-Mimic Foldamers Enable Structural Insights into INO80 Interactions*. **Manuel Loos**, Franziska Därr, Felix Metzner, Lisa Gourdon-Grünewaldt, Karl-Peter Hopfner, Ivan Huc.

Poster: Frontiers in Medicinal Chemistry 2024. *DNA Mimic Foldamers in Biological Systems*. **M. Loos**, J. Wu, V. Corvaglia, D. Deepak, F. Xu, I. Huc.

Poster: Munich Symposium on Foldamers 2023. *"DNA mimic foldamers to target protein-DNA interactions"*. **M. Loos**, D. Deepak, Tulika Chakrabortty, Felix Xu, Melis Cabbar, Valentina Corvaglia, Pradeep Mandal, Thomas Carell, Ivan Huc.

Poster: Ettal 2023 Retreat SFB1309. Postbiotic chemistry: *Combining DNA mimicry and DNA*.
Manuel Loos, Felix Xu, Pradeep Mandal, Valentina Corvaglia, Melis Cabbar, Thomas Carell, Ivan Huc.

Poster: SFB1309 Nikolaus symposium 2022. *DNA mimic foldamers to target protein-DNA interactions*. **Manuel Loos**, Deepak Deepak, Felix Xu, Valentina Corvaglia, Melis Cabbar, Ivan Huc.

Poster: Bordeaux Symposium on Foldamers 2021. *"Structural investigation of protein-foldamer complexes: a tethering approach"*. Deepak Deepak, **Manuel Loos**, Valentina Corvaglia, Jiaojiao Wu, Ivan Huc.

2 Abstract

Foldamers are abiotic oligomers that fold into defined three-dimensional structures, similarly to the biomolecules that inspired their design. Among them, aromatic oligoamide foldamers bearing negatively charged side chains can reproduce the double-helical array of negative charges characteristic of B-form DNA. They have been shown to inhibit several enzymes, including topoisomerase I and HIV-1 integrase in the presence of their canonical DNA ligand and outcompete DNA in binding affinity against other DNA binding proteins. These properties make them a versatile platform to interfere with protein-DNA interactions.

This thesis focuses on the design, synthesis, structural and functional characterization of these foldamers. Central to this work is the development of an expanded library of monomeric building blocks for a peptide-like solid-phase synthesis of oligomers. This includes novel dimeric macromonomers, which have enabled the efficient production of foldamers of sufficient length to engage in complex biomolecular interactions.

By using a combination of spectroscopic methods and computational techniques, the structural dynamics of these foldamers were investigated. The simulations revealed that the global flexibility parameters for twisting and bending of the foldamer helices are of similar magnitude to those of B-DNA, though distinct kinking events and motions are involved. We demonstrate that they are stable over a wide range of pH values, temperatures and salt concentrations. An assay was developed to quantitatively assess foldamer helix stability through measurement of the rate of interconversion between right-handed and left-handed diastereomeric conformers. Unexpectedly, suppressing some negatively charged side chains had a destabilizing effect on the helix, suggesting a more complex role of the side chains than electrostatic repulsions.

Through collaboration, we characterized the binding of a DNA mimic foldamer to a multiprotein complex that senses extranucleosomal DNA. Using cryogenic electron microscopy, we were able to visualize the foldamer both as aggregated clusters in vitreous ice and in complex with the A-module of the INO80 chromatin remodeling complex.

To target sequence selective DNA-binding proteins, we designed and synthesized a linker unit that interfaces DNA and DNA mimic foldamers. It serves both as a hairpin turn and as an anchor for ligating a foldamer, keeping their rims and grooves in register. These chimeric molecules may mutually benefit from the high affinities of DNA mimic foldamers and the sequence information provided by the DNA component. As such, they pave the way toward competitive inhibitors of protein-DNA interactions involving sequence-selective DNA-binding proteins as a new generation of DNA decoys.

3 Introduction

Deoxyribonucleic acid (DNA), often called the molecule of life, is the fundamental hereditary material in all cellular organisms, both in modern organisms and in the earliest traceable forms of cellular life.[1-3] One can derive from Francis Crick's central dogma of molecular biology – DNA is transcribed into ribonucleic acid (RNA) which is translated into primary protein structure – that it carries the instructions for growth, development, and functioning of life and its reproduction.[4] DNA's initial isolation from the cell nucleus, the discovery of the nucleobases as its components including their base pairing, the elucidation of its structure via X-ray crystallography and its hereditary role in biology all have marked major milestones in our understanding of life.[1, 2, 5-8] These discoveries among many others have founded the scientific field of molecular biology. Besides the elucidation of DNA's structure and functions, the progress in the ability to amplify, analyse, manipulate, and sequence DNA has led to many advancements that have since revolutionized the pharmaceutical and medicinal sciences. This has led to many advancements in diagnostic tools, therapies, personalized medicine and a deeper understanding of genetic diseases.[9] Some key scientific contributions through the decades since the determination of its structure and genetic functions are summarised in Figure 1.[2, 4, 6-8, 10-28] A culmination of this success story of science is the identification of the sequence of ca. 3.2 billion base pairs (bps) that constitutes the entire human genome,[23] an

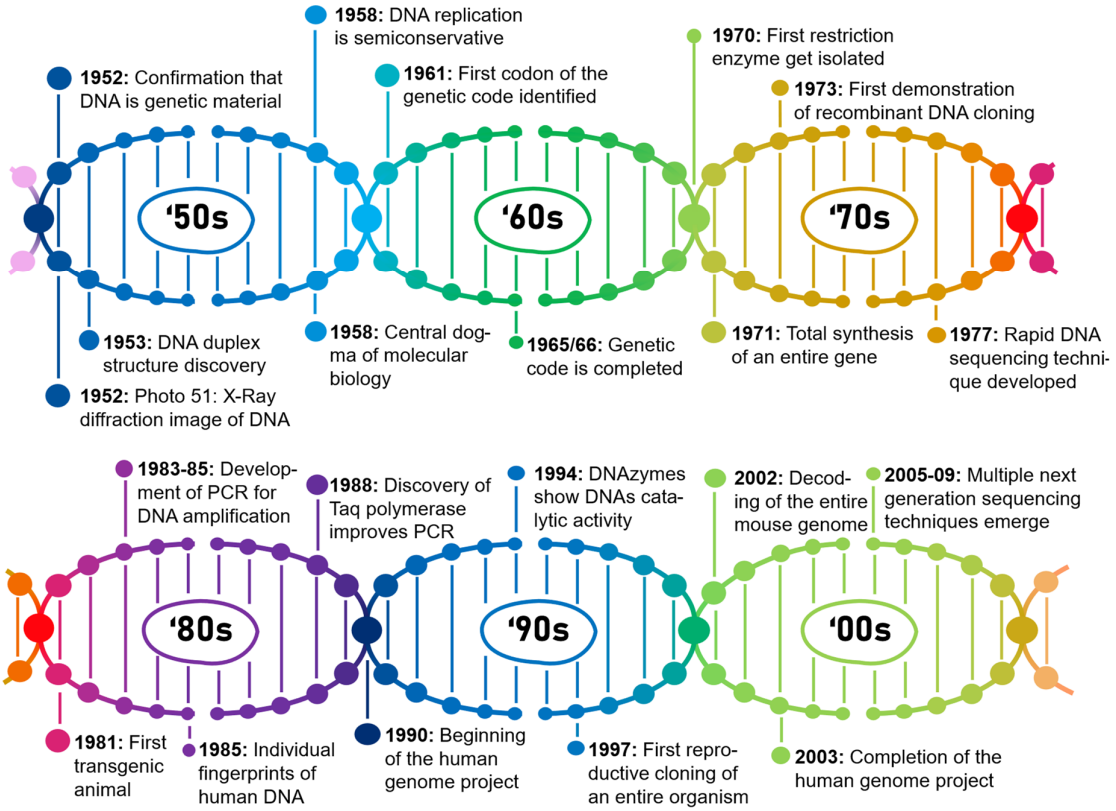


Figure 1 – Historic context Timeline of DNA related scientific breakthroughs from initial structural and functional discoveries in the 1950s until the late 2000s.

effort that undertook 13 years to complete.[29, 30] Even with these milestones reached, further research has not halted, e.g., the ENCODE project is, in spirit, the successor of the human genome project and is meant to reveal all functional elements of the human genome including its transcriptome.[31, 32] Since the original suggestion of the DNA duplex structure, advances in molecular biology, chemistry, biochemistry (Fig. 1) and technological progress (brighter beamlines for X-ray diffraction (XRD) or cryo-electron microscopy (cryo-EM)), have offered us a more detailed and comprehensive picture of DNA's structural features like its folding propensity, its complementarity and the functions that emerge from them.

3.1 World of Nucleic Acids

Since Rosalind Franklin's *Photo 51* revealed the helical nature of DNA's canonical B-isoform from hydrated fibers, it took another three decades to determine the first crystal structure of a DNA duplex in the B-form without a binding partner. This was obtained from a sequence ($d(CGCGAATTCGCG)_2$) now known as the Dickerson-Drew dodecamer.[33] The same sequence was later crystallized both with Mg^{2+} (Figure 2) and Ca^{2+} ions, both at or close to atomic resolution,[34-36] which, despite many technical advancements, still remain rare. This is especially interesting because other examples of DNA atomic resolution are either from shorter Z-DNA duplexes, racemic mixtures or shorter (decamers) duplexes.[37-40] DNA has been shown to adopt different folds in solution. Three of the most common structural forms are A-, B-, and Z-DNA, all of which are antiparallel duplexes.[41]

B-DNA has a helix pitch of 35.7 Å and 10.5 bps per turn. The axial rise of approximately 3.4 Å results in nucleobases that are oriented nearly perpendicular to the helix axis. Crucial structural and functional features of this are its major and minor grooves, which are spanned open by the glycosidic bonds of two nucleobases in a base pair not pointing directly at each other. This in itself is an expression of the sugar pucker as well as ribose angles δ and γ that define the DNA backbone and ultimately cause an asymmetry of the two strands in a DNA duplex relative to each other. The A-type fold of DNA resembles B-form, but the combination of a lower pitch and comparable residues per turn encodes an increased helix diameter. One effect of this is a more cavernous but narrower major groove with a wider, shallower minor groove.[41] This originates from a different orientation of the 5' oxygen in the phosphodiester backbone relative to the main helix axis due to the deoxyribose unit being in a C3' *endo* compared to C2' *endo* conformation.

Z-DNA, unlike A- and B-folds, forms a left-handed helix.[42] The alternation of *syn*- and *anti*-conformation of guanine (Figure 3) and cytosine of $d(CG)_n$ repeat motifs and resulting local chain reversal dictate its characteristic zig-zag backbone shape (Figure 2).

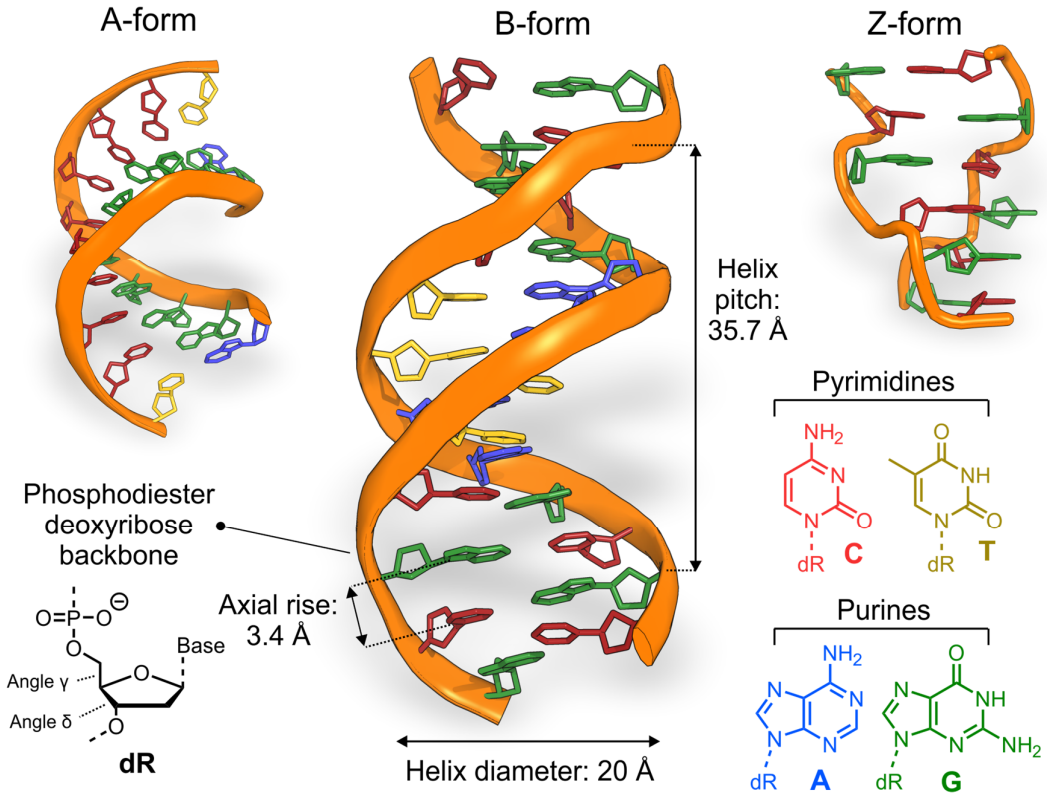


Figure 2 – The polymorphism of DNA folds Crystal structures of DNA in various folds, crystallized at atomic resolution.[34, 40, 43] DNA is displayed with its ribose unit and purine or pyrimidine heterocycles that are colored in blue, green red and yellow for dA, dG, dC and dT respectively. Orange ribbons or tubes represent the phosphodiester backbone.

The parameters of each fold are summarized in Table 1.[42] In principle, a switch between A- and B-forms can be achieved by flipping the sugar pucker in all nucleotides, whereas for Z-DNA, the transition of the base from *anti* to *syn* conformation requires a rotation around the glycosidic bond. Interestingly, the B–Z transition is generally considered more biologically relevant, in part because it is more structurally distinct.[44, 45]

Table 1 – Major helix parameters of A-, B- and Z-DNA.[42]

Parameter	A-DNA	B-DNA	Z-DNA
Helix sense	Right	Right	Left
Residue per turn	11.0	10.5	11.6
Axial rise	2.55 Å	3.4 Å	3.7 Å
Helix pitch	28.2 Å	35.7 Å	45.6 Å
Rotation per residue	32.7°	36°	–9°, –51°,
Diameter of helix	23 Å	20 Å	18 Å
Sugar pucker	C3' endo	C2' endo	dG = C3'*endo*; dC, dT, dA = C2'*endo*

While the various structural conformations of DNA have been well characterized, a remaining question is why nucleic acids adopt defined folded structures in the first place. One of the driving forces to stabilize DNA helices is base pairing, of which Watson-Crick bps are the most common. GC bps are composed of three hydrogen bonds (H-bonds), whereas AT bps consist of two H-bonds (Figure 3). Therefore, GC-rich regions confer higher stability to the DNA duplex, which is in parts responsible for the increase in the melting temperatures of polymeric strands from ca. 45 °C for poly(dA-dT) to above 90 °C for poly(dG-dC) duplexes.[41] The enthalpic gain from hydrogen bonding within base pairs must be considered in context, as unpaired nucleobases can form hydrogen bonds with water. Additional stabilization arises from the entropic gain associated with the release of these water molecules upon duplex formation. This, in turn, means that base pairing is more responsible for the specificity of the DNA duplex formation rather than the initial collapse into a folded structure.[46] The energy gain to overcome the entropic penalty of folding is rooted in the accumulation of non-base-pairing effects. This includes vertical interactions across nucleobases (mostly responsible for the high poly (dG-dC) melting temperature) such as their stacking to give rise to π-π stacking and other van der Waals contacts, as well as the hydrophobic effect, which displaces water and increase the systems entropy due to the shielding of the hydrophobic aromatic surfaces from the aqueous exterior.[41, 46, 47] For example, isolated purines were shown to associate extensively in solution, not by H-bonding, but by vertical stacking even in the absence of a backbone that confines their orientation.[48] Bps can deviate from their idealized orientation. To describe these deviations, a general reference frame has been established that distinguishes between the relative position within one bp (shear, buckle, stretch, propeller, stagger, and opening), for consecutive bps (shift, tilt, slide, roll, rise, and twist) and helical parameters (x/y-displacement, inclination, tip, helical rise, and helical twist).[49]

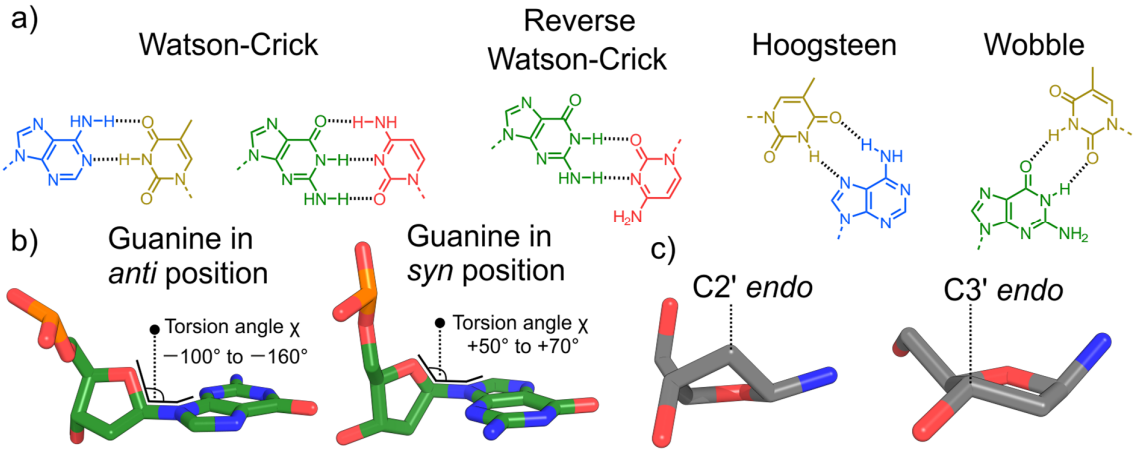


Figure 3 – DNA base pairing and structures a) Hydrogen bonding patterns of different base pairing with the same color code as in Fig. 2. B) Display of *syn*- and *anti*-positions of the base relative to the deoxyribose in one dG nucleotide derived from bases in crystal structures with the atoms colored in green, red, blue and orange for carbon, oxygen, nitrogen, and phosphorus respectively. c) Examples of sugar puckers for C2' *endo* and C3' *endo* conformations derived from solid-state structures at atomic resolution.[37, 40]

Watson-Crick bps are not the only geometries in which the nucleobases can H-bond to each other. In reverse Watson-Crick bps, one nucleobase is flipped 180° relative to its orientation in Watson-Crick bps. Further, base pairing is not exclusive to CG and AT pairs. Hoogsteen bps stabilize DNA triple helices and G-quadruplexes, and Wobble bps are prevalent in more kinked architectures such as in transfer RNAs (tRNAs) or aptamers.[42, 46, 50, 51]

G-quadruplexes are yet another folded structure of nucleic acids, which is not limited to DNA, but also prevalent in RNA folds. A self-association of four guanylic acids into quartets was first described as early as 1962, followed by their structural characterization both by nuclear magnetic resonance (NMR) spectroscopy in 1993 and single-crystal XRD in 2002.[52-55] They require the presence of a monovalent central metal cation (typically K^+ or Na^+) that is surrounded by four Hoogsteen base-paired guanines composing G-quadruplexes characteristic tetrads (Fig. 4b). Unlike A-, B-, and Z-DNA that are all antiparallel duplexes, G-quadruplexes can be formed *in vitro* from one to four separate strands that may be either parallel with propeller-type loops or antiparallel from lateral or diagonal loops as well as hybrid combinations of these orientations (Fig 4a). Additionally, the guanines can be either *syn* or *anti*, which greatly adds to their polymorphism and dictates the quadruplex helix sense. Initially, G-quadruplexes have been considered structural anomalies lacking a significant function in nature. Immunostaining of G-quadruplexes in human cells has revealed that these topologies exist and that they are dynamically sensitive towards the cell cycle, being particularly present in the S-phase when DNA is single-stranded.[56] Later, over 700,000 G4 structures were identified in the human genome and their prevalence at telomeres and promoter regions has been linked mostly to protecting chromosome ends, and regulating telomerase activity as well as gene transcription.[57-59]

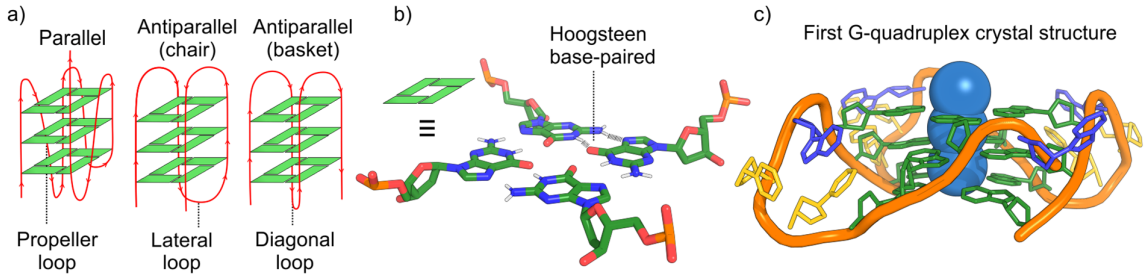


Figure 4 – G-quadruplexes a) Different folds of single stranded G-quadruplexes. b) One guanine tetrad. The central potassium cation and non-polar hydrogen atoms are omitted for clarity. Guanine carbons are colored green, other atoms are colored in red, blue, orange and white for oxygen, nitrogen, phosphorus and hydrogen respectively. Hydrogen bonds are displayed for one Hoogsteen pair as dashed lines. c) First crystal structure of a G-quadruplex. DNA is depicted as in Fig. 2 with the phosphodiester backbone represented as an orange tube and central potassium ions displayed as blue spheres. Water atoms are omitted for clarity.[55]

However, even though non-B DNA's contribute to a variety of critical organismal processes, DNA, *in cellulo*, is prevalent mostly in B-form and most of its functions, especially storage of the genetic information and the structural basis of gene regulation emerge from it.[60-62]

3.2 Protein Nucleic Acid Interactions (PNIs)

Genetic information is stored as DNA. However, this does not mean that DNA is dormant or static. Instead, it exists in a highly dynamic environment. It is duplicated by the replication machinery, transcribed to various RNAs for which it is constantly packaged and unpackaged. It is also rearranged and subjected to numerous repair mechanisms such as double-strand break, mismatch, or base excision repairs.[63] All these crucial processes are dependent on proteins and how they interact with DNA in their respective ways. Most interactions between DNA and proteins *in vivo* are non-sequence specific: in the cellular machinery, some processes must involve and be exerted by various proteins and enzymes on the entire genomic DNA. This includes packaging by histones to obtain the chromatin structure and its remodeling, DNA unwinding by helicases before transcription or replication, or relaxing supercoiling by topoisomerases after replication.[63] Non-sequence-selective interactions add a significant amount to respective sequence-selective DNA-binding proteins (DBP)-DNA binding. In one study investigating non-consensus transcription factor (TF)–DNA interactions, such non-consensus binding was estimated to account for as much as 45–66% of total TF–DNA binding.[64]

Many DBPs share similar properties, sequence or non-sequence selective, such as their high isoelectric points, as their basic residues match well to pair with the DNA phosphate negative charges to promote interactions via electrostatic attraction or salt bridges. They often occur with protein β-sheets or loops, with DNA's minor groove, as the close proximity of phosphates results in a higher local negative charge density compared to the major groove.[65, 66] These interactions can be supplemented by hydrophobic contacts, e.g. with ribose methylenes or sometimes hydrophobic residues intercalate between nucleobases causing DNA bending or kinking.[66-69]

These effects can be exemplified by the diverse ways in which different organisms package their DNA. Histones are rich in lysine and arginine and facilitate electrostatic interactions visualized from the crystal structure of the nucleosome core particle (NCP) in Fig. 5c. The accumulation of the positively charged residues on the outside of the histone octamer in contact with DNA are highlighted in red. The histone-like protein Sac7d, derived from a thermophilic archaeon, does increase DNA's thermal stability by intercalating amino acid side chains in between nucleobases (highlighted in Fig. 5a) to induce sharp kinks. As an example of extensive β-sheet-backbone contacts, a crystal structure of HU in complex with DNA shows chromatin-like compaction of DNA. This bending occurs via many ribose-phosphate contacts of the protein lateral loops and simultaneously forms many electrostatic interactions (Fig. 5b), demonstrating that these effects often act in tandem.

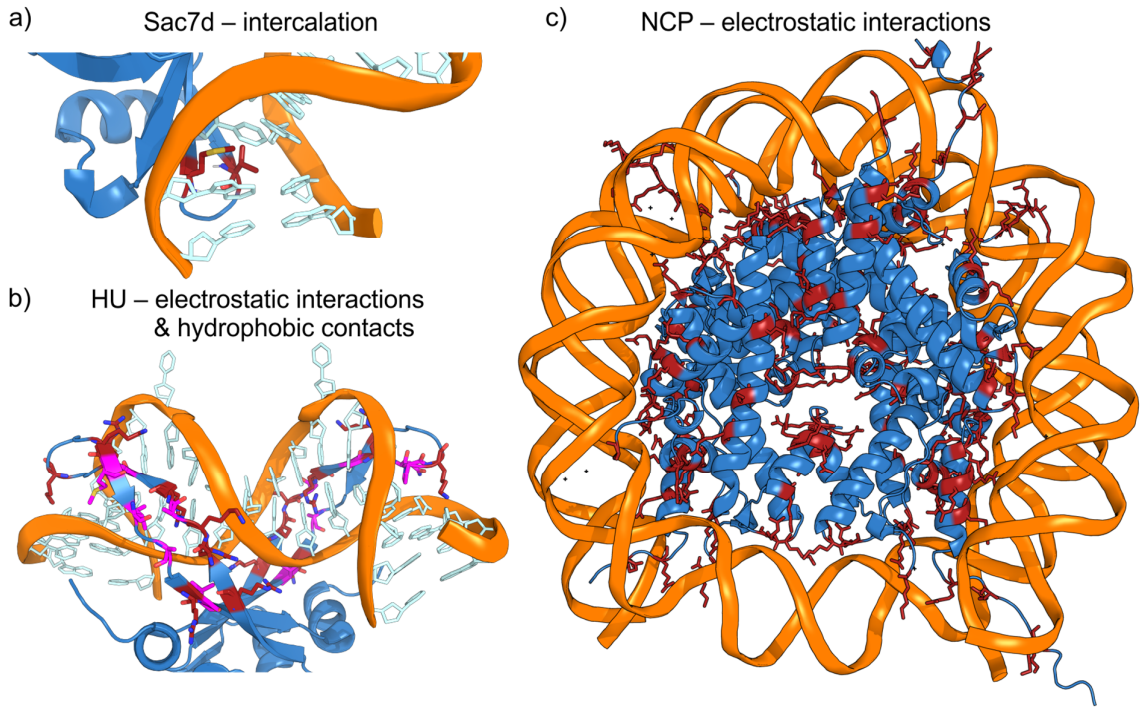


Figure 5 – **Non-sequence selective DNA protein interactions** a) Sac7d DNA co-crystal structure with intercalating Met26 and Val29 highlighted as red sticks. b) HU-DNA co-crystal structure with hydrophobic residues of DNA binding domain (Res. 50-76) highlighted as magenta sticks. Arginines and lysines are highlighted as red sticks. c) Crystal structure of the nucleosome core particle with arginines and lysines colored as red sticks. In all structures water atoms are omitted for clarity and DNA bases are represented as their deoxyribose unit and respective heterocycle as light blue sticks.[66, 70, 71]

However, for processes that are biologically decisive such as gene expression, the genetic information stored within the base sequence must be accessible by the cellular machinery sequence selectively. The process of how proteins specifically interact with DNA to gain this information is called DNA readout. This can be further dissected into direct and indirect readout.[63] Indirect readout is a blurred transition from non-selective interactions. It was coined as the selective recognition of DNA sequences that is driven by DNA's conformational polymorphism as a function of its base sequence. It can be described more technically as sequence-specific recognition where direct interactions are confirmed to be missing.[63, 72] These conformational changes can cross A-, B-, and Z-helical families, but more often are deviations of average B-DNA bp geometries as described above. They cause DNA deformation due to differences in base pair step parameters for either pyrimidine-purine (most deformable), purine-purine (same as pyrimidine-pyrimidine, intermediate deformable), and purine-pyrimidine (least deformable). Each of the ten possible base pair steps has a unique fingerprint, both in the presence and absence of bound protein. Proteins can sense variations of these parameters, e.g. their helical twist or roll. Most indirect readout mechanisms involve DNA minor groove contacts with β-sheets. For example, A-T rich regions cause a groove narrowing or bending and via higher electrostatic potential can attract arginine residues, interestingly

more so than lysines. Direct readout refers to the specific recognition of base sequence information through direct protein–nucleobase contacts by either H-bonding or hydrophobic contacts. These high-specificity mechanisms typically rely on interactions between protein α-helices and nucleobases exposed in the DNA major groove, because nucleobases are more sterically accessible in the major groove of DNA. Each bp displays its donor or acceptor moieties in a different array for protein readout as depicted in Fig. 6.[65, 73]

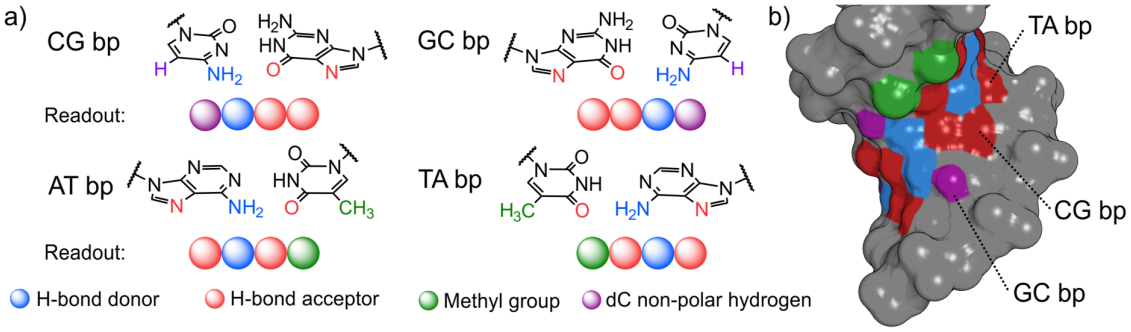


Figure 6 – DNA major groove readout. a) Different array of each bps hydrogen bond donor, acceptor, dT's methyl group and dC's non-polar hydrogen atom that is displayed in the major groove of DNA. b) Surface representation of a DNA major groove from an atomic resolution crystal structure with the bps moieties marked in the same color code as in a).[36]

H-bonds formed on the basis of these base-pair-specific signatures provide the highest degree of specificity, particularly for distinguishing purines, and their specificity depends both on the number of contacts made by protein residues and on the uniqueness of their geometric arrangement. Bidentate H-bonds generally contribute to the highest degree of binding specificity, followed by bifurcated H-bonds, with single hydrogen bonds providing little specificity. Bidentate H-bonds are so effective because geometries can be formed within one base, within two bases in one bp, or across bps, either vertically or diagonally across strands.[73] Complementing H-bonds, hydrophobic contacts predominantly contribute to pyrimidine recognition to distinguish between thymine and cytosine. [74-76]

The major groove can host a protein α-helix in both shape and diameter. These interactions are common in many DNA-binding motifs such as helix-turn-helix (HTH), winged HTH, basic leucine zippers (bZIP) and zinc fingers (Figure 7a-d). Unlike β-sheets, an α-helix projects its side chain outwards in all directions allowing for diverse interactions.[77] There are notable deviations from this general trend. For example, a ribbon-helix-helix motif like in the Met repressor operator that binds via a β-sheet major groove insertion (Fig. 7f) or the TATA box binding protein that binds through a β-strand minor groove interaction, but is accompanied by a significant groove widening (Fig. 7g).[78-80] Of particular interest for this work are proteins from the erythroblast transformation specific (ETS) family — a group of evolutionarily related transcription factors comprising 28 genes in humans and representing one of the largest TF families.[81, 82] They direct gene expression in diverse normal and disease states and have been associated with various types of cancer, like leukemia, prostate, or pediatric cancers through

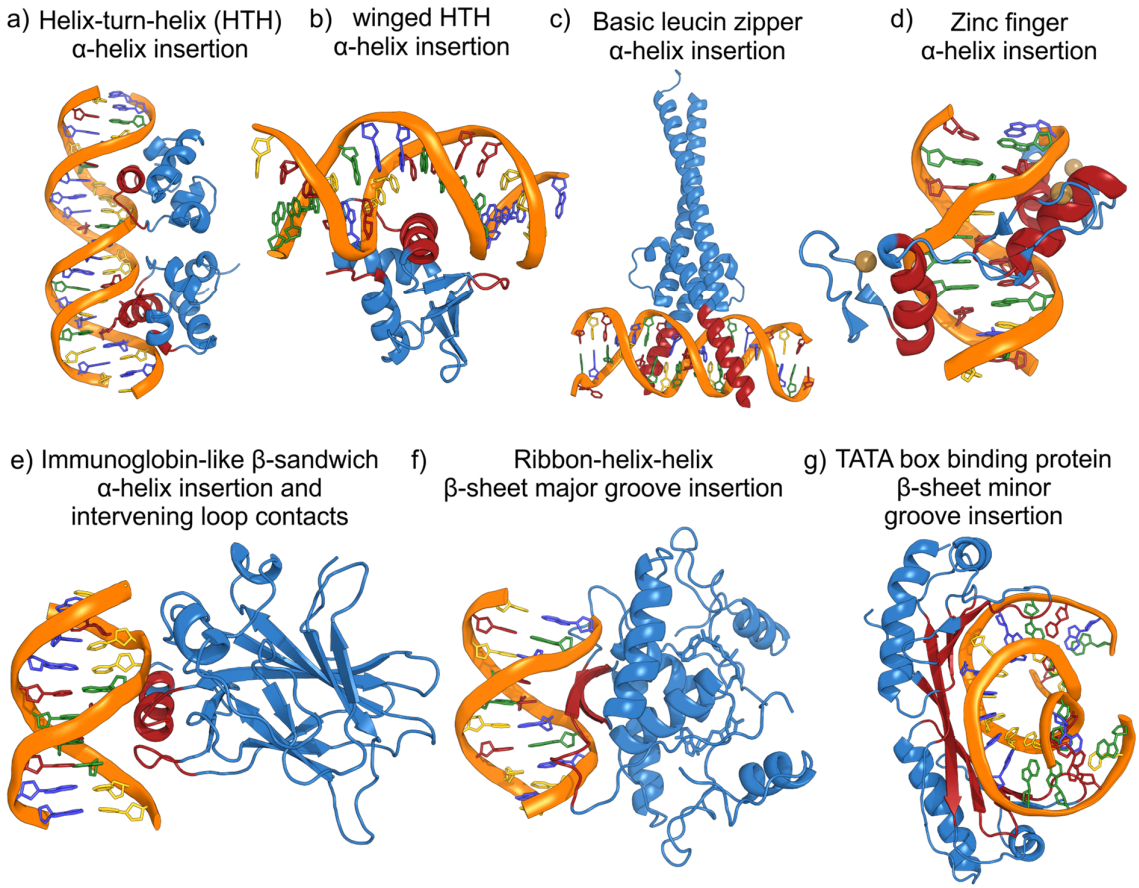


Figure 7 – Common motifs of DBP-DNA interactions – Crystal structures of different DNA binding motifs. a) The phage 434 Cro/OR1 complex.[76] b) ETS transcription factor PU.1. c) Heterodimeric bZIP Myc-Max.[83] d) Zinc finger Zif268. Zinc atoms are displayed as dark-yellow spheres.[77] e) P53 tumor suppressor complex.[84] f) Met repressor-operator complex. Additional protein copies in the asymmetric unit are omitted and DNA is truncated for clarity.[79] g) TATA box binding protein.[78] The DNA-binding domain of each protein is colored in red. DNA is displayed as established in Fig. 2. All water molecules are omitted.

gene fusion.[85-87] They bind to specific promoters and enhancers and facilitate the recruitment of other components of the transcriptional machinery to form cooperative ternary complexes, for example between the ETS TF serum accessory protein 1 (SAP-1) and MADS-box serum response factor (SRF), that create protein-protein contacts to optimize binding to the c-fos promoter.[63, 88, 89] These TFs are defined by their highly conserved ca. 85 residues winged helix-turn-helix (wHTH) ETS DNA-binding domain that specifically binds a 5'-GGA(A/T)-3' motif by the insertion of a recognition helix into DNA's major groove. wHTH folds consist of a three-helix HTH bundle that contains an antiparallel β-sheet that sits adjacent to the HTH motif (wing 1) to create backbone contacts with the minor groove located above (Fig. 8b). ETS folds are also composed of a second wing (loop between HTH helices 2 and 3) to form additional backbone contacts on the opposite site of the HTH core.[63] Specific interactions in the major groove consist of both direct and water-mediated H-bonds of two arginine residues with two guanines and one adenine exemplified by a crystal structure of PU.1 bound to its consensus

sequence.[90] Additionally, as common for ETS TFs, the wings adjacent to the consensus site create backbone contacts dominated by electrostatic interactions to complement the binding specificity of each TF (Fig. 8b). ETS TFs were further categorized into four different classes, based on contacts these flanking loops form with bases that lie outside their 5'-GGA(A/T)-3' core.[91] For example, in the case of PU.1, an ETS TF that plays a pivotal role in hematopoiesis, the wing 1 and 2 each form a notable hydrophobic contact above and below the consensus site (Fig. 8b, magenta). These contacts are lacking in other ETS TFs like SAP1, which creates specificity within the ETS family as they share their consensus sequence and can determine between high- and low- affinity interactions within one given TF/DNA complex.[92] More recent studies of human PU.1 have shown that it spans affinities of multiple orders of magnitude, depending on the flanking nucleotides, through direct interaction of its Glu226 side chain with the 5′ flank and indirect readout effects on the 3′ flank arising from sequence-dependent DNA flexibility.

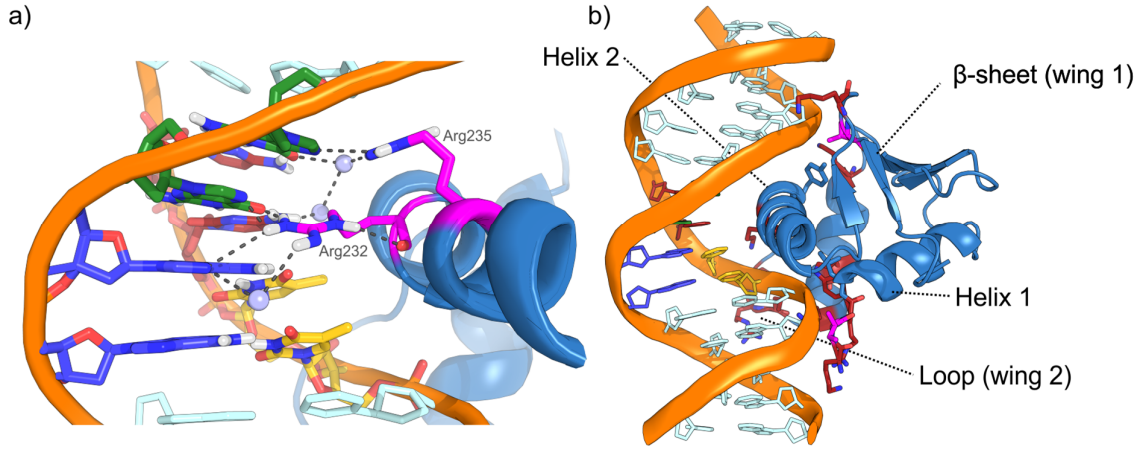


Figure 8 – PU.1-DNA interactions a) Binding interactions of mouse PU.1's α-helix 2 with its 5'-GGA-3' core consensus sequence.[90] Arg232 and Arg235 are highlighted with their carbons colored in magenta and their H-bonding interactions, both direct and water-mediated, are displayed as grey dashed bonds with water atoms colored as pale blue spheres. b) Contacts of protein side chain with the backbone of DNA. Polar and electrostatic interactions are colored in red and hydrophobic contacts are colored in magenta. DNA bases are colored as light-blue sticks, except consensus bases that are colored in the established color-code.

To summarize, DBP–DNA interactions are diverse, but the plentiful structural investigations allow for some trends to emerge. This is reflected in one study about the dissection of 62 DBP–DNA complex crystal structures, that has shown that even though direct vs. indirect readout proportions vary significantly between each unique DBP-DNA complex, both components are present in every interaction.[93]

3.3 Interfering in protein nucleic acid interactions – nucleic acid therapeutics

The therapeutic interest in interfering with protein nucleic acids is immense and many different strategies to achieve this have been pursued. Small molecule therapeutics to interfere in PNIs are relatively rare, but an expanding field, especially focusing on RNA-binding proteins (RBPs).[94] Some notable examples of small molecule drugs or drug candidates exists, albeit there is no general mechanism of action (MoA). Some bind the DNA minor groove to stall transcription (lurbinectedin, pyrrole-imidazole polyamides)[95, 96] and some intercalate between nucleobases to prevent TF readout (echinomycin, actinomycin D).[97-99] Others stall enzymes by poisoning ternary NA-protein-inhibitor complexes (raltegravir) and some molecules combine these features in a single molecule (camptothecin).[100, 101]

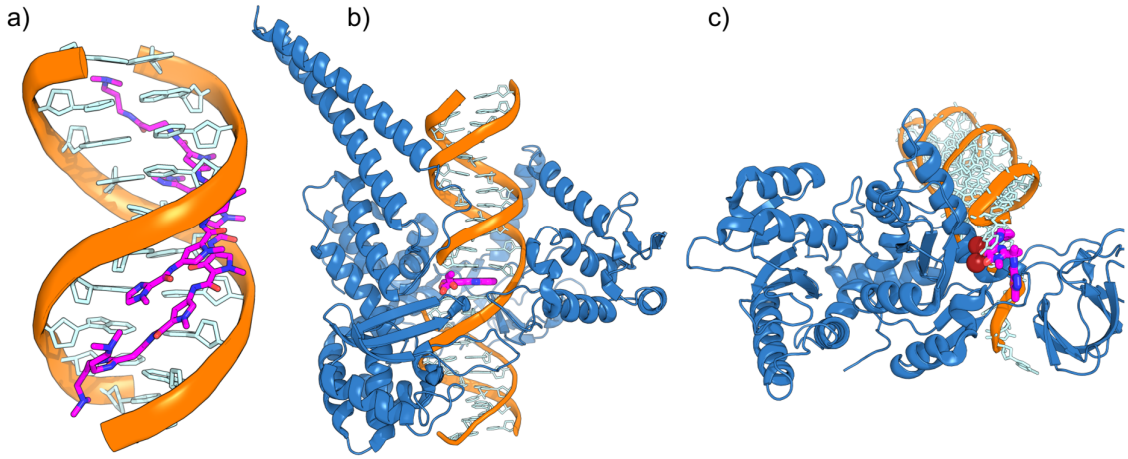


Figure 9 – Small molecule interference in PNIs a) Crystal structure of DNA in complex with a Py-Im polyamide bound in its minor groove. b) Camptothecin intercalated in a DNA-Top1 complex. c) Complex of prototype foamy virus in complex with its cognate DNA and raltegravir. Mn^{2+} ions are displayed in red spheres and proteins N-terminal residues (1-93) are omitted. DNA bases are displayed as pale blue rings representing their nucleobases and the ligands are colored in magenta.[95, 100, 101]

The existing small molecule therapeutics already project a trend. Firstly, some of them being cyclic depsipeptide antibiotics are not traditional small molecules and secondly, they primarily either target an existing DBP-DNA complex or specific DNA sequences. Thus, they prevent proteins from binding it, but rarely target the protein itself, especially at its DNA binding site. Small molecules (including cyclic peptides) struggle to resemble the binding motifs, especially the spatial arrangement of polar residues that dominate PNIs in direct readout interactions. To complement the successful small molecules therapies, targeting nucleic acids on a more genome-wide level, including non-coding regions, is necessary. Looking beyond proteins (typical small molecule targets), many therapeutic strategies to target PNIs or nucleic acids involve nucleic acids themselves.[102] The linearity and complementarity of the genetic code found in DNA and RNA makes them easier to design for targeted intervention. They can be

classified into antisense oligonucleotides (ASOs), small interfering RNA (siRNA), decoy oligonucleotides, aptamers and, to a lesser extent, triplex-forming oligonucleotides (TFOs).[102] Of these classes, antisense oligonucleotides are the most successful therapeutic with the most United States Food and Drug Administration (FDA) approved drugs, e.g. fomivirsen (first of its kind, later revoked), mipomersen, eteplirsen, nusinersen, inotersen and golodirsen.[103-107] They are engineered, single-stranded DNA, RNA, or hybrids and bind to their complementary sequences on mature messenger RNA (mRNA) by Watson-Crick base pairing. Their success stems from the diversity of their potential applications such as the inhibition of translation, RNase H1 induced mRNA cleavage or can even enhance translation by targeting upstream open reading frames (Fig. 10a).[108] siRNAs, with multiple approved drugs such as patisiran or givosiran,[109, 110] are double-stranded RNA molecules. In the endogenous process, a precursor double-stranded RNA or pre-miRNA is cut by dicer proteins to yield a 3' overhang siRNA duplex, which is not necessary for the therapeutic approach. This RNA and a multi-protein complex then form an RNA induced silencing complex (RISC) while the sense-strand is discarded in the process. The antisense strand guides the RISC to its target sequence and then degrades it with its Argonaute 2 endonuclease unit (Fig. 10b).

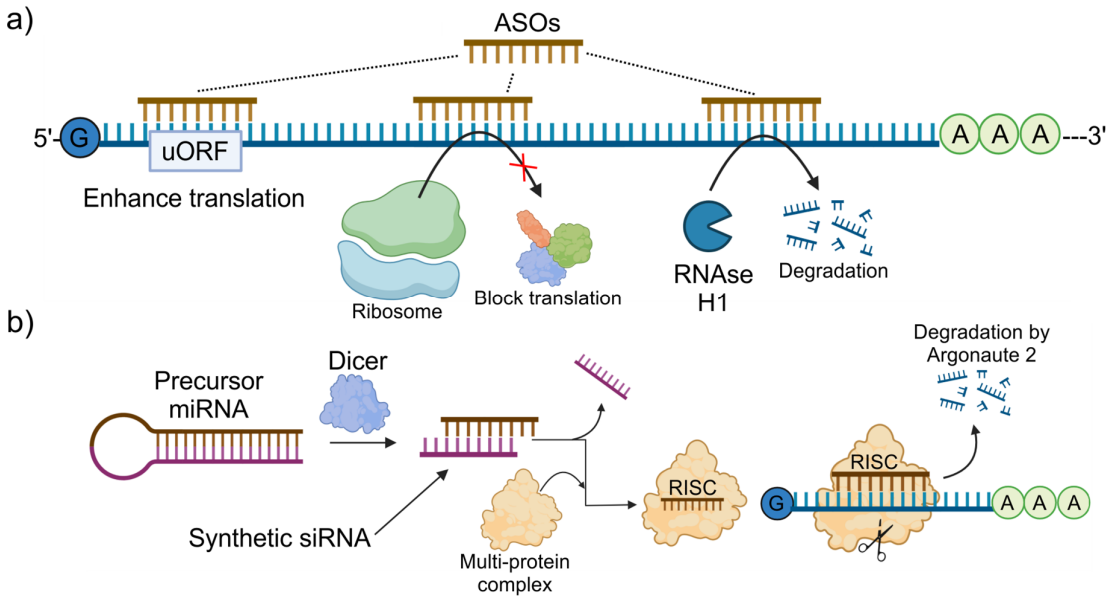


Figure 10 – ASOs and siRNA a) Examples of ASO applications. b) siRNAs guide RISC to complementary mRNA, where argonaute proteins cleave the target, leading to its degradation and gene silencing.

ASOs and siRNAs are typically extensively chemically modified for a variety of reasons (Fig. 11). Modifications enhance pharmacokinetic properties: they can increase target binding affinity, enhance therapeutic potential, the safety profile and cellular delivery. Most importantly, the oligonucleotide must be protected from degradation by cellular nuclease systems.[108] Among them, modifications of the 2'-ribose hydroxy group are present in almost all RNA ASOs and siRNAs. They typically consist of an etherification to simple ethers such as methoxy- or methoxy-ethoxy groups, or replacement of the hydroxyl group with a fluorine.

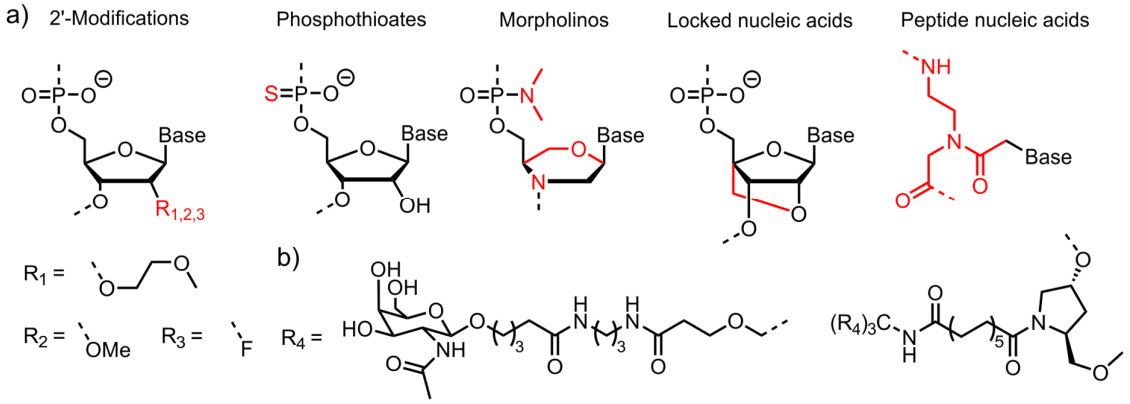


Figure 11 – Important nucleic acid modifications a) Common modification of nucleic acids in ASOs, siRNAs and other application. Differences from canonical nucleotides are colored in red. b) Tris-GalNAc cap commonly conjugated in siRNA-therapeutics.

This increases RNA's stability by both preventing 2′-OH-mediated intramolecular cleavage, and, in some cases, can increase duplex stability while increasing nuclease resistance especially when used in tandem with phosphothioates that are often used for RNase H1-dependent MoAs.[111] Other stabilizing backbone modifications are morpholinos. They have their ribose replaced by a morpholine ring, which is incorporated into a phosphordiamidate. As a result of this, the backbone linkage is uncharged. It is thus incompatible with RNase H1 recruitment and more typically used as a translation blocking ASO.[102] Other backbone modifications of nucleic acids, albeit with no approved drugs, are locked nucleic acids (LNAs) that lock the sugar in an C3' *endo* conformation by bridging 2' oxygen and 4' carbon and peptide nucleic acids (PNAs) that replace the entire backbone with an oligo-γ or -δ amino acid chain. Other modifications include linking siRNAs to other moieties, such as the conjugation to a branched tris-*N*-acetyl galactosamine cap, that has high affinity for receptors on hepatocytes in a targeted delivery approach (Fig. 11b).[112]

These mRNA targeting nucleic acid drugs represent the most therapeutically successful approaches available to date. Yet, they can only interfere with PNIs indirectly by silencing genes that express the relevant protein of interest. A more direct interference would involve targeting the interaction between nucleic acid and DBP where it is most specific and with the highest abundance, which are direct readout contacts between DNA's major grooves and protein α-helices. In principle, both components can be targeted individually. For DNA, a ligand must selectively bind inside a DNA's major groove competitively, preventing direct contacts. Interestingly, a good ligand for this interaction is another nucleic acid strand, albeit with some restrictions. In oligopurine strands, the bases display their Hoogsteen side in the major groove allowing a pyrimidine-rich TFO to bind by canonical base triads T:AT and C^+:GC (C^+ represents protonated C) in a parallel fashion (Fig. 12a). Conversely, purine-rich TFOs bind oligopyrimidine-strands anti-parallel via reverse-Hoogsteen bonds.[113, 114]

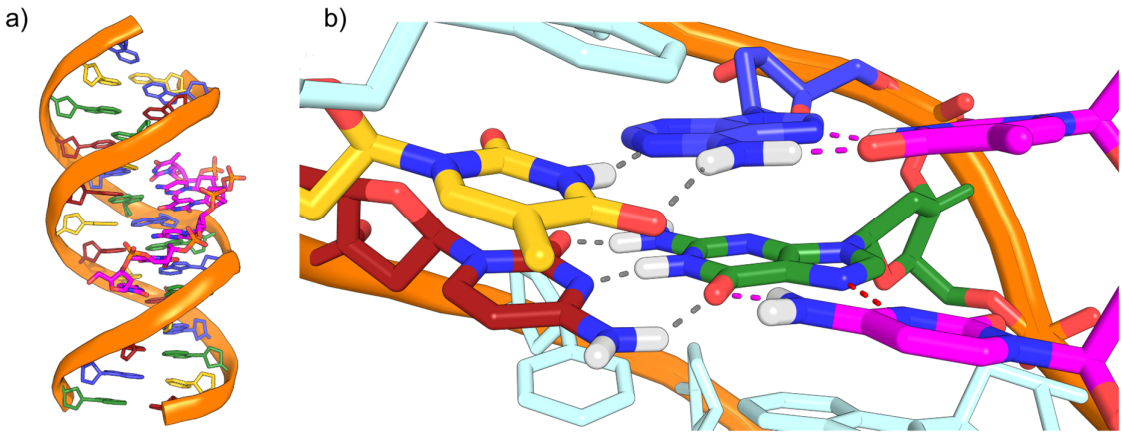


Figure 12 – TFO – major groove interactions a) NMR structure of a TFO bound by Hoogsteen base-pairing to the oligopurine strand of a B-DNA duplex. DNA is displayed in the established color code and the TFO is displayed with its carbons, oxygens, nitrogens, phosphorus and hydrogens in magenta, red, blue, orange and white respectively. [115] b) Base-pairing of the TFO with the DNA duplex for one T:AT triad (top) and one C:GC triad (bottom). Watson-Crick H-bonds are displayed as grey dashes, Hoogsteen H-bonds are displayed as purple dashes. The lacking H-bond of the unprotonated dC and dG is displayed as red dash.

The ability of TFOs to suppress gene expression by binding promoters has been investigated in several studies targeting c-MYC, erbB2, and BCL-2. However, TFO-based approaches face several major challenges, including inefficient nuclear uptake, susceptibility to nuclease degradation, and difficulty achieving high-affinity binding to target sequences. Most importantly, they are hampered by being limited to polypurine or polypyrimidine stretches that are intrinsically rare on a genome-wide level. Moreover, pyrimidine-rich TFOs require a protonated cytosine to form triplexes, whereas particularly guanine rich TFOs prefer to fold to G-quadruplexes at physiological K^+ concentrations.[113] The accumulation of these challenges has led to their pursuit being largely dropped in the pharmaceutical industry.

Targeting the protein-side of DBP-nucleic acid interactions shares some of these challenges, like nuclear delivery. However, since they are not limited by the stringent sequence constraints of TFOs, they offer far greater opportunities due to the intrinsic diversity of potential target proteins, e.g. TFs. TFs form the basis of selective gene control, and gene dysregulation is a common driver of many diseases. As such, efforts to modulate gene regulation for therapeutic purposes have been one of the most intractable endeavors in drug discovery.[116, 117] Their role in cellular processes demands them to recruit multiple co-regulators through a single short domain, which facilitates the necessity to adopt specific but highly disordered protein-protein interactions (PPIs).[116, 118] For this reason, TFs, unlike enzymes, do not typically involve targetable hydrophobic pockets, but rather shallow surfaces as displayed in the homo-dimerization of NF-κB (Fig 13a). Additionally, TF-DNA interactions often include many low-affinity interactions over a wide surface and only a few specific contacts for high specificity. Some TFs, including bZIPs, basic helix-loop-helix (bHLH) or some high mobility group (HMG) box proteins are not only disordered at their PPI interface, but also in their DNA binding

domain and their α-helices only become helical upon interaction with their cognate DNA.[119-122] Since interference in either TF PPIs or their PNIs was deemed to be difficult, TFs were considered undruggable, but significant progress has been made over the past decade with synthetic oligonucleotides and small molecules alike.[118]

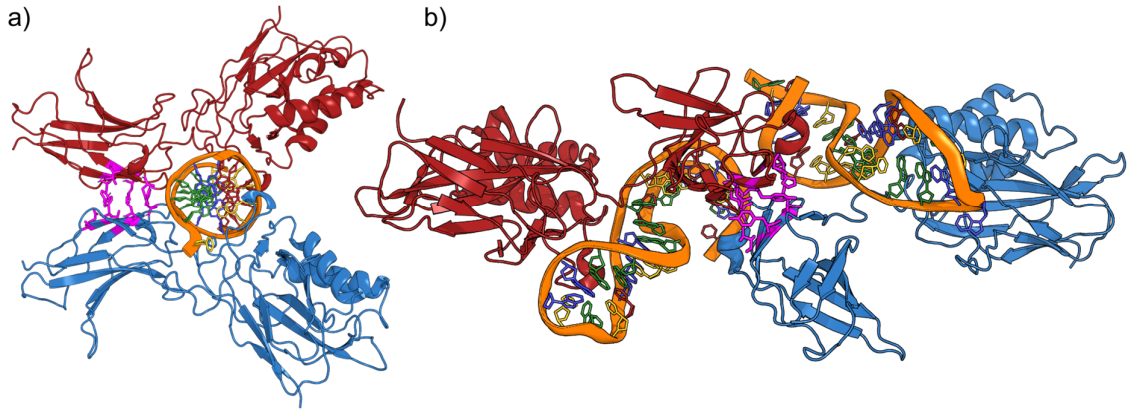


Figure 13 – Decoys and aptamers to target DBPs a) Crystal structure of NF-κB P50 homo-dimer bound to its palindromic cognate DNA strand with each of the respective protein chains colored in blue and red respectivly.[123] b) Crystal structure of an RNA aptamer bound to the same NF-κB P50 homo-dimer.[124] Protein dimerization domain residues are colored as magenta sticks.

Among nucleic acid therapeutics, aptamers are promising candidates due to their unique structural topologies.[125] They are composed of either DNA or RNA molecules with the length of 20-100 nucleotides and can fold into unique 3D structures, often unprecedented in nature. They initially emerged from systematic evolution of ligands by exponential enrichment (SELEX), an *in vitro* display-selection method to identify nucleic acid binders against a target from a library of 10^{13} to 10^{15} molecules.[126, 127] The only FDA-approved aptamer drug, pegaptanib, does not actually target a TF, but the vascular endothelial growth factor 165 instead.[127, 128] Aptamers have been selected against TFs, like against the P50 homo-dimer of NF-κB, that shows ca. 300 fold higher specificity for P50 over P65, another subunit of NF-κB.[129] From its XRD structure (Fig. 13b), the protein still dimerizes via the same interface, but one aptamer occupies the DNA binding site of each protein and heavily distorts the overall assembly of the complex.[124] Whereas aptamers leverage complex 3D folds that allow them to bind diverse protein targets, decoy oligonucleotides take a more direct approach by mimicking natural DNA binding sites to competitively inhibit TF-nucleic acid interactions. They are typically double-stranded DNA fragments that contain the consensus sequence of their target TF. There are no current FDA approved decoys, but unlike aptamers, several candidates are in clinical trials, including STAT3 for antitumor treatment, a decoy oligonucleotide designed to interfere with NF-κB-DNA interactions.[130, 131]

Nucleic acids have established themselves as drugs or drug candidates whose activity relies on their folded structures, much like certain peptides and carbohydrates, but this raises the questions if similar structure-based function could be derived from artificial abiotic backbones.

3.4 Going Beyond Nature's Toolkit – Foldamers

The three classes of polymeric biomacromolecules, peptides, nucleic acids, and oligosaccharides, all fold into defined structures in solution, and their functional diversity is expressed within their complex three-dimensional structure. Many of these properties can be replicated by synthetic oligomers. Sam Gellman first coined the term "foldamers", to loosely classify "any polymer with a strong tendency to adopt a specific compact conformation."[132] Biomacromolecules may be excluded from this compound class, not due to the absence of the relevant defining properties, but because they are already described as biopolymers.

However, unlike biomolecules, foldamers are not confined to nature's toolkit. In particular, backbones extending beyond α-peptides or phosphodiester linkages allow for topologies distinct from those found in nature.[133] Bio-oligomers are restricted to α-amino acids and nucleotides by evolutionary pressure, but foldamer science explores the broader, non-natural chemical space beyond these boundaries.[134, 135] The collapse of foldamers into their structurally defined folded states is governed by the same rules established in the folding of nucleic acids: attractive or repulsive interactions or contacts between sites in oligomeric sequences (phosphate charge repulsion, H-bonding, π-π-stacking), local conformational restrictions (sugar pucker) and solvent effects or a combination of all three.[136] Some foldamers are directly derived from their biotic counterparts such as the above outlined nucleic acid analogues morpholinos or LNAs or the class of peptidic foldamers, that expand upon canonical α-peptides. The latter achieve this by expanding the main chain repeat unit by methylenes to form β- (Fig. 14a, c, d), γ- or δ-peptides or hybrids of these.[137, 138] In both bio-oligomers and foldamers, helices are the most common type of secondary structure.[139-141] Most building blocks have an intrinsic curvature and thus adopt a constant translational and rotational relationship, that allows for a compact stack of atoms along a helix axis.

β-peptides can have their side chains attached at either methylene group of the backbone (Fig 14a) to give β_2- or β_3-isomers and typically form 3_{14}-helices (Fig. 14d) or can be locked in a specific conformation (conceptually comparable to LNAs) to adopt 2.6_{12} helices for building blocks like 2-aminocyclopentanecarboxylic acid (2-ACPC, Fig. 14c).[133, 142, 143] Instead of the canonical *i*, *i*+4 H-bonding pattern of a 3.6_{13} α-helix (Fig. 14b), these helices adopt alternative arrangements by interacting with different residues vertically (*i*, *i*+2 for 3_{14} helix or *i*, *i*+3, 2.6_{12} helix).[143] Conversely to canonical peptides and (*S*)-β_2 peptides, helices of (*S*)-β_3-peptides form *M*-helical arrays (Fig. 14d). When viewed from the N-terminus (top) to C-terminus (bottom), the carbonyl groups point upwards when compared with 3.6_{13}, 2.6_{12} and 2.5_{14} helix types leading to a reversal of the helix macrodipole (C- to N- in β-peptides).[144] N,N′-linked oligoureas also adopt a well-defined 2.5 helical structure (Fig. 14e).

Albeit their intramolecular H-bonding pattern is quite different. In oligoureas, the structure is held by H-bonds closing 12- and 14-membered rings formed between C=O(i) and N′H(*i*+2)

and NH(*i*+3) respectively. While ACPC-peptides achieve helical stability through backbone conformational locking, oligoureas attain comparable stability via bifurcated H-bonds that restrict backbone flexibility. Thus, they have been used as scaffolds for biomimicry, protein-like assemblies or biomimetic membrane channels.[143, 145-147] When further elongating the main chain length, e.g. to δ-peptides, constructs become significantly more flexible due to two additional rotatable bonds.[148] Nevertheless some structural data was obtained by conformationally restricting backbone conformation, by either combining α- with cyclohexyl δ-amino acids that form a 13/11 helix or from a homo-δ-peptide when locking some rotational bonds by cyclopropane rings.[149, 150]

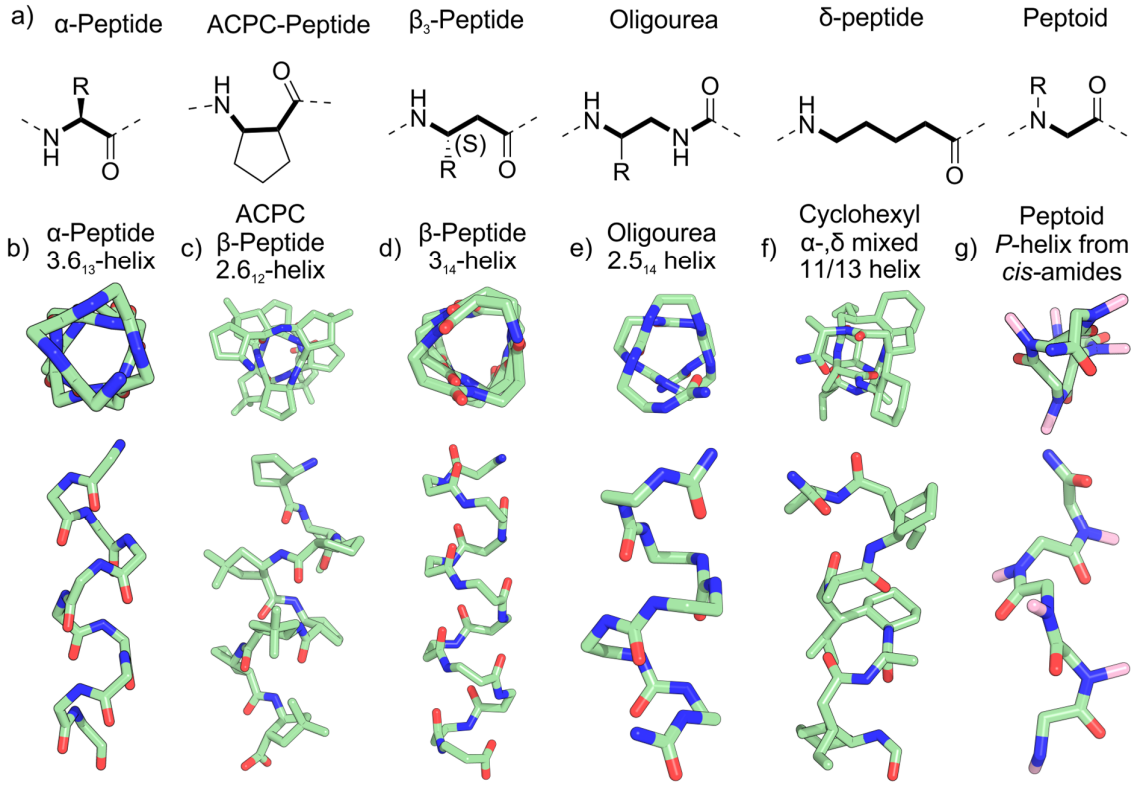


Figure 14 – Peptide derived foldamer backbones and helices a) Backbones of canonical α-amino acids and different peptide inspired foldamers. Crystal structures of: the α-helix of crambin (b),[151] ACPC β-peptide helix (c),[142] a β-peptide 14-helix (d),[152] an oligourea 2.5 helix (e),[145] an α-/δ- heteromeric sequence (f)[149] and a *P*-helical peptoid (g).[153] Atoms are colored light green, red and blue for carbon, oxygen and nitrogen respectively. Hydrogen, solvent and side chain atoms are omitted for clarity. Terminal groups or chain-elongations are cut. In (f) and (g), the first side chain carbon atom is colored in light pink.

Polyglycine main chains that have their side chains attached to the peptide nitrogen instead of the carbon in α-position are called peptoids. Their amides lack the hydrogen atoms to form innate H-bonding patterns and as such, their properties resemble those of polyproline chains, like the ability to form collagen triple helices.[154] Additionally, the lack of internal H-bonds causes the peptoid's conformation to predominantly rely on the side chain, e.g. the tertiary amides can adopt both *cis* and *trans* conformations much more easily than in canonical

peptides or can form either right-handed (*P*) or left-handed (*M*) helices depending on the side chains (Fig. 14g shows a *P*-helical example).[155] This side chain influence of the main chain secondary structure is particularly pronounced for peptoids, but is a property that is prevalent in all listed examples at least to some extent. The ability of e.g. α-amino acids to fold not only into helices but also loops and sheets shapes the diversity of protein folds in nature. Predicting which primary sequence adopts a given protein fold has been one of the most persistent challenges in structural biology and has culminated in two Nobel prizes of chemistry spanning almost 50 years for Christian Anfinsen for defining the initial sequence-structure paradigm and Demis Hassabis, John Jumper and David Baker for providing protein structure prediction.[156] The same prediction problem also exists in foldamer science but can be circumvented by increasing the rigidity of the foldamer backbone and in the process create molecules whose three-dimensional shape depends less on their side chains. This can be achieved by minimizing local conformational variety, by both restricting rotations around main chain bonds as well as creating energetically beneficial backbone contacts such as π-π stacking.[157] Aromatic oligoamide foldamers (AOF) combine these two features, especially when combined with heterocyclic or heteroatom-substituted building blocks. The partial charge repulsions of heteroatoms and their involvement in H-bonding networks further contributes to the energy non-degeneracy of their amides *syn*- and *anti* conformations (Fig. 15a). This conjugation is supported by the close-to-coplanar conformations of amide and aromatic ring.[157] Typical scaffolds are built from substituted benzenes like anisoles, pyridines, quinolines or diaza-anthracenes.[158-164] Among these examples, the aromatic δ-amino acid 8-aminoquinoline-2-carboxylic acid (Q) forms defined helices due to the angle of ca. 60° between the quinoline C2-carbonyl carbon and C8-amino-nitrogen bonds.[163]

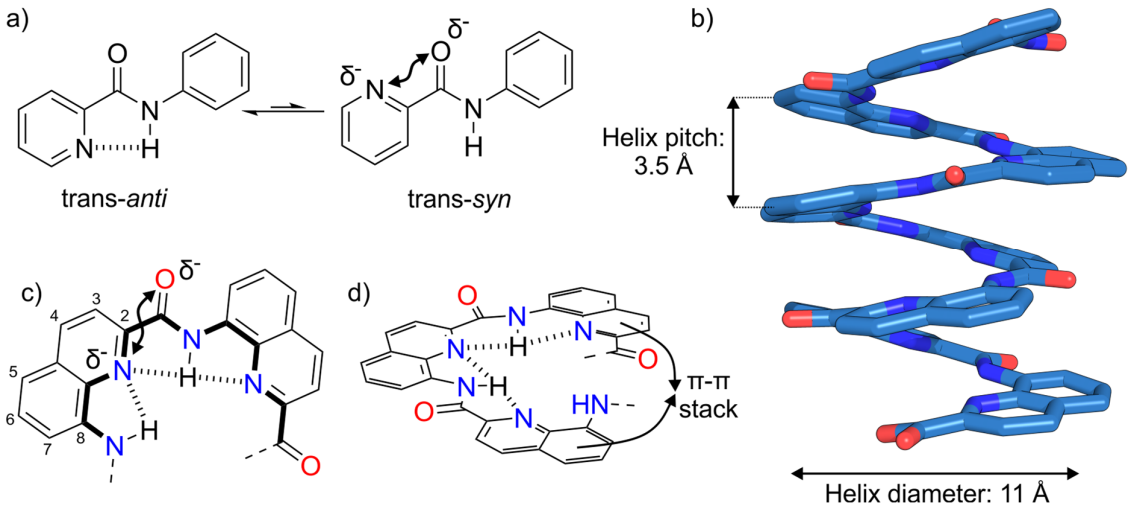


Figure 15 – Oligoamides and folding behavior of Q sequences a) Secondary aromatic amide stabilized in *anti* over *syn* conformation. b) Crystal structure of $(Q)_8$ with a nitro group at the N-terminus. Atoms are colored in light blue, dark blue and red for carbon, nitrogen and oxygen respectively. Hydrogen atoms and side chains are omitted for clarity.[163] c) Hydrogen bonding network of two adjacent Q-monomers and the repulsion of endocyclic nitrogen atoms as well as amide-oxygen.[162] d) 3D H-bonding network and π-π stacking of *i*, *i+2* quinoline residues.

This encodes for a helix with 2.5 monomers per turn, with the vertical rise per turn of 3.5 Å, the thickness of one aromatic ring, and a helix diameter of ca. 11 Å, measured between the outer rims of the quinoline aromatic rings.[163] The backbone is locked by the quinoline aromatic plane as well as rotationally fixed by the amide in trans-*anti* and thus stays helical for oligomers as small as four units.[165] Their helices are particularly stable both in polar and apolar solvents, especially when compared with their aliphatic δ-peptide analogues. They are highly resistant to thermal unfolding, show no denaturation up to 120 °C in DMSO, are resistant to proteolytic degradation and their *M*- to *P*-helix interconversion is kinetically locked in water for sequences of octamers or longer.[166-169] These properties make Q_n-helices an excellent scaffold for mimicking secondary structures of biopolymers[170] and thus have inspired the development of peptide-like solid-phase synthesis (SPS) protocols.[168, 171] To facilitate deeper side chain diversity than α-peptides, the synthesis of necessary monomers functionalized with biotic and abiotic side chains both at positions 4, 5 and 6 of the quinoline ring for rapid automated oligomer production was pursued.[170, 172, 173]

3.5 From Aromatic Oligoamide Foldamers to DNA Mimic Foldamers (DMFOs)

Even though natural DNA lookalikes, such as DNA mimic proteins, have been described and their *de novo* design has been proposed, the concept of functional DNA mimicry has not received much attention.[174-176] Potentially, this is because nucleic acids themselves provide an excellent platform for tailoring applications to interfere with PNIs. However, foldamers with a predictable shape may offer alternatives, or even outperform nucleic acids for some applications. Among the functionalizations of Q-based AOFs are phosphonic acids, which can be anchored at position 4 of the oligoquinoline backbone (Fig 16a). A polyphosphonic AOF constructed entirely from these units forms a helix with dimensions comparable to nucleic acids, while also featuring phosphorus atoms that bear negative charges.[177] However, as $(Q)_n$-helices consist of close to five units per two turns, the array of side chains forms five columns of phosphonic acids along the helix axis (Fig 16d) as the 2.5 helix is severely overtwisted compared to DNA.

This can be corrected by alternating Q units with M units (Fig 16a), that contain an additional methylene group in between the C8-carbon and the amine of Q. $(M)_n$-helices encode for almost one helix turn per two units, which causes a close to parallel alignment of two phosphonic acid strands (Fig 16e). However, an $(MQ)_n$ alternating sequence (Fig. 16f) is predicted for 0.9 helix turns per dimeric building block and as such encodes for an angular shift of one tenth of a turn with a twist of ca. 36° – both similar to B-DNA's helical parameters. This results in vertical π-π stacking of units exclusively within the same type, e.g. M units stack with other M residues in the i+2 *and* i−2 positions.[177] As neither Q nor M building blocks contain stereo information, foldamer *P*- and *M*-helices must be energetically degenerate and thus exist as a racemic

mixture in solution. To mimic B-DNA's charge topology, the array of negatively charged residues must be *P*-helical.[178] Interestingly, the left-handed foldamer main chain displays its so-called *exo*-helices in a B-DNA-like right-handed fashion (Fig. 16h). This correlation stems from the fact that an $(MQ)_n$ dimer encodes for 0.9 turns and the angular shift of the side chain is oriented "backwards" by 0.1 helix turns with respect to the previous dimer.[178]

In $(MQ)_n$-mixed sequences, the respective distances Q..M and M..Q along the helix exterior of the charge-bearing phosphonic acid residues are not identical. This is because the main chain length for Q (δ-amino acid) differs from M (ε-amino acid) by one methylene group. As such, a minor and major groove can be defined. The resulting groove width can be tuned by varying the distance between phosphonic acids (Fig 16b). One way to achieve this is to install the phosphonic acid side chain in position 5 instead of position 4 of the quinoline ring (named Q^5). This causes a widening of the major groove and a narrowing of the minor groove (Fig 16c) and overall, better matches the grooves found in canonical B-DNA.

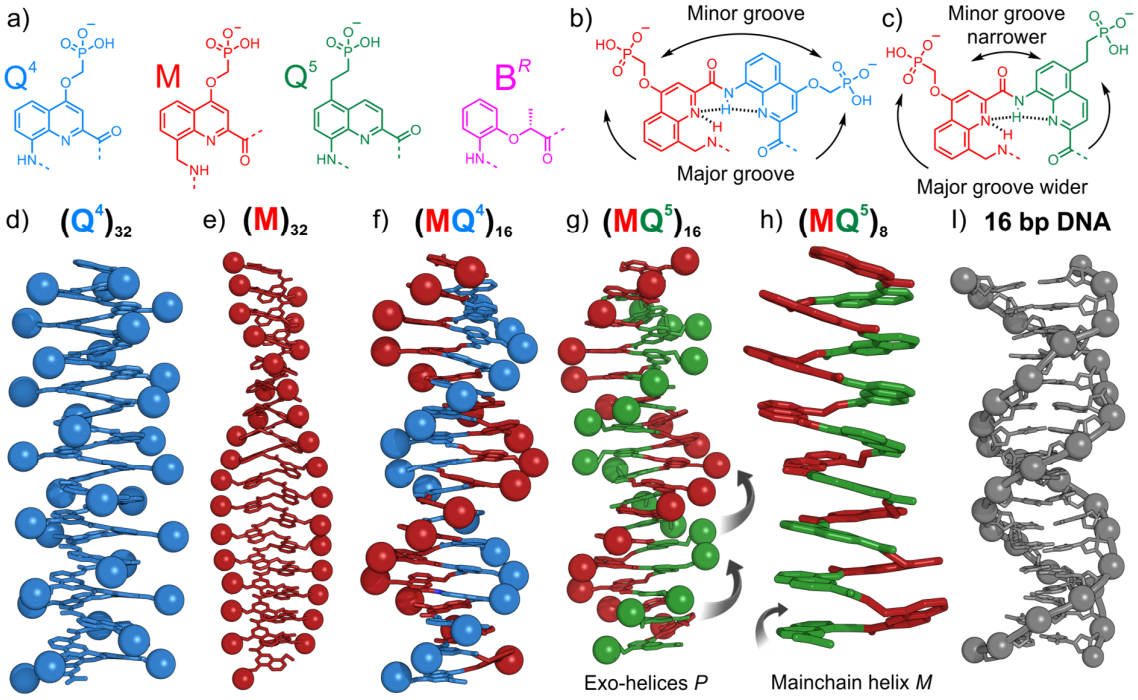


Figure 16 – Decoding DNA mimic foldamers a) Amino-acid DMFO building blocks. DMFOs repeat units and schematic depictions of DMFO grooves for MQ^4 (b) and MQ^5 (c). Molecular models of $(Q^4)_{32}$ (d) and $(M)_{32}$ (e). f) Crystal structure of $(MQ^4)_{16}$, grown from organic solvents with phosphonates protected as diethyl esters (omitted for clarity).[177] g) Molecular model of $(MQ^5)_{16}$, based on the crystal structure of $(MQ^4)_{16}$, with full positional constraints on the main chain atoms. The arrows depict the right-handed handedness of the two phosphonic acid *exo*-helices. h) Display of a DNA mimic $(MQ^5)_8$ sequence mainchain with a left-handed helix sense. For d-g), all molecules are achiral and exist as racemic mixture of *P*- and *M*- enantiomeric conformers. In all figures, the *M* helix of the DMFO is displayed. i) Crystal structure of a 16 bp double-stranded DNA in B-form.[179]

Q^4 and Q^5 also differ in the way the side chain is connected to the main chain, primarily due to synthetic accessibility. Q^4 linkage by an ether ($-OCH_2-$) causes the side chain to be more rigid than Q^5 linkage by ethylene ($-CH_2CH_2-$), whereas both are more flexible compared to DNA's

phosphates, as these are confined within the phosphodiester backbone.[177]

There are more notable differences between DNA and its mimic analogue such as the possibility for phosphonic acids to be doubly deprotonated, which causes them to have higher average charge density than nucleic acids under physiological conditions. Besides their shape and charge distribution, DMFOs based on AOFs differ from DNA in several aspects, which is a necessity if the goal is to outperform them in some properties.

Since their inception and first publication in 2018, these molecules were further developed chemically and their effects in biological systems were studied. Their initial synthesis was done by fragment couplings in solution, but has since shifted to SPS approaches, both for polyphosphonic as well as polycarboxylate $(MQ)_n$ sequences.[178, 180] This may pave the way towards sequence diversity as has been explored in the interactions of Q_n based AOFs with proteins. For example, this allowed for the incorporation of a chiral benzene-based building block B^R (Fig. 16a) within the sequence that induces a strong bias on helix handedness towards *M*-helicity.[178, 181-184] Other chemical advancements include the synthesis of DMFOs with C_2-symmetry by linking sequences by their respective C- and N-termini with C_2-symmetrical diamines or diacids and thus, they inherit some properties of DNA palindromes. C_2-symmetry can also be achieved by an artificial dimerization domain, causing stacking of foldamer C-terminal cross-sections, which conceptually can be compared to DNA sticky ends (Fig 17a).[178]

With these advancements, it was possible to co-crystallize a C_2-symmetrical DMFO with the DBP Sac7d in a 1:2 complex (Fig 17b).[178] Sac7d is required to stabilize its organism's genomic DNA and prevent thermal denaturation under the extreme growth temperatures of its host archaeon, which is achieved by compaction of DNA via sharp kinks introduced by the protein.[66, 185]

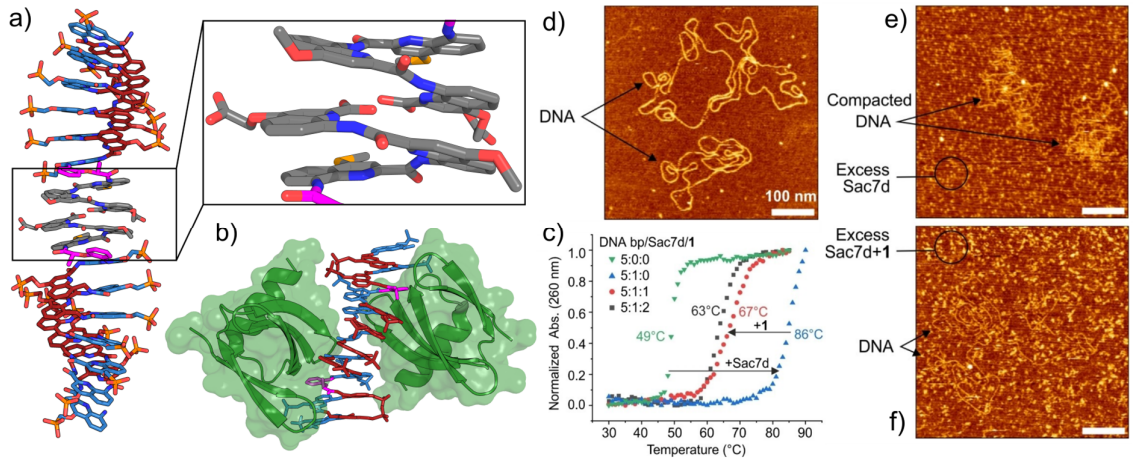


Figure 17 – Further DNA mimic development, structural and functional investigations a) Crystal structure of a polyphosphonic acid DMFO from water, bearing a C-terminal extension that functions as a DNA sticky ends mimic.[178] b) Crystal structure of a C2-symmetrical DMFO in complex with a DBP.[186] c) UV-VIS thermal denaturation assay shows that Sac7d initially stabilizes DNA melting, then loses this function partially upon binding to foldamer molecules. AFM assay shows the initial compaction of free DNA (d) into a collapsed state by Sac7d (e), and the loss of compaction upon addition of DMFO (f). Figs. c-f have been adopted from the following publication.[186]

The interference of DMFOs into this process was shown by both a thermal denaturation assay and an atomic force microscopy (AFM) assay (Figs. 17d-f). Further, a circular dichroism (CD) spectroscopy assay revealed that the DMFO outcompetes DNA in binding to the protein by one order of magnitude.[186]

DMFOs have further been explored as potent inhibitors of therapeutically DNA-binding enzymes, topoisomerase 1 (Top1) and human immunodeficiency virus integrase (HIV-IN). Similarly to Sac7d, DMFOs outcompeted DNA in binding and could inhibit the relaxation of supercoiled DNA (Top1, Fig 18b) or prevent DNA integration (HIV-In, Fig. 18d) in presence of the respective enzymes natural substrates.[177] From these initial studies, it was deduced that the MoA of DMFOs is based on competitive inhibition (Fig. 18a), opposed to typical MoAs of small molecule inhibitors that tend to inhibit the DNA-enzyme complex at one specific intermediate step, such as the strand transfer of viral cDNA into host DNA.

This allows for potential therapeutics based on DMFOs to be used as a combinatory treatment with currently existing drugs as proposed for ASOs together with immune checkpoint blockers.[187] These studies have been extended by $(MQ)_n$ sequences bearing carboxylates with different side chain-linkages on the foldamers surface. Remarkably, some carboxylate-functionalized DMFOs displayed inhibitory potency on par with or exceeding that of their phosphonate counterparts. Others lacked activity and some even stimulated enzymatic function underlining the critical impact of structural changes and charge distribution of indirect readout in PNIs. In cell studies, DMFOs showed no cytotoxicity in the absence of transfection reagents which was attributed to poor cell penetration, yet Top1 inhibition was still detected.[177] However, in the presence of a cationic amphiphile to transfect the cell, in both human embryonic kidney cells (HEK293) and human colon cancer cells (HCT116), cytotoxicity was greatly enhanced (Fig 18f). Besides cell delivery by transfection reagents, a DMFO was conjugated site-specifically as a payload to trastuzumab to form an antibody-drug conjugate (ADC).[188]

This strategy was developed to enable human epidermal growth factor receptor 2 (HER2)-dependent uptake. The resulting conjugate exhibited selective uptake and growth inhibition in human HER2-overexpressing breast and ovarian cancer cell lines, showing the highest efficacy in SK-BR-3 cell lines (Fig. 18g), slightly weaker effects in SK-OV-3 cell lines (ovarian cancer), while showing minimal effects in HER2-low or -negative lines (T-47D, MCF-7). This establishes a selective ADC-based strategy for intracellular delivery of DMFOs.[188]

In another study, the ability of DMFOs to influence the chromatin-bound proteome in a controlled *in vitro* system using a *Drosophila* embryo extract (DREX) chromatin assembly assay was investigated (Fig. 19a).[189] Chromatin assembly in presence of foldamer, and subsequent proteomic analysis revealed that foldamer treatment causes distinct changes in chromatin composition and the affected proteins were assigned to clusters based on their effect.

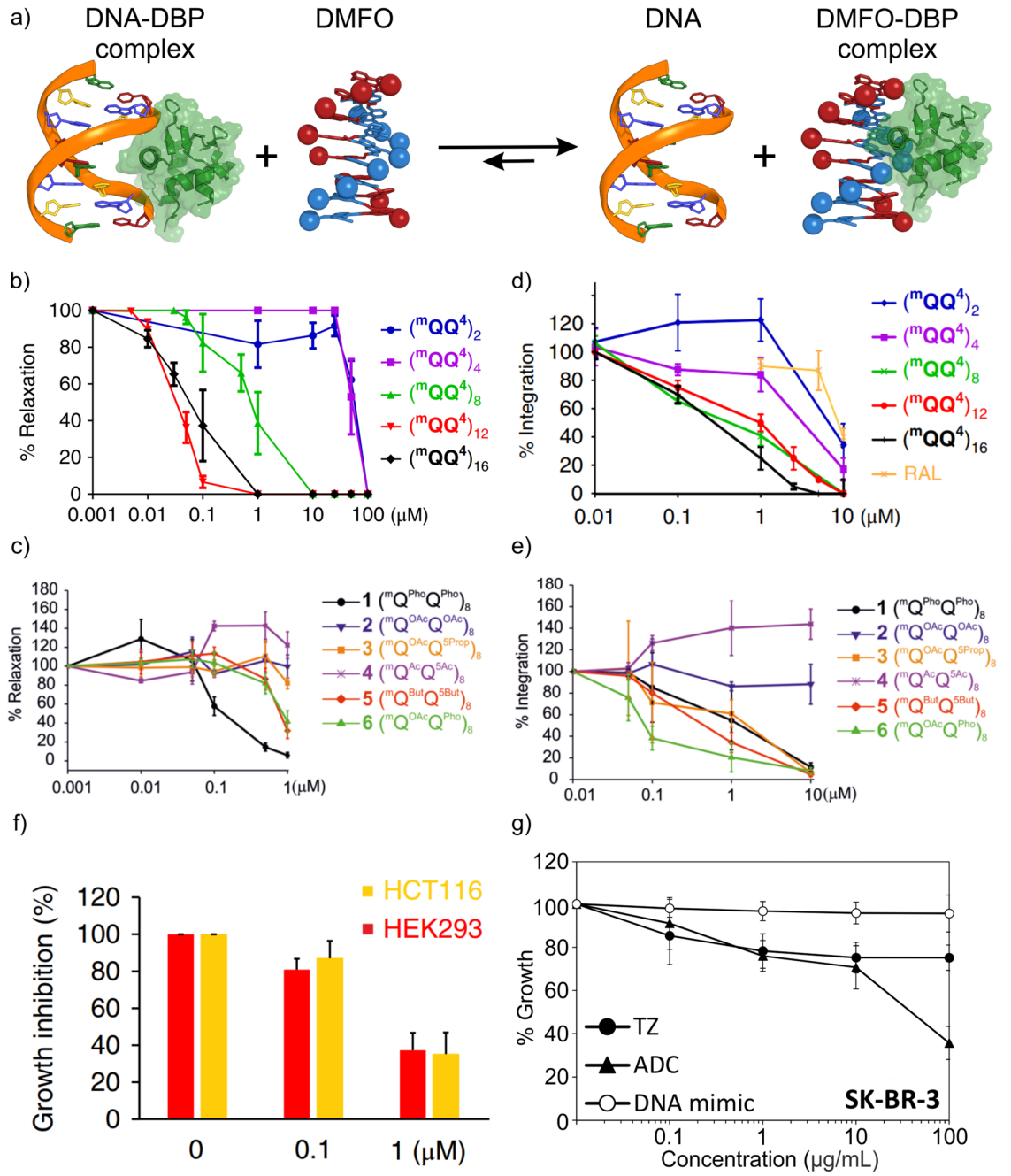


Figure 18 –DNA mimic competitive enzyme inhibition and investigations in biological systems a) Schematic display of DMFOs as competitive inhibitors. Inhibition of Top1 and HIV-1-In by DMFOs of different length (b) for Top1, (d) for HIV-1-In) and different polycarboxylate hexadecamers (c) for Top1, (e) for HIV-1-In) in presence of their natural substrates.[177, 180] f) Effect of polyphosphonic $(MQ^4)_8$ on HEK293 and HCT116 cells in the presence of PULSin as transfection reagent after 4h exposure time. g) Effect on cell growth of ADC, naked antibody (TZ) and naked DMFOs on SK-BR-3 cell lines.[188] Figures b-g) are adapted from the publications cited within this caption.

In parallel, a pulldown assay was used on DREX to identify proteins that bind directly to the DMFO (Fig 19b). Proteins that appeared in both sample sets (Fig. 19c) were deemed of particular interest such as proteins required for replication initiation, including most of the

subunits of the origin recognition complex (ORC). This clustering of affected proteins highlighted that the DMFOs act selectively on ORC and its functional partners, thus forming the most significantly disrupted group. Cell-based experiments confirmed that DMFOs hinder ORC loading *in vivo*, leading to impaired DNA replication and arrest cell cycle progression in the S-phase (Fig 19d).

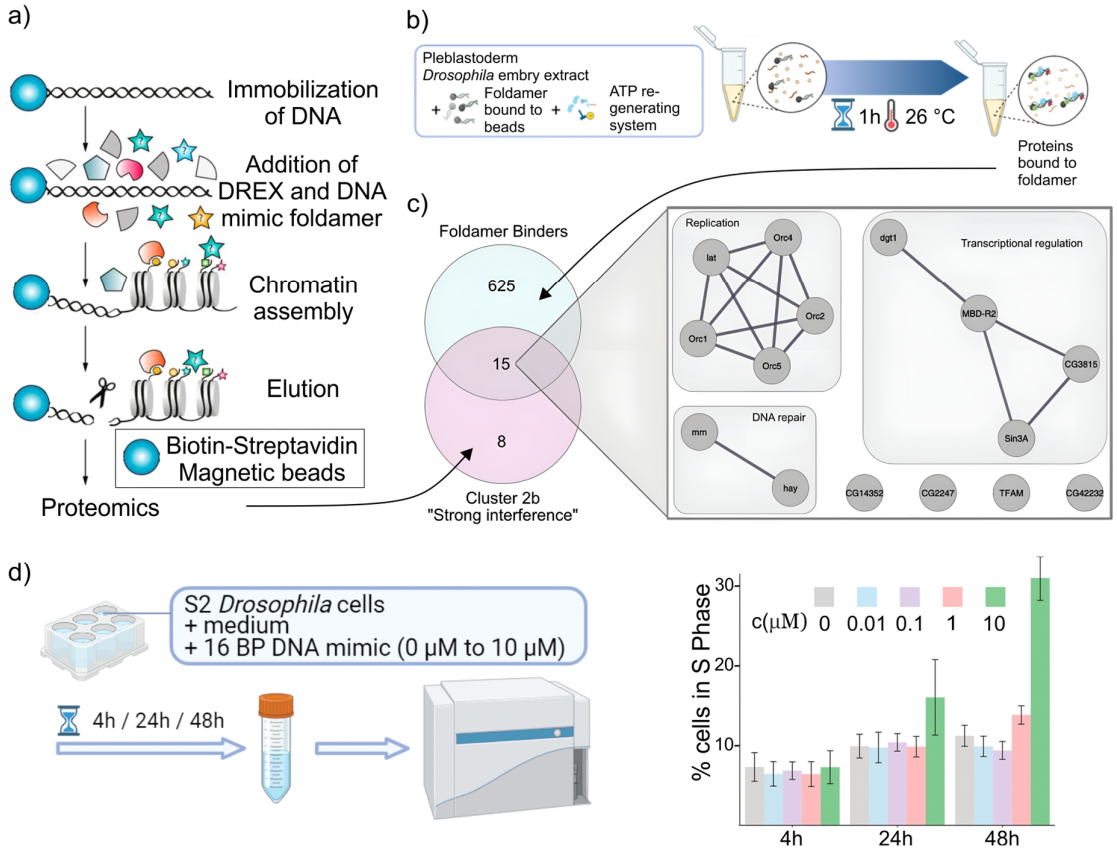


Figure 19 –DNA mimic foldamers interference in chromatin composition and cell cycle progression
a) Chromatin assembly assay with DREX in the presence of DMFO. b) Pulldown with biotinylated DMFO against DREX to identify foldamer binders. c) Comparison of proteins binding foldamer to proteins interfered in binding of chromatin fiber in presence of DMFO. Venn diagram depicting the overlap of the two sets. Cytoscape network plot after string analysis for proteins shown to be binding to DMFOs and interfered with strongly by DMFOs in chromatin assembly in DREX. d) Schematic diagram of experimental flow of cell cycle analysis experiment and bar graph representing percentage of cells in S-phase after DMFO treatment with indicated concentration after the indicated time. Figures are adopted from the following publication.[189]

These studies demonstrate that DMFOs can be recognized by DBPs, sometimes much better than DNA as their natural substrate and are potent inhibitors of replication origin licensing and other enzymes of therapeutic interest. We hypothesize that they offer a versatile and designable platform that may achieve similar pharmacological potency than e.g. nucleic acid drugs or small molecule inhibitors of PNIs. Their future development may have the potential to elevate them to the next generation of drugs.

3.6 References

1. O. T. Avery, C. M. MacLeod and M. McCarty, *J. Exp. Med.*, 1944, **79**, 137.
2. A. D. Hershey and M. Chase, *J. Gen. Physiol.*, 1952, **36**, 39-56.
3. J. W. Schopf, *Science*, 1993, **260**, 640-646.
4. F. H. C. Crick, *Nature (London)*, 1970, **227**, 561.
5. J. F. Miescher, *Ueber die chemische Zusammensetzung der Eiterzellen*, 1871.
6. J. D. Watson and F. H. C. Crick, *Nature (London, U. K.)*, 1953, **171**, 737.
7. R. E. Franklin and R. G. Gosling, *Nature (London, U. K.)*, 1953, **172**, 156.
8. R. E. Franklin and R. G. Gosling, *Nature (London, U. K.)*, 1953, **171**, 740.
9. E. A. Ashley, *Nat. Rev. Genet.*, 2016, **17**, 507-522.
10. M. Meselson and F. W. Stahl, *Proc. Natl. Acad. Sci. U. S. A.*, 1958, **44**, 671.
11. M. H. F. Wilkins, A. R. Stokes and H. R. Wilson, *Nature (London, U. K.)*, 1953, **171**, 738.
12. M. W. Nirenberg and J. H. Matthaei, *Proc. Natl. Acad. Sci. U. S. A.*, 1961, **47**, 1588.
13. F. Jacob and J. Monod, *J. Mol. Biol.*, 1961, **3**, 318.
14. M. Nirenberg and P. Leder, *Science (Washington, DC, U. S.)*, 1964, **145**, 1399.
15. H. G. Khorana, H. Buechi, H. P. Ghosh, N. K. Gupta, T. M. Jacob, H. Koessel, R. Morgan, S. A. Narang, E. Ohtsuka and R. D. Wells, *Cold Spring Harbor Symp. Quant. Biol.*, 1967, **31**, 39-49.
16. S. N. Cohen, A. C. Y. Chang, H. W. Boyer and R. B. Helling, *Proc. Natl. Acad. Sci. U. S. A.*, 1973, **70**, 3240.
17. A. J. Jeffreys, V. Wilson and S. L. Thein, *Nature (London)*, 1985, **316**, 76.
18. R. K. Saiki, S. Scharf, F. Faloona, K. B. Mullis, G. T. Horn, H. A. Erlich and N. Arnheim, *Science (Washington, D. C., 1883-)*, 1985, **230**, 1350.
19. R. K. Saiki, D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis and H. A. Erlich, *Science*, 1988, **239**, 487-491.
20. R. R. Breaker and G. F. Joyce, *Chem. Biol.*, 1994, **1**, 223.
21. I. Wilmut, A. E. Schnieke, J. McWhir, A. J. Kind and K. H. S. Campbell, *Nature (London)*, 1997, **385**, 810-813.
22. E. S. Lander, L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh*, et al.*, *Nature (London, U. K.)*, 2001, **409**, 860-921.
23. F. S. Collins, M. Morgan and A. Patrinos, *Science (Washington, DC, U. S.)*, 2003, **300**, 286-290.
24. M. Margulies, M. Egholm, W. E. Altman, S. Attiya, J. S. Bader, L. A. Bemben, J. Berka, M. S. Braverman, Y.-J. Chen, Z. Chen*, et al.*, *Nature (London, U. K.)*, 2005, **437**, 376-380.
25. D. R. Bentley, S. Balasubramanian, H. P. Swerdlow, G. P. Smith, J. Milton, C. G. Brown, K. P. Hall, D. J. Evers, C. L. Barnes, H. R. Bignell*, et al.*, *Nature (London, U. K.)*, 2008, **456**, 53-59.
26. J. Eid, A. Fehr, J. Gray, K. Luong, J. Lyle, G. Otto, P. Peluso, D. Rank, P. Baybayan, B. Bettman*, et al.*, *Science (Washington, DC, U. S.)*, 2009, **323**, 133-138.
27. S. B. Ng, K. J. Buckingham, C. Lee, A. W. Bigham, H. K. Tabor, K. M. Dent, C. D. Huff, P. T. Shannon, E. W. Jabs, D. A. Nickerson*, et al.*, *Nat. Genet.*, 2010, **42**, 30-35.
28. R. D. Palmiter, R. L. Brinster, R. E. Hammer, M. E. Trumbauer, M. G. Rosenfeld, N. C. Birnberg and R. M. Evans, *Nature (London)*, 1982, **300**, 611.
29. L. Hood and L. Rowen, *Genome Med*, 2013, **5**, 79.
30. R. A. Gibbs, *Nat. Rev. Genet.*, 2020, **21**, 575-576.
31. E. Birney, J. A. Stamatoyannopoulos, A. Dutta, R. Guigo, T. R. Gingeras, E. H. Margulies, Z. Weng, M. Snyder, E. T. Dermitzakis, J. A. Stamatoyannopoulos*, et al.*, *Nature (London, U. K.)*, 2007, **447**, 799-816.
32. I. Dunham, A. Kundaje, S. F. Aldred, P. J. Collins, C. A. Davis, F. Doyle, C. B. Epstein, S. Frietze, J. Harrow, R. Kaul*, et al.*, *Nature (London, U. K.)*, 2012, **489**, 57-74.
33. H. R. Drew, R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura and R. E. Dickerson, *Proc. Natl. Acad. Sci. U. S. A.*, 1981, **78**, 2179.
34. V. Tereshko, G. Minasov and M. Egli, *J. Am. Chem. Soc.*, 1999, **121**, 470-471.

35. J. Liu and J. A. Subirana, *J. Biol. Chem.*, 1999, **274**, 24749-24752.
36. L. Lercher, M. A. McDonough, A. H. El-Sagheer, A. Thalhammer, S. Kriaucionis, T. Brown and C. J. Schofield, *Chem. Commun. (Cambridge, U. K.)*, 2014, **50**, 1794-1796.
37. C. L. Kielkopf, S. Ding, P. Kuhn and D. C. Rees, *J. Mol. Biol.*, 2000, **296**, 787-801.
38. V. Tereshko, C. J. Wilds, G. Minasov, T. P. Prakash, M. A. Maier, A. Howard, Z. Wawrzak, M. Manoharan and M. Egli, *Nucleic Acids Res.*, 2001, **29**, 1208-1215.
39. P. K. Mandal, G. W. Collie, B. Kauffmann and I. Huc, *Angew. Chem., Int. Ed.*, 2014, **53**, 14424-14427.
40. M. Egli, L. D. Williams, Q. Gao and A. Rich, *Biochemistry*, 1991, **30**, 11388.
41. W. Saenger, *Principles of Nucleic Acid Structure*, Springer New York, 2013.
42. P. Belmont, J. F. Constant and M. Demeunynck, *Chem. Soc. Rev.*, 2001, **30**, 70-81.
43. Y. G. Gao, H. Robinson and A. H. Wang, *Eur J Biochem*, 1999, **261**, 413-420.
44. A. Rich and S. Zhang, *Nat. Rev. Genet.*, 2003, **4**, 566-572.
45. M. A. Fuertes, V. Cepeda, C. Alonso and J. M. Perez, *Chem. Rev. (Washington, DC, U. S.)*, 2006, **106**, 2045-2064.
46. H. Lonnberg, *Chemistry of Nucleic Acids*, Walter de Gruyter GmbH, 2020.
47. P. Yakovchuk, E. Protozanova and M. D. Frank-Kamenetskii, *Nucleic Acids Res.*, 2006, **34**, 564-574.
48. A. D. Broom, M. P. Schweizer and P. O. P. Ts'o, *J. Am. Chem. Soc.*, 1967, **89**, 3612.
49. W. K. Olson, M. Bansal, S. K. Burley, R. E. Dickerson, M. Gerstein, S. C. Harvey, U. Heinemann, X.-J. Lu, S. Neidle, Z. Shakked*, et al.*, *J. Mol. Biol.*, 2001, **313**, 229-237.
50. F. H. C. Crick, *J. Mol. Biol.*, 1966, **19**, 548.
51. P. F. Agris, F. A. P. Vendeix and W. D. Graham, *J. Mol. Biol.*, 2007, **366**, 1-13.
52. M. F. Gellert, M. N. Lipsett and D. H. Davies, *Proc. Natl. Acad. Sci. U. S. A.*, 1962, **48**, 2013.
53. D. Sen and W. Gilbert, *Nature (London)*, 1988, **334**, 364.
54. Y. Wang and D. J. Patel, *J. Mol. Biol.*, 1993, **234**, 1171.
55. G. N. Parkinson, M. P. H. Lee and S. Neidle, *Nature (London, U. K.)*, 2002, **417**, 876-880.
56. G. Biffi, D. Tannahill, J. McCafferty and S. Balasubramanian, *Nat. Chem.*, 2013, **5**, 182-186.
57. J. Tu, M. Duan, W. Liu, N. Lu, Y. Zhou, X. Sun and Z. Lu, *Nat. Commun.*, 2021, **12**, 6014.
58. J. Spiegel, S. Adhikari and S. Balasubramanian, *Trends Chem.*, 2020, **2**, 123-136.
59. R. Hansel-Hertsch, D. Beraldi, S. V. Lensing, G. Marsico, K. Zyner, A. Parry, M. Di Antonio, J. Pike, H. Kimura, M. Narita*, et al.*, *Nat. Genet.*, 2016, **48**, 1267-1272.
60. F. Kouzine, D. Wojtowicz, L. Baranello, A. Yamane, S. Nelson, W. Resch, K.-R. Kieffer-Kwon, C. J. Benham, R. Casellas, T. M. Przytycka*, et al.*, *Cell Syst.*, 2017, **4**, 344-356.e347.
61. A. G. W. Leslie, S. Arnott, R. Chandrasekaran and R. L. Ratliff, *J. Mol. Biol.*, 1980, **143**, 49-72.
62. J. Liquier, M. C. Gadenne, E. Taillandier, N. Defer, F. Favatier and J. Kruh, *Nucleic Acids Res.*, 1979, **6**, 1479.
63. *Protein-Nucleic Acid Interactions: Structural Biology*, The Royal Society of Chemistry, 2008.
64. A. Afek, J. L. Schipper, J. Horton, R. Gordan and D. B. Lukatsky, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 17140-17145.
65. R. Rohs, S. M. West, A. Sosinsky, P. Liu, R. S. Mann and B. Honig, *Nature (London, U. K.)*, 2009, **461**, 1248-1253.
66. H. Robinson, Y.-G. Gao, B. S. McCrary, S. P. Edmondson, J. W. Shriver and A. H. J. Wang, *Nature (London)*, 1998, **392**, 202-205.
67. F. Spyrakis, P. Cozzini, C. Bertoli, A. Marabotti, G. E. Kellogg and A. Mozzarelli, *BMC Struct. Biol.*, 2007, **7**.
68. P. A. Rice, S.-w. Yang, K. Mizuuchi and H. A. Nash, *Cell (Cambridge, Mass.)*, 1996, **87**, 1295-1306.
69. S. Su, Y.-G. Gao, H. Robinson, Y.-C. Liaw, S. P. Edmondson, J. W. Shriver and A. H. J. Wang, *J. Mol. Biol.*, 2000, **303**, 395-403.

70. K. K. Swinger, K. M. Lemberg, Y. Zhang and P. A. Rice, *EMBO J.*, 2003, **22**, 3749-3760.
71. K. Luger, A. W. Mader, R. K. Richmond, D. F. Sargent and T. J. Richmond, *Nature (London)*, 1997, **389**, 251-260.
72. Z. Otwinowski, R. W. Schevitz, R. G. Zhang, C. L. Lawson, A. Joachimiak, R. Q. Marmorstein, B. F. Luisi and P. B. Sigler, *Nature (London)*, 1988, **335**, 321.
73. R. Rohs, X. Jin, S. M. West, R. Joshi, B. Honig and R. S. Mann, *Annu. Rev. Biochem.*, 2010, **79**, 233-269.
74. S. C. Harrison and A. K. Aggarwal, *Annu. Rev. Biochem.*, 1990, **59**, 933.
75. A. K. Aggarwal, D. W. Rodgers, M. Drottar, M. Ptashne and S. C. Harrison, *Science (Washington, D. C., 1883-)*, 1988, **242**, 899-907.
76. A. Mondragon and S. C. Harrison, *J. Mol. Biol.*, 1991, **219**, 321.
77. M. Elrod-Erickson, T. E. Benson and C. O. Pabo, *Structure (London)*, 1998, **6**, 451-464.
78. D. B. Nikolov, H. Chen, E. D. Halay, A. Hoffmann, R. G. Roeder and S. K. Burley, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 4862-4867.
79. W. S. Somers and S. E. V. Phillips, *Nature (London)*, 1992, **359**, 387.
80. W. S. Somers, J. B. Rafferty, K. Phillips, S. Strathdee, Y. Y. He, T. McNally, I. Manfield, O. Navratil and I. G. Old, *Ann. N. Y. Acad. Sci.*, 1994, **726**, 105.
81. D. Leprince, A. Gegonne, J. Coll, C. De Taisne, A. Schneeberger, C. Lagrou and D. Stehelin, *Nature (London)*, 1983, **306**, 395.
82. A. Gutierrez-Hartmann, D. L. Duval and A. P. Bradford, *Trends Endocrinol. Metab.*, 2007, **18**, 150-158.
83. S. K. Nair and S. K. Burley, *Cell (Cambridge, MA, U. S.)*, 2003, **112**, 193-205.
84. Y. Cho, S. Gorina, P. D. Jeffrey and N. P. Pavletich, *Science (Washington, D. C.)*, 1994, **265**, 346.
85. K. Ida, S. Kobayashi, T. Taki, R. Hanada, F. Bessho, S. Yamamori, T. Sugimoto, M. Ohki and Y. Hayashi, *Int. J. Cancer*, 1995, **63**, 500.
86. S. A. Tomlins, D. R. Rhodes, S. Perner, S. M. Dhanasekaran, R. Mehra, X.-W. Sun, S. Varambally, X. Cao, J. Tchinda, R. Kuefer*, et al.*, *Science (Washington, DC, U. S.)*, 2005, **310**, 644-648.
87. P. Peeters, S. D. Raynaud, J. Cools, I. Wlodarska, J. Grosgeorge, P. Philip, F. Monpoux, L. Van Rompaey, M. Baens, H. Van den Berghe*, et al.*, *Blood*, 1997, **90**, 2535-2540.
88. P. C. Hollenhorst, L. P. McIntosh and B. J. Graves, *Annu. Rev. Biochem.*, 2011, **80**, 437-471.
89. Y. Mo, W. Ho, K. Johnston and R. Marmorstein, *J. Mol. Biol.*, 2001, **314**, 495-506.
90. R. Kodandapani, F. Pio, C.-Z. Ni, G. Piccialli, M. Klemsz, S. McKercher, R. A. Maki and K. R. Ely, *Nature (London)*, 1996, **380**, 456.
91. G.-H. Wei, G. Badis, M. F. Berger, T. Kivioja, K. Palin, M. Enge, M. Bonke, A. Jolma, M. Varjosalo, A. R. Gehrke*, et al.*, *EMBO J.*, 2010, **29**, 2147-2160.
92. J. R. Terrell, S. J. Taylor, A. L. Schneider, Y. Lu, T. N. Vernon, S. Xhani, R. H. Gumpper, M. Luo, W. D. Wilson, U. Steidl*, et al.*, *Cell Rep.*, 2023, **42**, 112671.
93. S. Ahmad, H. Kono, M. J. Arauzo-Bravo and A. Sarai, *Nucleic Acids Res.*, 2006, **34**, W124.
94. P. Wu, *Nat. Rev. Chem.*, 2020, **4**, 441-458.
95. C. L. Kielkopf, R. E. Bremer, S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird, P. B. Dervan and D. C. Rees, *J. Mol. Biol.*, 2000, **295**, 557-567.
96. G. Santamaria Nunez, C. M. G. Robles, C. Giraudon, J. F. Martinez-Leal, E. Compe, F. Coin, P. Aviles, C. M. Galmarini and J.-M. Egly, *Mol. Cancer Ther.*, 2016, **15**, 2399-2412.
97. H. Robinson, Y.-G. Gao, X.-l. Yang, R. Sanishvili, A. Joachimiak and A. H. J. Wang, *Biochemistry*, 2001, **40**, 5587-5592.
98. S. Kamitori and F. Takusagawa, *J. Am. Chem. Soc.*, 1994, **116**, 4154.
99. J. A. Cuesta-Seijo and G. M. Sheldrick, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2005, **D61**, 442-448.
100. S. Hare, S. S. Gupta, E. Valkov, A. Engelman and P. Cherepanov, *Nature (London, U. K.)*, 2010, **464**, 232-236.
101. B. L. Staker, M. D. Feese, M. Cushman, Y. Pommier, D. Zembower, L. Stewart and A. B. Burgin, *J. Med. Chem.*, 2005, **48**, 2336-2345.

102. C. I. E. Smith and R. Zain, *Annu. Rev. Pharmacol. Toxicol.*, 2019, **59**, 605-630.
103. S. T. Crooke, *Antisense Nucleic Acid Drug Dev.*, 1998, **8**, vii.
104. S. T. Crooke and R. S. Geary, *Br. J. Clin. Pharmacol.*, 2013, **76**, 269-276.
105. M. Scoto, R. Finkel, E. Mercuri and F. Muntoni, *Lancet Child Adolesc Health*, 2018, **2**, 600-609.
106. M. D. Benson, M. Waddington-Cruz, J. L. Berk, M. Polydefkis, P. J. Dyck, A. K. Wang, V. Plante-Bordeneuve, F. A. Barroso, G. Merlini, L. Obici*, et al.*, *N. Engl. J. Med.*, 2018, **379**, 22-31.
107. F. Muntoni, D. Frank, V. Sardone, J. Morgan, F. Schnell, J. Charleston, C. Desjardins, R. Phadke, C. Sewry, L. Popplewell*, et al.*, *Neurology*, 2018, **90**.
108. S. Naeem, J. Zhang, Y. Zhang and Y. Wang, *Mol. Ther.--Nucleic Acids*, 2025, **36**, 102440.
109. D. Adams, A. Gonzalez-Duarte, W. D. O'Riordan, C. C. Yang, M. Ueda, A. V. Kristen, I. Tournev, H. H. Schmidt, T. Coelho, J. L. Berk*, et al.*, *N. Engl. J. Med.*, 2018, **379**, 11-21.
110. M. Balwani, E. Sardh, P. Ventura, P. A. Peiro, D. C. Rees, U. Stolzel, D. M. Bissell, H. L. Bonkovsky, J. Windyga, K. E. Anderson*, et al.*, *N. Engl. J. Med.*, 2020, **382**, 2289-2301.
111. T. P. Prakash and B. Bhat, *Curr. Top. Med. Chem. (Sharjah, United Arab Emirates)*, 2007, **7**, 641-649.
112. J. K. Nair, J. L. S. Willoughby, A. Chan, K. Charisse, R. M. Alam, Q. Wang, M. Hoekstra, P. Kandasamy, A. V. Kel'in, S. Milstein*, et al.*, *J. Am. Chem. Soc.*, 2014, **136**, 16958-16961.
113. C. Li, Z. Zhou, C. Ren, Y. Deng, F. Peng, Q. Wang, H. Zhang and Y. Jiang, *Front. Pharmacol.*, 2022, **13**, 1007723.
114. P. A. Beal and P. B. Dervan, *Science (Washington, D. C., 1883-)*, 1991, **251**, 1360.
115. J. L. Asensio, T. Brown and A. N. Lane, *Structure (London)*, 1999, **7**, 1-11.
116. B. G. Su and M. J. Henley, *Front. Mol. Biosci.*, 2021, **8**, 795743.
117. C. Yan and P. J. Higgins, *Biochim. Biophys. Acta, Rev. Cancer*, 2013, **1835**, 76-85.
118. Z. Tao and X. Wu, *Methods Mol. Biol. (N. Y., NY, U. S.)*, 2023, **2594**, 107-131.
119. T. E. Ellenberger, C. J. Brandl, K. Struhl and S. C. Harrison, *Cell (Cambridge, Mass.)*, 1992, **71**, 1223.
120. J. Liu, N. B. Perumal, C. J. Oldfield, E. W. Su, V. N. Uversky and A. K. Dunker, *Biochemistry*, 2006, **45**, 6873-6888.
121. J. Zhou, C. J. Oldfield, W. Yan, B. Shen and A. K. Dunker, *Protein Sci.*, 2019, **28**, 1652-1663.
122. N. B. Phillips, J. Racca, Y.-S. Chen, R. Singh, A. Jancso-Radek, J. T. Radek, N. P. Wickramasinghe, E. Haas and M. A. Weiss, *J. Biol. Chem.*, 2011, **286**, 36787.
123. G. Ghosh, G. Van Duyne, S. Ghosh and P. B. Sigler, *Nature (London)*, 1995, **373**, 303.
124. D.-B. Huang, D. Vu, L. A. Cassiday, J. M. Zimmerman, L. J. Maher, III and G. Ghosh, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 9268-9273.
125. N. Pavlova, V. Dyakova, M. Traykovska, D. Kaloudas and R. Penchovsky, *ChemistrySelect*, 2025, **10**, e02954.
126. A. D. Ellington and J. W. Szostak, *Nature (London)*, 1990, **346**, 818.
127. A. D. Keefe, S. Pai and A. Ellington, *Nature Reviews Drug Discovery*, 2010, **9**, 537-550.
128. E. S. Gragoudas, A. P. Adamis, E. T. Cunningham, Jr., M. Feinsod and D. R. Guyer, *N. Engl. J. Med.*, 2004, **351**, 2805-2816.
129. L. A. Cassiday and L. J. Maher, III, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 3930-3935.
130. M. Sen, S. M. Thomas, S. Kim, J. I. Yeh, R. L. Ferris, J. T. Johnson, U. Duvvuri, J. Lee, N. Sahu, S. Joyce*, et al.*, *Cancer Discovery*, 2012, **2**, 694-705.
131. M. Mehta, K. R. Paudel, S. D. Shukla, V. S. R. R. Allam, V. K. Kannaujiya, N. Panth, A. Das, V. K. Parihar, A. Chakraborty, K. M. Ali*, et al.*, *J. Controlled Release*, 2021, **337**, 629-644.
132. S. H. Gellman, *Acc. Chem. Res.*, 1998, **31**, 173-180.
133. C. M. Goodman, S. Choi, S. Shandler and W. F. DeGrado, *Nat. Chem. Biol.*, 2007, **3**, 252-262.

134. M. Pasco, C. Dolain and G. Guichard, 2017.
135. C. A. Lipinski, *Adv. Drug Delivery Rev.*, 2016, **101**, 34-41.
136. I. Huc and L. Cuccia, in *Foldamers*, 2007, DOI: https://doi.org/10.1002/9783527611478.ch1, pp. 1-33.
137. W. S. Horne and S. H. Gellman, *Acc. Chem. Res.*, 2008, **41**, 1399-1408.
138. S. Rinaldi, *Molecules*, 2020, **25**, 3276.
139. B. A. F. Le Bailly and J. Clayden, *Chem. Commun. (Cambridge, U. K.)*, 2016, **52**, 4852-4863.
140. A. R. Sanford and B. Gong, *Curr. Org. Chem.*, 2003, **7**, 1649-1659.
141. C. N. Pace and J. M. Scholtz, *Biophys. J.*, 1998, **75**, 422-427.
142. S. H. Choi, I. A. Guzei, L. C. Spencer and S. H. Gellman, *J. Am. Chem. Soc.*, 2010, **132**, 13879-13885.
143. T. A. Martinek and F. Fueloep, *Chem. Soc. Rev.*, 2012, **41**, 687-702.
144. D. Seebach, D. F. Hook and A. Glattli, *Biopolymers*, 2006, **84**, 23-37.
145. C. Dutta, P. Krishnamurthy, D. Su, S. H. Yoo, G. W. Collie, M. Pasco, J. K. Marzinek, P. J. Bond, C. Verma, A. Grelard*, et al.*, *Chem*, 2023, **9**, 2237-2254.
146. C. M. Lombardo, G. W. Collie, K. Pulka-Ziach, F. Rosu, V. Gabelica, C. D. Mackereth and G. Guichard, *J. Am. Chem. Soc.*, 2016, **138**, 10522-10530.
147. L. Cussol, L. Mauran-Ambrosino, J. Buratto, A. Y. Belorusova, M. Neuville, J. Osz, S. Fribourg, J. Fremaux, C. Dolain, S. R. Goudreau*, et al.*, *Angew. Chem., Int. Ed.*, 2021, **60**, 2296-2303.
148. C. Baldauf, R. Guenther and H.-J. Hofmann, *J. Org. Chem.*, 2004, **69**, 6214-6220.
149. Q. Lin, H. Lan, C. Ma, R. T. Stendall, K. Shankland, R. A. Musgrave, P. N. Horton, C. Baldauf, H.-J. Hofmann, C. P. Butts*, et al.*, *Angew. Chem., Int. Ed.*, 2023, **62**, e202305326.
150. M. Nagata, M. Watanabe, R. Doi, M. Uemura, N. Ochiai, W. Ichinose, K. Fujiwara, Y. Sato, T. Kameda, K. Takeuchi*, et al.*, *Org. Biomol. Chem.*, 2023, **21**, 970-980.
151. C. Jelsch, M. M. Teeter, V. Lamzin, V. Pichon-Pesme, R. H. Blessing and C. Lecomte, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 3171-3176.
152. D. S. Daniels, E. J. Petersson, J. X. Qiu and A. Schepartz, *J. Am. Chem. Soc.*, 2007, **129**, 1532-1533.
153. C. W. Wu, K. Kirshenbaum, T. J. Sanborn, J. A. Patch, K. Huang, K. A. Dill, R. N. Zuckermann and A. E. Barron, *J. Am. Chem. Soc.*, 2003, **125**, 13525-13530.
154. J. L. Kessler, G. Kang, Z. Qin, H. Kang, F. G. Whitby, T. E. Cheatham, III, C. P. Hill, Y. Li and S. M. Yu, *J. Am. Chem. Soc.*, 2021, **143**, 10910-10919.
155. D. Kalita, B. Sahariah, S. Pravo Mookerjee and B. Kanta Sarma, *Chem. - Asian J.*, 2022, **17**, e202200149.
156. R. Pearce and Y. Zhang, *J. Biol. Chem.*, 2021, **297**, 100870.
157. I. Huc, *Eur. J. Org. Chem.*, 2004, DOI: 10.1002/ejoc.200300495, 17-29.
158. I. Saraogi and A. D. Hamilton, *Chem. Soc. Rev.*, 2009, **38**, 1726-1743.
159. I. Saraogi, C. D. Incarvito and A. D. Hamilton, *Angew. Chem., Int. Ed.*, 2008, **47**, 9691-9694.
160. Q. Gan, C. Bao, B. Kauffmann, A. Grelard, J. Xiang, S. Liu, I. Huc and H. Jiang, *Angew. Chem., Int. Ed.*, 2008, **47**, 1715-1718.
161. I. Huc, V. Maurizot, H. Gornitzka and J.-M. Leger, *Chem. Commun. (Cambridge, U. K.)*, 2002, DOI: 10.1039/b111612f, 578-579.
162. H. Jiang, J.-M. Leger, C. Dolain, P. Guionneau and I. Huc, *Tetrahedron*, 2003, **59**, 8365-8374.
163. H. Jiang, J.-M. Leger and I. Huc, *J. Am. Chem. Soc.*, 2003, **125**, 3448-3449.
164. E. Berni, C. Dolain, B. Kauffmann, J.-M. Leger, C. Zhan and I. Huc, *J. Org. Chem.*, 2008, **73**, 2687-2694.
165. C. Dolain, A. Grelard, M. Laguerre, H. Jiang, V. Maurizot and I. Huc, *Chem. - Eur. J.*, 2005, **11**, 6135-6144.
166. M. Kudo, V. Maurizot, B. Kauffmann, A. Tanatani and I. Huc, *J. Am. Chem. Soc.*, 2013, **135**, 9628-9631.
167. D. Bindl, P. K. Mandal and I. Huc, *Chem. - Eur. J.*, 2022, **28**, e202200538.

168. X. Hu, S. J. Dawson, Y. Nagaoka, A. Tanatani and I. Huc, *J. Org. Chem.*, 2016, **81**, 1137-1150.
169. S. J. Dawson, A. Meszaros, L. Petho, C. Colombo, M. Csekei, A. Kotschy and I. Huc, *Eur. J. Org. Chem.*, 2014, **2014**, 4265-4275.
170. M. Zwillinger, P. S. Reddy, B. Wicher, P. K. Mandal, M. Csekei, L. Fischer, A. Kotschy and I. Huc, *Chem. - Eur. J.*, 2020, **26**, 17366-17370.
171. V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat and I. Huc, *Chem. - Eur. J.*, 2023, **29**, e202300898.
172. X. Hu, S. J. Dawson, P. K. Mandal, X. de Hatten, B. Baptiste and I. Huc, *Chem. Sci.*, 2017, **8**, 3741-3749.
173. M. Zwillinger, P. Soregi, F. Sanchez, C. Douat, M. Csekei, I. Huc and A. Kotschy, *J. Org. Chem.*, 2025, **90**, 3043-3052.
174. H.-C. Wang, C.-H. Ho, K.-C. Hsu, J.-M. Yang and A. H. J. Wang, *Biochemistry*, 2014, **53**, 2865-2874.
175. D. T. F. Dryden, *Trends Biotechnol.*, 2006, **24**, 378-382.
176. D. Yuksel, P. R. Bianco and K. Kumar, *Mol. BioSyst.*, 2016, **12**, 169-177.
177. K. Ziach, C. Chollet, V. Parissi, P. Prabhakaran, M. Marchivie, V. Corvaglia, P. P. Bose, K. Laxmi-Reddy, F. Godde, J.-M. Schmitter*, et al.*, *Nat. Chem.*, 2018, **10**, 511-518.
178. V. Corvaglia, J. Wu, D. Deepak, M. Loos and I. Huc, *Chem. - Eur. J.*, 2024, **30**, e202303650.
179. N. Narayana and M. A. Weiss, *J. Mol. Biol.*, 2009, **385**, 469-490.
180. V. Corvaglia, D. Carbajo, P. Prabhakaran, K. Ziach, P. K. Mandal, V. Dos Santos, C. Legeay, R. Vogel, V. Parissi, P. Pourquier*, et al.*, *Nucleic Acids Res.*, 2019, **47**, 5511-5521.
181. D. Bindl, E. Heinemann, P. K. Mandal and I. Huc, *Chem. Commun. (Cambridge, U. K.)*, 2021, **57**, 5662-5665.
182. L. Wang, C. Douat, J. Sigl, P. Sai Reddy, L. Fischer, B. L. d'Estaintot, Z. Liu, V. Pophristic, Y. Yang, Y. Zhang*, et al.*, *Chem. Sci.*, 2025, **16**, 12385-12396.
183. M. Vallade, M. Jewginski, L. Fischer, J. Buratto, K. Bathany, J.-M. Schmitter, M. Stupfel, F. Godde, C. D. Mackereth and I. Huc, *Bioconjugate Chem.*, 2019, **30**, 54-62.
184. J. Buratto, C. Colombo, M. Stupfel, S. J. Dawson, C. Dolain, B. Langlois d'Estaintot, L. Fischer, T. Granier, M. Laguerre, B. Gallois*, et al.*, *Angew. Chem., Int. Ed.*, 2014, **53**, 883-887.
185. J. G. McAfee, S. P. Edmondson, P. K. Datta, J. W. Shriver and R. Gupta, *Biochemistry*, 1995, **34**, 10063.
186. D. Deepak, J. Wu, V. Corvaglia, L. Allmendinger, M. Scheckenbach, P. Tinnefeld and I. Huc, *Angew. Chem., Int. Ed.*, 2025, **64**, e202422958.
187. M. Mahjoubin-Tehran, S. Rezaei, A. H. Eid, P. Kesharwani and A. Sahebkar, *Biochem. Pharmacol. (Amsterdam, Neth.)*, 2025, **239**, 117027.
188. V. Corvaglia, I. A. M. Amar, V. Garambois, S. Letast, A. Garcin, C. Gongora, M. Del Rio, C. Denevault-Sabourin, N. Joubert, I. Huc*, et al.*, *Pharmaceuticals*, 2021, **14**, 624.
189. V. Kleene, V. Corvaglia, E. Chacin, I. Forne, D. B. Konrad, P. Khosravani, C. Douat, C. F. Kurat, I. Huc and A. Imhof, *Nucleic Acids Res*, 2023, **51**, 9629-9642.

4 Objectives and outline

Envisioning foldamers as potential drug candidates is tempting because they share many key properties that also define bio-oligomers. In principle, one should be able to replicate features of e.g., oligonucleotide drugs with abiotic molecules. Nucleic acids and peptides have decades of research behind them – not only their investigations in biological systems but also their chemical synthesis. Before tackling questions such as the inhibition of crucial biochemical processes with AOFs, the development and optimization of reproducible, robust protocols for their synthesis are necessary.

Initially, DMFOs were synthesized exclusively in solution phase by fragment condensation (doubling their length) to yield molecules of defined 2^n length with a maximum reported foldamer length of 32 units, mimicking 16 bp DNA.[1] This strategy has two key advantages: the fragment doubling rapidly increases the length (n=6 for 2^n means 64 units) and scalability to produce significant quantities of material. Although key advantages, this synthetic strategy is severely hampered by the inflexibility of rapidly changing single units in the designed molecule. As a potential platform for designed therapeutics, these foldamers need to be tailored towards a specific target and therefore they must be variable. For these reasons, the synthesis of AOFs (including that of DMFOs) has mostly been shifted to solid phase-based approaches inspired by peptide chemistry. With well-established solid-phase foldamer synthesis (SPFS) protocols, new building blocks may demand more chemical creativity than oligomer production itself.[2, 3]

One of the objectives of this thesis is the development of all the necessary monomers for an Fmoc-based SPS strategy (Fig. 20), including the synthesis of Fmoc-protected dimers that can be used as macro monomers (Fig. 20d, chapter 6), which will serve for the synthesis of DMFOs of unprecedented length. A further objective is the synthesis phosphonic acid monomers for

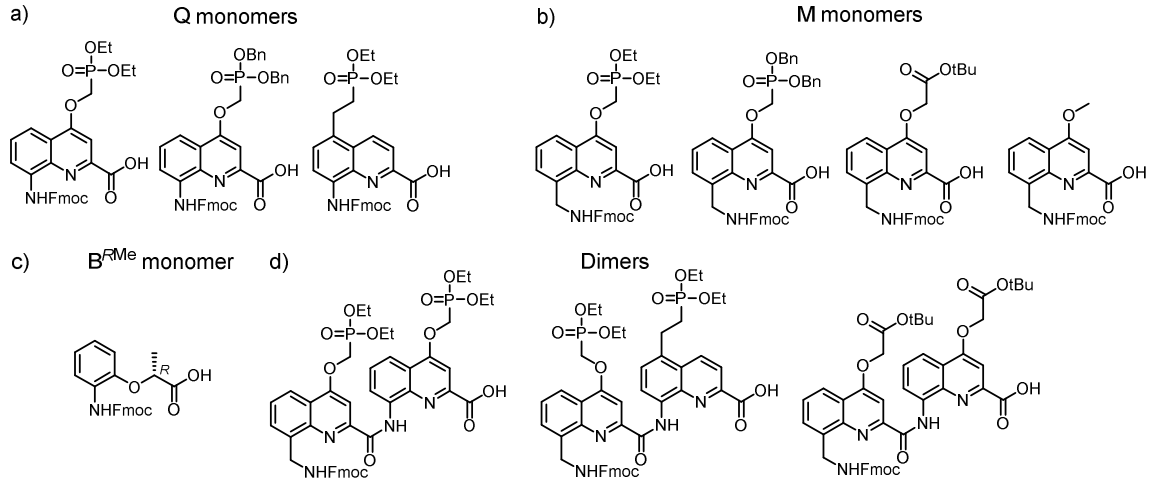


Figure 20 – Fmoc protected amino acid building blocks for DMFOs a) Derivatives of Q with different side chain position and PGs. b) Derivatives of M with different side chains. c) Chiral monomer to bias handedness.[4] d) Different Fmoc-(MQ)-OH dimers to be used as macro-monomers in SPFS.

Q^4 and M with protecting groups (PGs) that allow for trifluoroacetic acid-mediated (TFA) deprotection (see chapter 8). TFA, commonly used in peptide synthesis, enables compatibility of phosphonic acids with monomers bearing other side chains and sets the synthetic baseline for sequence diversity of DMFOs in the future. This includes a previously described chiral benzene monomer (Fig. 20c) that biases foldamer handedness to mimic B-DNA's *P*-helical array of negative charges.[4, 5]

Optimizing SPFS will be essential to achieve high crude purity and good yields, especially given the complexity of monomer synthesis (6–13 steps). SPFS must balance oligomer purity with efficient monomer consumption and recovery. These novel strategies for synthesis will give access to foldamers with unprecedented length, therefore analytical techniques will have to be developed to ensure both accurate determination of identity as well as purity prior to moving on for further experiments. Longer DMFOs will allow for the opportunity to investigate proteins or multi-domain protein complexes that typically bind DNA segments much longer than typical DMFO length. We intend to analyze protein-DMFO complexes structurally and biophysically, for which an effect on their function is already established, such as the ORC complex or full-length Top1.[6, 7] Promising targets for an initial proof-of-concept of a cryo-EM structure of a DMFO–protein complex include systems with extensive literature supporting successful structural characterization with this technique. This represents a methodological investigation into whether it is feasible to capture interactions or structural arrangements between foldamers and associated proteins and may offer valuable insights into their binding interfaces and potential mechanisms of molecular recognition. Potential protein targets include chromatin remodelers (INO80) and gene regulators like heterochromatin protein 1α (HP1α).[8-10] Beyond structural investigations, long DMFOs may also lead to insights into DNA-sensing inflammasomes that cause recurrent atherosclerotic stroke.[11]

B-DNA has been thoroughly described in the past decades. This includes its structural parameters from various crystal structures as outlined in the introduction of this thesis. Additionally, numerous studies have been pursued to answer the questions about its structural flexibility and deformability both *in silico* and experimentally.[12-19] Naturally, compared to nucleic acids, knowledge about these parameters is underexplored for AOFs, yet for poly-$(Q)_n$ some information, such as a limited effect of temperature on helix stability and *P*/*M*-helix interconversion rates in water, is established.[20, 21] For DMFOs, there is little prior knowledge about their helical integrity compared to B-DNA and oligo-$(Q)_n$ foldamers. Their torsional parameters and deformability in different conditions such as pH and salt concentration might provide crucial information for their interaction mechanisms with DBPs. This is especially important in the context of indirect readout, as DMFOs show specificity even though they lack sequence information, at least in the disruption of chromatin composition.[22] Selectivity by direct readout is unlikely due to lack of sequence diversity, but torsional flexibility and deformability may confer selectivity by indirect readout mechanisms. Parts of this thesis will

be the investigation of DMFOs in different conditions, including the design of assays to answer this question experimentally. This will set an overall benchmark and will allow us to quantify some of these properties of DMFOs. Combining experimental data and molecular dynamics (MD) simulations could provide valuable tools for guiding and refining DMFO design in the future (chapter 5).

Binding DBPs and interfering into PNIs in a non-sequence-selective manner with DMFOs is a milestone in foldamer science. As a designable platform with access to a far greater variety of side chains than those found in nature, the goal remains the ability to interfere with biological processes in a sequence-selective manner. As an underlying objective, this extends beyond the scope of this thesis, but progress in design and synthesis may be initiated. To interfere sequence-selectively into PNIs, an intrinsic necessary property of an oligomer must inherently contain sequence information itself. Given that phosphonates are currently the only available side chain for M-units, additional building blocks will need to be synthesized to deepen the available chemical space. Conveniently, some intermediates (Chapter 8), also necessary for TFA-mediated deprotection, of phosphonic acid units represent suitable starting points to allow for a divergent synthetic strategy toward M monomers with new side chains. However, even with a more diverse DMFO, there is no guarantee that it will bind a target selectively via side chain interactions. Also, it might impose a genetic-like code for their interactions that will have to be solved either iteratively by experiments or computationally using emerging design tools to ultimately produce sequence-selective binders of relevant protein targets.

As outlined in the introduction, nucleic acids themselves possess the ability for selectivity, both when interacting with other nucleic acids or when interacting with proteins. Additionally, since many sequence-selective PNIs have been analyzed structurally, designing nucleic acids to interfere sequence-selectively in PNIs is more feasible than with a *de novo* designed molecule. One could think of a molecule that combines the properties of DMFOs such as the resistance towards nucleases and their typical high binding affinities as well as the selectivity of nucleic acids in a single chimeric entity. We aim to design such chimeras to obtain a single coalescent molecular architecture that includes foldamer handedness control as well as the continuation of each other's grooves vertically across the molecule's main helix axis. If successful, this could pose an extension towards concurrent work by the nucleic acid community in the field of DNA decoys. The design and the convergent synthesis of these two very distinct backbones as well as the development of assays that allow correct assessment of DBP binding is reported in chapter 7.

References

1. K. Ziach, C. Chollet, V. Parissi, P. Prabhakaran, M. Marchivie, V. Corvaglia, P. P. Bose, K. Laxmi-Reddy, F. Godde, J.-M. Schmitter*, et al.*, *Nat. Chem.*, 2018, **10**, 511-518.
2. B. Baptiste, C. Douat-Casassus, K. Laxmi-Reddy, F. Godde and I. Huc, *J. Org. Chem.*, 2010, **75**, 7175-7185.
3. V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat and I. Huc, *Chem. - Eur. J.*, 2023, **29**, e202300898.
4. D. Bindl, E. Heinemann, P. K. Mandal and I. Huc, *Chem. Commun. (Cambridge, U. K.)*, 2021, **57**, 5662-5665.
5. V. Corvaglia, J. Wu, D. Deepak, M. Loos and I. Huc, *Chem. - Eur. J.*, 2024, **30**, e202303650.
6. M. J. Jaremko, K. F. On, D. R. Thomas, B. Stillman and L. Joshua-Tor, *eLife*, 2020, **9**, e58622.
7. D. T. Takahashi, D. Gadelle, K. Agama, E. Kiselev, H. Zhang, E. Yab, S. Petrella, P. Forterre, Y. Pommier and C. Mayer, *Nat. Commun.*, 2022, **13**, 59.
8. S. Eustermann, K. Schall, D. Kostrewa, K. Lakomek, M. Strauss, M. Moldt and K.-P. Hopfner, *Nature (London, U. K.)*, 2018, **556**, 386-390.
9. F. Kunert, F. J. Metzner, J. Jung, M. Hopfler, S. Woike, K. Schall, D. Kostrewa, M. Moldt, J.-X. Chen, S. Bantele*, et al.*, *Sci. Adv.*, 2022, **8**, eadd3189.
10. V. Sokolova, J. Miratsky, V. Svetlov, M. Brenowitz, J. Vant, T. S. Lewis, K. Dryden, G. Lee, S. Sarkar, E. Nudler*, et al.*, *Structure (Oxford, U. K.)*, 2024, **32**, 2094-2106.e2096.
11. J. Cao, S. Roth, S. Zhang, A. Kopczak, S. Mami, Y. Asare, M. K. Georgakis, D. Messerer, A. Horn, R. Shemer*, et al.*, *Nature (London, U. K.)*, 2024, **633**, 433-441.
12. P. Yakovchuk, E. Protozanova and M. D. Frank-Kamenetskii, *Nucleic Acids Res.*, 2006, **34**, 564-574.
13. J. Lipfert, G. M. Skinner, J. M. Keegstra, T. Hensgens, T. Jager, D. Dulin, M. Koeber, Z. Yu, S. P. Donkers, F.-C. Chou*, et al.*, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 15408-15413.
14. S. Neidle, *J. Biol. Chem.*, 2021, **296**, 100553.
15. A. Singh, A. Maity and N. Singh, *Entropy*, 2022, **24**, 1587.
16. R. M. Wartell and A. S. Benight, *Phys. Rep.*, 1985, **126**, 67-107.
17. S. Forth, M. Y. Sheinin, J. Inman and M. D. Wang, *Annu. Rev. Biophys.*, 2013, **42**, 583-604.
18. X. Gao, Y. Hong, F. Ye, J. T. Inman and M. D. Wang, *Phys. Rev. Lett.*, 2021, **127**, 028101.
19. D. Duzdevich, S. Redding and E. C. Greene, *Chem. Rev. (Washington, DC, U. S.)*, 2014, **114**, 3072-3086.
20. S. J. Dawson, A. Meszaros, L. Petho, C. Colombo, M. Csekei, A. Kotschy and I. Huc, *Eur. J. Org. Chem.*, 2014, **2014**, 4265-4275.
21. F. Devaux, X. Li, D. Sluysmans, V. Maurizot, E. Bakalis, F. Zerbetto, I. Huc and A.-S. Duwez, *Chem*, 2021, **7**, 1333-1346.
22. V. Kleene, V. Corvaglia, E. Chacin, I. Forne, D. B. Konrad, P. Khosravani, C. Douat, C. F. Kurat, I. Huc and A. Imhof, *Nucleic Acids Res*, 2023, **51**, 9629-9642.

5 Structural dynamics of DNA mimic foldamers

DBPs interact with DNA through indirect readout mechanisms (chapter 3.2). This recognition is conformation-dependent, because it depends on shape, not solely on sequence. DNA bends, twists, and undergoes local breathing motions that may expose some of its sequence features or alter groove widths.[1] For example, nature has developed these mechanisms to differentiate between transcription factors within a family recognizing the same core-consensus sequence.[2] Because the $(MQ)_n$ repeat motif of DMFOs lacks inherent sequence diversity, these molecules cannot achieve the same level of selectivity as DNA. Yet during chromatin assembly disruption, certain proteins are disproportionately affected.[3] Concomitantly, enzymes such as Top1 show inhibition with IC_{50} values comparable to their classical inhibitors, whereas others are only moderately inhibited, and some remain entirely unaffected.[4] This apparent selectivity might originate in effects like those prevalent in indirect readout DNA recognition. Thus, they may depend on the shape, the structural dynamics of DMFOs and parameters like their bending or twist persistence, which are well-established for DNA.[5] The dynamic properties of $(Q)_n$ foldamers are well understood, as their conformational (in-) flexibility arises from their fixed geometry.[6] However in DMFOs, the additional benzylic methylene of M that enables B-duplex mimicry also adds an additional rotatable bond within the molecule's backbone. This causes DMFOs to be much more structurally dynamic, a property that was exploited to probe for binding by a helix-handedness inversion assay.[7] To gain information on how DMFOs might interact with some DBPs based on their structural deformability we investigated these parameters both computationally and experimentally.

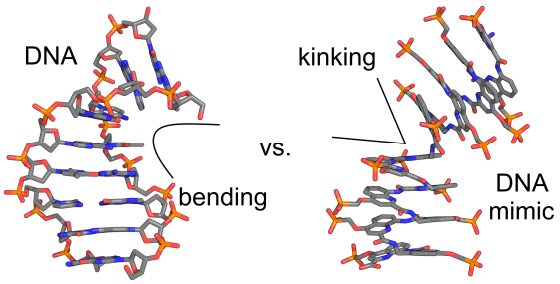


Figure 21 Double-stranded DNA shows a high degree of bendability compared to DNA mimic foldamers, which are prone to kinking.

Contributions

The project was planned in collaboration with I. Huc and M. Zacharias. Synthetic routes and characterization for Fmoc-protected building blocks were developed and provided by me. M. Rogovoi provided synthetic intermediates. Oligomer synthesis and analysis was conducted by me, J. Wu and V. Corvaglia. pH and salt-dependent NMR spectra were conducted by V. Corvaglia and VT-NMR experiments were conducted by me. L. Allmendinger was involved in the design and execution of all NMR experiments. VT-UV-Vis and CD-spectroscopic measurements were conducted by me. Lion Thurecht and M. Zacharias performed all parts of

the computational study. Z. Liu and V. Pophristic provided the custom GAFF forcefield parameters. The manuscript was written by me in collaboration with I. Huc and M. Zacharias.

References

1. J. P. Peters and L. J. Maher, III, *Q. Rev. Biophys.*, 2010, **43**, 23-63.
2. R. Rohs, X. Jin, S. M. West, R. Joshi, B. Honig and R. S. Mann, *Annu. Rev. Biochem.*, 2010, **79**, 233-269.
3. V. Kleene, V. Corvaglia, E. Chacin, I. Forne, D. B. Konrad, P. Khosravani, C. Douat, C. F. Kurat, I. Huc and A. Imhof, *Nucleic Acids Res*, 2023, **51**, 9629-9642.
4. K. Ziach, C. Chollet, V. Parissi, P. Prabhakaran, M. Marchivie, V. Corvaglia, P. P. Bose, K. Laxmi-Reddy, F. Godde, J.-M. Schmitter*, et al.*, *Nat. Chem.*, 2018, **10**, 511-518.
5. J. Lipfert, G. M. Skinner, J. M. Keegstra, T. Hensgens, T. Jager, D. Dulin, M. Koeber, Z. Yu, S. P. Donkers, F.-C. Chou*, et al.*, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 15408-15413.
6. D. Bindl, P. K. Mandal and I. Huc, *Chem. - Eur. J.*, 2022, **28**, e202200538.
7. D. Deepak, J. Wu, V. Corvaglia, L. Allmendinger, M. Scheckenbach, P. Tinnefeld and I. Huc, *Angew. Chem., Int. Ed.*, 2025, **64**, e202422958.

5.1 Publication (accepted): Structural dynamics of DNA mimic foldamers

Manuel Loos,[a] Lion Thurecht,[‡b] Jiaojiao Wu,[‡a] Valentina Corvaglia,[‡¶a] Zhiwei Liu,[c] Vojislava Pophristic,[c] and Martin Zacharias*[b], Ivan Huc,*[a]

[a] *Department Pharmazie, Ludwig-Maximilians-Universität München, Butenandtstr. 5–13, 81377 München (Germany). E-mail: ivan.huc@cup.lmu.de*

[b] *Physik-Department and Center of Protein Assemblies, Technische Universität München, James Franck Str. 1, 85748 Garching (Germany). E-mail: martin.zacharias@mytum.de*

[c] *Department of Chemistry & Biochemistry, Rowan University, 201 Mullica Hill Road, 08028 Glassboro, New Jersey (USA).*

¶ Present affiliation: 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, Milan, Rotondo, Italy.

‡ Contributed equally

Abstract

DNA mimic foldamers are helically folded aromatic oligoamides bearing negatively charged side chains that mimic the shape and charge distribution of double-stranded B-DNA. They have been shown to bind to some DNA-binding proteins better than DNA itself and thus have potential to interfere with DNA-protein interactions. Their structure has been previously characterized in detail by X-ray crystallography. We have now investigated their structural dynamics both computationally and experimentally. The force field parameters of the building blocks required for DNA mimicry were optimized and implemented in AMBER to perform molecular dynamics simulations of the foldamer helices. The position of the negatively charged side chains on the helix, the charge state of the side chains, and the presence of salt were systematically varied. The simulations revealed that the global flexibility parameters for twisting and bending of the foldamer helices are of similar magnitude to those of B-DNA, though distinct kinking events and motions are involved. A range of sequences were then prepared for experimental investigations using ^{1}H NMR, UV-vis absorption and circular dichroism spectroscopies. Measurements revealed that the foldamer helices are stable over a broad range of temperature, pH and salt conditions in aqueous solutions, but that they nevertheless undergo structural changes when conditions are modified. An assay was developed to quantitatively assess foldamer helix stability through the measurement of the rate of interconversion between right-handed and left-handed diastereomeric conformers. Unexpectedly, suppressing some negatively charged side chains had a destabilizing effect on the helix, suggesting a more complex role of the side chains than electrostatic repulsions.

Introduction

The molecular recognition properties of DNA include its ability to form duplex structures via A:T and G:C base pairs (bp) and the interactions of these duplexes with a large number of proteins required for the repair, packaging, regulation, transcription, and duplication of genetic information. Molecules that mimic the shape and surface features of DNA duplexes

may engage interactions with these DNA-binding proteins, interfere with protein-nucleic acid interactions (PNIs), and serve as pharmacological, diagnostic, or even therapeutic tools. These mimics are to be distinguished from other analogues of DNA that reproduce its base pairing ability, such as PNAs and LNAs, whose target is DNA, not DNA-binding proteins.[1,2] The most obvious analogues of DNA's overall shape are derived from DNA itself and called DNA decoys.[3-5] In nature, some proteins that mimic DNA have been shown to interfere with PNIs.[6-8] It is actually remarkable that the peptide backbone may generate shapes similar to a DNA duplex. Subsequently, coiled-coil peptides have been proposed to serve for that purpose,[9] and some anionic polymers such as heparin are known to bind to DNA-binding proteins so much so that heparin-chromatography is used for their purification.[10] Along this line, we have developed helically folded aromatic oligoamide foldamer-based DNA mimics that reproduce the overall shape and negative charge distribution of double-stranded B-DNA.[11-16] These DNA mimic foldamers are chemically remote from DNA – they contain no sugar and no nucleobase – but they are able to bind to some DNA binding proteins better than DNA itself and to inhibit PNIs.[11,12] The structural resemblance between DNA and DNA mimic foldamers is essential to explain why the latter bind to DNA-binding proteins. Conversely, only structural differences between DNA and DNA mimic foldamers may explain why the mimics may outcompete DNA and inhibit PNIs. We have reported several solid-state structures of DNA mimic foldamers either alone or bound to a chromosomal protein,[11,12,14,17] shedding some light on their structural features and gathering useful information to understand and further improve DNA-mimic foldamer inhibition of PNIs. However, information about their structural dynamics and how they compare with those of DNA is lacking. Here, we report a computational and experimental investigation of the structural dynamics of DNA mimic foldamers. We analyze parameters such as overall stability, shape, kinking, and bending properties, as a function of the charge state (pH), the presence of salt, and temperature and highlight both resemblances and differences between DNA mimic foldamers and DNA.

The original DNA mimic foldamer design consists of single-stranded oligoamides in which two types of amino acid monomers alternate: 8-amino-2-quinolinecarboxylic acid Q and 8-aminomethyl-2-quinolinecarboxylic acid ^{m}Q (Fig. 1a). Q carries an aromatic amine whereas ^{m}Q carries an aliphatic (benzylic) amine. Both types of monomers may be functionalized with a phosphonic acid-containing side chain that can be mono or dianionic in aqueous solution depending on pH, in contrast with phosphodiesters in DNA which cannot be dianionic. Q^4 and Q^5 have their side chain in position 4 and 5 of the quinoline ring, respectively, and also differ by an oxygen *vs.* a carbon atom as the first exocyclic atom of the side chain (Fig. 1a). Q_n oligomers are well known to adopt stable single-helical conformations both in the solid state and in solution,[18-21] and so do $(^{m}QQ)_n$ DNA mimic oligomers. In both cases, helical structures are stabilized by hydrogen bonds between consecutive units in the sequence (Fig. 1b) and, in protic solvents, by hydrophobic effects associated with aromatic stacking. However, a unique

feature of the $(^{m}QQ)_n$ single helix is that its curvature is such that side chains borne by ^{m}Q and Q units form a double helical array of *exo*-helices[22] and that their positions match the positions of phosphate groups in B-DNA (Fig. 1c). Thus, the single helix of $(^{m}QQ)_n$ features two grooves that resemble the minor and major grooves of B-DNA. The grooves are defined as shown in Fig. 1b which also highlights that the major groove is wider and the minor groove narrower in $(^{m}QQ^5)_n$ than in $(^{m}QQ^4)_n$. Solid state structures of $(^{m}QQ)_n$ single helices largely explained their resemblance with the shape of the B-DNA double helix.[11,12,14] Each ^{m}QQ dimer is similar in size to a DNA base pair, defining a foldamer helix diameter comparable to the diameter of B-DNA, and contributes about 0.9 helix turn, meaning consecutive ^{m}QQ dimers are twisted by *ca.* 0.1 turn (36°), as are base pairs in B-DNA. Furthermore, the helix pitch in a foldamer helix, *i.e.,* the vertical rise per turn, equals the thickness of an aromatic ring (*ca.* 3.5 Å), like the base pair distance in B-DNA. In solution, the good solubility (absence of aggregation even at mM concentrations) and folding of $(^{m}QQ)_n$ gives rise to sharp ^{1}H nuclear magnetic resonance (NMR) spectra in which signals are spread over a wide range of chemical shift values despite the repetitive nature of the sequence (Fig. 1d).[21] This spreading results from the upfield shifts of NMR signals due to ring current effects between intramolecularly stacked aromatic rings. The upfield shifts become more pronounced as the helix length increases (Fig. 1d), indicating a cumulative effect. Because ring current effects quickly decay with distance,[23,24] cumulative effects are possible only if the helix is conformationally very well defined.[21]

As shown in the following, computations and experiments concur to support that $^{m}QQ^4$ and $^{m}QQ^5$ oligomers adopt stable helical conformations over a wide range of temperature and pH. Their global flexibility parameters for twisting and bending were found to be of similar magnitude as those of B-DNA. However, their structural dynamics reflect the properties of the aromatic amide backbone and differ from that of DNA. For example, reversible unstacking events occur within the aromatic helix, leading to temporary kinks with transient redirection of the helical axis. Some of these events can last longer than 100 ns. We report that pH, *i.e.* the charge state of the phosphonate side chains appears to have the greatest effect on these structural dynamics, but also that the effect of the charged phosphonate side chains is more complex than simple helix-destabilizing electrostatic repulsions. In case of singly-anionic phosphonates, hydrogen bonding between neighboring phosphonates may also stabilize close contacts. Eventually, removing a charged side chain was found to have a stabilizing effect on the helix. Finally, experiments and simulations revealed that helices in the $^{m}QQ^5$ series are more stable than those in the $^{m}QQ^4$ series.

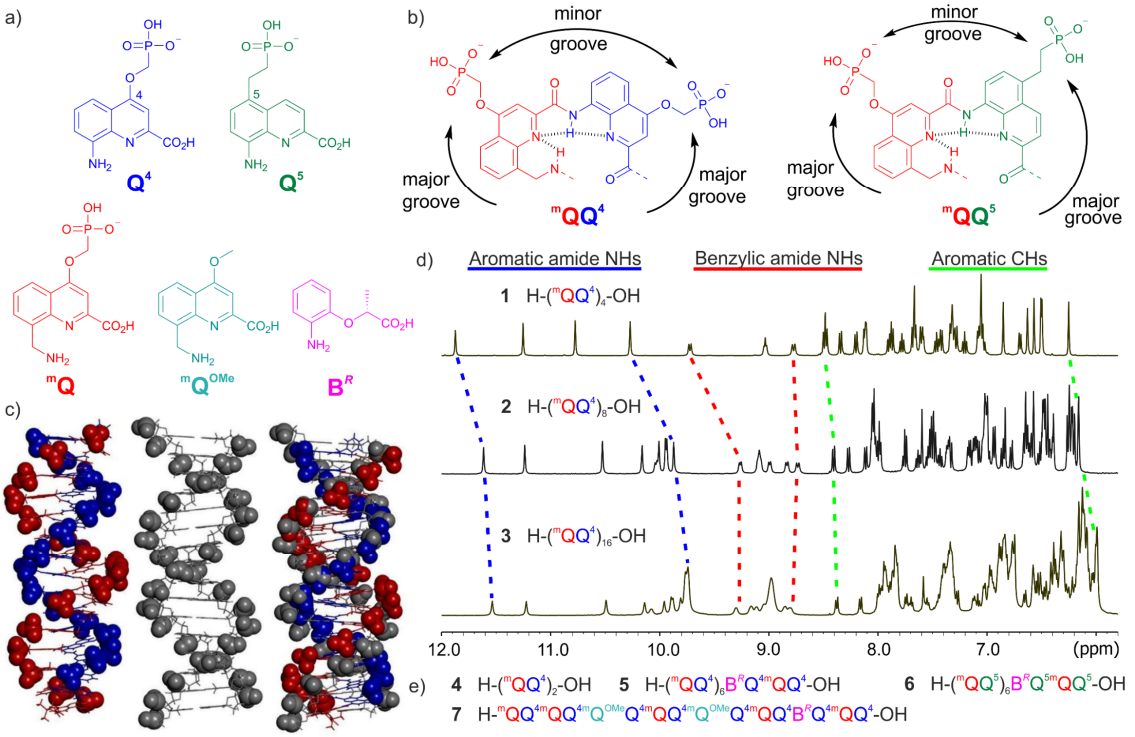


Fig. 1 (a) Structural formulae of amino acid monomers Q^4, Q^5, ^mQ, $^mQ^{OMe}$, and B^R. (b) Formulae of $^mQQ^4$ and $^mQQ^5$ repeat units. (c) Left: solid state structure of the side chain protected $(^mQQ^4)_{16}$ (^mQ units in red, Q^4 units in blue, stick representation except the phosphonate groups shown in space filling representation, ethyl ester protection of the phosphonate omitted for clarity),[11] center: structure of the 16-bp B-DNA duplex d(ACTGAACGGCTACGTA)•d(TGACTTGCCGATGCAT) (gray, stick representation except the phosphonate groups shown in space filling representation), right: overlay of the two structures. (d) Part of the ^1H NMR spectra of $(^mQQ^4)_4$ (**1**), $(^mQQ^4)_8$ (**2**) and $(^mQQ^4)_{16}$ (**3**) in 50 mM ammonium bicarbonate in H_2O/D_2O (9:1 v/v) at 25 °C showing amide NH and aromatic CH resonances. (e) Foldamer sequences **4-7** used to investigate the kinetics of helix handedness reversal.

Computational methods

Construction of arylamide residues for computations

The arylamide residues were constructed using model compounds shown in Fig. 2 and following a protocol similar to the creation of amino acid residues in protein simulations. The partial atomic charges were derived using the multi-conformational fitting method provided by the R.E.D. tools[25] for restrained electrostatic potential (RESP) charge derivation.[26]

Optimization of torsional parameters

Several backbone torsional parameters were reparametrized by taking into consideration monomer specific effects such as the intramolecular hydrogen bonds between the quinoline endocyclic N and amide NH groups (Fig. 2, bonds shown in color). Specifically, the torsional parameters for the rotations about the aryl-amide bonds (blue and red bonds in Fig. 2), and about the aryl-aliphatic C (green) and aliphatic C-amide N (purple) bonds were re-parameterized using methods developed previously.[27,28] Briefly, torsional energy profiles were

calculated by a high level quantum mechanical (QM) method on model compounds bearing the same backbone functional groups. Next, non-bonded energies were evaluated using the general AMBER force field (GAFF)[29] and RESP charges and subtracted from the QM torsional profile. Last, least square fitting was applied to fit the GAFF torsional term against the torsional profile with non-bonded corrections to obtained optimized parameters $V_{n,i}$.

$$E_{torsion} = \sum_{i} V_{n,i}(1 + \cos(n\emptyset - \gamma))$$

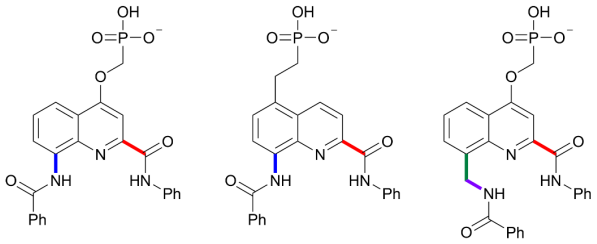


Fig. 2 Model compounds used for RESP charge fitting. The terminal Ph-CO- and Ph-NH- groups have a sum of charges set to 0 and are removed to build the arylamide residues. Depicted in this Figure are residues with a −1 charged side chains. Residues with a −2 charged side chain (further deprotonated) were also created following the same procedure. Bold bonds in color indicate re-parameterized torsions.

Molecular Dynamics Simulations

For Molecular Dynamics (MD) simulations, we used $(^{m}QQ^{4})_{18}$ and $(^{m}QQ^{5})_{18}$ sequences consisting of thirty-six ^{m}Q and Q^{4} or Q^{5} residues. Initial coordinates were generated by energy minimization to yield a starting straight helix geometry (without any kinking or bending) with a helical pitch of 3.5 Å. In the energy minimized structure, the pitch of the exo-helices spans ~12 ^{m}QQ units (~24 residues), instead of ~10 ^{m}QQ in crystal structures. Sequence length was chosen to span more than one full exo-helix turn and, at the same time, kept short enough to limit the size of the simulated system in order to perform simulations in the microsecond regime. All phosphonate groups within a chain were either in a singly- or doubly-anionic state. In solution, the second pKa of an isolated phosphonic acid is expected to be near 6.5 but we have measured that the second pKa of $(^{m}QQ^{4})_{n}$ oligomers is shifted to higher values centered above 8.5[11], probably due to the negative charge density of the helix. We performed MD simulations with both possible protonation states resulting in a total of four simulation systems. All simulations were performed with the Amber18 package.[30] The systems were solvated with TIP3P water[31] in an octahedral box with a minimum of 12 Å (for monoanionic phosphonates) and 25 Å (for dianionic phosphonates) between the DNA mimic foldamers and the box boundaries. A larger box size was chosen for the double charged system because we noticed a possible stretching of the molecule during test simulations. The ion concentration was adjusted to ~0.1 M with added Na^{+} and Cl^{-} ions (neutral system) for most simulations. After energy minimization (2500 steps) using the sander module of the Amber18 package, the

systems were heated in steps of 100 °C (each 0.1 ns) up to a temperature of 26.85 °C (300 K), keeping positional restraints on all non-hydrogen atoms with respect to the start structure. The positional restraints were removed within another 2 ns equilibration at 26.85 °C and constant pressure of 1 bar. Data gathering production simulations were extended to up to 2-3 μs for each system. All MD-simulations were performed using the pmemd.cuda program of Amber18 in combination with hydrogen mass repartition[32] allowing a time step of 4 fs. Trajectory analysis for recording root-mean-square deviation (RMSD) *vs.* time, root-mean-square-fluctuations relative to the mean (RMSF) and analysis of dihedral angles was performed using the cpptraj module of Amber18. For the analysis of bending, twisting and stretching of the DNA mimics in-house scripts were used based on coarse graining of the chain. The averages over heavy atoms in two consecutive units were used to define the coarse-grained centers that follow the helical axis of the systems. The distance between these centers defines the helical rise (pitch) along the helical axis. The twist was calculated as the dihedral angle between the segments linking the helical axis and the geometric centers of aromatic rings. It can also be used to calculate the bending angle along the chain at each chain unit. The bending persistence length was calculated from the mean scalar product of the unit vector along the helical axis at the beginning of the chain and the axis vector at the end of the chain. Together with the geometric centers for each of the aromatic ring segments of each residue it is possible to define a rotation angle per unit along the chain and to calculate an overall periodicity or twist of the DNA mimics.

Computational study

General outcome of MD simulations

The simulations of $(^{m}QQ^{4})_{18}$ and $(^{m}QQ^{5})_{18}$ (also named $^{m}QQ^{4}$ and $^{m}QQ^{5}$ oligomers in the following paragraphs) with singly charged phosphonates resulted in rapid equilibration to conformations that remained within an RMSD of ~3-4 Å relative to the start structure (Fig. 3). In the case of $^{m}QQ^{4}$, two transitions to conformations with larger RMSD of ~6 Å were observed that will be discussed in a separate paragraph below (Fig. 3c). The mean equilibrium twist as defined in the Methods section was larger for $^{m}QQ^{5}$ than for $^{m}QQ^{4}$ (28.8° *vs.* 25.5°, Table 1) and also indicated some unwinding relative to the experimental solid-state structure where the twist angle is closer to 36°,[11] as between base pairs in B-DNA.

Table 1 Mean helix parameters and global flexibility of foldamers and DNA.

Molecule	Twist	Rise	Bending persistence	Twist persistence	Rise fluctuations
$(^{m}QQ^{4})_{18}$	25.5°	3.66 Å	377 Å	365 Å	0.035 Å^2/step
$(^{m}QQ^{5})_{18}$	28.8°	3.65 Å	620 Å	489 Å	0.025 Å^2/step
DNA[35]	35.2°	3.30 Å	450 Å	1090 Å	

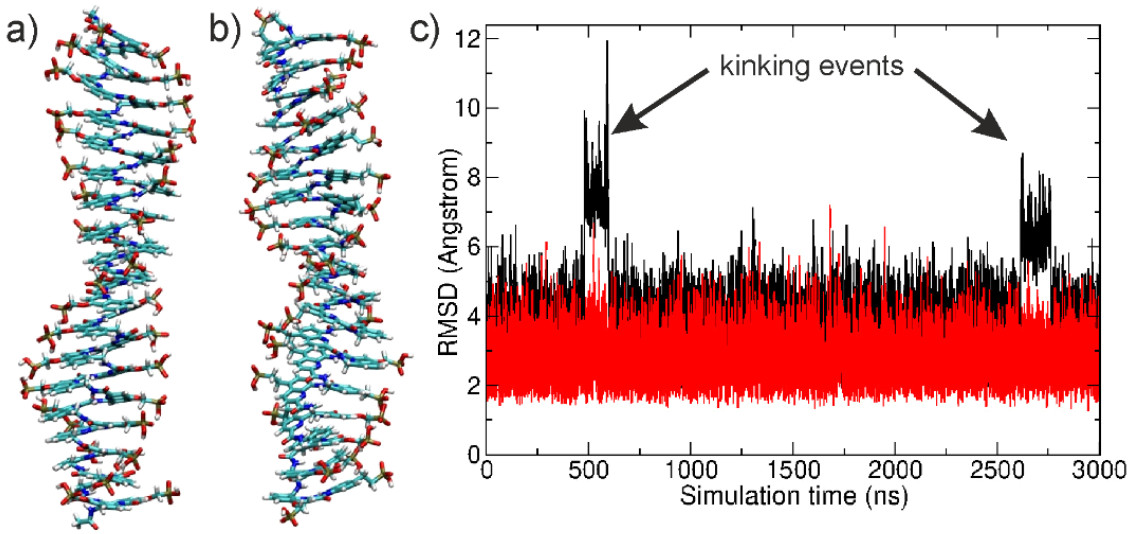


Fig. 3 (a) Equilibrated start structure of the $(^{m}QQ^{4})_{18}$ DNA-mimic (36 residues = 18 repeat units) in stick representation. (b) same as (a) for the $(^{m}QQ^{5})_{18}$ DNA-mimic. (c) RMSD of non-hydrogen atoms from the start structure *vs.* simulation time for the $^{m}QQ^{4}$ (black) and $^{m}QQ^{5}$ (red) DNA-mimics.

We are not aware of such unwinding in simulations of B-DNA. In the foldamers, it results in a mean *exo*-helical periodicity of the structures of 14 units for $^{m}QQ^{4}$ and 12.5 for $^{m}QQ^{5}$ instead of 10 in the solid-state structures. A possible explanation for the unwinding is the frequent formation of hydrogen bonds between neighboring phosphonate groups that may influence the mean twist (Fig. 4c, d). On average 10-60% of the phosphonate groups were involved in hydrogen bonding (with $d_{H..O}$ < 2.5 Å). Such stabilization of close contacts between phosphonate groups is not possible in the crystal structures where phosphonates were protected as ethyl esters. The mean helical pitch of 3.65 Å observed during the simulations is close to the pitch found in the crystal structure (3.5 Å). For the RMSF of all heavy atoms, a regular pattern was observed with the terminus showing on average larger fluctuations than the central part of the DNA mimics (Fig. 4a, b).

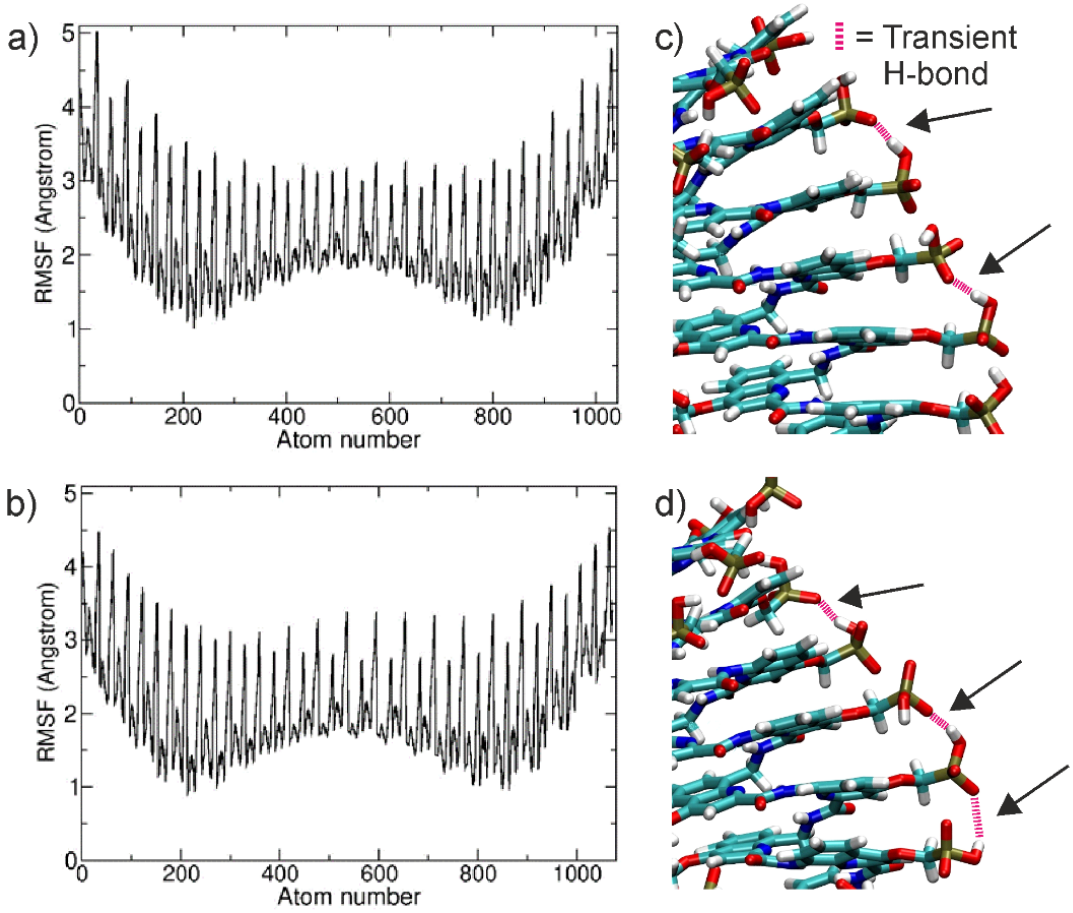


Fig. 4 (a) Root-mean-square fluctuations (RMSF) of non-hydrogen atoms *vs.* atom number observed during $(^{m}QQ^{4})_{18}$ simulations (excluding the simulation parts with significant kinking). (b) same as (a) for simulations on $(^{m}QQ^{5})_{18}$. The maxima in the plots indicate fluctuations of the phosphonate groups and the less fluctuating regions correspond to the main chain aromatic amide atoms. (c, d) Transient hydrogen bond formation (pink dashed lines and black arrows) from two snapshots of a segment of the $(^{m}QQ^{4})_{18}$ structure.

The phosphonate groups of the side chains form the peaks in the RMSF distribution indicating larger mobility compared to atoms belonging to the aromatic and amide groups of the main chain helix (Fig. 4a, b). The phosphonate groups move mostly in a direction perpendicular to the helical axis, resulting in distance fluctuations due to transient hydrogen bond formation and longer-range repulsion due to the negative charge of the neighboring phosphonate residues (Fig. 4c, d). However, in the case of unstacking events, the distance between phosphonate groups also increases in the direction of the helical axis (see also next paragraph). In addition, there is no complete equivalence of adjacent phosphonate groups, indicating a nearest-neighbor influence of the fluctuations.

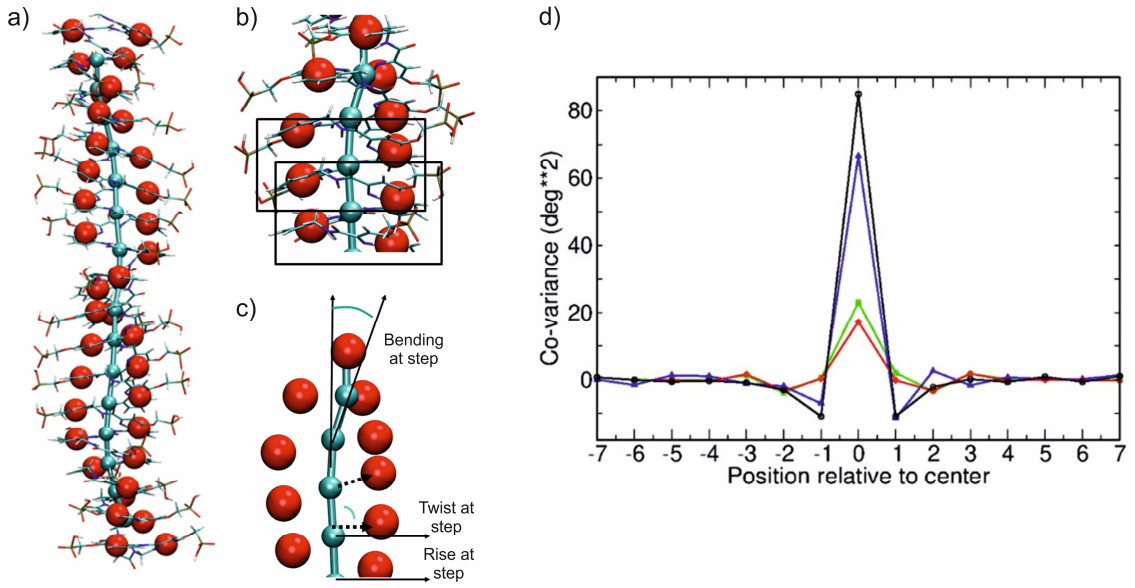


Fig. 5 Illustration of a coarse-grained representation of the DNA mimic foldamers to calculate twisting, helical axis bending, and helical rise along the helix. (a) Superposition of an atomistic structure of $(^{m}QQ^{4})_{18}$ and of its coarse-grained representation with red spheres representing the geometric centers of the quinoline rings and cyan spheres delineating the helical axis. (b) Definition of the helical axis spheres as geometric centers of four consecutive residues, equivalent to two consecutive ^{m}QQ repeat units. (c) The helical rise was calculated as the distance between consecutive spheres along the helical axis. The twist was calculated as the dihedral angle between the segments linking the helical axis and the geometric centers of aromatic rings. The bending persistence length was calculated as the scalar product of the helical direction vector at the end of the chain relative to the start of the chain. (d) Covariation of local twist and bending fluctuations. The covariance between the twist fluctuation at the central unit step (0 on the x-axis) of $(^{m}QQ^{4})_{18}$ (black line) and $(^{m}QQ^{5})_{18}$ (blue line) chain and the twist fluctuations at neighboring steps is plotted. Positive numbers indicate covariation in forward direction and negative numbers in reverse direction of the chain. The covariation at the center indicates the squared twist fluctuation. Local bending covariation is also indicated ($^{m}QQ^{4}$: green line; $^{m}QQ^{5}$: red line). Local bending angles are calculated from the position of three consecutive coarse-grained centers defining the helical axis.

Global bending, twisting and stretching flexibility

To calculate the global bending, twisting and stretching of the foldamer helices, we used a coarse-grained model illustrated in Fig. 5 (see also Methods section). Each quinoline ring was simplified by its geometric center (red spheres in Fig. 5). The helix axis was generated by creating coarse-grain centers (cyan spheres) each defined as the geometric center of four consecutive residues (Fig. 5b). The chain contour axis formed by these centers can be used to calculate the bending angle along the chain at each ^{m}QQ unit and the distance (helical rise or pitch) between consecutive ^{m}QQ units. Using the geometric centers of the quinoline rings, it is possible to define a rotation angle per ^{m}QQ unit along the chain and to calculate an overall periodicity or twist of the DNA mimics. The average twist, rise (along the contour of the chain) and bending, of contour elements were calculated for each trajectory (Table 1). Note that phases of the trajectories with significant structural transitions (RMSD > 6 Å, *e.g.*, as seen for $^{m}QQ^{4}$) were excluded from the analysis.

Transient kinking of the foldamer helix

Inspection of the trajectories indicates that even within the simulation parts with no major RMSD change, there are occasional local short-lived (< 0.1 ns) kinking events (minor kinks) illustrated in Fig. 6b. When they are not hydrogen-bonded, the negative phosphonate groups possibly repel each other, and one can observe significant distance fluctuations of adjacent phosphonate residues. Typically, the motion occurs perpendicular to the helix axis. However, the phosphonate groups occasionally separate in the direction of the helical axis, and this leads to a local unstacking and kinking of the chain (Fig. 6b, c). The structure of the DNA mimics is mainly determined by dihedral angle rotations associated with the bonded geometry of the connections between the repeating units. These variable dihedral angles are termed α, β, γ, δ, ε, and ζ (Fig. 6d) and adopt characteristic mean values associated with modest fluctuations. An analysis of the associated changes in all dihedral torsion angles along the chain indicates that this is mostly due to a change in the dihedral angle δ coupled to smaller changes of the dihedral angle γ both in the $^{m}QQ^{4}$ and $^{m}QQ^{5}$ simulations (Fig. 6f). Counting all kinks (by considering all corresponding changes in γ) that occurred during the MD simulations resulted in 120 and 95 kinks per 1 μs simulation time for $^{m}QQ^{4}$ and $^{m}QQ^{5}$, respectively. Fewer kinks were observed for $^{m}QQ^{5}$, hinting at a higher stability that was confirmed by experiments (see below).

During the $^{m}QQ^{4}$ simulation, two reversible conformational transitions to longer lived kinked states (major kinks) were observed (large shift in the RMSD for ~100-150 ns at time point 480 ns and 2500 ns). The associated structural changes are illustrated in Fig. 6c,g,h. It leads to a local break in the periodic structure, giving rise to two helix segments that have different directions with a kink angle of ~40°. The observed local transient kink distortion involves coupled transitions of several dihedral angles (mostly γ and δ and smaller changes in other dihedral angles) and leads to an increased distance between phosphonate groups for three consecutive phosphonate groups along the chain (Fig. 6c). The significant dihedral angle flips mediate a transient yet longer-lived conformational state (lifetime ~100 ns) separated by an energy barrier from the otherwise regular structure of the $^{m}QQ^{4}$ chain. It also leads to a different partially disrupted stacked structure at the kink location (Fig. 6c). However, it eventually reverts to the original continuously stacked $^{m}QQ^{4}$ helix. No such transition was observed during the $^{m}QQ^{5}$ simulation, suggesting a higher stability that was verified experimentally (see below).

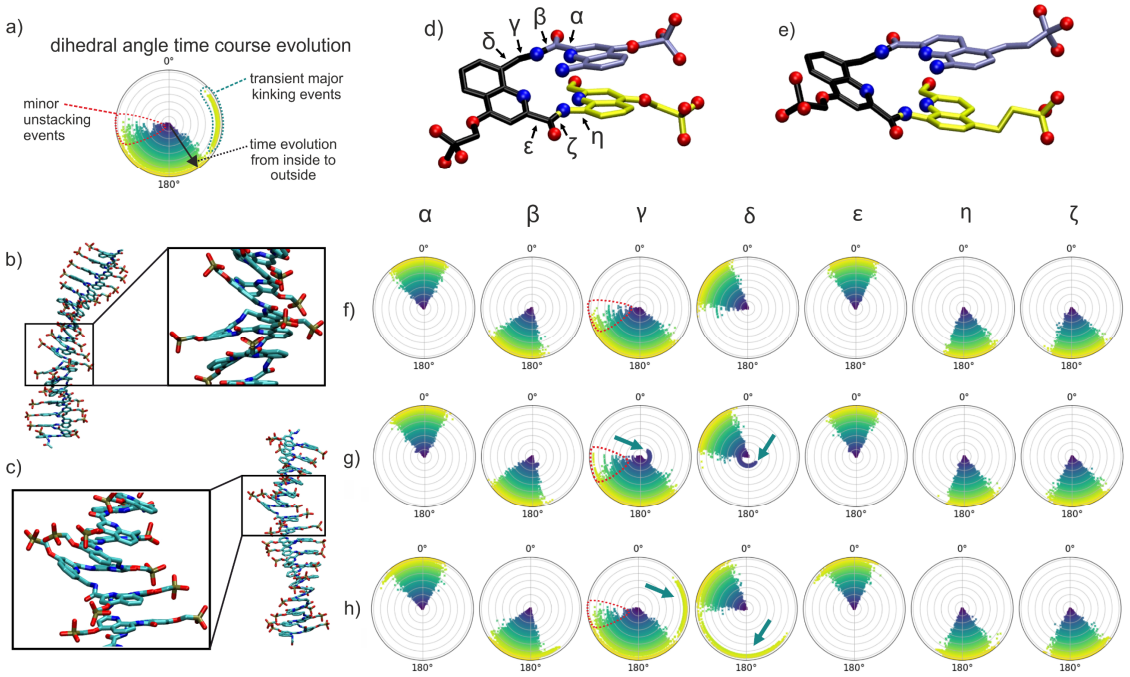


Fig. 6 (a) Depiction of the time evolution of dihedral angle distributions during MD simulations (0° at top and 180° at bottom of the circular diagram). The start corresponds to a zero radius in each circular distribution (blue dots) and the final distribution (at end of simulation) corresponds to the largest radius (yellow dots). (b, c) Kink conformations observed during MD simulations. (b) Local unstacking event between neighboring repeating units (indicated by a rectangular box) that approximately doubles the distance between adjacent phosphonate groups along the chain. An enlarged view of the local unstacking is shown in the inset of (b). Such minor kinking events occurred transiently with lifetimes of tens of picoseconds and involved small shifts (< 60°) of dihedral angles. (c) Major kinking events occurred in the simulations of the $^{m}QQ^{4}$ DNA mimic, resulting in a transient disruption of the stacking of adjacent units (inset of (c)) and strong kinking with a break in the helical axis. Two such events with lifetimes > 100 ns were observed in the simulation of $(^{m}QQ^{4})_{18}$ with singly charged phosphonates. (d,e) Chemical structures (stick model of three consecutive residues, hydrogens omitted) of a $Q^{4m}QQ^{4}$ segment (d) and a $Q^{5m}QQ^{5}$ segment (e). Relevant dihedral torsion angles along one repeat unit are indicated. Oxygen atoms are shown as red spheres and nitrogens as blue spheres. (f) Time evolution of dihedral angle distributions for a $^{m}QQ^{4}$ unit without a major kink event during the simulation of $(^{m}QQ^{4})_{18}$ with singly charged phosphonates. The occurrence of minor kinking events is mainly associated with changes in the γ dihedral angle (dots encircled in red). (g) Same as (f) but for an $^{m}QQ^{4}$ unit with a major kinking event early in the simulation that involves coupled changes of δ and γ dihedral angles (indicated by teal arrows). (h) Same as (f) for an $^{m}QQ^{4}$ unit that underwent a major kinking event towards the final stages of the simulation (marked with teal arrows). The major kinks involved also small changes in the ζ dihedral angle.

MD simulation of $(^{m}QQ^{4})_{18}$ and $(^{m}QQ^{5})_{18}$ oligomers with doubly-charged phosphonates

In contrast to the MD simulations with singly charged phosphonate groups, the simulations with doubly charged phosphonate residues resulted in many strong kinking events along the chain and eventually in the formation of partially collapsed states with several “unstacked” connections within the $^{m}QQ^{4}$ as well as $^{m}QQ^{5}$

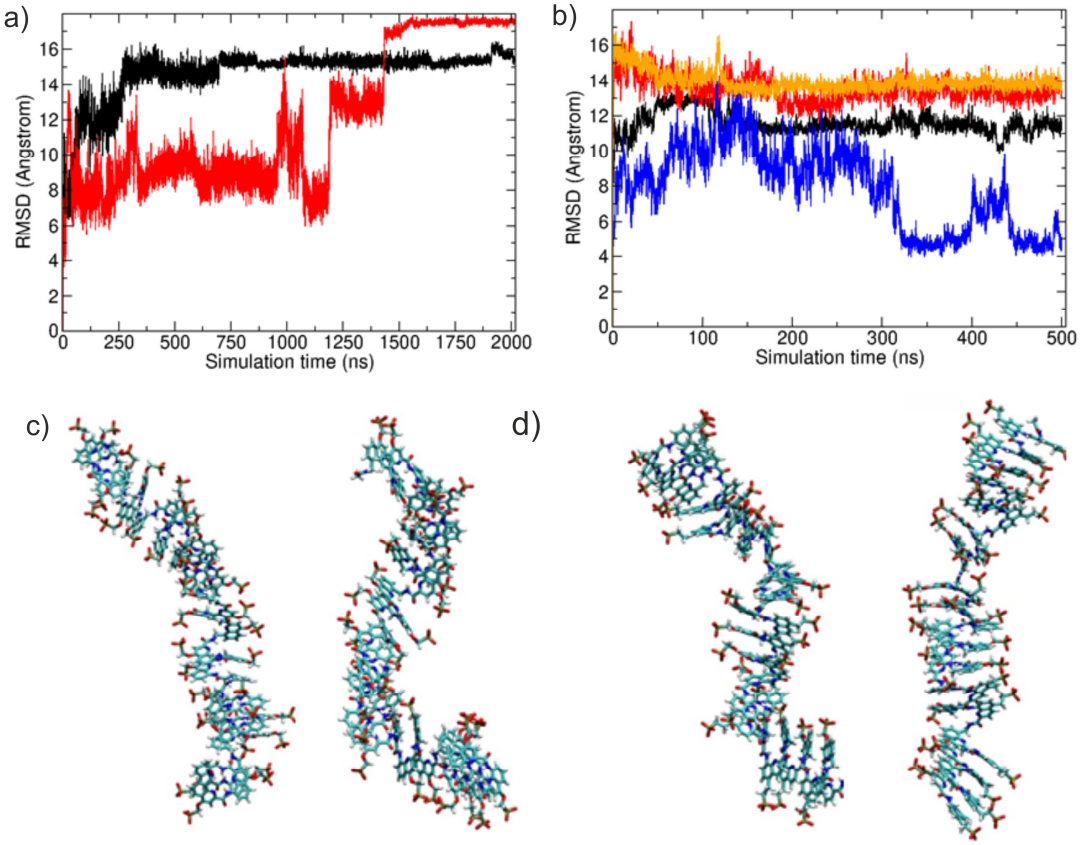


Fig. 7 MD simulations of mQQ4 and mQQ5 with doubly-charged phosphonate groups. (a) RMSD (non-hydrogen atoms) from the start structure vs. simulation time for the ^{m}QQ^{4} (black line) and for the mQQ5 (red line) oligomers with two negative charges per phosphonate group. (b) RMSD from start structure vs. simulation time for mQQ4 at high salt (0.5 M NaCl, black line), in 0.1 M LiCl (red line), ^{m}QQ^{5} at high salt (blue line) and in 0.1 M LiCl (orange line). (c) Representative snapshots of mQQ4 conformations (2 negative charges per phosphonate) sampled during the simulations shown in (a). (d) same as (c) for the mQQ5 oligomer.

sequences (Fig. 7a, c, d). The RMSD *vs* time shows much larger deviations from the initial regular structure compared to the simulations using single charges on the phosphonate groups. Yet again, RMSD values appeared to be lower for ^{m}QQ^{5}. Also, the RMSF plots indicate a much less regular pattern than for mono-protonated phosphonate residues in the chains (Fig. S1). To further investigate the potential electrostatic origin of the observed conformation transitions, we further performed simulations at higher salt concentration (500 mM NaCl) or using smaller highly hydrated monovalent Li^{+} ions (100 mM LiCl). Still, for both conditions large deviations from the start structure were sampled already in the equilibration phase. However, the average deviation from the start structure was slightly smaller compared to the standard conditions (Fig. 7b). It is likely that the increased repulsion between phosphonate groups along the chain drives the unstacking and kinking events. The final often collapsed conformations are stabilized by ion-mediated interactions between phosphonate groups (Fig. S1).

Experimental study

Foldamer synthesis

Foldamers **1-4** were synthesized as previously described.11 New chiral foldamer sequences 5-7 (Fig. 1e) were synthesized on solid phase and purified by reversed-phase high-performance liquid chromatography (HPLC) following published procedures15,34 from previously described building blocks.[14,35] For sequences **6** and **7**, two new building blocks – the variants of $^{m}QO^{Me}$ and Q^{5} suitably protected for solid phase synthesis – were necessary. Their synthesis along with the characterization of all new compounds are presented in the supplementary information.

pH and salt dependence of NMR spectra

To correlate the MD results with the experimental data, 1D and 2D NMR experiments were carried out on H-$(^{m}QQ^{4})_{8}$-OH (sequence **2**, with sodium as counterions) in water at pH 8 where the phosphonate side chains are in part dianionic and in a 50 mM sodium hydroxide solution at pH 12.5 in which the phosphonate side chains are all dianionic.[11] This sequence does not contain any stereogenic center and thus exists as a racemic mixture of right-handed (*P*) and left-handed (*M*) conformers. A full assignment of the NMR spectra was not attempted but the spectra were compared to those of shorter sequence **1** for which a full assignment was reported previously.[11] In addition, some resonances of **2** could be partly assigned using 2D COSY NMR spectra. For example, the benzylic NH resonances can unequivocally be distinguished from the aromatic NH resonances because only the former are coupled to benzylic methylene protons (Figs. S2, S3). These scalar couplings also allow one to distinguish the latter from methylene protons belonging to the side chains.

The ^{1}H NMR spectra of **2** at pH 8 and 12.5 show significant differences (Fig. 8a), but both are indicative of a folded helical structure. They are sharp; NMR signals are spread over a broad range of chemical shift values despite the repetitive nature of the sequence. Benzylic C$\underline{H}_{2}$ are anisochronous, consistent with their diastereotopic nature in a chiral helical structure and indicating slow exchange on the NMR timescale between *P* and *M* helices. In contrast, the benzylic methylene protons of short sequence **4** appear as a singlet indicating fast exchange between *P* and *M* helices (Fig. S4). Thus, in agreement with the MD simulations, different charge states of the side chains may give rise to changes in the helix structure and structural dynamics, but the overall helix folding is preserved, unlike in α-helical peptides that are destabilized by charge repulsions.[36]

Differences between the NMR spectra of **2** at pH 8 and 12.5 include both upfield and downfield shifts of the resonances and cannot be interpreted simply in terms of a reduction of ring current effects at higher pH due to more frequent kinking events that temporarily disrupt aromatic stacking. More aryl-C$\underline{H}$ resonances are found above 8 ppm at pH 12.5 than at pH 8 (Fig. 8a).

Concomitantly, the benzylic methylene protons are upfield-shifted by over 0.5 ppm at pH 12.5, some even resonate as low as 2.2 ppm as a result of strong ring current effects (Fig. 8b). The benzylic amide NH resonances, but not the aromatic NH resonances, are downfield-shifted at pH 12.5. We also note that increasing pH results in a decrease of the intensity of some amide NH signals (arrows in Fig. 8a).

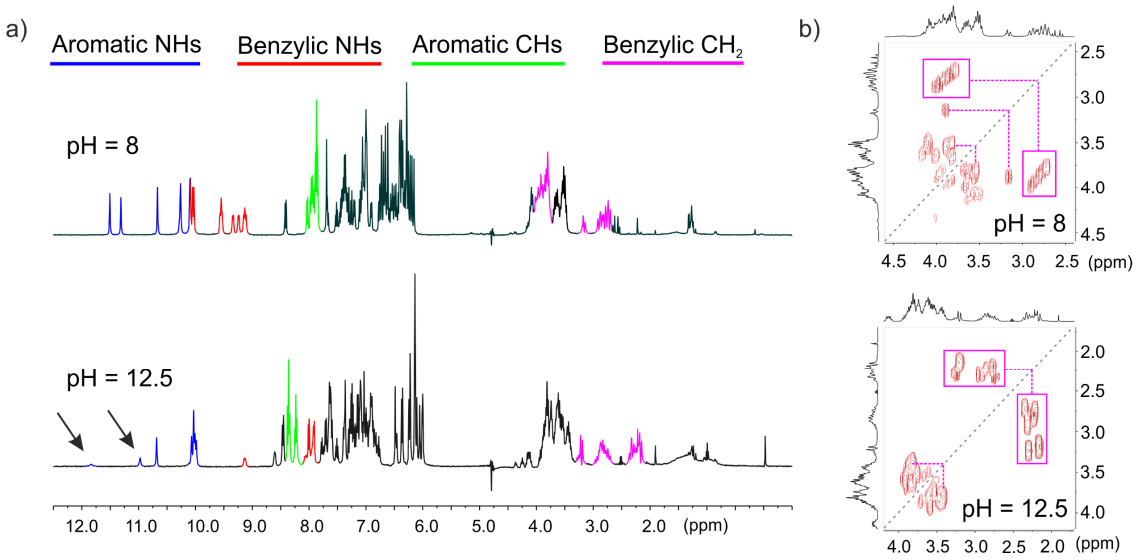


Fig. 8 ^{1}H NMR spectra of 2 recorded with water suppression at 25 °C in H_2O/D_2O (9:1 v/v) at pH 8 and 50 mM sodium hydroxide H_2O/D_2O (9:1 v/v) at pH 12.5. Blue, red, pink and green lines indicate the NMR resonances of aromatic amide protons, benzylic amide protons, aromatic CH-protons, and benzylic CH_2 protons, respectively. (b) Excerpts of 2D COSY NMR spectra of 2, same solution as in (a). In both cases the diagonal suppression was applied. The spectra show strong anisochronicity ($\Delta\delta$ > 1ppm for some signals) of the main chain benzylic CH_2 protons of the ^{m}Q monomers that are marked with pink boxes and are upfield shifted at pH 12.5.

This is due to an enhanced exchange with water,yielding suppression of the NH resonances by the water suppression part of the pulse sequence. The decrease of signal intensity is stronger with aliphatic and aromatic NHs found at lower field in both series. These signals can reasonably be assigned to amide protons near the ends of the helix, where fraying can enhance exposure to, and exchange with, water.

The effect of monovalent ions was also assessed. Spectra at pH 8 and 12 were measured in the presence of increasing concentrations of NaCl and LiCl (from 12.5 mM to 100 mM). For both salts, changes in chemical shift values were marginal compared to the changes observed upon changing pH (Figs. S5-S8). The stabilizing effects suggested by MD simulations did not apparently translate into notable changes in chemical shift values of the DNA mimic foldamers.

Effect of temperature

We investigated the effect of temperature. Even though single-stranded foldamers cannot dissociate in the way double-stranded B-DNA does, increasing temperature should enhance structural dynamics, *e.g.* the kinking events observed in MD simulations, and this may be reflected in their spectroscopic properties. We therefore analyzed **1** by variable temperature (VT) UV-Vis and NMR measurements (Fig. 9). NMR spectra showed some chemical shift variations suggesting structural changes such as differences in helix dynamics or altered helix curvature, but no sign of unfolding at temperatures up to 85 °C (Fig. 9a). The signals of some amide protons shift more than others (Δδ ranging from 0.1 ppm or less to *ca.* 0.5 ppm), which may reflect fraying at the end of the helix. The reduced intensity of some NH resonances was assigned to enhanced exchange with water at higher temperature together with the water signal suppression. Concomitantly, UV-Vis spectra were essentially unchanged over the same temperature range (Fig. 9b), in great contrast with the typical melting curves of B-DNA duplexes. Finally, we measured the CD spectra of chiral sequence **5** (the helix of **5** is left-handed, see next section) at various temperatures (Fig. 9c). CD is a particularly sensitive method to detect changes in the chiral environment of chromophores, and

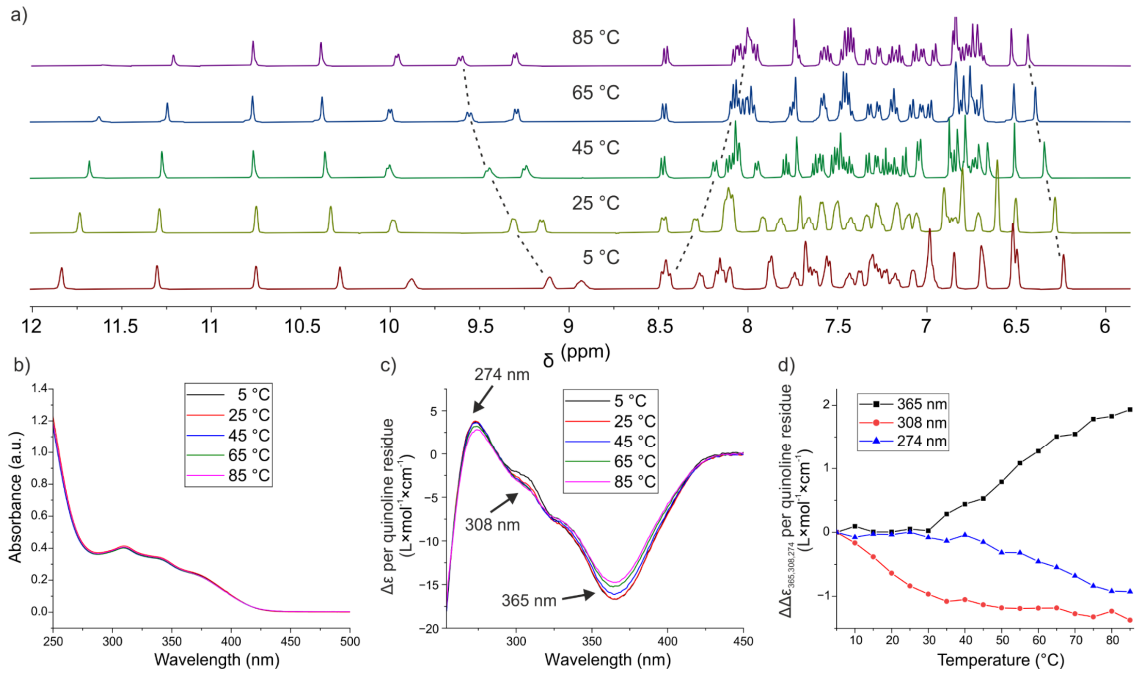


Fig. 9 (a) Excerpts of the ^{1}H NMR spectra from 12.0 to 6.0 ppm of H-(^{m}QQ^4)$_4$-OH (compound **1**) recorded in 50 mM ammonium bicarbonate in H_2O/D_2O (9:1 v/v) from 5 °C to 85 °C. (b) Variable temperature UV-Vis spectra of (^{m}QQ^4)$_4$ in 10 mM KH_2PO_4 (pH = 7.0), 20 mM KCl buffer. c) Variable temperature CD spectra of compound 5 in 20 mM Tris (pH = 7.5), 100 mM NaCl. d) Variations of absorbance as a function of temperature extracted from the experiment shown in (c) for three different wavelengths. Variations are with respect to values recorded at 5 °C.

some small variations of CD bands intensities were observed, but the trends were different at different wavelengths (Fig. 9d), suggesting that the changes reflect local variations rather than a transition from one state to another.

Quantitative assessment of helix stability

The NMR and UV-vis measurements presented above gave qualitative insights about how some of the structural dynamics observed in MD simulations translate into variations of the spectroscopic properties of the DNA mimic foldamers. We next sought for a quantitative assessment of foldamer helix stability. We have previously assessed the stability of aromatic foldamer helices by measuring the rate of interconversion between *P* and *M* helical conformations.[20,21,37] For organic-soluble Q_n oligomers, it was possible to separate the *P* and *M* enantiomeric conformers of an achiral sequence using HPLC on a chiral stationary phase at low concentration. Racemization of a one-handed conformer was then monitored by observing the decay of circular dichroism absorption bands.[20,21] For water soluble oligomers containing Q units, another approach was developed where a chiral group is introduced in the sequence that biases helix handedness in favor of *P* or *M* diastereomeric conformations in a solvent dependent manner.[38,39] Upon incubating the sequence in a first solvent, a certain *P*/*M* ratio was reached at equilibrium that translates into CD bands of a certain intensity. When that solvent was removed and the sequence dissolved in a new solvent, the CD bands intensity changes can be monitored as the *P*/*M* ratio readjusts to another equilibrium value in that new solvent.

We thus developed such an assay for $(^{m}QQ)_n$ helices. It was previously shown that introducing a single chiral B^R unit (Fig. 1a) in an $(^{m}QQ)_n$ sequence strongly biases helix handedness in water.[14,34] For example, in an ammonium bicarbonate buffer, the ^{1}H NMR spectrum of sequence **5**, (Fig. 1e, $^{m}QQ^4$ series) shows two distinctive sets of signals corresponding to the *P* and *M* diastereomeric conformers in a ratio of 98:2 (Fig. 10d). This strong bias is reflected in the CD spectrum with an intense negative band with a maximum at 365 nm as a result of the preferred *M* helicity (Fig. 10a). We screened a small number of water-miscible organic solvents and found that in a DMF/H_2O mixture (9:1, v/v), that is, with enough water to maintain solubility, the CD band is much less intense. Dissolving a sample of **5** in this solvent and monitoring the CD spectrum showed a clear drop in intensity (Fig. 10a, c), an effect that was not observed by 10-fold dilution in water (Fig. S9). The NMR spectrum in DMF-d_7/H_2O showed that the proportion between the two diastereomeric conformers is reduced to *ca.* 2:1 in favor of the *M*-helix (Fig. 10d), consistent with the intensity of the CD bands. It is interesting to point that similar effects were not observed with Q_n oligomers with which a single B^R unit was shown to quantitatively bias handedness in water, methanol, DMSO, and chloroform.[34,40] We hypothesize that the helices of $^{m}QQ^4$ oligomers are destabilized in organic solvents when hydrophobic effects are weakened (here in 90% DMF) due to the flexibility of the main chain benzylic methylene groups.[41,42] The extent of

destabilization is sufficient to lower the effectiveness of B^R at biasing helix handedness. We repeated these experiments with sequence **6** (Fig. 1e, $^{m}QQ^5$ series) and found that the diastereomeric ratio was 9:1 in favor of the *M*-helix in DMF/H_2O (9:1, v/v) *vs.* at least 98:2 in

water. The drop in CD intensity was thus much weaker (Figs. S9-S10). Consistent with the observations presented both above and below, $^{m}QQ^{5}$ oligomers appear to be more stable than $^{m}QQ^{4}$ oligomers. Here, this is reflected in a weaker destabilizing effect of DMF that results in a stronger handedness bias by B^{R}. Nevertheless, even if the drop in CD intensity was weaker in the $^{m}QQ^{5}$ series, resulting in noisier data and lower quality fits (Fig. S12), it was sufficient to monitor helix handedness inversion in water

Experiments were thus implemented where a solution of B^{R}-containing foldamers **5** or **6** was first let to equilibrate in DMF/H_2O (9:1 v/v), then diluted in water and immediately frozen with liquid nitrogen and lyophilized to obtain a mixture of the *P* and

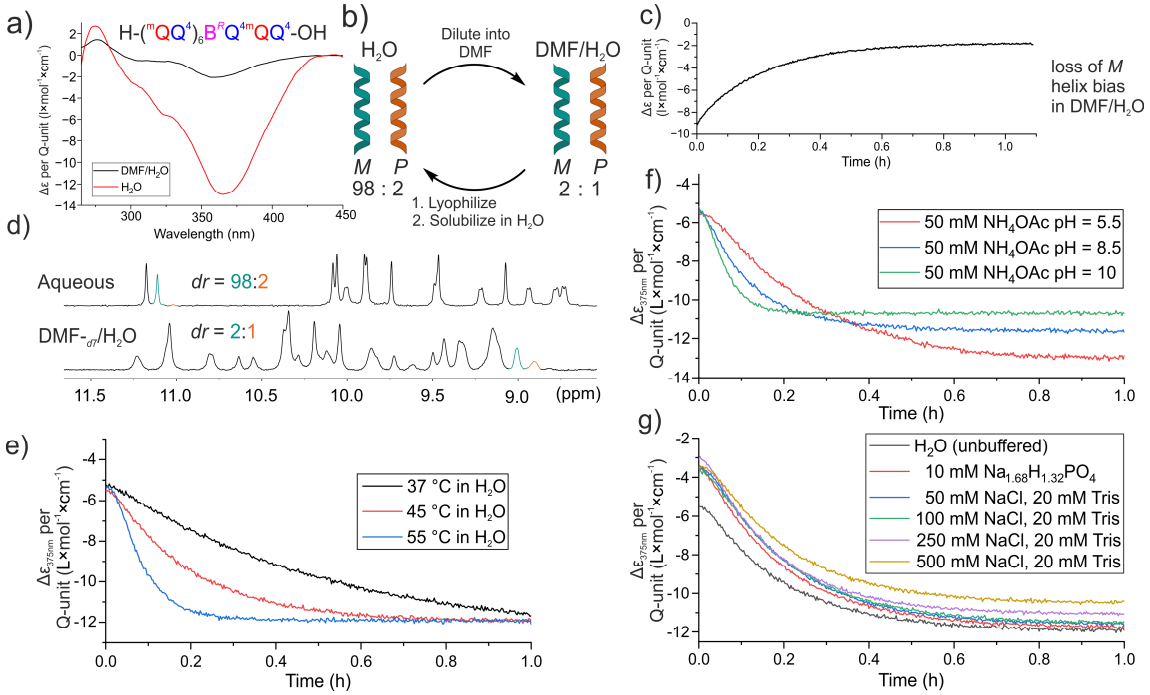


Fig. 10 (a) Chiral DNA mimic foldamer sequence **5** with the chiral unit incorporated (top) and its per quinoline normalized CD spectra in aqueous solution and DMF/H_2O at 25 °C. (b) Schematic display of diastereomeric enrichment assay to quantify foldamer helix dynamics. (c) Time-dependent CD monitoring at 375 nm of a water solubilized sample that has been freshly diluted with 9 times its volume of DMF displaying enrichment of the foldamer *P* helix. (d) Excerpts of ^{1}H NMR spectra in NH_4HCO_3 (50 mM in H_2O/D_2O, 9:1; v/v) buffer (top) and DMF-d_7/H_2O (9:1; v/v). of the amide region from 8.5 to 11.5 ppm. (e) Monitoring *M* helix enrichment in H_2O at varying temperatures. (f) Monitoring *M* helix enrichment in an NH_4OAc buffer system at different pH at 45 °C. (g) Monitoring *M* helix enrichment at different salt concentrations. All samples were buffered at pH = 7.5, unless stated otherwise.

M diastereomeric conformers in solid form. After dissolving this material in water at different pH, salt concentration, and temperature, the increase of CD intensity was monitored as a function of time as helix handedness inversion takes place to reach the new equilibrium diastereomeric ratio (Fig. 10). The data were fitted to a single-exponential decay function (1) to obtain their respective half-lives by equation (2).

$$y\,(x) = y_0 + A_1\, e^{\frac{-(x-x_0)}{t_1}} \qquad (1)$$

$$t\left(\frac{1}{2}\right) = x_0 + t_1 \ln (2) \qquad (2)$$

First, we analyzed the effect of temperature on helix handedness dynamics (Figs. 10e, S12, Table 2). Half-lives ranged from minutes at 55 °C to hours at 20 °C, highlighting the considerable stability of ^{m}QQ helices, and explaining why the equilibrium between *P* and *M* helices is slow on the NMR timescale unless the oligomer is very short (as for **4**, see: Fig. S4). The results also consistently showed the higher stability of ^{m}QQ^{5} sequences with respect to ^{m}QQ^{4} sequences. To keep balance between measurement time and accuracy, further measurements were carried out at 45 °C. Upon fitting the data, the initial time for the sample to reach 45 °C was accounted for by cutting the data off until an exponential decay was observed.

We then varied the pH from 5.5 to 10 (Figs. 10f, S13, Table 3) using adequate buffer systems. Increasing pH led to significantly faster helix handedness inversion kinetics, which appears to be in line with the stronger conformation dynamics observed in simulations. The destabilizing effect of increasing pH probably reflects an increase of electrostatic repulsions between phosphonate side chains when they are doubly charged.

Table 2 Half-lives of helix handedness inversion of 5 and 6 at different temperatures in H_2O.

T (°C)	20	37	45	55
5 (^{m}QQ^{4} series)	207 min	22.3 min	9.4 min	4.6 min
6 (^{m}QQ^{5} series)	526 min	66.1 min	18.5 min	8.1 min

Table 3 Half-lives of helix handedness inversion of 5 at different pH at 45 °C.

pH	5.5	8.5	10
5 (^{m}QQ^{4} series)	207 min	22.3 min	9.4 min

Table 4 Half-lives of helix handedness inversion of 5 at different salt concentrations at 45 °C.

[Na^+] in mM	0	16.8	50	100	250	500
5 (^{m}QQ^{4} series)	9.4 min	9.8 min	9.4 min	10.7 min	9.5 min	10.5 min

Nevertheless, one should keep in mind that the foldamer helices are still fully folded at high pH as reflected by their NMR spectra (Fig. 8) and their CD spectra (Fig. S14). Furthermore, the results presented below show that the effect of charges is more complex than just electrostatic repulsions. For instance, adding salt (NaCl up to 500 mM) had little effect on the helix handedness kinetics (Fig. 10g, Table 4), in line with the small effects seen on NMR spectra. Of note, adding Ca^{2+} or Mg^{2+} divalent cations led to precipitation of the foldamers.

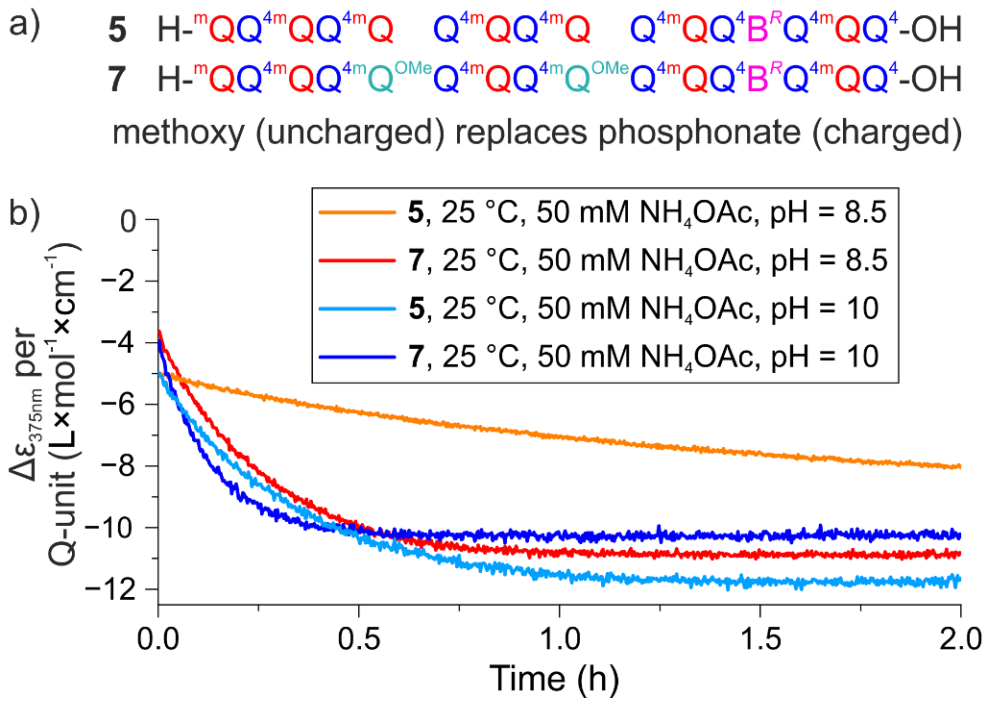


Fig. 11 (a) Chiral DNA mimic foldamer sequences 5 and 7. Compared to 5, 7 has two side-chain phosphonic acids replaced by methoxy side chains that lack charged residues. (b) Monitoring of *M* helix enrichment of 5 and 7 in NH_4OAc-buffer at 25 °C, pH 8.5 and pH 10.

To further investigate how the side chains and their charge state influence helix conformational dynamics, we prepared sequence **7**, an analogue of **5** in which two phosphonic acid-containing ^{m}Q residues have been replaced by $^{m}Q^{OMe}$ residues, whose side chains consist of a simple methoxy group (Figs. 1, 11a). Helix handedness inversion was eventually found to be faster with **7** than with **5**, so that measurements had to be carried out at 25 °C because they were too fast to measure at 45 °C. The half-life of helix handedness inversion at pH = 8.5 was calculated to be 11.2 min for **7** compared to 54.6 min for **5** (Figs. 11b, S15). The same trend was observed at pH = 10, though kinetics were overall faster (**7**: $t_{1/2}$ = 5.5 min, **5**: $t_{1/2}$ = 14.0 min). The fact that removing charges gives rise to faster helix handedness inversion kinetics shows that the negatively charged side chains do not have an exclusively destabilizing effect. Whether this directly relates to hydrogen bonding between side chains observed in MD simulations (Fig. 4) remains to be demonstrated.

Discussion

Several notable trends emerge from the combined computational and experimental investigations presented above. One important trend is that DNA mimic foldamers have an overall (in)flexibility, *e.g.*, bending persistence and twist persistence, similar to that of B-DNA despite their completely different chemical constitution (Table 1). The distinct backbone of DNA mimic foldamers is associated with local fluctuations and kinks that are unique to them,

but their overall shape persistence – not just their overall shape – nevertheless resembles that of DNA. Yet important differences include the absence of melting behavior in the DNA mimics, beyond the fact that they are single-stranded and cannot dissociate like DNA duplexes. Heating certainly enhances their internal dynamics but no transition to an unfolded state was observed and both CD and NMR indicate that the DNA mimics are still helically folded at 85 °C.

A second trend is the consistent observation that the $^{m}QQ^{5}$ oligomers are more stable than the $^{m}QQ^{4}$ oligomers. This is reflected in a lesser occurrence of kinks during MD simulations, and in longer bending and twist persistence along the $^{m}QQ^{5}$ helix (Table 1). It is also reflected in the lower destabilizing effect of an organic solvent such as DMF, resulting in a higher helix handedness bias in chiral $^{m}QQ^{5}$ sequence **6** compared to $^{m}QQ^{4}$ sequence **5** in that solvent. In addition, helix handedness inversion is slower in $^{m}QQ^{5}$ helices, regardless of the temperature considered (Table 2). Though it is not easy to directly relate these effects to the structural differences between Q^{4} and Q^{5}, it is possible to highlight these differences. The reason why the phosphonic acids are not connected to the quinoline ring with the same linker in Q^{4} (-OCH_2-) and in Q^{5} (-CH_2CH_2-) is not a design consideration but the result of different synthetic approaches. Yet it gives rise to distinct structural properties. The ethylene linker is less polar and more hydrophobic than the oxymethylene linker, possibly decreasing affinity for water molecules that fill the space between neighboring aromatic units during unstacking events. Furthermore, the two linkers have different conformational preferences. We have obtained the solid-state structures of numerous foldamer helices with Q units bearing a side chain in position 4 connected by an -OCH_2- linker. The general observation is that the $\underline{C}H_2$ carbon is found in the plane of the quinoline ring and close to position 3 (opposite to position 5). In contrast, the same carbon in -CH_2CH_2- linkers is generally found out of the plane of the quinoline ring (Fig. 12a).[43] Another difference between Q^{4} and Q^{5} concerns the protrusion of their side chain from the helix. As mentioned in the introduction, the design of Q^{5} was driven by the need for DNA mimic foldamers with a wider major groove and a narrower minor groove (Fig. 1).[11] Yet a difference that was not highlighted before is that phosphonic acids lie closer to the helix axis in Q^{4} monomers than in Q^{5} and ^{m}Q monomers. This is clearly visible in top views of solid state structures and molecular models (Fig. 12b,c). That the side chains of Q^{4} residues are somewhat closer to the helix axis implies that they are also slightly closer to each other.

A third trend is the strong effect of pH on the DNA mimic foldamer structural dynamics. Increasing the proportion of doubly charged phosphonate residues – *i.e.*, increasing pH – results in more frequent kinks in MD simulations. It is also reflected in changes in NMR spectra though these confirm that helix folding is maintained even at pH > 12. Increasing pH also enhances the kinetics of handedness inversion. However, these effects cannot all be interpreted simply in terms of electrostatic repulsions between negatively charged side chains. First, adding monovalent salts had little effect on the NMR spectra and on the rate of helix handedness inversion. Second, removing some negatively charged side chains resulted in

faster helix handedness inversion. To interpret this fact, one must consider that the kinetics of helix handedness inversion are not directly the reflection of helix stability as would be the transition between a folded and an unfolded state. Instead, they reflect the energy difference between the folded state and a higher energy state enabling handedness inversion. Modeling studies suggested that this higher energy state is a helix with a local misfold – a local reversal of handedness can be generated by flipping two residues – that can propagate along the helix.[21,44] It may well be that these misfolded states involve energetically costly contacts between phosphonate side chains and hydrophobic quinoline faces. When using a short nonpolar side chain as in $^{m}Q^{OMe}$, more favorable contacts may occur, leading to a stabilization of these intermediate states and thus to faster kinetics.

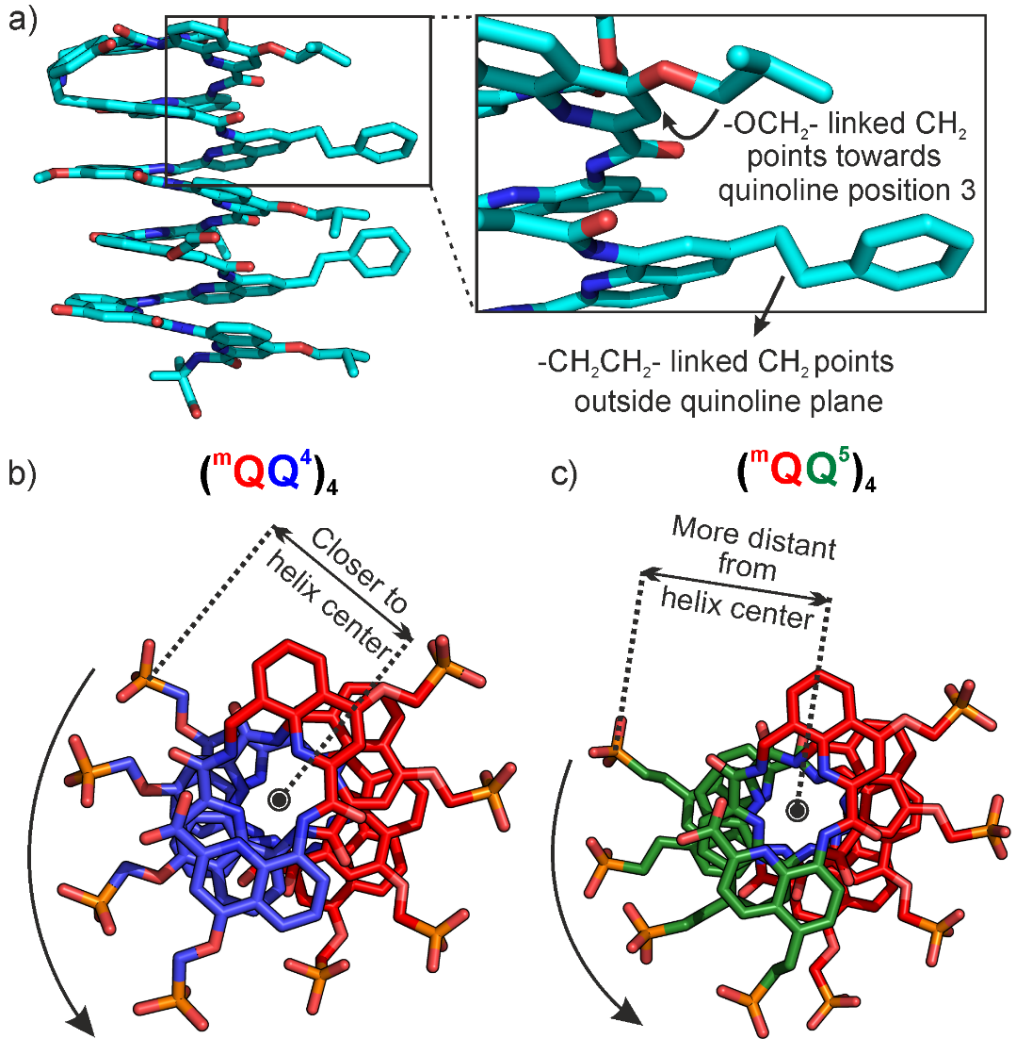


Fig. 12 (a) Crystal structure of a foldamer[43] that has side chain attached by both -CH_2CH_2- and -OCH_2- linkages. Typical orientation of $\underline{C}H_2$ bridges are depicted in the close-up. Crystal structures of diethyl phosphonate ester protected Q^4 and Q^5 based octamers (b) and (c).[11] Diethyl esters as well as C- and N-terminal protecting groups are omitted for clarity. The distance between Q^4 side chains at i, at i+2 and i–2 positions is smaller than the distance between Q^5 side chains at i, at i+2 and i–2 positions.

Conclusion

DNA mimic foldamers, aromatic oligoamides composed of $^{m}QQ^{4}$ or $^{m}QQ^{5}$ repeat motifs that adopt helical conformations reproducing the structure and charge distribution of B-type double-stranded DNA, have recently been designed. These molecules are promising candidates as competitive inhibitors of protein-DNA complexes and have potential to serve as pharmacological tools. We have investigated their structural dynamics both computationally and experimentally. For this purpose, force field parameters were optimized and an assay to experimentally assess helix stability was developed.

The helical conformations of the DNA mimic foldamers are stable over a wide range of temperature and pH in aqueous solutions and no melting behavior was observed upon heating. The $^{m}QQ^{5}$ oligomers show a higher stability than the $^{m}QQ^{4}$ oligomers. MD simulations revealed different types of kinking in the foldamer helices associated with more or less long-lived unstacking events between the aromatic monomers. At higher pH, when the side chains are doubly negatively charged, kinking events become more frequent and lasting. This resulted in a chain containing small helical segments interrupted by kinked steps and an overall irregular conformation but in which the overall helix integrity – its handedness – was preserved, unlike in polyanionic or polycationic α-helical peptides. The structural dynamic parameters of the DNA mimic helices, including bending, twisting and stretching flexibility, were found to be in the same range as those of DNA. This result extends the resemblance between double stranded B-DNA and DNA mimic foldamer beyond structure to some dynamic properties. This resemblance is remarkable considering that they are based on completely different backbones and that local conformational changes associated with these backbones are also completely different.

Progress is being made both in the structural elucidation of DNA mimic foldamer-protein complexes[17] and in the production of hybrid molecules combining a foldamer and a DNA segment.[16] In this context, the results and methods reported here will be useful to design distinct foldamers that specifically recognize DNA-binding proteins as well as other backbones that may also mimic B-DNA or other nucleic acid structures.

Author contributions

ML, JW, and VC performed syntheses and experimental studies. LT, ZL, and MZ performed computations. VP, MZ and IH supervised the research. ML, VC, LT, MZ and IH wrote the manuscript. All authors reviewed and edited the manuscript and approved its final version.

Conflicts of interest

There are no conflicts to declare.

Data availability

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

Acknowledgements

We acknowledge financial support from the Deutsche Forschungsgemeinschaft (DFG) via projects SFB863/A10 and SFB1309/C07, and from European Union's Horizon Europe Framework Program under Grant Agreement No. ERC-2021-ADG-320892. We thank the Leibniz super computer (LRZ) center for computer resources under grant pr27za and Dr. Lars Allmendinger for assistance with NMR measurements and DR. Maksim Rogovoi for providing some synthetic precursors.

Notes and references

1. M. J. Kamali, M. Salehi, S. Fatemi, F. Moradi, A. Khoshghiafeh and M. Ahmadifard, *Exp. Cell Res.*, 2023, **423**, 113442.
2. B. Hyrup and P. E. Nielsen, *Bioorg. Med. Chem.*, 1996, **4**, 5-23.
3. G. Casas, F. Perche, P. Midoux, C. Pichon and J.-M. Malinge, *Mol. Ther. Nucleic Acids*, 2022, **29**, 162-175.
4. B. Johari and M. Moradi, *Methods Mol. Biol. (N. Y., NY, U. S.)*, 2022, **2521**, 207-230.
5. S. Yasuda, K. Morihiro, S. Koga and A. Okamoto, *Angew. Chem., Int. Ed.*, 2025, **64**, e202424421.
6. D. T. F. Dryden, *Trends Biotechnol.*, 2006, **24**, 378-382.
7. H.-C. Wang, C.-C. Chou, K.-C. Hsu, C.-H. Lee and A. H. J. Wang, *IUBMB Life*, 2019, **71**, 539-548.
8. H.-C. Wang, C.-H. Ho, K.-C. Hsu, J.-M. Yang and A. H. J. Wang, *Biochemistry*, 2014, **53**, 2865-2874.
9. D. Yuksel, P. R. Bianco and K. Kumar, *Mol. BioSyst.*, 2016, **12**, 169-177.
10. S. N. Bolten, U. Rinas and T. Scheper, *Appl. Microbiol. Biotechnol.*, 2018, **102**, 8647-8660.
11. 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 and I. Huc, *Nat. Chem.*, 2018, **10**, 511-518.
12. V. Corvaglia, D. Carbajo, P. Prabhakaran, K. Ziach, P. K. Mandal, V. Dos Santos, C. Legeay, R. Vogel, V. Parissi, P. Pourquier and I. Huc, *Nucleic Acids Res.*, 2019, **47**, 5511-5521.
13. V. Corvaglia, I. A. M. Amar, V. Garambois, S. Letast, A. Garcin, C. Gongora, M. Del Rio, C. Denevault-Sabourin, N. Joubert, I. Huc and P. Pourquier, *Pharmaceuticals*, 2021, **14**, 624.
14. V. Corvaglia, J. Wu, D. Deepak, M. Loos and I. Huc, *Chem. - Eur. J.*, 2024, **30**, e202303650.
15. V. Kleene, V. Corvaglia, E. Chacin, I. Forne, D. B. Konrad, P. Khosravani, C. Douat, C. F. Kurat, I. Huc and A. Imhof, *Nucleic Acids Res.*, 2023, **51**, 9629-9642.
16. M. Loos, F. Xu, P. K. Mandal, T. Chakrabortty, C. Douat, D. B. Konrad, M. Cabbar, J. Singer, V. Corvaglia, T. Carell, and I. Huc, *Angew. Chem., Int. Ed.*, 2025, **64**, e202505273.
17. D. Deepak, J. Wu, V. Corvaglia, L. Allmendinger, M. Scheckenbach, P. Tinnefeld and I. Huc, *Angew. Chem., Int. Ed.*, 2025, **64**, e202422958.
18. H. Jiang, J.-M. Léger and I. Huc, *J. Am. Chem. Soc.*, 2003, **125**, 3448-3449.
19. C. Dolain, A. Grélard, M. Laguerre, H. Jiang, V. Maurizot and I. Huc, *Chem. - Eur. J.*, 2005, **11**, 6135-6144.

20. T. Qi, V. Maurizot, H. Noguchi, T. Charoenraks, B. Kauffmann, M. Takafuji, H. Ihara, I. Huc, *Chem. Commun.*, 2012, **48**, 6337-6339.
21. N. Delsuc, T. Kawanami, J. Lefeuvre, A. Shundo, H. Ihara, M. Takafuji and I. Huc, *ChemPhysChem*, 2008, **9**, 1882-1890.
22. For an illustration of the concept of exo-helices in peptides, see: J. M. Martinez-Parra, R. Gomez-Ojea, G. A. Daudey, M. Calvelo, H. Fernandez-Caro, J. Montenegro and J. Bergueiro, *Nat. Commun.*, 2024, **15**, 6987.
23. N. H. Martin, N. W. Allen and J. C. Moore, *J. Mol. Graphics Modell.*, 2000, **18**, 242-246.
24. M. B. Ferraro, P. Lazzeretti, R. G. Viglione and R. Zanasi, *Chem. Phys. Lett.*, 2004, **390**, 268-271.
25. F.-Y. Dupradeau, A. Pigache, T. Zaffran, C. Savineau, R. Lelong, N. Grivel, D. Lelong, W. Rosanski and P. Cieplak, *Phys. Chem. Chem. Phys.*, 2010, **12**, 7821-7839.
26. J. Wang, P. Cieplak and P. A. Kollman, *J. Comput. Chem.*, 2000, **21**, 1049-1074.
27. Z. Liu, A. M. Abramyan and V. Pophristic, *New J. Chem.*, 2015, **39**, 3229-3240.
28. Z. Liu, A. Teslja and V. Pophristic, *J. Comput. Chem.*, 2011, **32**, 1846-1858.
29. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.
30. D. A. Case, I. Y. Ben-Shalom, S. R. Brozell, D. S. Cerutti, T. E. Cheatham, III, V. W. D. Cruzeiro, T. A. Darden, R. E. Duke, D. Ghoreishi, M. K. Gilson, H. Gohlke, A. W. Goetz, D. Greene, R Harris, N. Homeyer, Y. Huang, S. Izadi, A. Kovalenko, T. Kurtzman, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. J. Mermelstein, K. M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. R. Roc, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C. L. Simmerling, J. Smith, R. Salomon-Ferrer, J. Swails, R. C. Walker, J. Wang, H. Wei, R. M. Wolf, X. Wu, L. Xiao, D. M. York, and P. A. Kollman, AMBER2018, University of California, San Francisco, CA, 2018.
31. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926.
32. C. W. Hopkins, S. Le Grand, R. C. Walker and A. E. Roitberg, *J. Chem. Theory Comput.*, 2015, **11**, 1864-1874.
33. J. Lipfert, G. M. Skinner, J. M. Keegstra, T. Hensgens, T. Jager, D. Dulin, M. Köber, Z. Yu, S. P. Donkers, F. Chou, R. Das and N.H. Dekker, *Proc. Natl. Acad. Sci. U.S.A.*, 2014, **111**, 15408-15413.
34. V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat and I. Huc, *Chem. - Eur. J.*, 2023, **29**, e202300898.
35. D. Bindl, E. Heinemann, P. K. Mandal and I. Huc, *Chem. Commun.*, 2021, **57**, 5662-5665.
36. E. A. Gooding, S. Sharma, S. A. Petty, E. A. Fouts, C. J. Palmer, B. E. Nolan and M. Volk, *Chem. Phys.*, 2013, **422**, 115-123.
37. M. Vallade, P. Sai Reddy, L. Fischer and I. Huc, *Eur. J. Org. Chem.*, 2018, **2018**, 5489-5498.
38. C. Dolain, H. Jiang, J.-M. Leger, P. Guionneau and I. Huc, *J. Am. Chem. Soc.*, 2005, **127**, 12943-12951.
39. A. M. Kendhale, L. Poniman, Z. Dong, K. Laxmi-Reddy, B. Kauffmann, Y. Ferrand and I. Huc, *J. Org. Chem.*, 2011, **76**, 195-200.
40. F. S. Menke, B. Wicher, L. Allmendinger, V. Maurizot and I. Huc, *Chem. Sci.*, 2023, **14**, 3742-3751.
41. N. Delsuc, F. Godde, B. Kauffmann, J.-M. Léger and I. Huc, *J. Am. Chem. Soc.*, 2007, **129**, 11348-11349.
42. D. Sánchez-García, B. Kauffmann, T. Kawanami, H. Ihara, M. Takafuji, M.-H. Delville and I. Huc, *J. Am. Chem. Soc.*, 2009, **131**, 8642-8648.
43. M. Zwillinger, P. S. Reddy, B. Wicher, P. K. Mandal, M. Csekei, L. Fischer, A. Kotschy and I. Huc, *Chem. - Eur. J.*, 2020, **26**, 17366-17370.
44. A. M. Abramyan, Z. Liu and V. Pophristic, *Chem. Commun.*, 2016, **52**, 669-672.

5.2 Supporting information

5.2.1 Supplementary figures

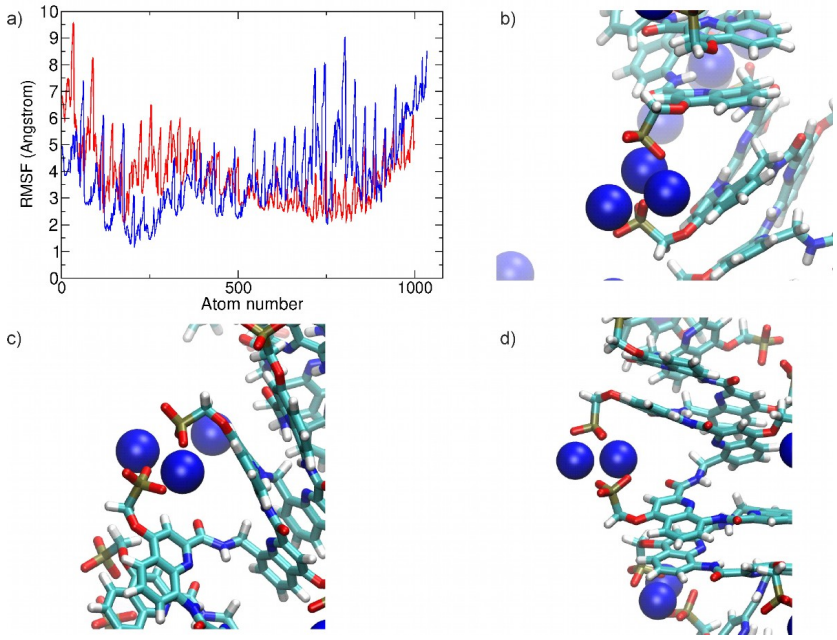


Figure S1 (a) Root-mean-square fluctuations (RMSF) of non-hydrogen atoms *vs*. atom number observed during $(^{m}QQ^{4})_{18}$ simulations (blue line) and $(^{m}QQ^{5})_{18}$ simulations (red) with doubly charged phosphonate groups. (b-d) Simulation snapshots indicating trapped binding of several sodium ions (blue spheres) between double anionic phosphonate groups (stick models, $^{m}QQ^{4}$ case).

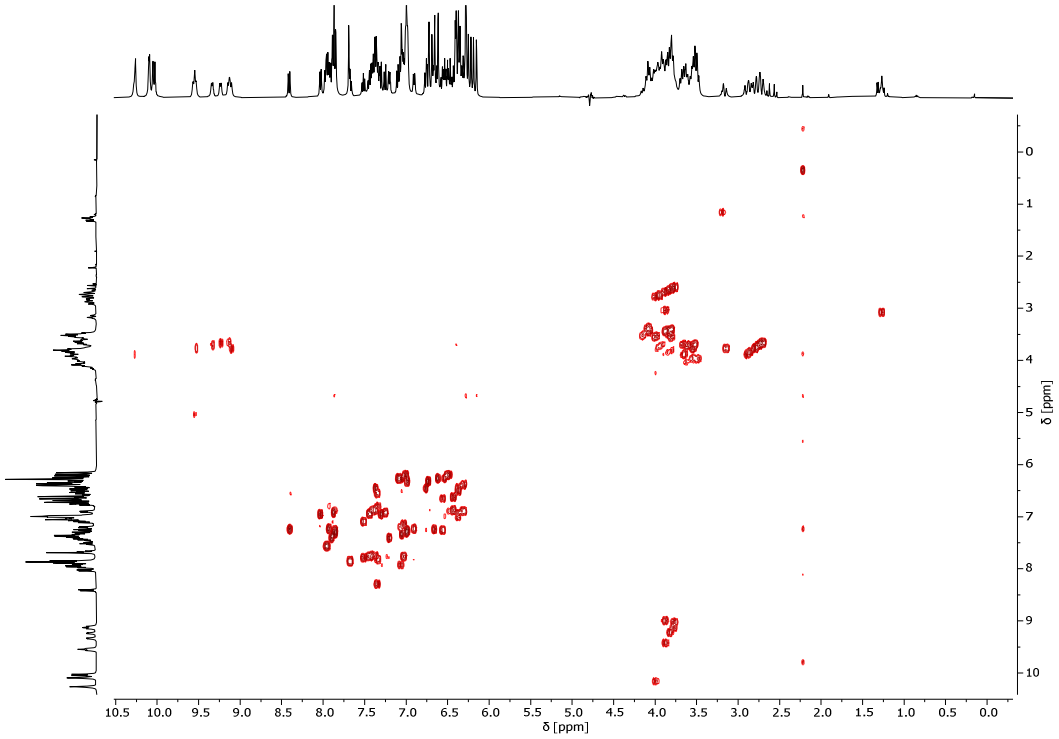


Figure S2 2D COSY NMR spectrum of **2** recorded with water suppression at 25 °C in H_2O/D_2O (9:1, v/v). After dissolving, the solution was measured to be at pH 8.

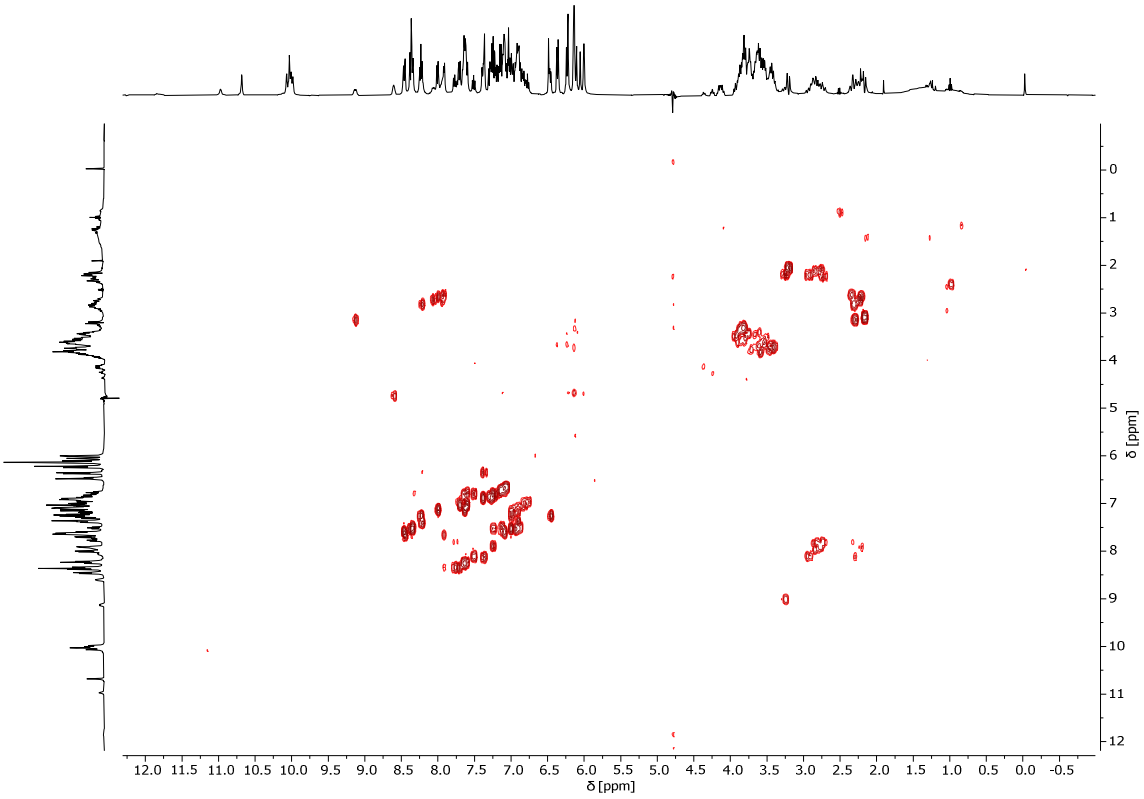


Figure S3 2D COSY NMR spectrum of **2** recorded with water suppression at 25 °C in 50 mM sodium hydroxide H_2O/D_2O (9:1, v/v) at pH 12.5.

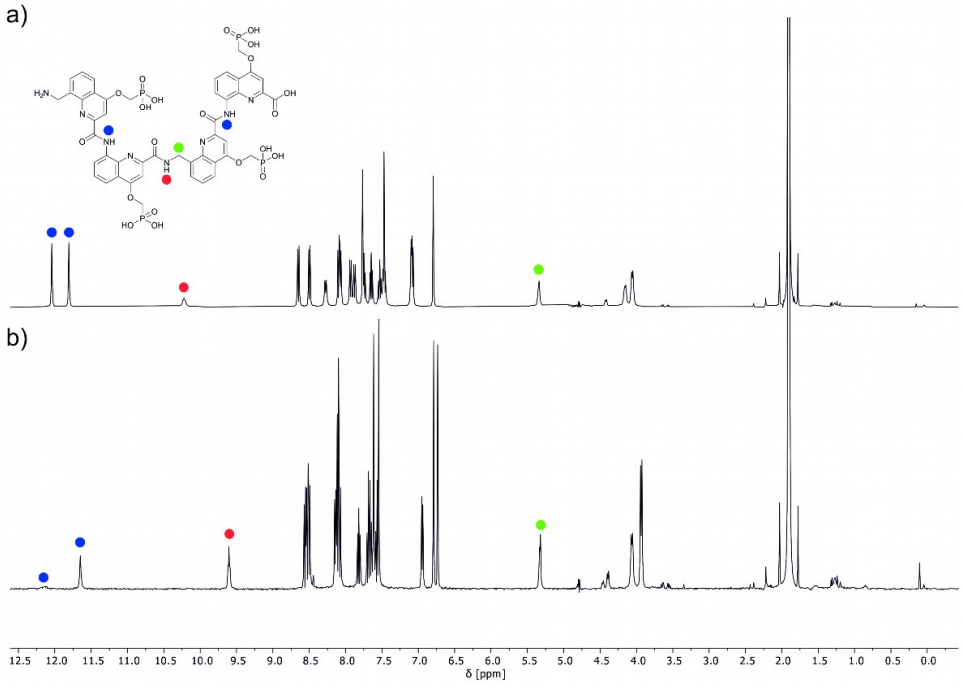


Figure S4 ^1H NMR spectra of **4** recorded with water suppression at 25 °C in H_2O/D_2O (9:1, v/v) at pH 8 (a) and 50 mM sodium hydroxide H_2O/D_2O (9:1 v/v) at pH 12.5 (b). Blue, red and green balls indicate aromatic amide protons, benzylic amide protons and benzylic CH_2 protons, respectively.

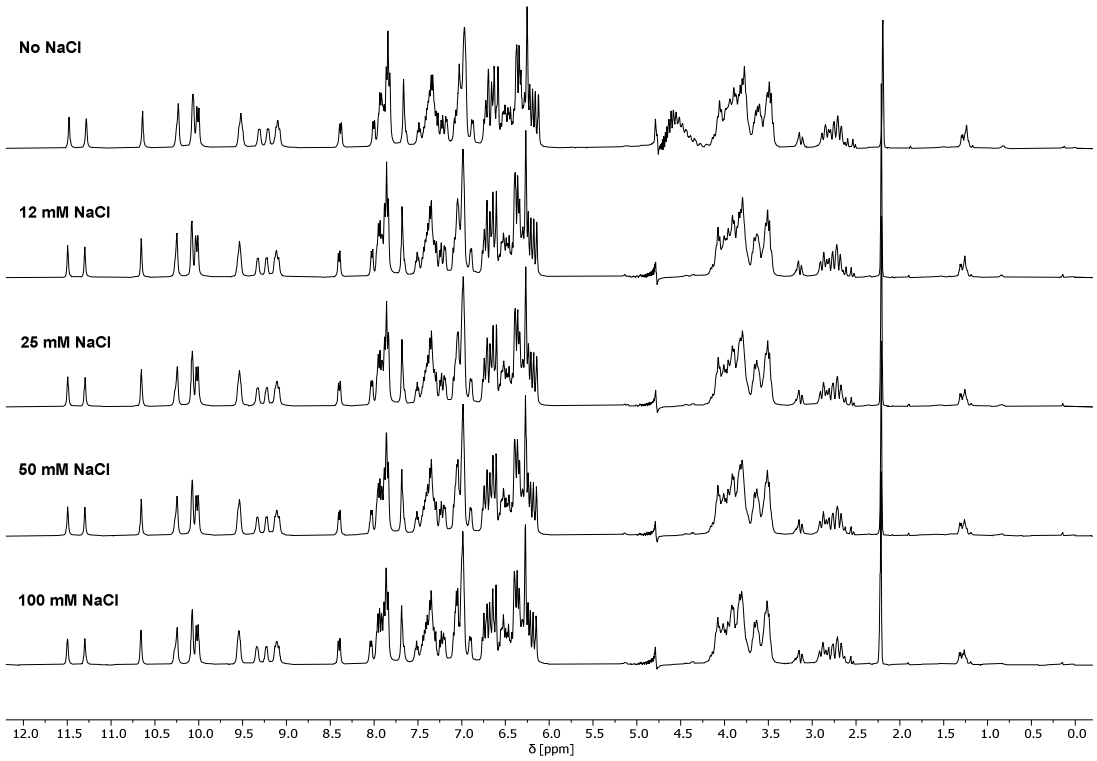


Figure S5 ^{1}H NMR spectra of **2** recorded with water suppression at 25 °C in H_2O/D_2O (9:1, v/v) at pH 8, with increasing NaCl concentration.

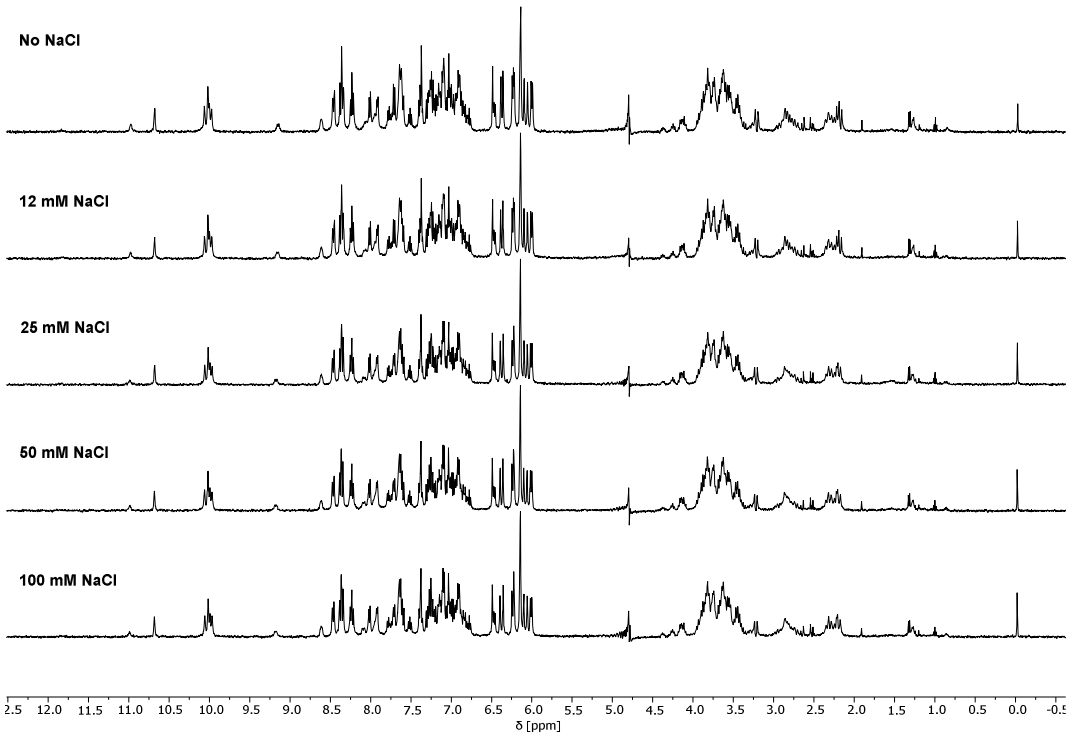


Figure S6 ^{1}H NMR spectra of **2** recorded with water suppression at 25 °C in H_2O/D_2O (9:1, v/v) at pH 12.5, with increasing NaCl concentration.

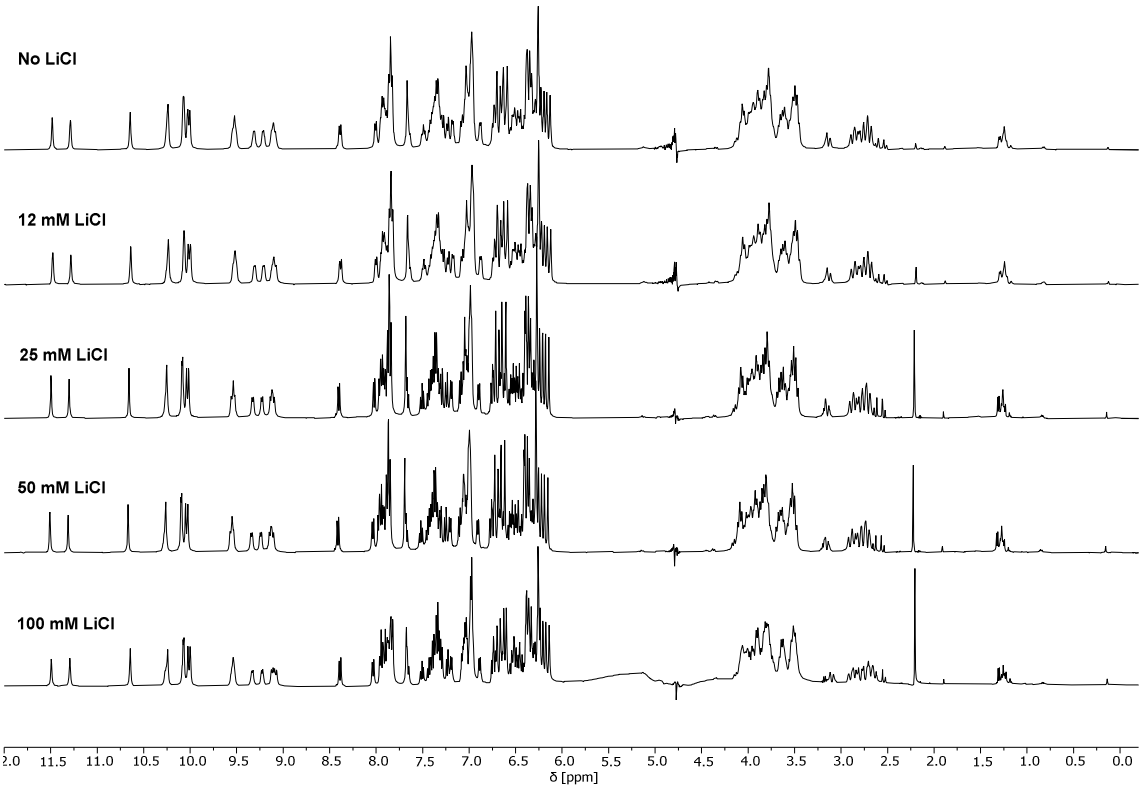


Figure S7 ^{1}H NMR spectra of **2** recorded with water suppression at 25 °C in H_2O/D_2O (9:1, v/v) at pH 8, with increasing NaCl concentration.

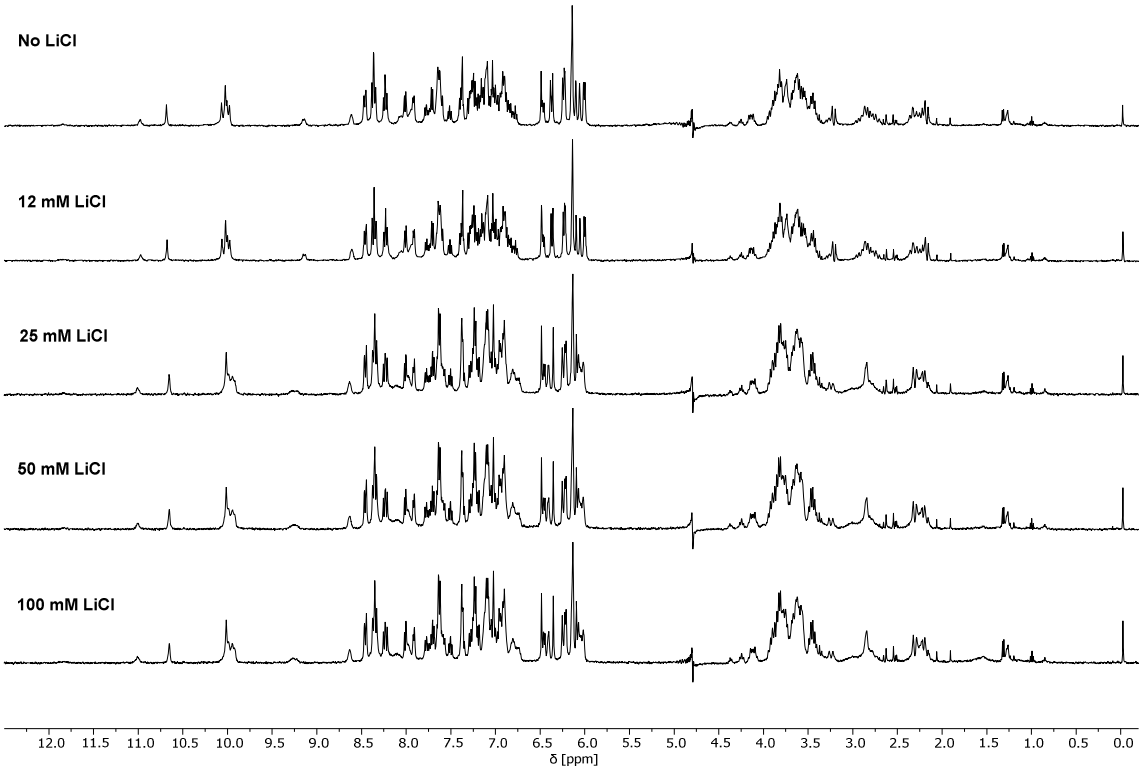


Figure S8 ^{1}H NMR spectra of **2** recorded at 25 °C in H_2O/D_2O (9:1 v/v) at pH 12.5 with increasing LiCl concentration.

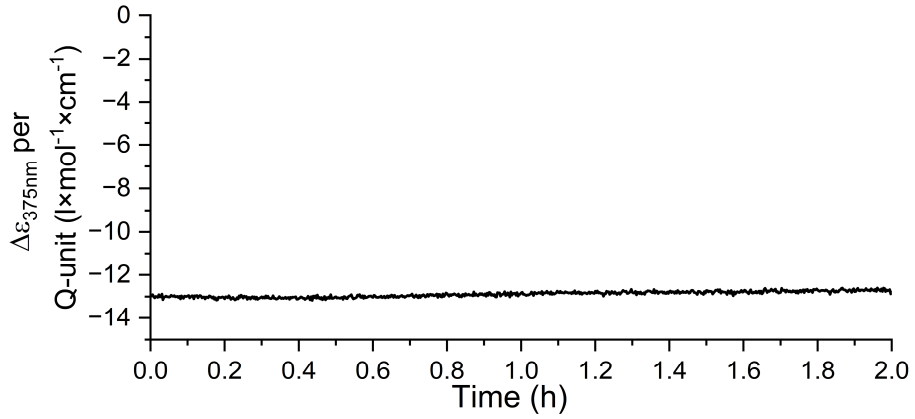


Figure S9 Time-dependent monitoring of the CD-band at 375 nm of compound **5**, which was dissolved in H_2O and freshly diluted into nine times its volume of H_2O.

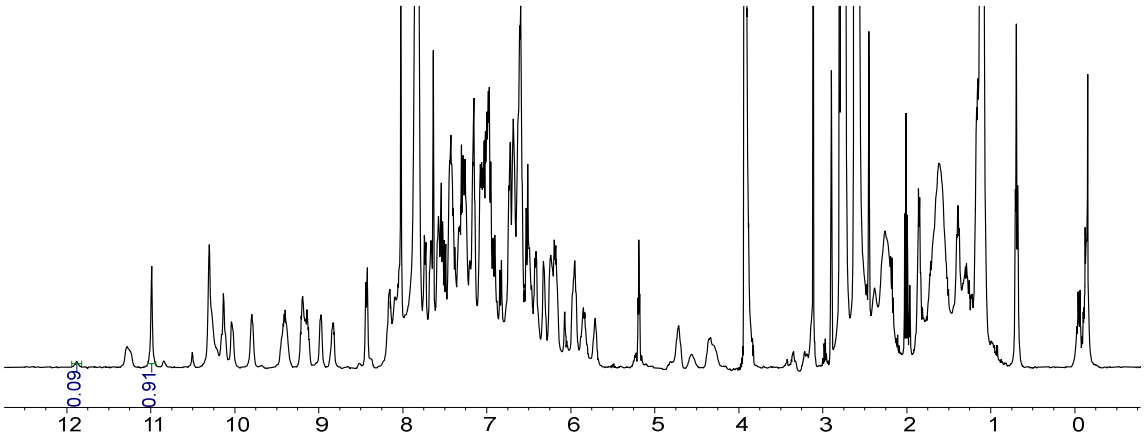


Figure S10 ^{1}H NMR of compound **6** in (DMF-d_7/H_2O, 9:1, v/v) at 25 °C.

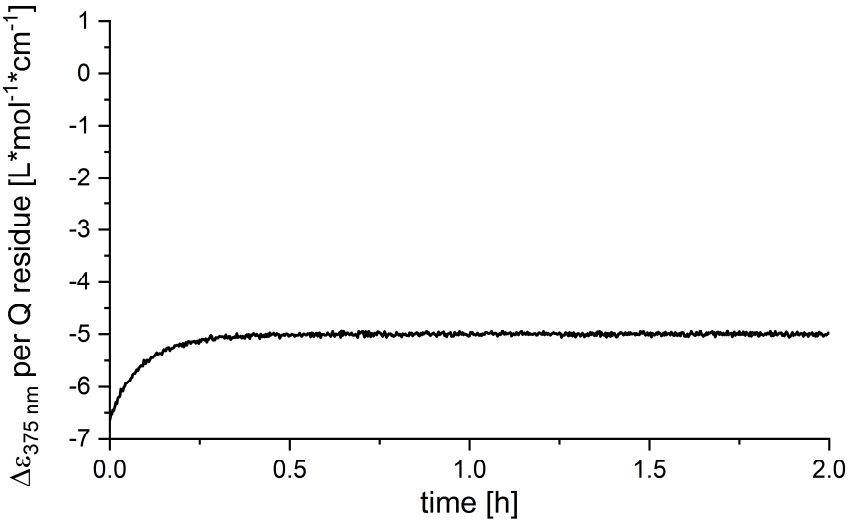


Figure S11 Time-dependent monitoring of the CD-band at 375 nm of compound **6**, which was dissolved in H_2O and freshly diluted into nine times its volume of DMF, showing enrichment of the foldamer P-helix.

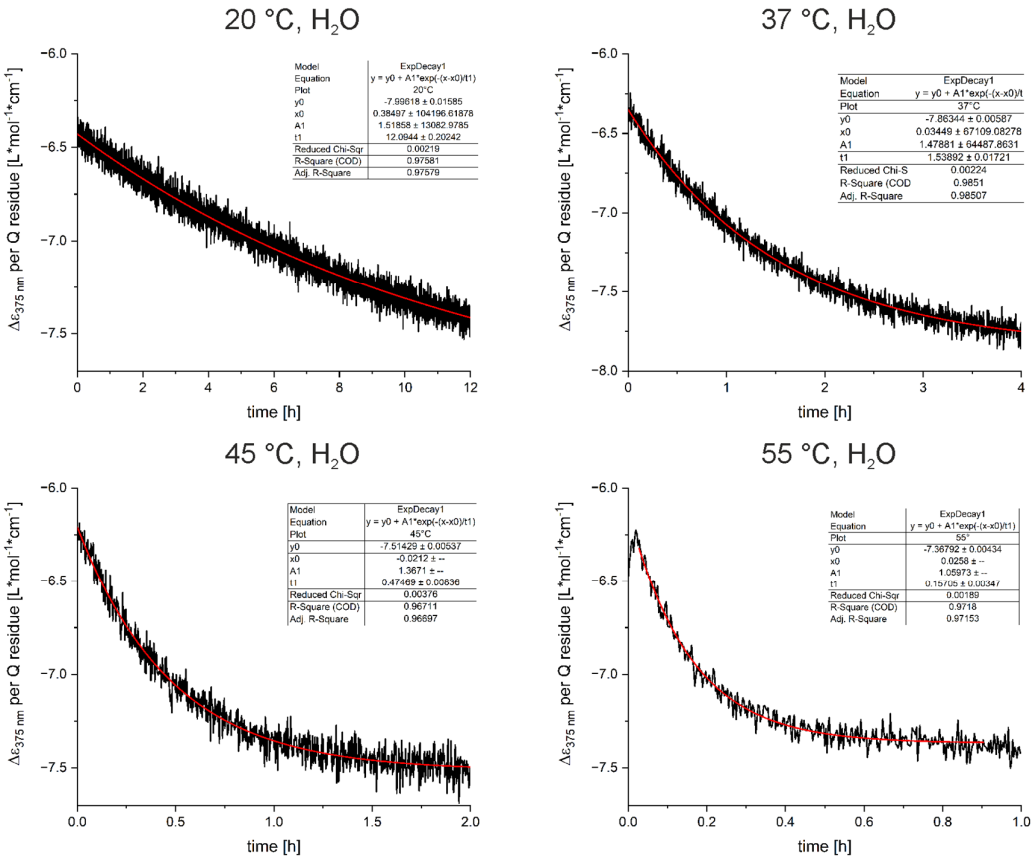


Figure S12 Helix-handedness enrichment of **6**, showing conversion of excess *P*-helix to *M*-helix (black), and a single-exponential decay fit to the corresponding data (red) in H_2O at different temperatures.

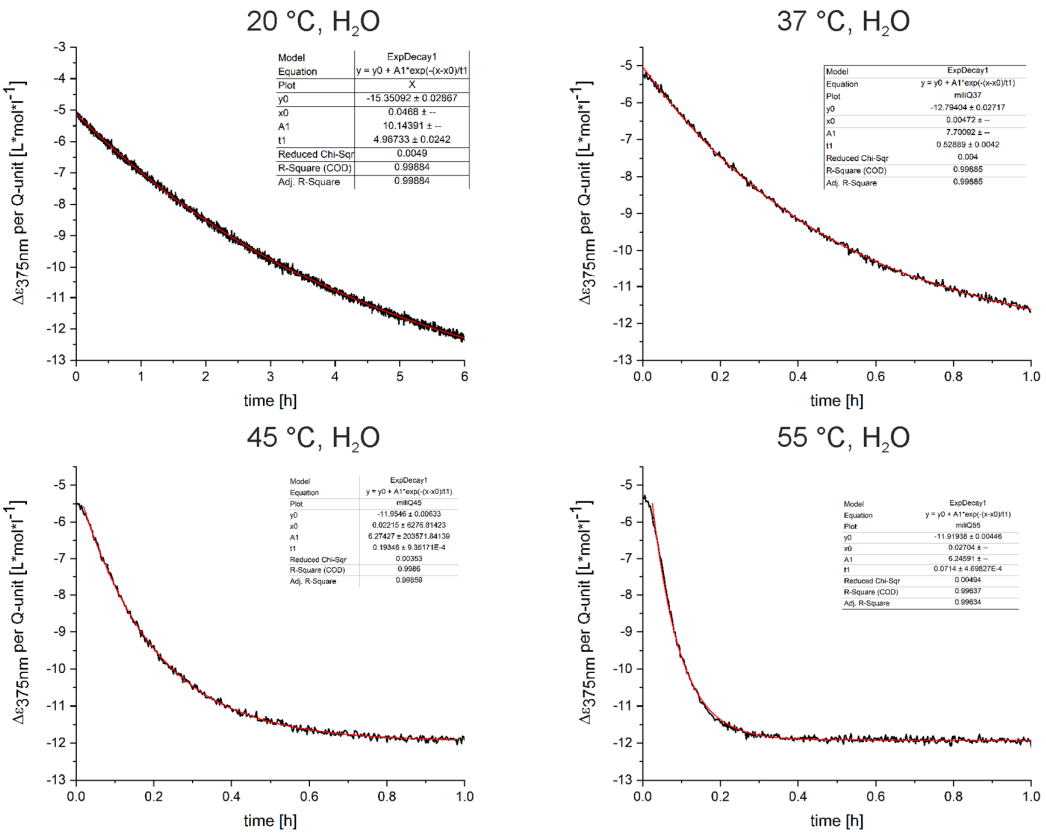


Figure S13 Helix-handedness enrichment of **5**, showing conversion of excess *P*-helix to *M*-helix (black), and a single-exponential decay fit to the corresponding data (red) in H_2O at different temperatures.

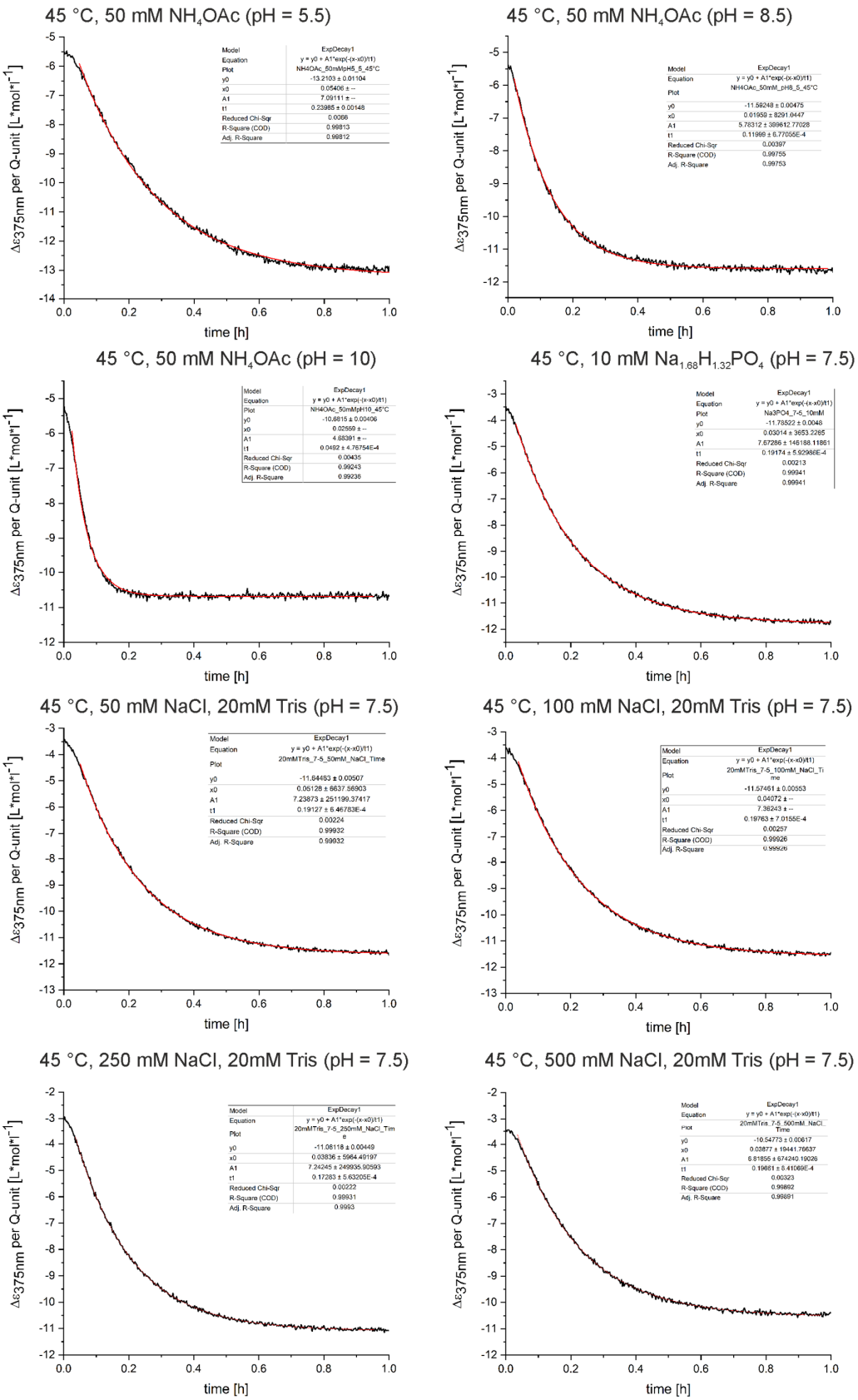


Figure S14 Helix-handedness enrichment of **5**, showing conversion of excess *P*-helix to *M*-helix (black), and a single-exponential decay fit to the corresponding data (red) in H_2O at different pH values and salt-concentrations.

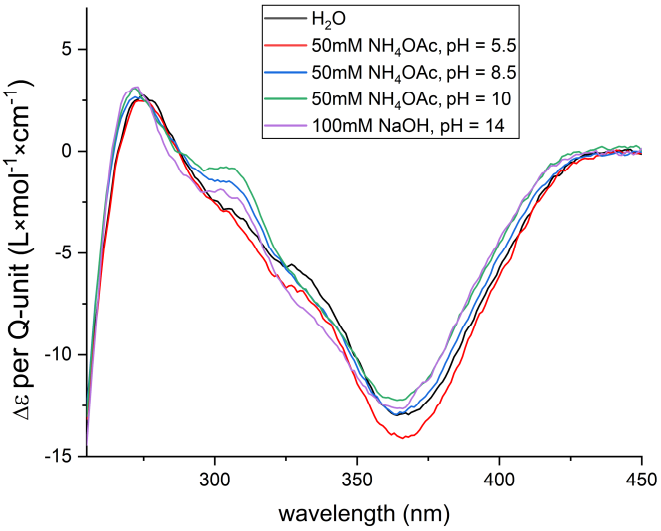


Figure S15 CD-spectra of **5** at different pH values at 25 °C.

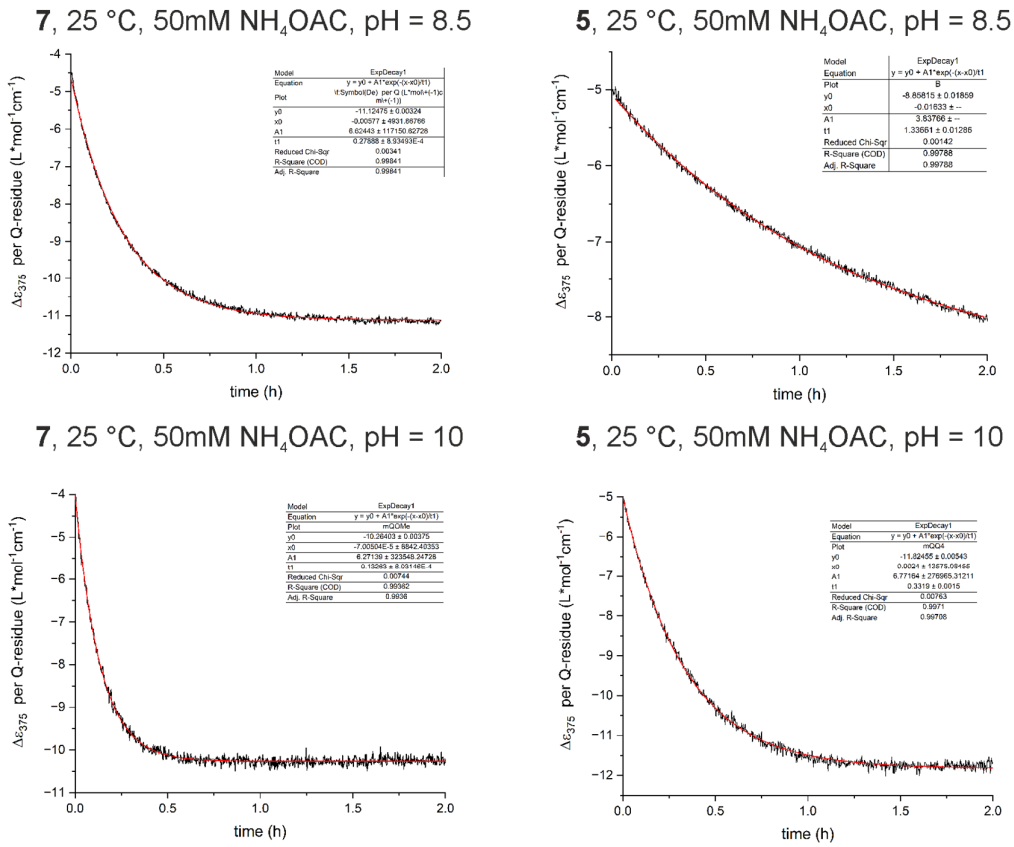


Figure S16 Helix-handedness enrichment of **5** and **7**, showing conversion of excess *P*-helix to *M*-helix (black), and a single-exponential decay fit to the corresponding data (red) in H_2O at different pH values.

5.2.2 Materials and Methods

5.2.2.1 Materials

Reagents were used as purchased from commercial sources without further purification. Column chromatography purifications were performed on silica gel (230-400 mesh, 40-63 µm, Merck). Thin-layer chromatography was performed on silica gel plates (60-F254, Merck). Reactions requiring anhydrous conditions were performed under nitrogen with commercial anhydrous solvents unless stated otherwise. Anhydrous THF for solid-phase synthesis was dispensed from a *MBRAUN SPS-800* solvent purification system using alumina columns for drying. $CHCl_3$ was freshly distilled over CaH_2 under N_2-atmosphere. Ultrapure water was obtained via a Stakpure OmniaPure-T UV-TOC ultrapure water system.

5.2.2.2 Methods

5.2.2.2.1 Small molecule nuclear magnetic resonance spectroscopy and mass spectrometry

NMR spectra were recorded on AVANCE NEO NMR spectrometer 500 MHz (Bruker BioSpin) with CryoProbe™ Prodigy and a BCU II. $CDCl_3$ (δ_H: 7.26, δ_C: 77.0), DMSO-d_6 (δ_H: 2.50, δ_C: 39.4) were used as solvents and their residual solvent signals were used as internal standards.[1] The derived data signals are stated with chemical shift in ppm, their multiplicity (s = singlet, bs = broad singlet, d = doublet, t = triplet, q = quadruplet, m = multiplet, or a combination of these), their coupling constant in Hz and their integrated values. Small molecule mass spectra were recorded on a micrOTOF II mass spectrometer by Bruker Daltonics and ionized by ESI.

5.2.2.2.2 Oligomer NMR spectroscopy

One- and two-dimensional spectra were recorded at 25 °C in NH_4HCO_3 (50 mM, H_2O/D_2O (9:1 v/v)) at pH 8 or adjusted to pH 12.5 with NaOH on an Avance III HD 500 MHz Bruker BioSpin spectrometer equipped with a CryoProbe™ Prodigy. For small molecule analysis and measurements with variable temperature (VT) control, an AVANCE NEO NMR spectrometer 500 MHz (Bruker BioSpin) with CryoProbe™ Prodigy and a BCU II were used. Chemical shifts are reported in ppm and are referenced against an internal standard. Data processing was performed with MestReNova NMR processing software (v.12.0.0) from Mestrelab Research. 2D homonuclear correlation spectroscopy (COSY) spectra were recorded with a phase-sensitive pulse sequence with water suppression employing a Watergate pulse scheme from the Bruker pulse program library (cosygpphwpp5). Data acquisition was performed with 1K (F2) x 256 (F1) data points. The recycling delay was 1.0 s and 64 transients per increment were applied at a sweep width of 8 kHz in both dimensions resulting in an acquisition time of 0.1204 s. Automatic phase correction as well as baseline correction was applied in both dimensions.

5.2.2.2.3 RP-HPLC chromatography and LC-MS

Analytical RP-HPLC analysis and semi-preparative purifications were performed with 4 different buffer systems. TEAA- and NH_4OAc buffer systems were adjusted to the given pH-values by adding NEt_3 for TEAA and aqueous NH_3 (1 M) NH_4OAc on a Mettler ToledoTM SevenCompact pH meter.

Analytical RP-HPLC analysis and semi-preparative purifications (eg. Compound **5a** and **6a**) were performed with a TFA buffered system 0.1% TFA in ultra-pure water (referred to as mobile phase A) and 0.1% TFA in HPLC-grade acetonitrile (referred to as mobile phase B) were used on a Thermo Fisher Scientific Ultimate 3000 HPLC system equipped with Macherey-Nagel Nucleodur C18 columns (4.6 × 100 mm, 5 µm) and a flowrate of 1 mL/min for analysis and a Waters system equipped with a 2545 Quaternary Gradient Module and a XBridge® Prep C8 OBD™ column (19 × 150 mm, 5 µm) and a flowrate of 25 mL/min for purifications. LC-MS analysis of these compounds were performed on a a Thermo Fisher Scientific Ultimate 3000 HPLC system equipped with Macherey-Nagel Nucleodur C18 gravity (2× 50 mm, 1.8 µm) and 0.18% formic acid and 0.02% TFA in ultra-pure water (referred to as mobile phase A) and 0.18% formic acid and 0.02% TFA in LCMS-grade acetonitrile (referred to as mobile phase B) and a flowrate of 0.33 mL/min. Analytical RP-HPLC analysis and semi-preparative purifications in basic conditions (e.g. compound **5** and **6**) were performed on a Thermo Fisher Scientific Ultimate 3000 HPLC system using Macherey-Nagel Nucleodur C18 columns (4.6 × 150 mm, 5 µm) with a triethylammonium acetate buffer system with a flowrate of 1 mL/min. The mobile phase was composed of 12.5 mM TEAA in water at pH 8.5 (A) and 12.5 mM TEAA in water: acetonitrile mixture (1:2, v/v) at pH = 8.5 (B). Purifications in basic conditions were done on a Thermo Fisher Scientific Ultimate 3000 HPLC system equipped with a a Kinetex C18 EVO column (10 x 100 mm, 5 µm) and a flowrate of 5 mL/min. LCMS analysis of these compounds were conducted on a Thermo Fisher Scientific Ultimate 3000 HPLC system with an NH_4OAc buffer system consisting of 12.5 mM NH_4OAc dissolved in ultra-pure water and adjusted to pH = 8.5 with aqueous ammonia (referred to as mobile phase A) and LCMS-grade acetonitrile (referred to as mobile phase B) on a Kinetex C18 EVO column (2.1 x 50 mm, 1.8 µm) column and a flowrate of 0.33 mL/min. In all cases, elution was monitored by UV detection at 254 and 300 nm with a diode array detector. TEAA- and NH_4OAc buffer systems were adjusted to the given pH-values by adding NEt_3 for TEAA and aqueous NH_3 (1 M) NH_4OAc on a Mettler ToledoTM SevenCompact pH meter. Formic acid/TFA and TFA buffer systems were not adjusted. For LCMS analysis, the LC system was coupled to a micrOTOF II mass spectrometer by Bruker Daltonics and molecules were ionized by ESI.

5.2.2.2.4 Circular dichroism (CD) and UV-Vis absorption spectroscopy

CD-spectra were recorded on a Jasco J-1500 CD spectrometer. Full CD-spectra were recorded from 450 to 250 nm with a continuous scanning rate of 50 nm/min, a Digital Integration Time (D.I.T.) of 0.5 seconds and a bandwidth of 1.00 nm. The data shown are the mean of two

measurements and were smoothed using a Savitzky-Golay filter with a polynomal order of 3. Time-course measurements were recorded at 375 nm with a D.I.T of 2 seconds and a data pitch of 10.0 seconds with a Peltier element for temperature control. UV-Vis spectra were measured on a Jasco V-750 spectrophotometer with a peltier element for temperature control. Spectra were recorded from 500 to 250 nm, a bandwidth of 2.00 nm, a continuous scanning mode with a scanning speed of 400 nm/min and a UV-Vis response of 0.06 s. All spectra were recorded in 2 mm quartz glass cuvettes at a concentration range of 30-40 µM for CD spectra and 50 µM for UV spectra. Baseline correction with the respective solvent or buffer used was implemented. DNA mimic foldamers are readily soluble in water. Concentrations were determined by UV-absorbance using an average ε value at 375 nm per monomer of 2506 $Lmol^{-1}cm^{-1}$ for $^{m}QQ^{4}$ sequences and 2140 $Lmol^{-1}cm^{-1}$ for $^{m}QQ^{5}$ sequences.[2]

5.2.2.3 Foldamer building blocks

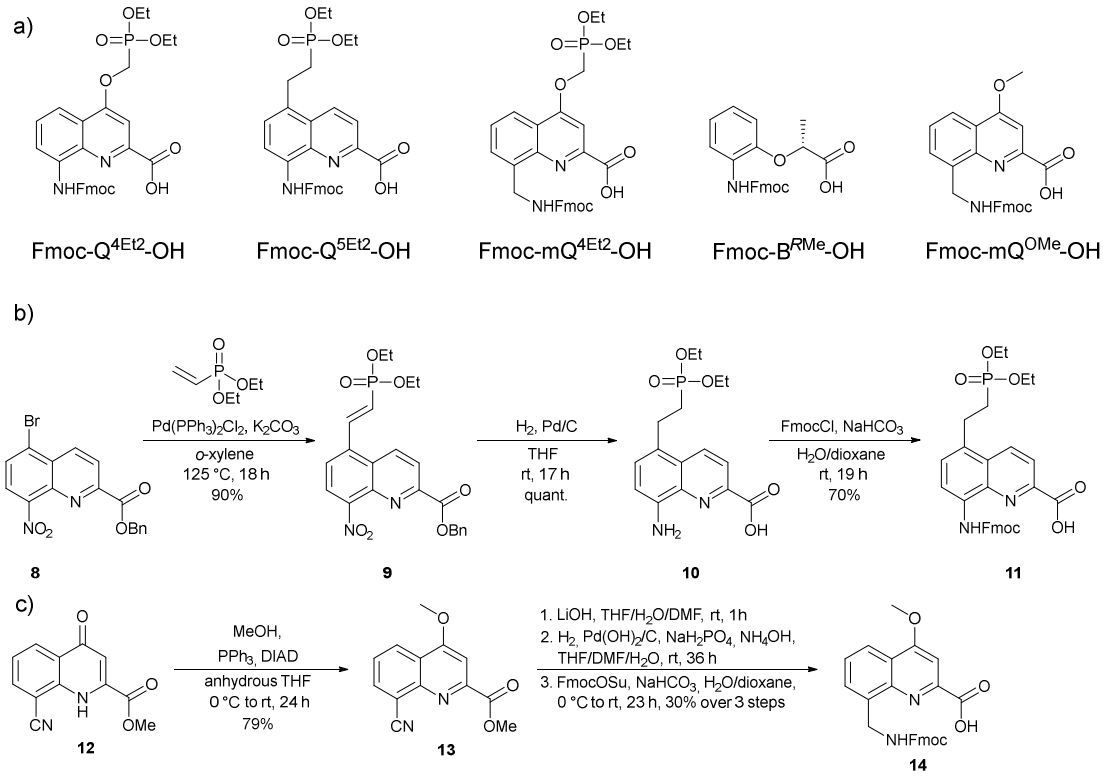
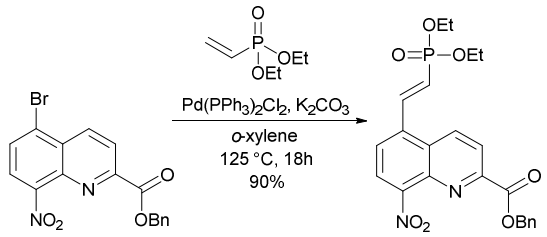


Figure S17 a) Respective diethyl phosphonate Fmoc-protected monomers for solid-phase synthesis. The synthesis of monomers Fmoc-Q^{4Et2}-OH, Fmoc-mQ^{4Et2}-OH and B^{RMe} are described elsewhere.[3, 4] b) Synthetic route towards monomer Fmoc-Q^{5Et2}-COOH **11** from described intermediate **8**, that was synthesized according to published procedures.[5] c) Synthesis of Fmoc-mQ^{OMe}-COOH monomer **14** from described intermediate **12**, that was synthesized according to published protocols.[6]

5.2.2.4 Synthetic procedures

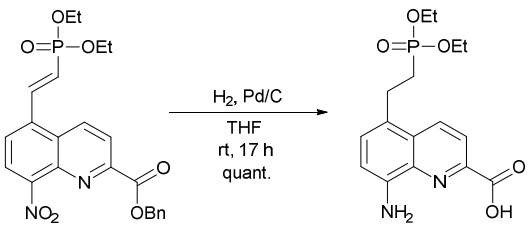
Synthesis of **9** (Synthesis of precursor **8** is described here.[5])



In a dry flask, freshly vacuum oven-dried (60 °C, overnight) **8** (1.50 g, 3.87 mmol, 1.0 eq.) and oven-dried (120 °C) K_2CO_3 (535 mg, 3.87 mmol, 1.0 eq.) were suspended in anhydrous o-xylene (15 mL) under N_2-atmosphere. To this, diethyl vinyl phosphonate (763 mg, 714 µL, 4.65 mmol, 1.2 eq.) and $Pd(PPh_3)_2Cl_2$ (81.6 mg, 116 µmol, 0.03 eq.) were added and the solution was stirred at 125 °C for 18 h under N_2-atmosphere until the HPLC analysis of an aliquot showed full conversion of the starting material. The black solution was diluted with EtOAc (200 mL) and washed with citric acid (2x, 150 mL, 5% w/v). The combined aqueous phases were extracted with EtOAc (2 x 100 mL) and the combined organic phases were washed with brine (200 mL), dried over Na_2SO_4, filtered and the solvent was removed in vacuo and dried at a high-vacuum line. The crude product was purified by column chromatography (dry load, 60% → 100% EtOAc in cyclohexane, starting R_F = 0.2) to give the title compound (1.64 g, 3.50 mmol, 90%) as yellow solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ (ppm) = 8.71 (d, *J* = 8.9 Hz, 1H), 8.35 (d, *J* = 8.9 Hz, 1H), 8.17 (dd, *J* = 22.1, 17.3 Hz, 1H), 8.11 (d, *J* = 7.9 Hz, 1H), 7.90 (d, *J* = 7.9 Hz, 1H), 6.52 (t, *J* = 17.0 Hz, 1H), 5.50 (s, 2H), 4.27 – 4.16 (m, 4H), 1.40 (t, *J* = 7.1 Hz, 6H).
^{13}C NMR (126 MHz, $CDCl_3$) δ (ppm) = 164.30, 150.35, 149.28, 141.77 (d, *J* = 7.3 Hz), 139.29, 137.10 (d, *J* = 23.6 Hz), 135.48, 133.62, 128.82, 128.57, 128.33, 127.89, 125.54 (d, *J* = 2.0 Hz), 124.23, 123.29 (d, *J* = 188.9 Hz),123.15, 68.00, 62.53 (d, *J* = 5.6 Hz), 16.63 (d, *J* = 6.1 Hz).
^{31}P NMR (202 MHz, $CDCl_3$) δ (ppm) = 16.11.
HRMS: (ESI$^+$) *m/z* calc. for $C_{23}H_{24}N_2O_7P^+$: 471.1316 $[M+H]^+$; found: 471.1312.

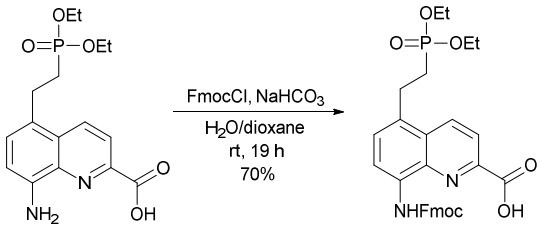
Synthesis of **10**



Compound **9** (1.93 g, 4.10 mmol, 1 eq.) was suspended in THF (contained no stabilizer and was freshly filtered over an alumina column from a solvent dispenser system, 20 mL, peroxide

containing THF might undergo addition during hydrogenation) under N_2- atmosphere and the solvent was degassed for 10 min by bubbling N_2-gas through the solution while sonicating. Pd/C (200 mg) was added and the solvent was degassed for an additional 5 min. The reaction mixture was put under H_2-atmosphere and vigorously stirred at rt for 17 h. The crude product was filtered over celite and washed with THF (500 mL). The solvent was removed in vacuo, co-evaporated with DCM, and dried at the high-vacuum line to give the title compound (1.44 g, 4.10 mmol, quant.) as a yellow solid that was used in the next step without further analysis

Synthesis of **11**



Amino acid **10** (1.44 g, 3.92 mmol, 1.0 eq., 96% purity by HPLC) was suspended in dioxane (80 mL) and $NaHCO_3$ (6.92 g, 82.4 mmol, 21 eq.) dissolved in water (80 mL) was added. The suspension was cooled to 0 °C and FmocCl (4.71 g, 4.71 mmol, 1.2 eq.) dissolved in dioxane (150 mL) was added dropwise over 1 h at 0 °C. The reaction mixture was stirred at 0 °C for 2 h and then stirred at rt for 19 h. The solvents were removed *in vacuo* and the resulting solid was suspended in water (150 mL). The suspension was acidified to pH = 3 with a KHSO4-solution (saturated), extracted with DCM (3 x 200 mL) and dried over Na_2SO_4. After filtration, the solvent was removed *in vacuo*, co-evaporated with DCM, and the resulting solid was dried at the high-vacuum line to remove residual dioxane. The crude product was precipitated from MeCN (20 mL), sonicated shortly to allow full precipitation, filtered, washed with cold MeCN (−14°C) and lyophilized to give the title compound (1.57 g, 2.74 mmol, 70%) as a yellow solid in a yield over 2 steps.

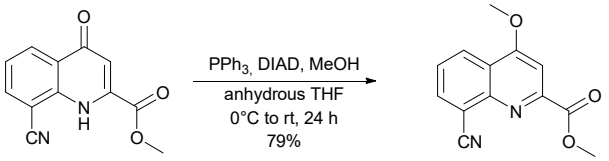
^{1}H NMR (500 MHz, DMSO-d$_6$) δ (ppm) = 13.59 (s, 1H), 10.40 (s, 1H), 8.65 (d, J = 8.8 Hz, 1H), 8.25 (d, *J* = 8.7 Hz, 1H), 7.93 (d, *J* = 7.5 Hz, 2H), 7.78 (d, *J* = 7.4 Hz, 2H), 7.57 (d, *J* = 8.1 Hz, 1H), 7.43 (t, *J* = 7.4 Hz, 2H), 7.36 (t, *J* = 7.5 Hz, 2H), 4.61 (d, *J* = 6.8 Hz, 2H), 4.44 (t, *J* = 6.8 Hz, 1H), 3.22 (tt, *J* = 11.4, 6.8 Hz, 2H), 2.17 – 2.07 (m, 2H), 1.21 (t, *J* = 7.0 Hz, 6H).

^{13}C NMR (126 MHz, DMSO-d$_6$) δ (ppm) = 165.35, 153.48, 145.07, 143.70, 140.80, 137.03, 134.70, 134.40, 130.58 (d, *J* = 15.5 Hz), 128.93, 127.77, 127.42, 127.21, 125.14, 120.63, 120.24, 115.90, 66.34, 61.04 (d, *J* = 6.2 Hz), 46.60, 25.62 (d, *J* = 137.0 Hz), 24.05 (d, *J* = 4.0 Hz), 16.26 (d, *J* = 5.7 Hz).

^{31}P NMR (202 MHz, DMSO-d$_6$) δ (ppm) = 30.62.

HRMS: (ESI$^+$) *m/z* calc. for $C_{31}H_{32}N_2O_7P^+$: 575.1942.1472 $[M+H]^+$; found: 575.1937.

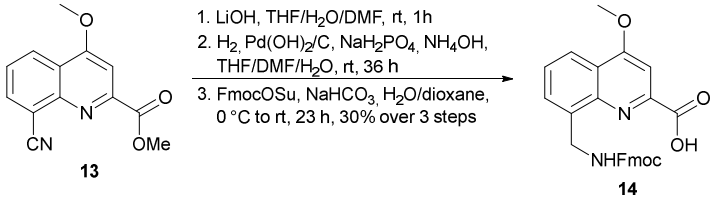
Synthesis of **13** (Synthesis of precursor **12** is described here.[6])



Synthesis of compound **12** is described here.[5] Compound **12** (8.00 g, 35.1 mmol, 1.0 eq.), PPh_3 (11.95 g, 45.6 mmol, 1.3 eq.) and MeOH (1.68 g, 2.13 mL, 52.6 mmol, 1.5 eq.) were suspended in anhydrous THF (70 ml) under N_2-atmosphere and cooled to 0 °C. DIAD (9.22 g, 8.95 mL, 45.6 mmol, 1.3 eq.) was added dropwise over 20 min at 0 °C. The resulting solution was stirred at 0 °C for 1 h, at rt for a further 2 h, and at 50 °C for 25 h. THF was removed *in vacuo* and the residuals taken up in MeOH and sonicated until full precipitation. The resulting solid was filtered off and dried overnight over under reduced pressure. The crude product was dissolved in $CHCl_3$, and a layer of MeOH was added on top. It was crystallized at −24 °C overnight, filtered and washed with cold MeOH. Crystallization was repeated twice, until the HPLC analysis of an aliquot showed no residual PPh_3O, to yield the title compound **2** (6.67 g, 27.54 mmol, 79%) as an off-white crystalline solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ (ppm) = 8.48 (dd, *J* = 8.5, 1.4 Hz, 1H), 8.17 (dd, *J* = 7.2, 1.4 Hz, 1H), 7.69 (s, 1H), 7.66 (dd, *J* = 8.5, 7.2 Hz, 1H), 4.17 (s, 3H), 4.08 (s, 3H).
^{13}C NMR (126 MHz, $CDCl_3$) δ (ppm) = 165.94, 163.86, 151.31, 147.86, 136.85, 127.25, 126.76, 122.51, 117.01, 113.78, 101.87, 56.74, 53.62.
HRMS: (ESI^+) m/z calc. for $C_{13}H_{11}N_2O_3^+$: 243.0764 $[M+H]^+$; found: 243.0765.

Synthesis of **14**



In a 2 liter round bottom flask **13** (5.5 g, 22.7 mmol, 1.0 eq.) was dissolved in 850 ml THF/DMF (90:10 v/v). LiOH (816 mg, 34.1 mmol, 1.5 eq.) dissolved in 310 ml H_2O was added and stirred at rt for 1 h until TLC showed full consumption of the starting material. $NaH_2PO_4{\cdot}(H_2O)_2$ (35.4 g, 227 mmol, 10 eq.) was added as a solid too the ternary solvent mixture. Then, the solvent was degassed for 15 min by bubbling N_2-gas through the solution while sonicating. Afterwards, $Pd(OH)_2/C$ (542 mg) and NH_4OH (20%, 12.93 mL, 68.13 mmol, 3.0 eq.) were added and the black suspension was further degassed for 15 min. Under vigorous stirring, a H_2-balloon was placed on top of the round bottom flask and the reaction mixture was stirred

for 36 h at rt under H_2 atmosphere until no nitrile-containing intermediate product was detectable by HPLC in an aliquot. To the mixture (82% HPLC purity at 300 nm), $NaHCO_3$ (9.53 g, 114 mmol, 5 eq.) dissolved in H_2O (200 mL) was added, and FmocOSu (8.04 g, 23.8 mmol, 1.05 eq.) was added dissolved in THF (200 mL) at 0 °C over 1 h, stirred at 0 °C for 1 h and 21 h at rt. The mixture was filtered through a paper filter, washed with THF and then concentrated *in vacuo* until a mostly aqueous suspension remained, which was then acidified with $KHSO_4$ (1.8 M) to pH = 2–3. Water (800 mL) and DCM (600 mL) were added and stirred until the mixture was dissolved. The organic phase was separated, and the aqueous phase was extracted with DCM (3x 500 ml). The combined organic phases were dried over $MgSO_4$ and filtered through a paper filter and the solvent was removed *in vacuo*. The resulting crude oil was purified by column chromatography (silica gel, 13% acetone in DCM + 0.1% AcOH (v/v). The solvent was removed *in vacuo* and the resulting solid was recrystallized from boiling MeCN and then recrystallized from boiling EtOAc to yield the title compound as white solid (2.51 g, 5.52 mmol, 30%).

^{1}H-NMR shows 2 conformers in a ratio of 85:15. Integrals are given with their respective integration. Overlapping NMRs signals are integrated as m and their integration is given. For ^{13}C NMR, only the major species is listed.

^{1}H NMR (500 MHz, DMSO-d_6): δ [ppm] = 8.11 (dd, *J* = 7.9, 1.9 Hz, 1H), 7.92 – 7.85 (m, 2.6H), 7.80 (d, *J* = 7.5 Hz, 0.3H), 7.69 (d, *J* = 7.5 Hz, 1.7H), 7.67 – 7.59 (m, 2.7H), 7.58 – 7.54 (m, 0.15H), 7.52 – 7.48 (m, 0.15H), 7.45 – 7.36 (m, 2H), 7.34 – 7.27 (m, 2H), 7.10 (t, *J* = 7.6 Hz, 0.3H), 4.84 (d, *J* = 6.2 Hz, 1.7H), 4.73 (d, *J* = 6.3 Hz, 0.3H), 4.40 (d, *J* = 6.8 Hz, 1.7H), 4.33 (d, *J* = 6.3 Hz, 0.3H), 4.25 (t, *J* = 6.8 Hz, 1H), 4.13 (s, 3H).

^{13}C NMR (126 MHz, DMSO-d_6): δ [ppm] = 166.09, 163.17, 156.53, 148.55, 145.18, 143.86, 140.76, 137.84, 128.27, 127.58, 127.42, 127.02, 125.10, 121.34, 120.31, 120.11, 99.90, 65.27, 56.47, 46.83, 40.59.

HRMS: (ESI$^+$) m/z calc. for $C_{27}H_{23}N_2O_5$: 455.1602 $[M+H]^+$; found: 455.1613.

5.2.2.5 Foldamer synthesis

5.2.2.5.1 General solid-phase foldamer synthesis

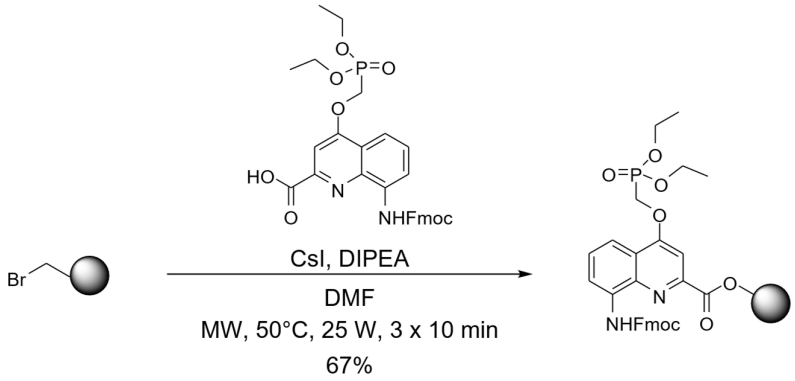
Solid phase synthesis (SPS) was performed manually under MW-irradiation on a CEM Discover (Liberty Bio) microwave oven using a reaction vessel and an internal fiber optic probe for temperature control as described below, or with a fully automated synthesizer followed by previously reported protocol.[7]

- Bromination of low loading Wang resin

Low loading-(LL) Wang resin (Novabiochem, 100-200 mesh, 1.00 g, 0.37 mmol, 1.0 eq.) was swollen in DMF (6 mL) for 30 min. PPh_3 (970 mg, 3.70 mmol, 10.0 eq.) and CBr_4 (1.23 g, 3.70 mmol, 10 eq.) were quickly added and the suspension was stirred slowly for 20 h at RT.

The resin was filtered, washed with DMF (3 × 3 mL) and DCM (3 × 3 mL), and dried by passing N_2 through the resin and stored at 4 °C until usage.

- Loading of the 1st monomer unit & resin loading estimation



LL-brominated Wang resin (Novabiochem, 100-200 mesh, 100 mg, 37.0 µmol, 1.0 eq.) was swollen in anhydrous DMF (3 mL) for 30 min under N_2. After washing with anhydrous DMF (3 mL), Fmoc-Q_4-COOH (42.5 mg, 74.0 µmol, 2.0 eq.) dissolved in anhydrous DMF (2 mL), CsI (19.2 mg, 74.0 µmol, 2.0 eq.) and freshly distilled DIPEA (9.56 mg, 12.9 µL, 74 µmol, 2.0 eq.) were quickly added and the reaction vessel was placed under microwave irradiation (25 W, ramp to 50 °C over 5 min, hold at 50 °C for 10 min) while bubbling N_2 through the solution. The resin was washed with anhydrous DMF (3 × 3 mL) and the process was repeated twice. The resin was washed with DMF (3 × 3 mL) and DCM (3 × 3 mL) and dried by passing N_2 through the resin. For resin loading estimation, a solution of 20% piperidine in DMF (v/v, 3 mL) was added to a known mass of the previously dried resin (1–2 mg) and agitated for 5 min. Meanwhile, the absorbance at 290 nm of the piperidine/DMF-solution was measured. After agitation, the absorbance of the solution that contained the resin was measured at 290 nm.

$$Resin\ loading\ \left[\frac{mmol}{g}\right] = \frac{[Abs_{final} - Abs_{initial}]}{[2.00 \times m_{resin}]}$$

- Fmoc deprotection

To the pre-swollen loaded Wang resin (75.0 mg, 0.26 mmol g^{-1}, 19.5 µmol), a 20% solution of piperidine in DMF (3 mL, v/v) was added and the resin was mixed by bubbling N_2-gas through the solution for 3 min. The resin was filtered and washed with DMF (2 x 3 mL) and the deprotection was repeated once for 7 min to give the respective amine NH_2-Q-Wang resin. The resin was filtered, washed with DMF (5 x 3 mL), and washed with anhydrous THF (3 x 3 mL) prior to coupling.

- *In situ* acid chloride activation, coupling, and capping

The NH_2-Wang resin (75.0 mg, 0.26 mmol g^{-1}, 19.5 µmol, 1.0 eq.) was suspended in anhydrous THF (0.9 mmol) and collidine (23.0 µL, 176 µmol, 9.0 eq.) was added. Concurrently, in a glass vial, **Monomer** (35 mg, 58.6 µmol, 3.0 eq.) and PPh_3 (41 mg, 156 µmol, 8 eq.) were mixed and dissolved in anhydrous $CHCl_3$ (0.9 mL). Subsequently, trichloro acetonitrile (18.0 µL, 176 µmol, 9 eq.) was added to the vial, which was quickly shaken, and the mixture was added to the pre-swollen resin. After mixing, the reaction vessel was placed under microwave irradiation (25 W, ramp to 50 °C over 5 min, hold at 50 °C for 15 min). The resin was filtered off and washed with anhydrous THF (3 x 3 mL). The coupling step was repeated once. The resin was filtered off and washed with anhydrous THF (3 x 3 mL) and DCM (3 x 3 mL) prior to the capping step. The resin was suspended in a 50% solution of Ac_2O in DCM (v/v) and mixed by bubbling N_2-gas through the solution for 10 min. The resin was washed with DCM (2 x 3 mL) and DMF (2 x 3 mL).

- Resin cleavage

The resin-bound foldamer was placed in a syringe equipped with a filter and then suspended in TFA (3 mL). The resin was shaken for 2 h at rt. The resin was then filtered off and washed twice with TFA. The TFA was combined removed *in vacuo* and the resulting oil was precipitated by sonication in cold Et_2O. The precipitate was centrifuged, and the solvent was decanted to give a yellow solid. The decanted Et_2O was concentrated by rotary evaporation and the precipitation was repeated. The combined precipitates were dissolved in water/MeCN and then lyophilized to give the crude protected foldamer as a yellow solid and further purified as a diethyl-phosphonate protected compound.

5.2.2.5.2 Removal of the diethyl-phosphonate protecting groups

Purified protected foldamers were dissolved in anhydrous chloroform (1 mL per 10 mg compound) and cooled to 0 °C. TMSBr (0.2 mL per 10 mg compound) was mixed with anhydrous chloroform (0.8 mL per 10 mg compound). The TMSBr-solution was added dropwise over 10 min to the reaction mixture. The reaction mixture was allowed to warm to room temperature and stirred under N_2-atmosphere for 1–3 d until an HPLC- and LCMS aliquot of the reaction mixture showed complete cleavage of all phosphonate-diethyl esters. The reaction mixture was evaporated *in vacuo* (40 °C water bath) to give a yellow oil, then co-evaporated with DCM (2x) to give a yellow solid. The solid was suspended in water, basified to pH > 12 with triethylamine and stirred for 30 min. The suspension was filtered through nylon syringe filters (pore size: 0.22 µm) to give a pale-yellow solution that was freeze-dried to give the crude deprotected foldamers as yellow solids.

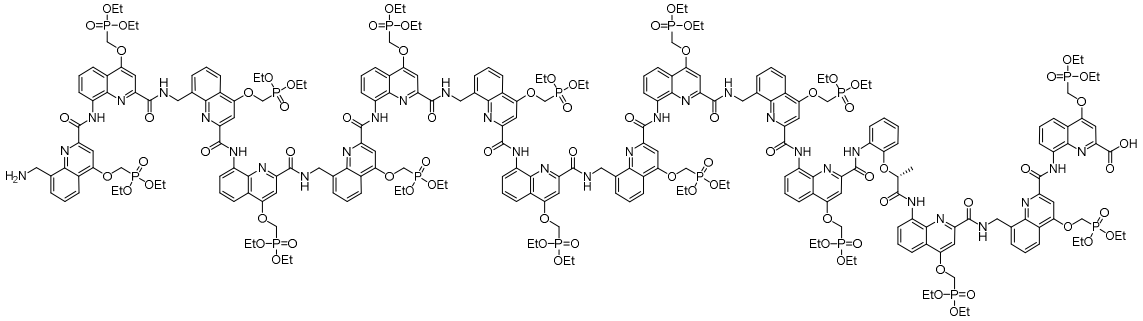
5.2.2.5.3 Cation exchange chromatography

Cation exchange chromatography was performed on Dowex© 50WX4 200-400 (H) resin. The resin was swollen in H_2O the orange solution was decanted. The resin was transferred into a

glass column (0.5 cm diameter, height 15–20 cm) and washed with H_2O (gravity flow). It was further washed with two column volumes (CV) of 2 M HCl solution, then H_2O until pH = 6-7 (ca. 5-10 CV), then two CV of 2 M NH_4OAc solution, then 5 CV of H_2O. Purified triethylammonium salts of foldamers were dissolved in water and loaded on the column. The column was closed without flow for 2 h. The compound was eluted with water (20 mL) and lyophilized to give the purified foldamer NH_4^+-salt as yellow solid.

5.2.2.5.4 Synthesized foldamers

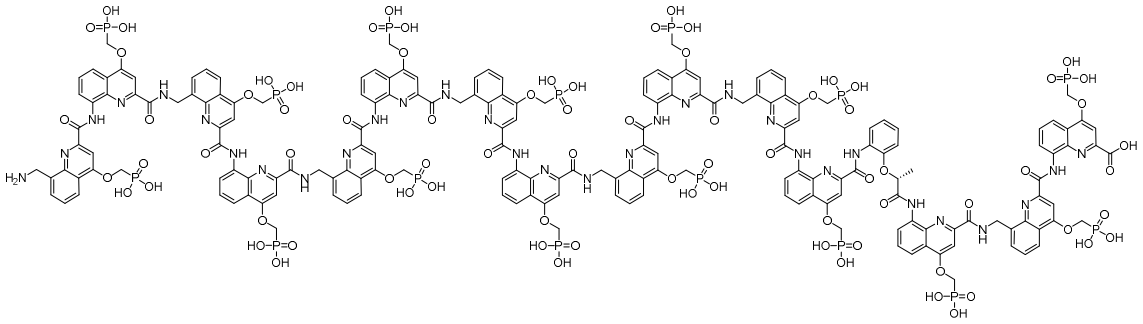
5.2.2.5.4.1 Compound **5a** protected chiral mQQ4 based 16mer



Oligomer **5a** was synthesized on LL-Wang resin (25.0 µmol) according to section 3.3.1. The compound was isolated by semi-preparative RP-HPLC (acidic conditions, linear gradient 33-48% B in A, column XBridge® Prep C8 OBD™) to give the title compound (59.0 mg, 11.1 µmol, 44%) as a yellow solid.

(ESI$^+$) *m/z* calc. for $C_{241}H_{283}N_{31}O_{78}P_{15}$: 1774.5060 $[M+3H]^{3+}$; found: 1774.5610.

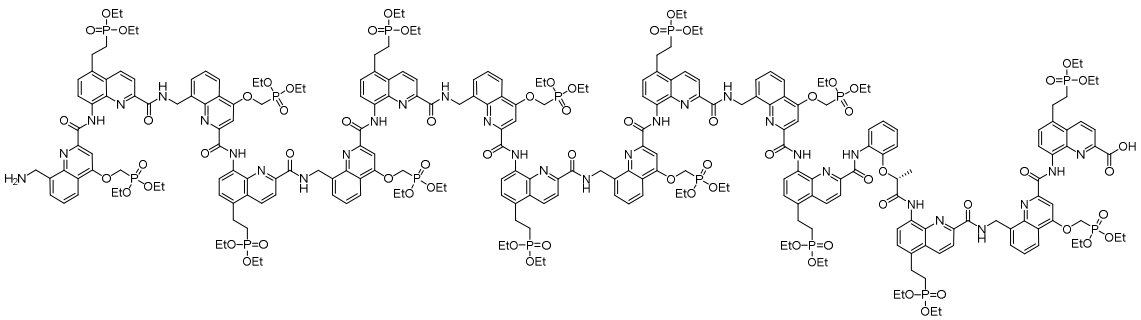
5.2.2.5.4.2 Compound **5** deprotected chiral ^{m}QQ^4 based 16mer



Oligomer **5a** was deprotected based on the protocol given in 3.3.2 (scale 27.0 mg, 5.10 µmol) and purified by semi-preparative RP-HPLC (TEAA buffer system, linear gradient 0-30% B in A (Kinetex C18 EVO column). From the lyophilized powder, the NEt_3 cations were exchanged with NH_4^+ (Section 3.3.3) to give the poly-ammonium salt of compound **5** as yellow solid (18.9 mg, 4.02 µmol, 78%).

(ESI$^-$) *m/z* calc. for $C_{181}H_{156}N_{31}O_{78}P_{15}$: 1118.8820 $[M-4H]^{4-}$; found: 1118.9263.

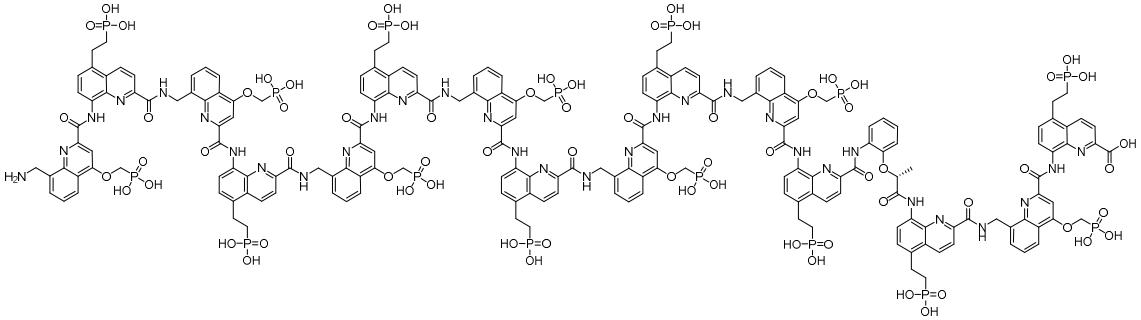
5.2.2.5.4.3 Compound **6a** protected chiral $^{m}QQ^{5}$ based 16mer



Oligomer **6a** was synthesized on LL-Wang resin (25.0 µmol) according to section 3.3.1. The compound was isolated by semi-preparative RP-HPLC (acidic conditions, linear gradient 35-55% B in A, column XBridge® Prep C8 OBD™) to give the title compound (82.5 mg, 15.5 µmol, 62%) as a yellow solid.

(ESI^{+}) *m/z* calc. for $C_{249}H_{299}N_{31}O_{70}P_{15}$: 1769.2279 $[M+3H]^{3+}$; found: 1769.2901.

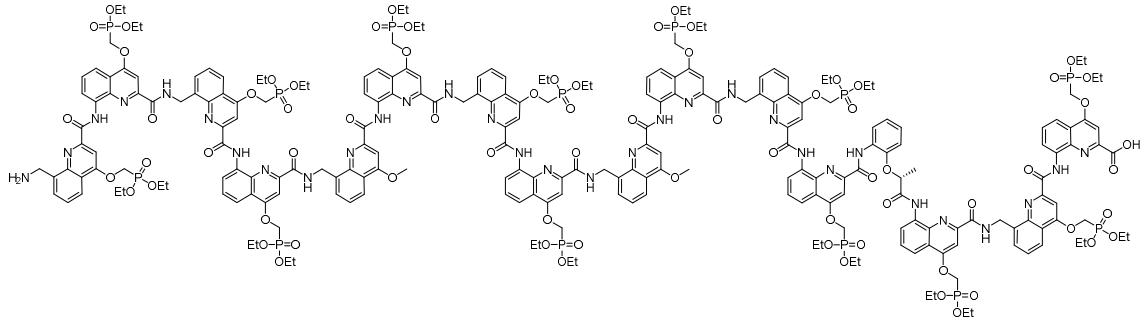
5.2.2.5.4.4 Compound **6** deprotected chiral $^{m}QQ^{5}$ based 16mer



Oligomer **6a** was deprotected based on the protocol given in 3.3.2 (scale 25.0 mg, 4.70 µmol) and purified by semi-preparative RP-HPLC (TEAA buffer system, linear gradient 0-30% B in A (Kinetex C18 EVO column). From the lyophilized powder, the NEt_3 cations were exchanged with NH_4^{+} (Section 3.3.3) to give the polyammonium-salt of compound **6** as yellow solid (18.2 mg, 3.87 µmol, 82%).

(ESI^{-}) *m/z* calc. for $C_{189}H_{172}N_{31}O_{70}P_{15}$: 1114.9235 $[M-4H]^{4-}$; found: 1114.9294.

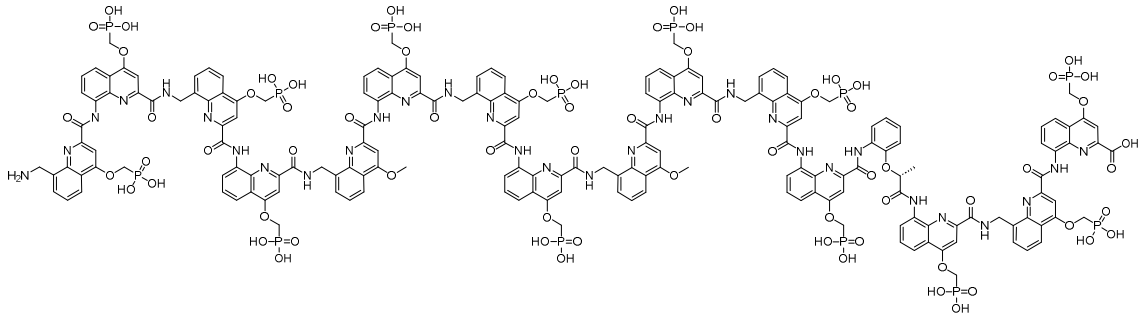
5.2.2.5.4.5 Compound **7a** protected chiral $^{m}QQ^{4}$-OMe 16mer



Oligomer **7a** was synthesized on TG-Wang resin (15.0 µmol) according to section 3.3.1. The compound was isolated by semi-preparative RP-HPLC (acidic conditions, linear gradient 35-55% B in A, column XBridge® Prep C8 OBD™) to give the title compound (40 mg, 8.0 µmol, 53%) as a yellow solid.

(ESI^{+}) *m/z* calc. for $C_{233}H_{265}N_{31}O_{72}P_{13}$: 1683.8200 $[M+3H]^{3+}$; found: 1683.8404.

5.2.2.5.4.6 Compound **7** deprotected chiral $^{m}QQ^{4}$-OMe 16mer



Oligomer **7a** was deprotected based on the protocol given in 3.3.2 (scale 11 mg, 2.20 µmol) and purified by semi-preparative RP-HPLC (TEAA buffer system, linear gradient 0-30% B in A (Kinetex C18 EVO column) and the cation was exchanged to NH_4^{+} (Section 3.3.3) to give polyammonium-salt of compound **7** as yellow solid (5.2 mg, 1.2 µmol, 55%).

(ESI^{-}) *m/z* calc. for $C_{181}H_{155}N_{31}O_{72}P_{13}$: 1078.8988 $[M-4H]^{4-}$; found: 1078.9162.

5.2.3 Spectra & Chromatograms

5.2.3.1 NMR spectra

5.2.3.1.1 Small molecules spectra

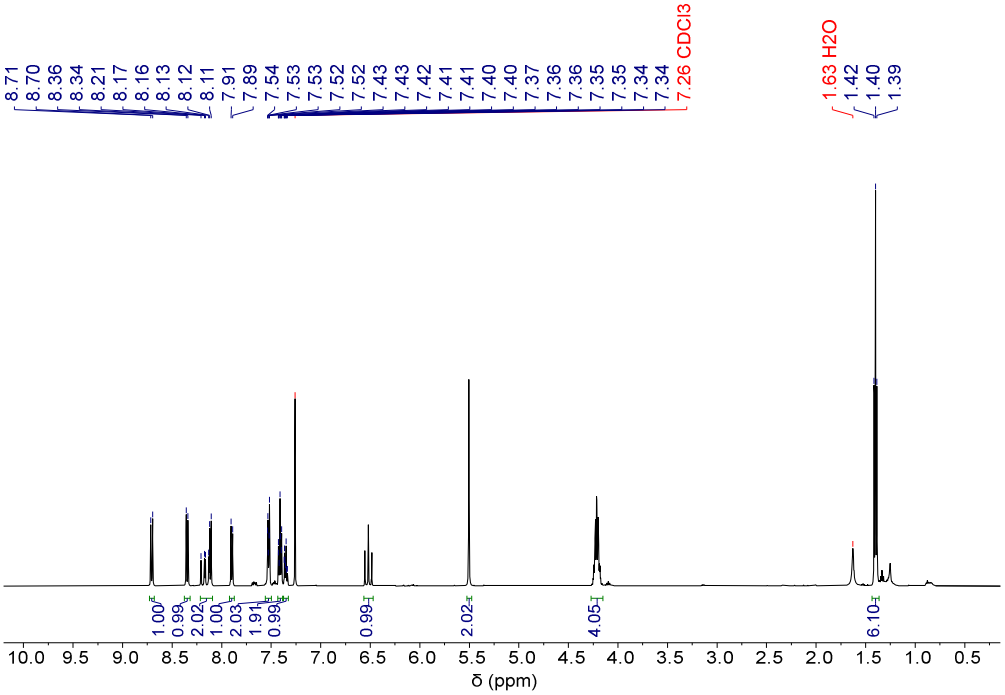


Figure S 18 ^{1}H NMR spectrum of compound **9** in $CDCl_3$.

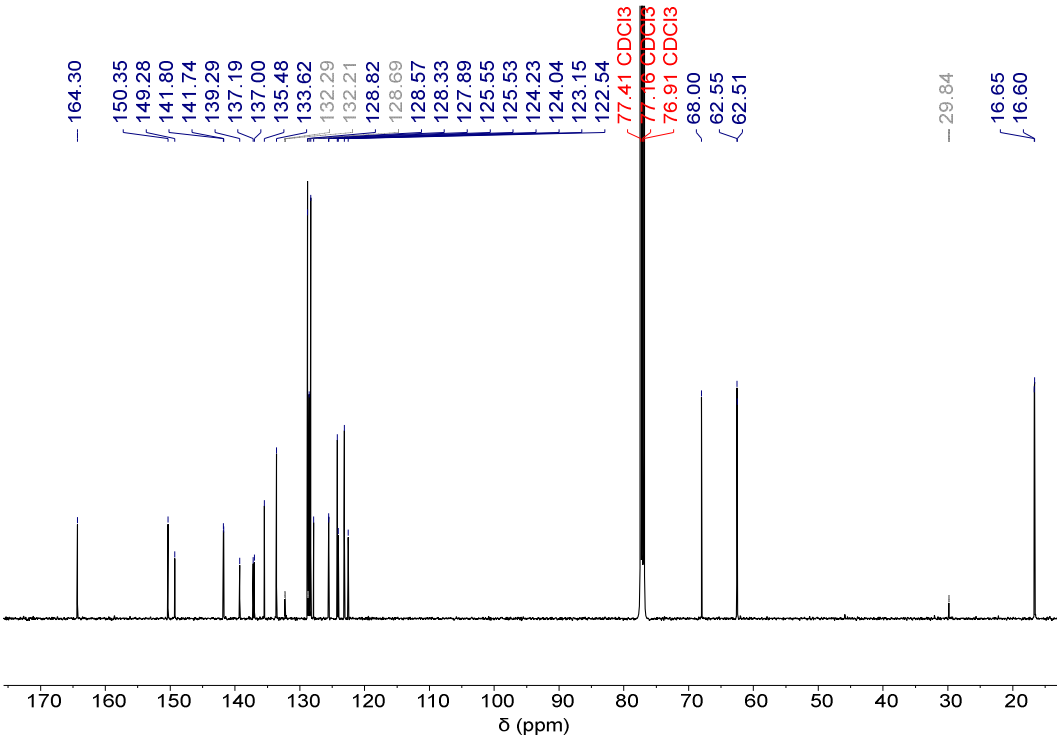


Figure S 19 ^{13}C NMR spectrum of compound **9** in $CDCl_3$.

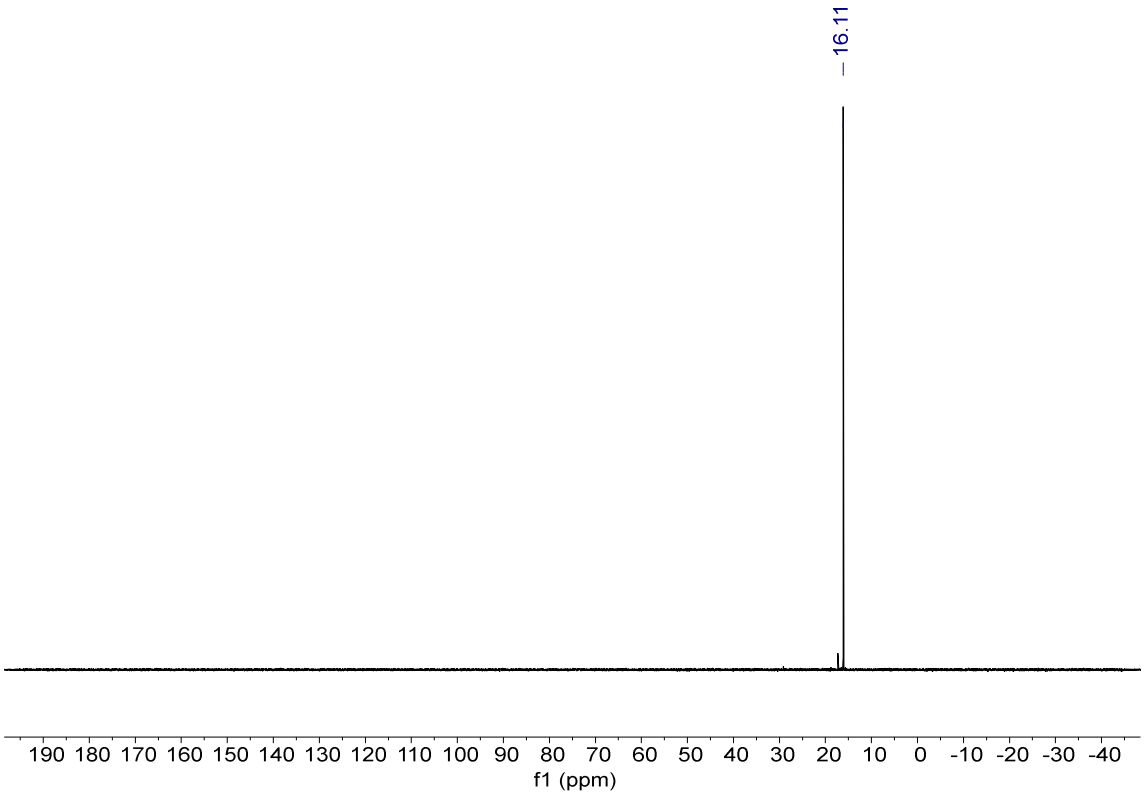


Figure S 20 ^{31}P NMR spectrum of compound **9** in $CDCl_3$.

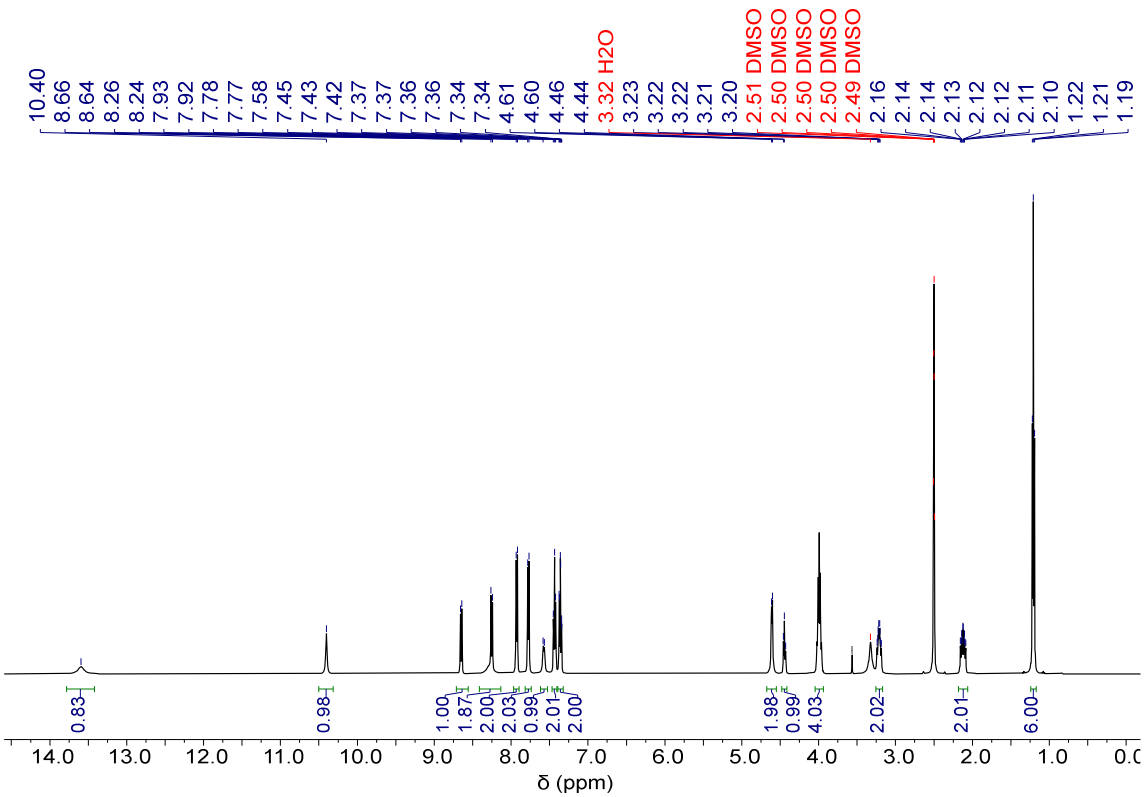


Figure S 21 ^{1}H NMR spectrum of compound **11** in DMSO-*d*$_6$.

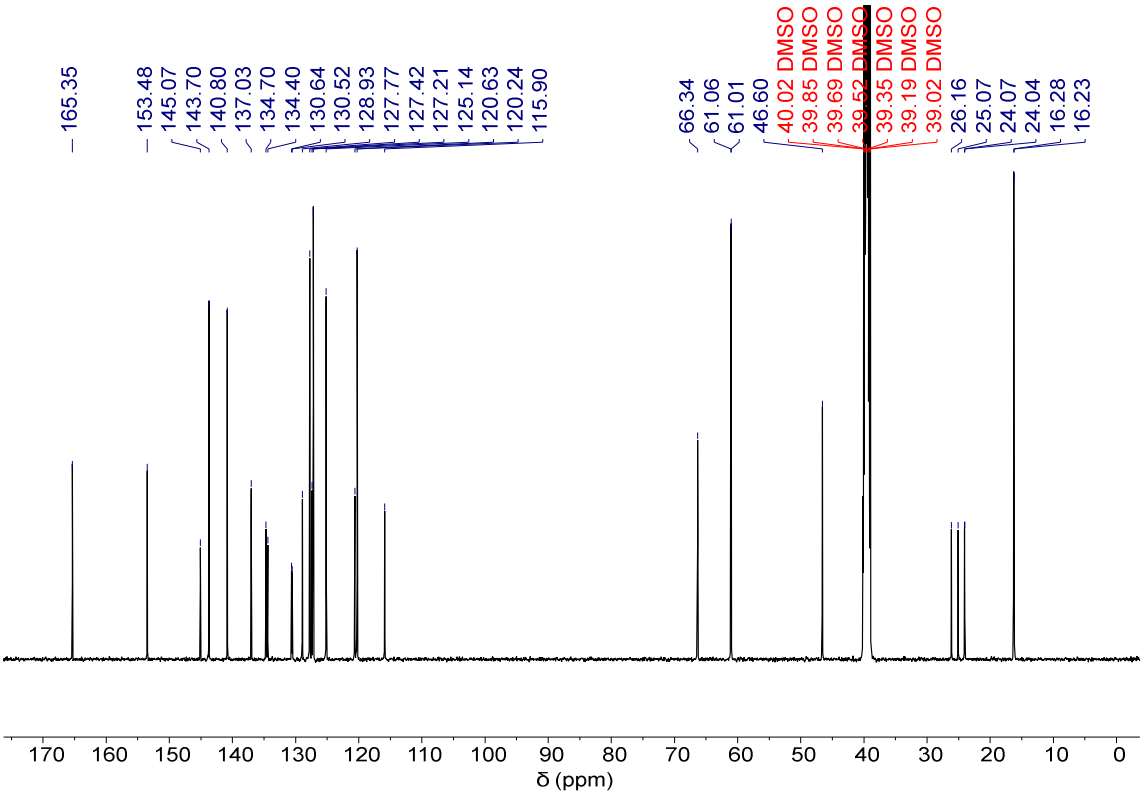


Figure S 22 ^{13}C NMR spectrum of compound **11** in DMSO-*d*$_6$.

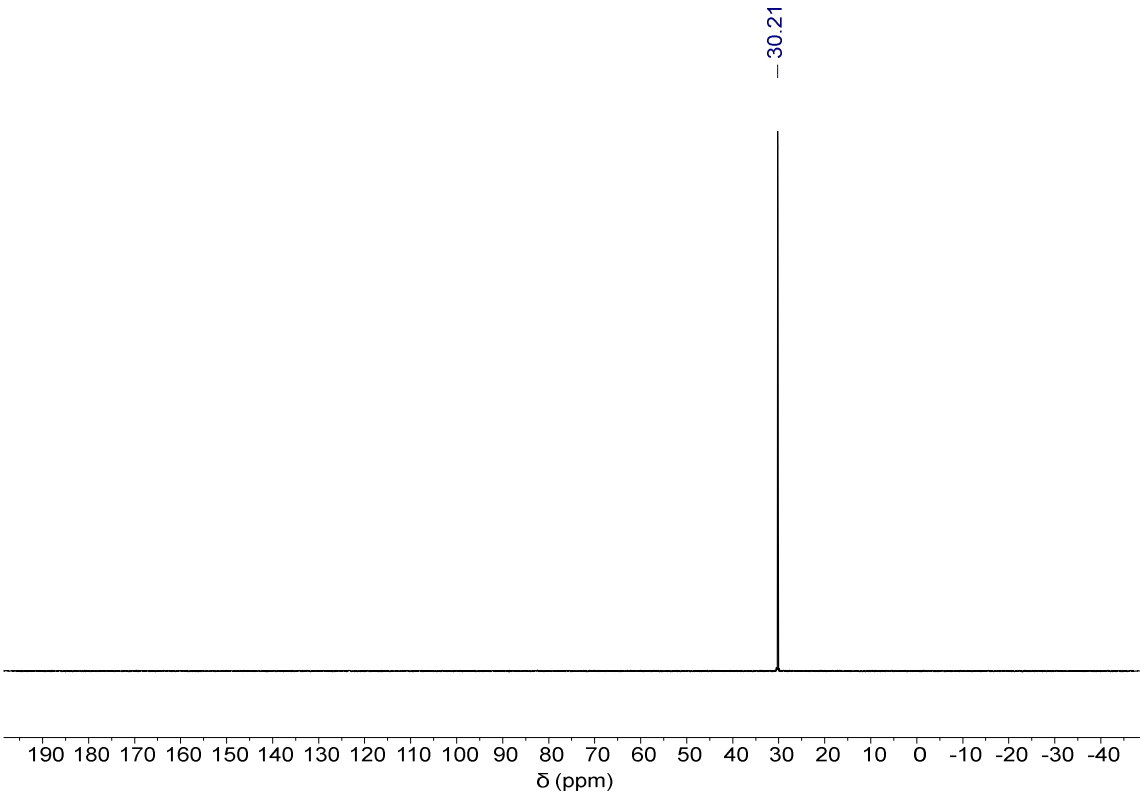


Figure S 23 ^{31}P NMR spectrum of compound **11** in DMSO-*d*$_6$.

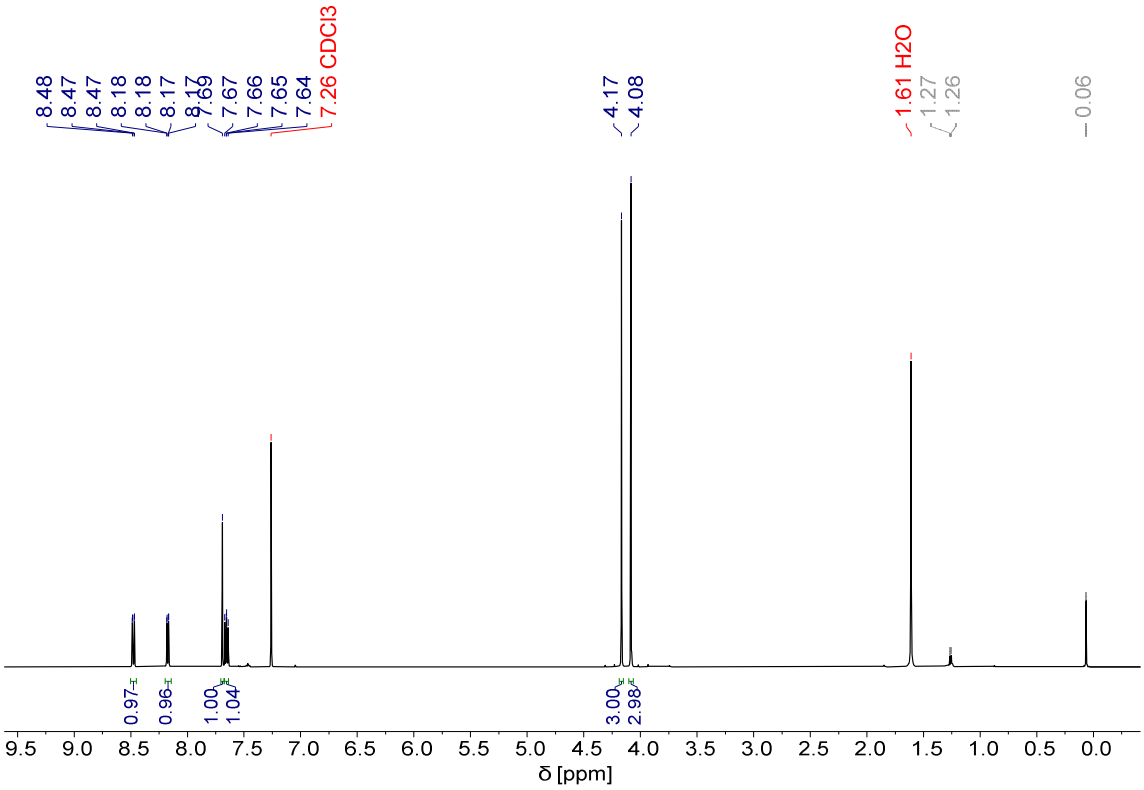


Figure S 24 ^{1}H NMR spectrum of compound **13** in $CDCl_3$.

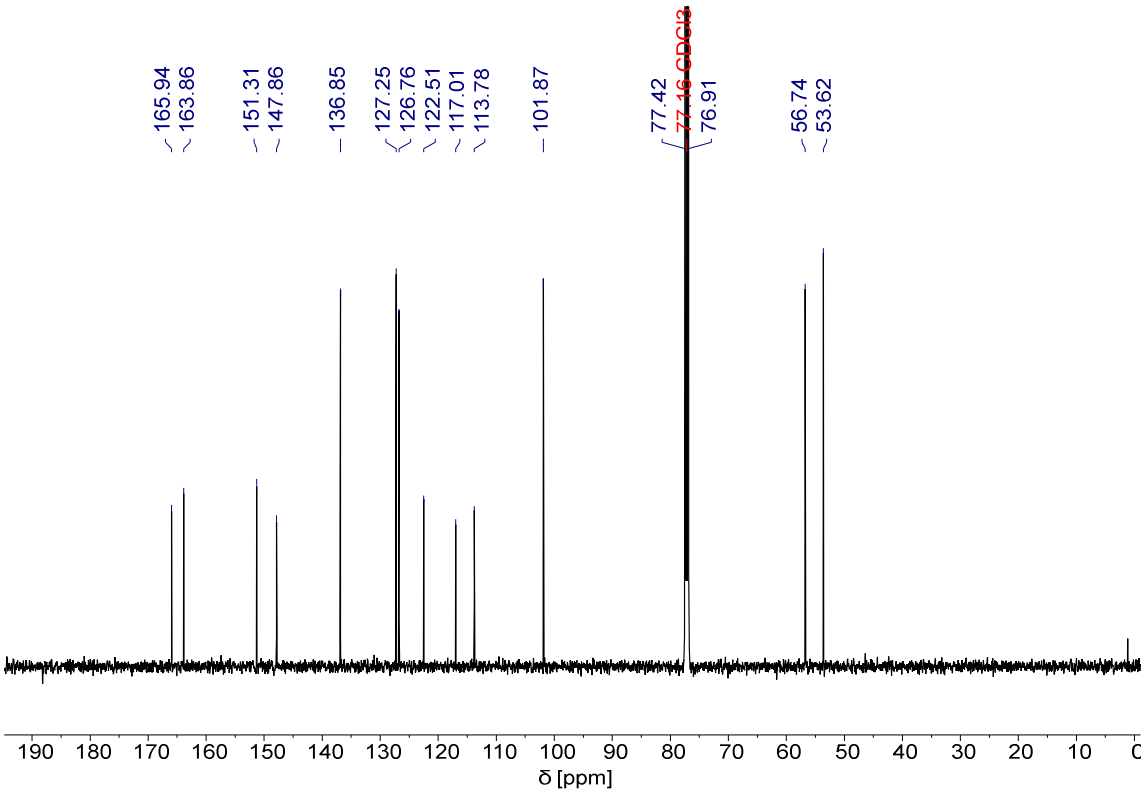


Figure S 25 ^{13}C NMR spectrum of compound **13** in $CDCl_3$.

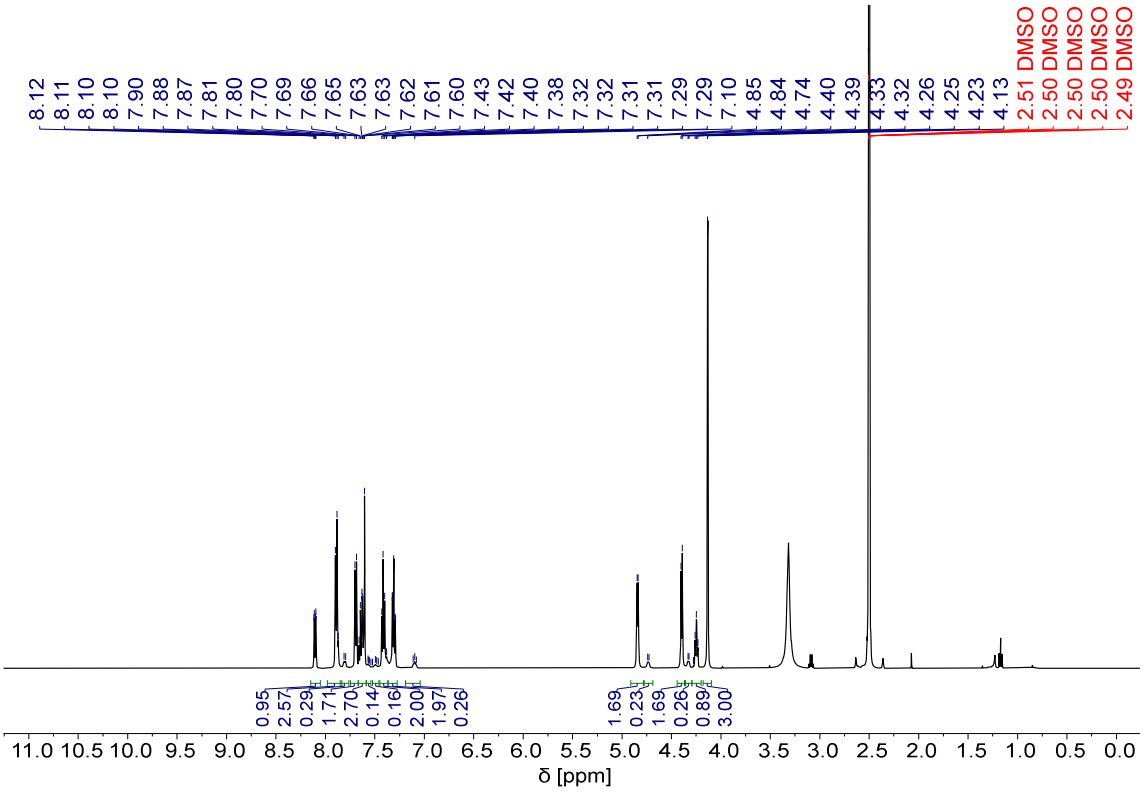


Figure S 26 ^{1}H NMR spectrum of compound **14** in DMSO-*d*$_6$.

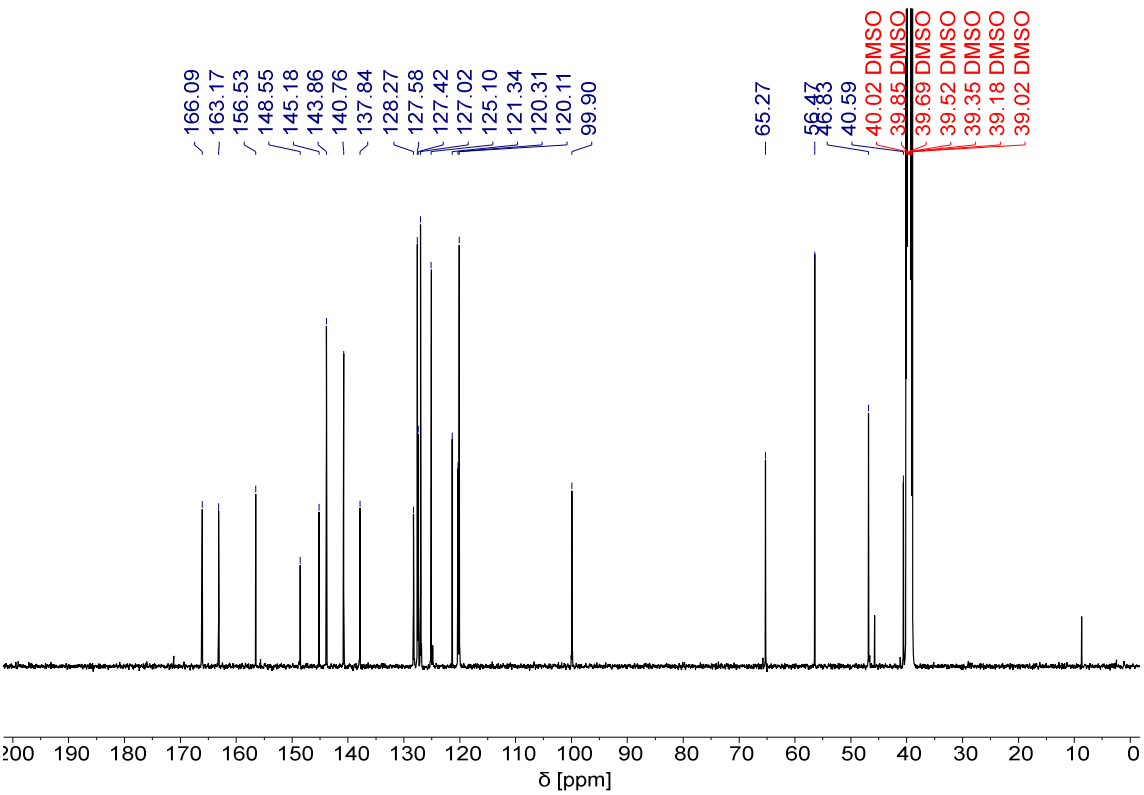


Figure S 27 ^{1}H NMR spectrum of compound **14** in DMSO-*d*$_6$.

5.2.3.1.2 Oligomer NMR spectra

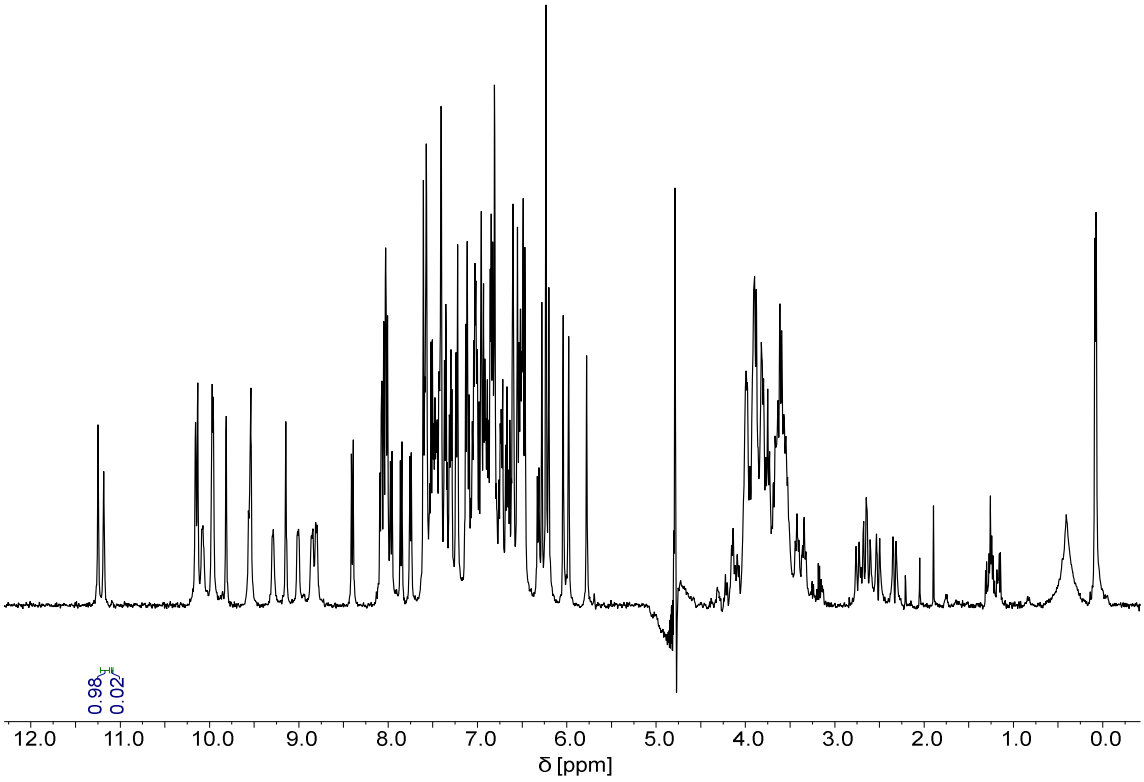


Figure S 28 ^{1}H NMR of spectrum compound **5** in 50 mM NH_4HCO_3 (H_2O/D_2O, 9:1, v/v) at 25 °C**.**

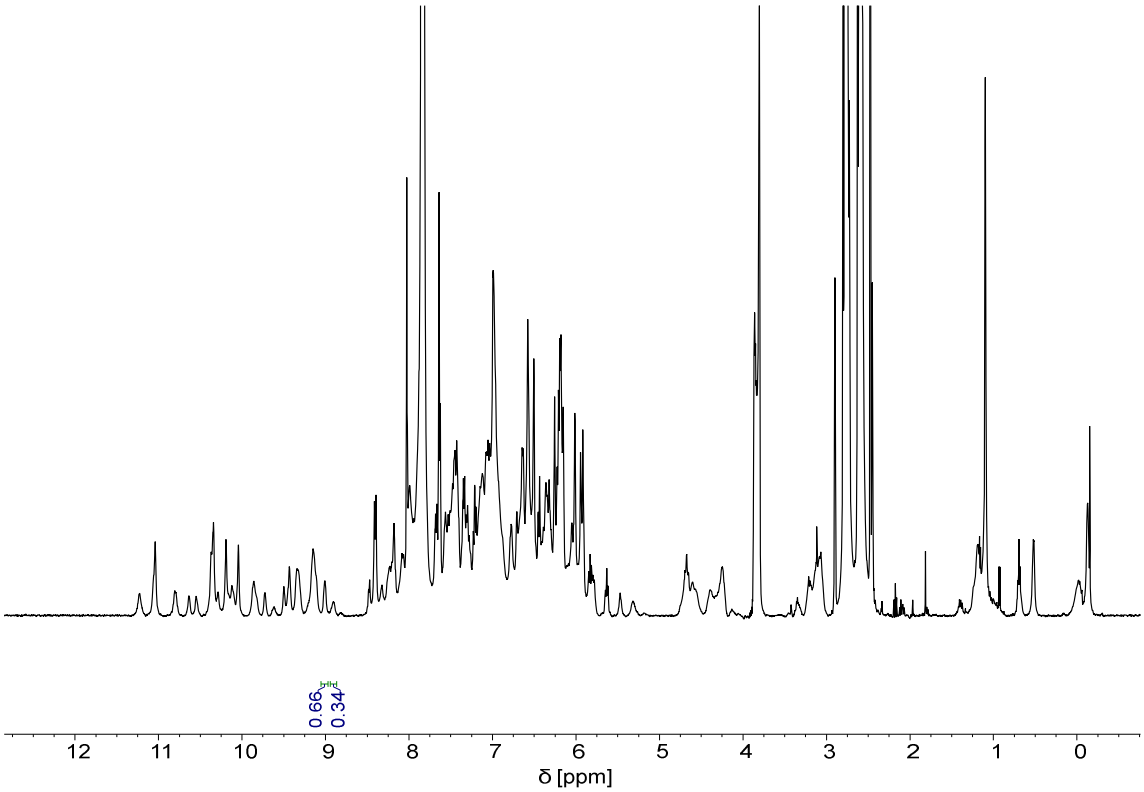


Figure S 29 ^{1}H NMR spectrum of compound **5** in (DMF-d_7/H_2O, 9:1, v/v) at 25 °C.

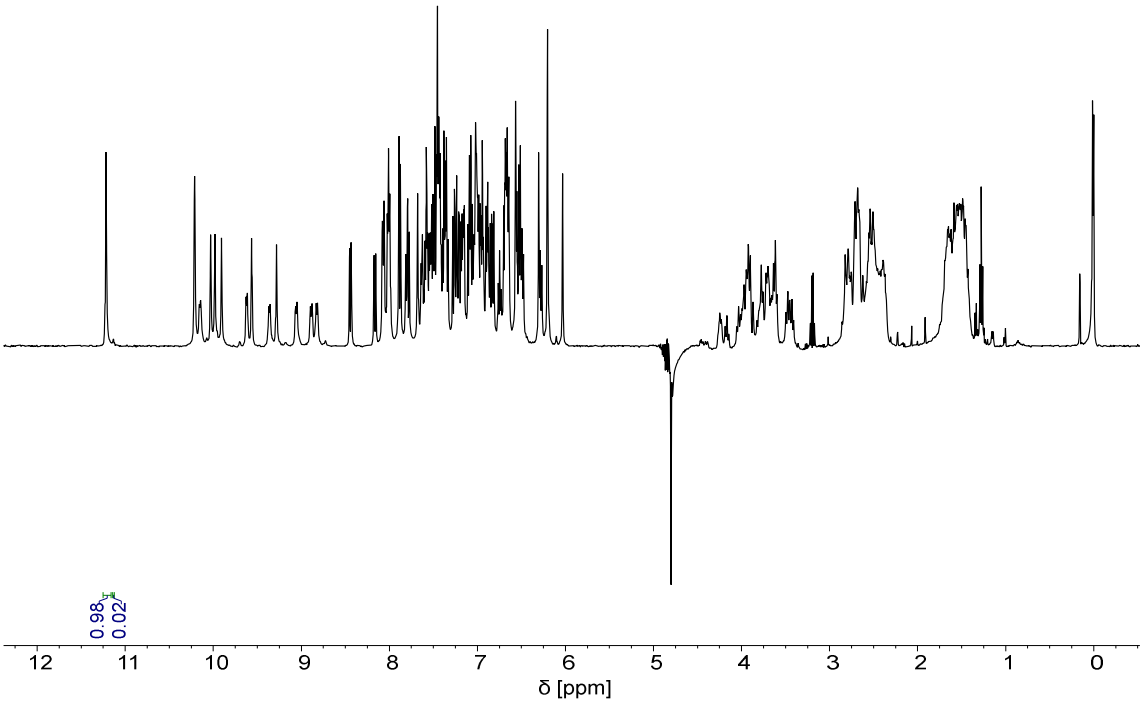


Figure S 30 ^{1}H NMR spectrum of compound **6** in 50 mM NH_4HCO_3 (H_2O/D_2O, 9:1, v/v) at 25 °C.

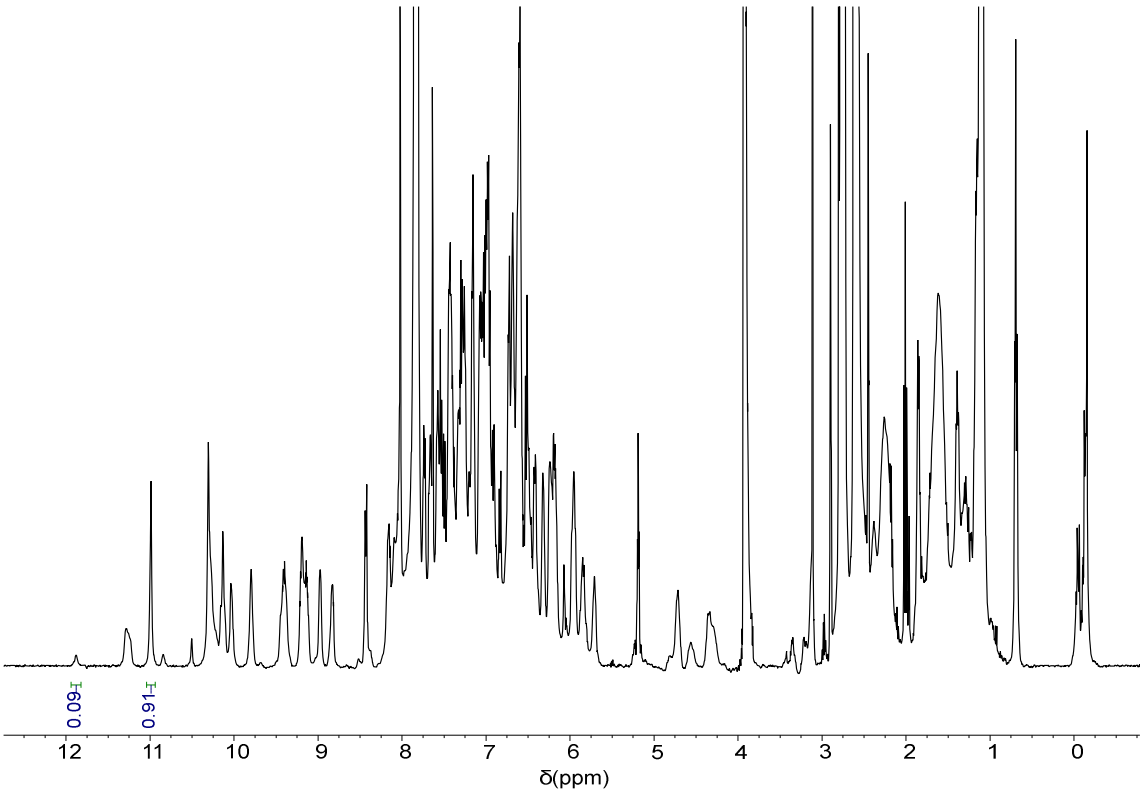


Figure S 31 ^{1}H NMR spectrum of compound **6** in (DMF-d_7/H_2O, 9:1, v/v) at 25 °C.

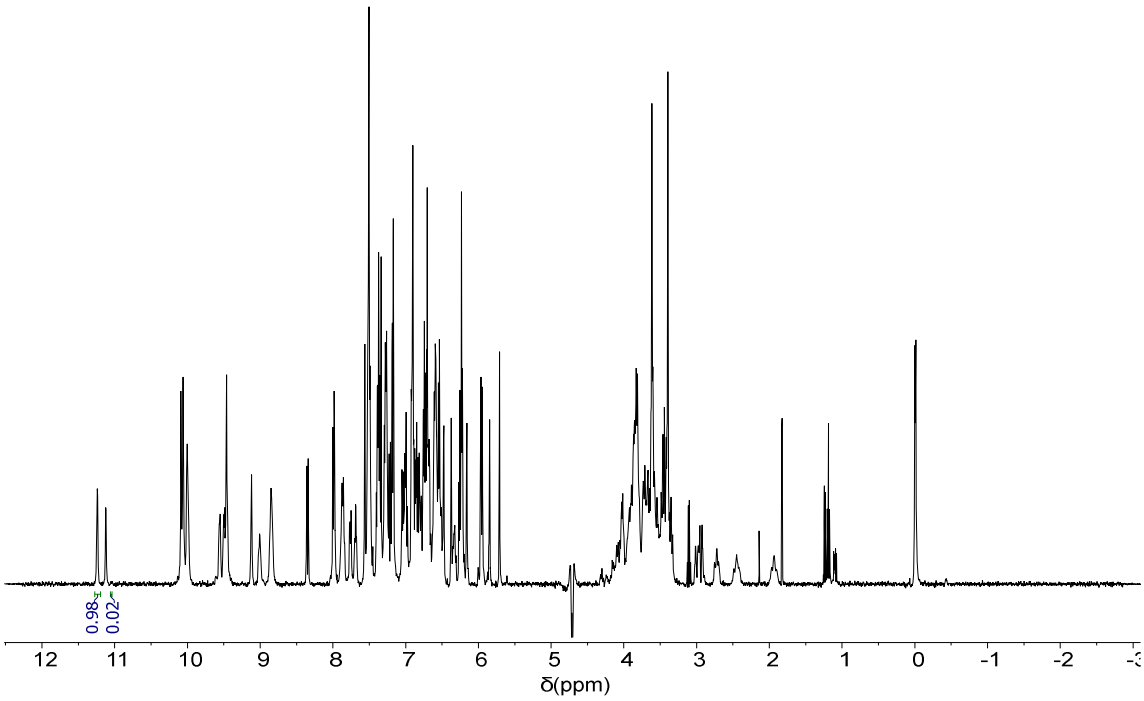


Figure S 32 ^{1}H NMR spectrum of compound **7** in 50 mM NH_4HCO_3 (H_2O/D_2O, 9:1, v/v) at 25 °C.

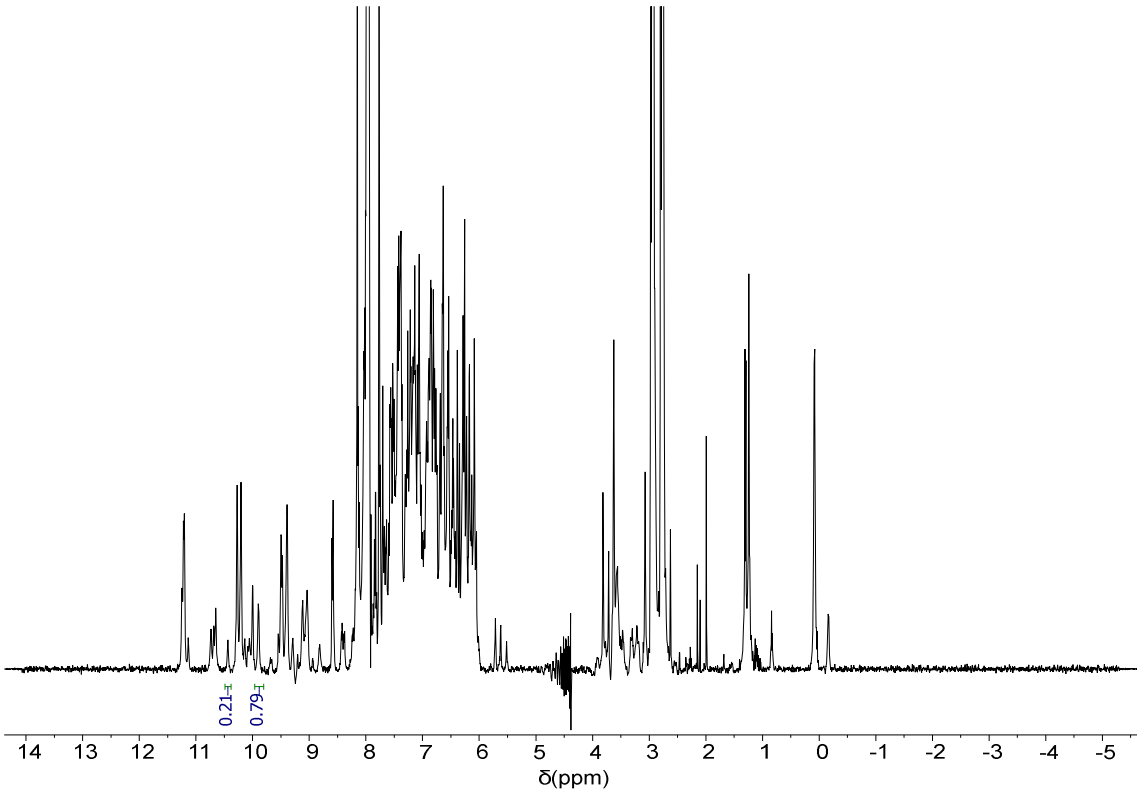
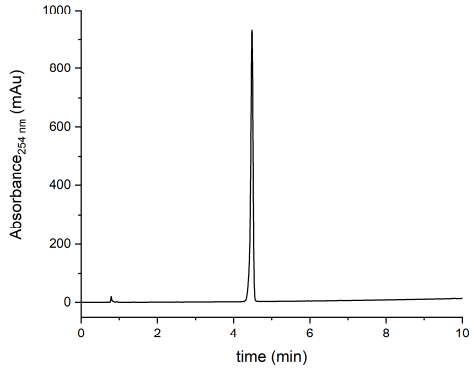


Figure S 33 ^{1}H NMR spectrum of compound **7** in (DMF-d_7/H_2O, 9:1, v/v) at 25 °C.

5.2.3.2 HPLC chromatograms

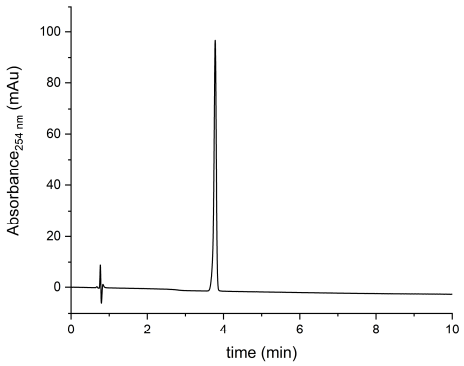
Compound **11**

Gradient: 50-100 % B in A, 0.1% TFA buffer



Compound **14**

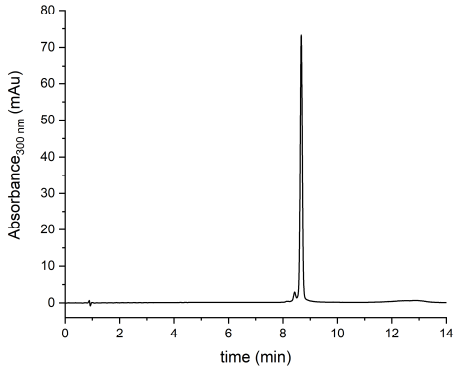
Gradient: 50-100 % B in A, 0.1% TFA buffer



Compound **5a**

Gradient: 30-70 % B in A, 0.1% TFA buffer

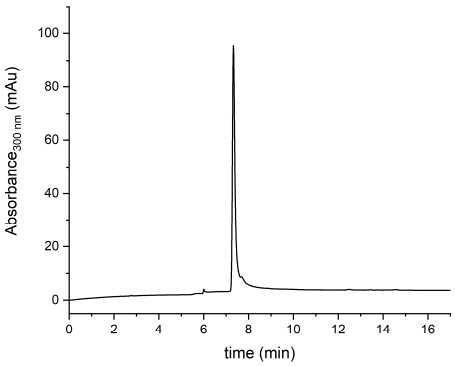
Column: Nucleodur C18



Compound **5**

Gradient: 0-100 % B in A, TEAA buffer

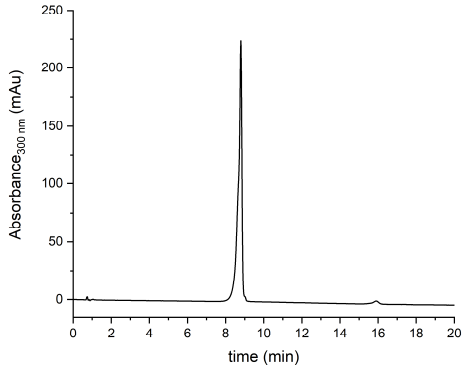
Column: Nucleodur C18



Compound **6a**

Gradient: 30-100 % B in A, 0.1% TFA buffer buffer

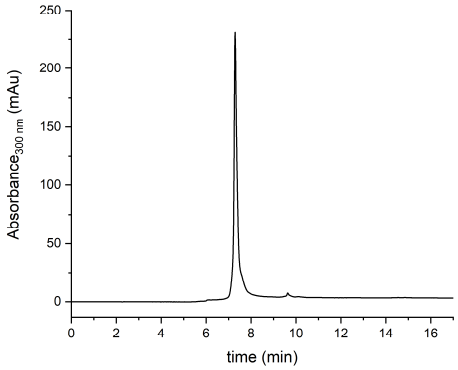
Column: Nucleodur C18



Compound **6**

Gradient: 0-100 % B in A, TEAA buffer

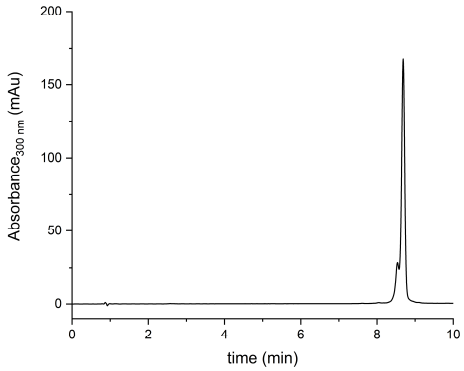
Column: Nucleodur C18



Compound **7a**

Gradient: 30-100 % B in A, 0.1% TFA buffer buffer

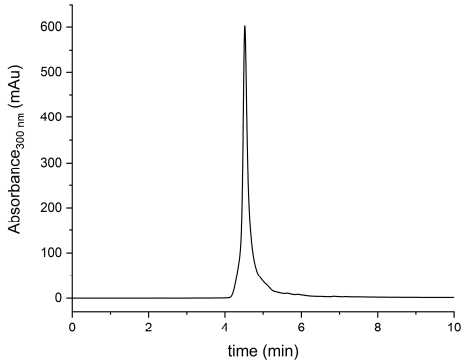
Column: Nucleodur C18



Compound **7**

Gradient: 0-100 % B in A, TEAA buffer

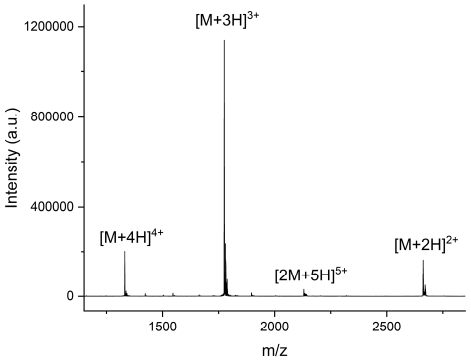
Column: Nucleodur C18



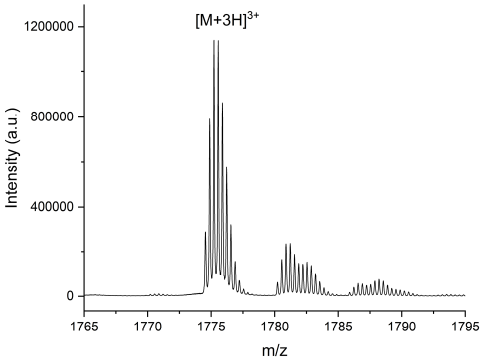
5.2.3.3 Mass spectra

5.2.3.3.1 Foldamer mass spectra

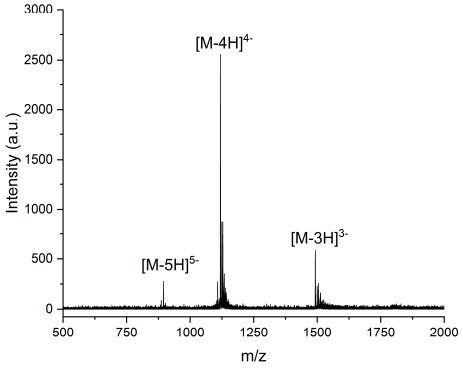
Compound **5a**



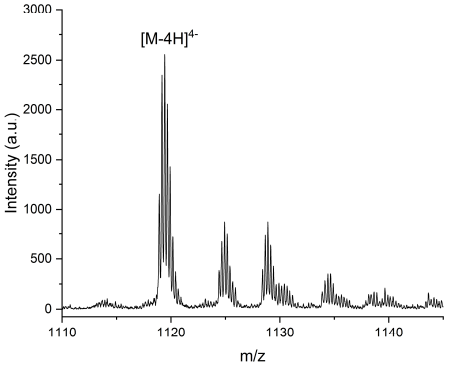
Compound **5a** Zoom



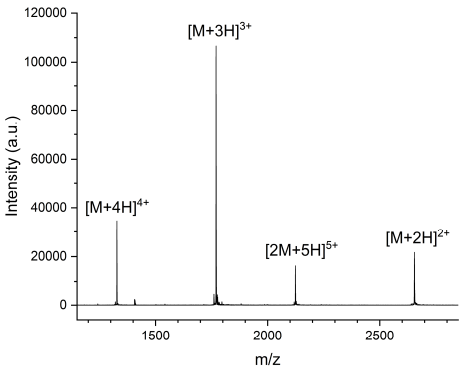
Compound **5**



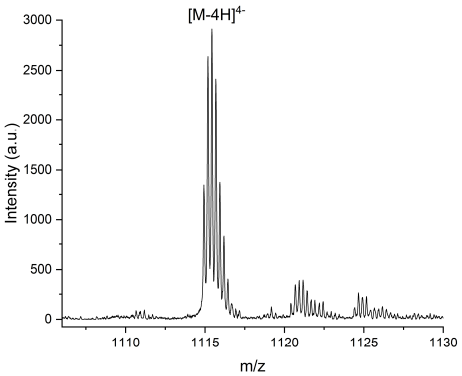
Compound **5** Zoom



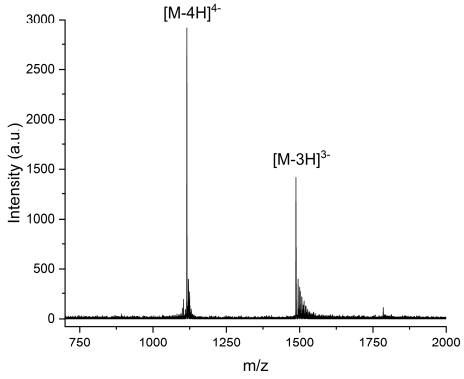
Compound **6a**



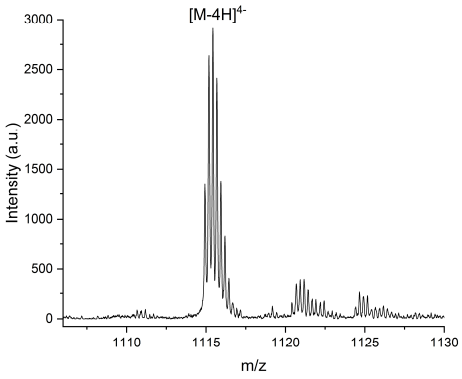
Compound **6a** Zoom



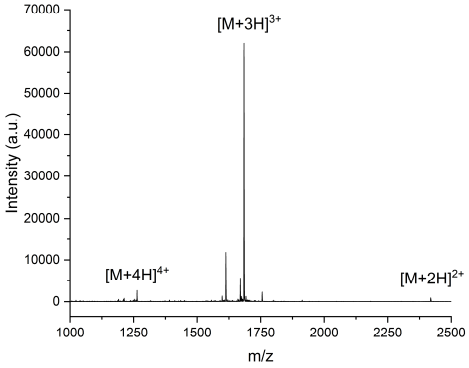
Compound **6**



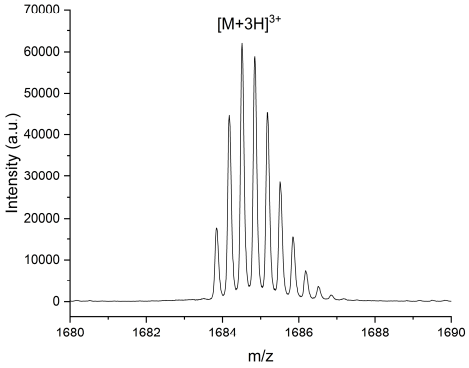
Compound **6** Zoom



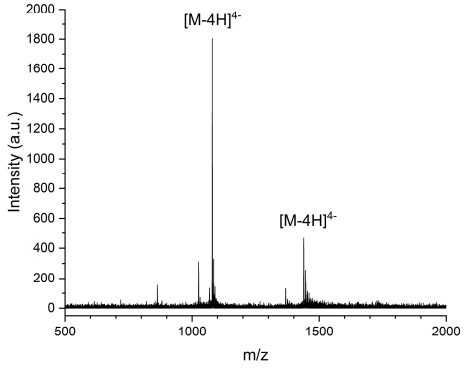
Compound **7a**



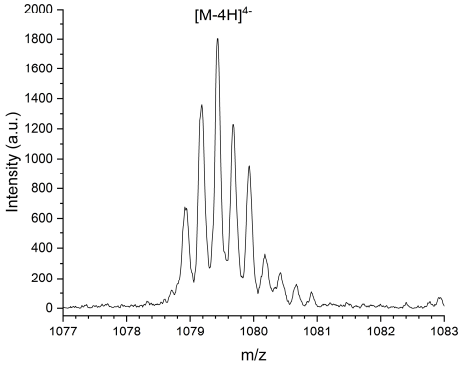
Compound **7a** Zoom



Compound **7**



Compound **7** Zoom



5.2.4 References

1. G. R. Fulmer, A. J. M. Miller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw and K. I. Goldberg, *Organometallics*, 2010, **29**, 2176-2179.
2. V. Corvaglia, D. Carbajo, P. Prabhakaran, K. Ziach, P. K. Mandal, V. Dos Santos, C. Legeay, R. Vogel, V. Parissi, P. Pourquier*, et al.*, *Nucleic Acids Res.*, 2019, **47**, 5511-5521.
3. V. Corvaglia, J. Wu, D. Deepak, M. Loos and I. Huc, *Chem. - Eur. J.*, 2024, **30**, e202303650.
4. D. Bindl, E. Heinemann, P. K. Mandal and I. Huc, *Chem. Commun. (Cambridge, U. K.)*, 2021, **57**, 5662-5665.
5. X. Hu, S. J. Dawson, P. K. Mandal, X. de Hatten, B. Baptiste and I. Huc, *Chem. Sci.*, 2017, **8**, 3741-3749.
6. K. Ziach, C. Chollet, V. Parissi, P. Prabhakaran, M. Marchivie, V. Corvaglia, P. P. Bose, K. Laxmi-Reddy, F. Godde, J.-M. Schmitter*, et al.*, *Nat. Chem.*, 2018, **10**, 511-518.
7. V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat and I. Huc, *Chem. - Eur. J.*, 2023, **29**, e202300898.

6 Long DNA mimic foldamers enable binding of multi-protein complexes

Molecular mimicry does not entail copying the parent biomolecule. Instead, some of the features and properties of the mimic arise from the very differences that distinguishes it from biomolecule that has inspired its design.[1] Thus, DMFOs may serve as unique tools for molecular biology precisely because they differ from DNA and it must be these differences, that cause them to outcompete DNA in binding affinities and enzyme inhibition of many DBPs.[1,2,3] However, DMFOs are typically being compared with DNA duplex of the same length, disregarding that DNA exists *in vivo* as an extended biopolymer rather than isolated short fragments. DMFOs, synthesized bottom up from simple anilines are reported up to a maximum of 32 quinolines mimicking 16 bps of DNA, albeit only as racemic mixtures of *P*-and *M*-helices.[2] Thus, they are not available in the size range of dsDNA, that can be produced in the length of several kilobases.[4] Reaching this size with DMFOs is unrealistic, but multiple intermediate milestones are of interest. We envision synthetic advances enabling DMFOs in 20-90 bp range, equivalent to extranucleosomal linker DNA in eukaryotes.[5] In the future, sequences equivalent to 147 bps are realistic, which is equivalent to the length of DNA wrapped around the histone octamer in the NCP.[6] This length regime is where DNA engages in structurally complex interactions with multi-protein assemblies. In the following manuscript, we report on the synthesis of chiral, one-handed DMFOs that start to reach some of these milestones. The impact of this was rewarded immediately, as we could visualize one DMFO in complex with a DNA sensing multi-protein complex by cryo-EM.

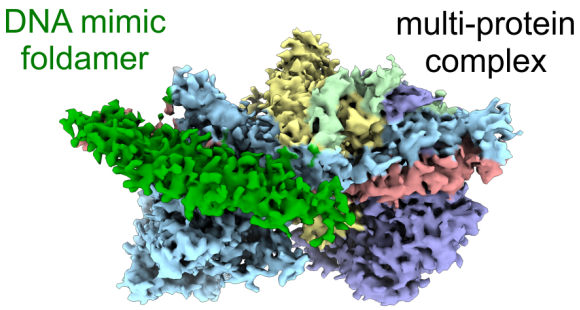


Figure 22 Electron density of a DMFO that mimics 36 bp B-DNA bound to the A-module of INO80.

Contributions

The project was planned in collaboration I. Huc and K-P- Hopfner. Synthetic routes and characterization for Fmoc-protected building blocks were developed and provided by me. Scale-up was conducted by L. Stanislawski. Oligomer synthesis was conducted with C. Douat and L. Bodero. Oligomer analysis was conducted by me. L. Allmendinger was involved in the design and execution of NMR experiments. F. Därr and A. López-Francos López-Romero did express and purify the proteins, performed biochemical studies and measured cryo-EM of complexes. F. Metzger characterized foldamer aggregates by cryo-EM. F. Därr processed the cryo-EM data. Lisa Gourdon-Grünewaldt performed BLI measurements The manuscript was written by me in collaboration with F. Därr, L. Gourdon-Grünewaldt and I. Huc.

References

1. D. Deepak, J. Wu, V. Corvaglia, L. Allmendinger, M. Scheckenbach, P. Tinnefeld and I. Huc, *Angew. Chem., Int. Ed.*, 2025, **64**, e202422958.
2. K. Ziach, C. Chollet, V. Parissi, P. Prabhakaran, M. Marchivie, V. Corvaglia, P. P. Bose, K. Laxmi-Reddy, F. Godde, J.-M. Schmitter*, et al.*, *Nat. Chem.*, 2018, **10**, 511-518.
3. V. Corvaglia, D. Carbajo, P. Prabhakaran, K. Ziach, P. K. Mandal, V. Dos Santos, C. Legeay, R. Vogel, V. Parissi, P. Pourquier*, et al.*, *Nucleic Acids Res.*, 2019, **47**, 5511-5521.
4. S. J. Reisinger, K. G. Patel and D. V. Santi, *Nat. Protoc.*, 2006, **1**, 2596-2603.
5. H. J. Szerlong and J. C. Hansen, *Biochem. Cell Biol.*, 2011, **89**, 24-34.
6. K. Luger, A. W. Mader, R. K. Richmond, D. F. Sargent and T. J. Richmond, *Nature (London)*, 1997, **389**, 251-260.

6.1 Long DNA mimic foldamers enable binding of multi-protein complexes (to be submitted)

Manuel Loos,[a] Franziska Därr,[b] Lisa Gourdon-Grünewaldt,[a] Céline Douat,[a] Alberto López-Francos López-Romero,[b] Karl-Peter Hopfner[b] and Ivan Huc.[a]

[a] Department of Pharmacy, Ludwig-Maximilians-Universität, Munich, Bavaria, 81377, Germany.

[b] Gene Center, Ludwig-Maximilians-Universität, Munich, Bavaria, 81377, Germany.

ABSTRACT

DNA mimic foldamers (DMFOs) are helically folded aromatic oligoamides bearing negatively charged side chains that mimic the shape and charge distribution of double-stranded B-DNA, which arises from the alternation of two different quinoline backbone moieties. Here we present an advancement of their synthesis, by combining them into a dimeric macromonomer. Three different Fmoc-protected dimers, one for each side chain pattern, have been produced and tested for chain elongation on a solid-support utilizing standard peptide conditions. This has allowed us to synthesize molecules previously inaccessible, rivaling the size of the synthetically available peptides. These longer oligomers were analyzed by reversed phase high-performance liquid chromatography, mass spectrometry, circular dichroism, and nuclear magnetic resonance spectroscopy for their purity, identity, and helix-handedness bias. Clusters of foldamers could be visualized at low resolution in cryogenic electron microscopy (cryo-EM), validating their dimensions. Lastly, we demonstrate that DMFOs of this length enable binding of the A-module of the chromatin remodeling complex INO80. A cryo-EM structure of their complex unveiled its binding mode to be alongside Ino80's HSA domain, albeit in a slightly different orientation than extranucleosomal straight or curved DNA.

INTRODUCTION

Many cellular processes which organize and process genetic information are executed by multidomain protein assemblies.[1] This for example includes the packaging of deoxyribonucleic acid (DNA) by wrapping around histone octamers,[2] the reorganization of chromatin by remodeling complexes,[3, 4] and the whole translational machinery revolving around the ribosome.[5] The regulation of these processes usually cannot be achieved by small molecules alone, but *in cellulo* is executed by other large protein assemblies or multi-subunit complexes.[4] Thus, in molecular biology, the interaction with large molecular complexes on a functional level requires other large molecular entities.

For these reasons, many efforts have been made to increase the size of chemically synthesized peptides or oligonucleotides. While, the practical size limit for standard oligonucleotide synthesis is considered to be around 200 nucleotides,[6] the chemical synthesis of an entire gene with 1728 nucleotides has been reported.[7] Peptides have been synthesized with lengths over 200 amino acids using automated flow systems,[8, 9] but are more typically produced in the range of 30-50 units, especially when using standard synthesizers or working on larger scales.[10, 11]

Oligomers based on abiotic backbones such as aromatic oligoamide foldamers (AOFs) are promising candidates to interfere in protein-protein or protein-nucleic acid interactions (PNIs), especially because they are resistant to proteolytic and nuclease degradation when compared with their biological counterparts.[12] As they can be functionalized with a diverse selection of side chains, a chemical space beyond the 20 biogenic amino acids is available.[13-15] Opposed to peptides or other oligomers based on abiotic backbones such as β-peptides, oligoureas, or peptoids, the backbone conformation of AOFs based on 8-aminoquinoline-2 carboxylic acid (Q) (Fig. 1A) is largely conserved and independent of their side chains.[16] This confers a high degree of structural predictability, which is a feature that is critical for effective molecular mimicry and recognition.

When AOFs are functionalized with anionic side chains like phosphonates or carboxylates,[17-19] they can serve as mimics of B-form DNA duplexes (Fig. 1B). They can replicate typical features like DNA's grooves and its charge surface distribution from a single helical array by alternating Q with M (Fig. 1B), a derivative of Q with an extra methylene group between the quinoline ring and its amine (Fig. 1C). One MQ dimer is the structural equivalent to one base pair (bp) in DNA, as they contribute likewise to vertical rise and helix curvature. Similarly to DNA mimic proteins and anionic polymers, DMFOs show the ability to interfere with PNIs,[20-22] including the inhibition of therapeutically relevant enzymes.[17, 18, 22, 23] Synthetically, DMFOs were not readily available in the size range that was achieved with peptides and oligonucleotides.

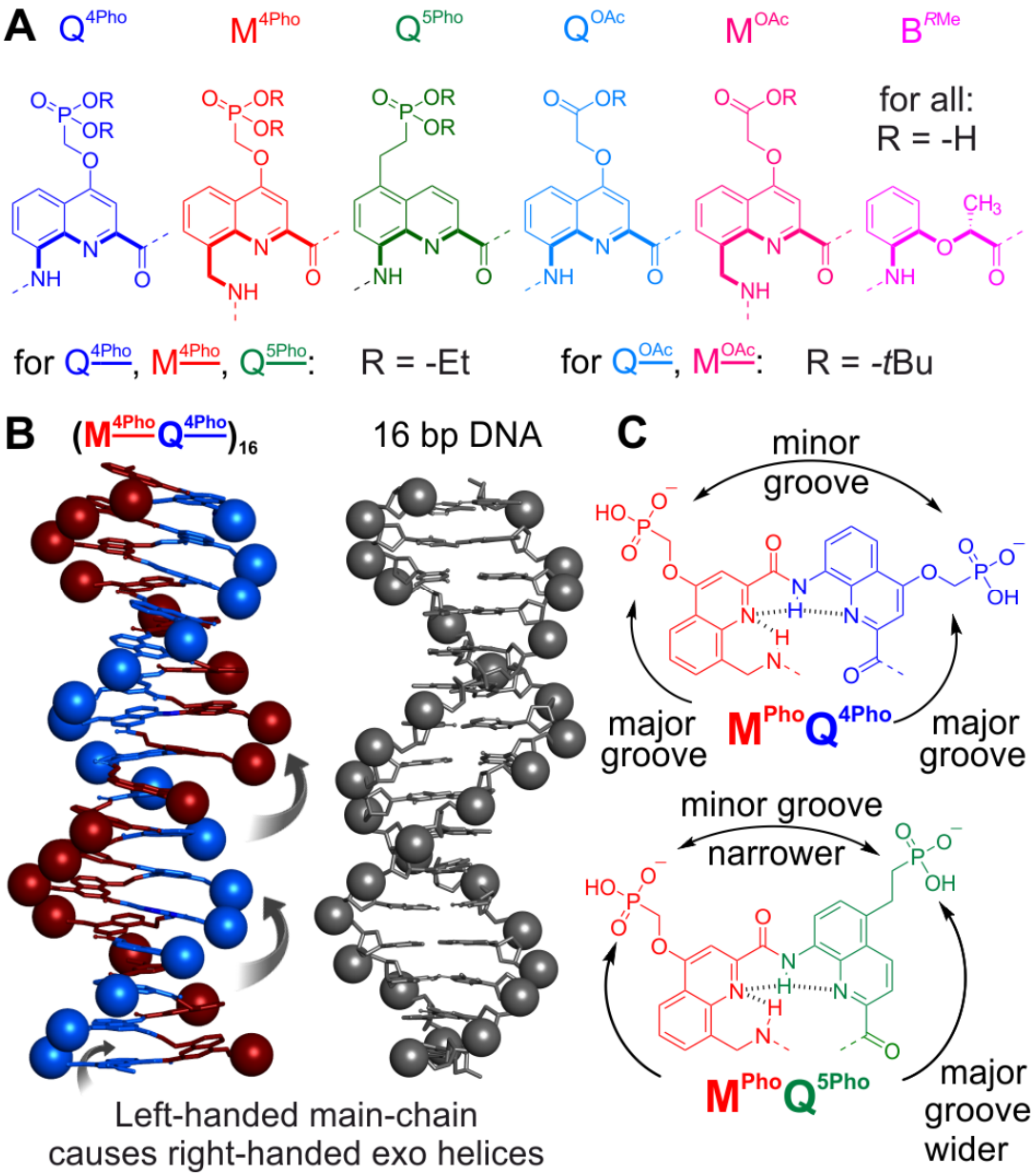


Figure 1 A) Monomeric building blocks of DMFOs in either side chain protected form (underlined) or free acid form (not underlined). The bold bonds highlight the main chain. B) Left: solid state structure of the side chain protected $(M^{4Pho}Q^{4Pho})_{16}$ (M units in red, Q units in blue, in stick representation except the phosphorus atom, that is shown in space filling representation, ethyl ester protection of the phosphonate omitted for clarity), center: structure of the 16-bp B-DNA duplex 5'-d(ACTACAATGTTGCAAT)-3' (gray, stick representation except the phosphorus atom shown in space filling representation).[24] C) Repeat unit for $(M^{4Pho}Q^{4Pho})_n$ (top) and $(M^{4Pho}Q^{5Pho})_n$ (bottom) dimers.

The production of longer DMFOs may allow interfering in the PNIs between long DNA segments and multi-protein complexes that dominate key processes of the cellular regulatory machinery.

INO80 is a >1 megadalton highly conserved chromatin remodelling complex,[25, 26] which slides nucleosomes to space them correctly,[27] especially the placement of the +1 nucleosome downstream of the transcription start site.[28] It consists of three major modules: "N", "A" and "C". The structures of A- and C-modules have been characterized both in the presence and absence of their respective substrates.[29-32] Among others, the C-module is composed of the central ATPase motor containing unit, a hexameric Rvb1–2 structural scaffold that is connected to the INO80 motor by Ino Eighty Subunit 2 (Ies2) and Arp5 (actin-related protein 5). Together with Ies6, they interact with the nucleosome by its unique grappler insertion element.[30] The A-module acts as a sensor for extranucleosomal DNA and is composed of helicase-SANT–associated (HSA) domain, along with actin (Act1), Arp4, Arp8, Ies4, and Taf14 of which Arp8 and the HSA domain were identified to be the core DNA binding elements.[32]

Several of the proteins that compose INO80 (Rvb1/Rvb2, Arp8 and Arp5) were identified as significant DMFO binders in pulldown assay against drosophila embryo extract (DREX), but were not affected when chromatin was assembled in vitro in the presence of a DMFO.[22] However, the DMFOs utilized in this study were mimics of 16 base pairs bps of DNA, much shorter than the immobilized duplex DNA (ca. 1.4 kbps) used for chromatin assembly.

Thus, we aim to produce one-handed DMFOs, which are significantly longer than previously available. These DMFOs will be utilized to investigate their binding propensities against multi-protein complexes. As a sensor of straight and curved extranucleosomal DNA,[31] we envisioned the A-module of INO80 as a potential target to probe binding interactions.

MATERIALS AND METHODS

Reagents for synthesis

Preparation of reagents for synthesis is given in the Supplementary Information. Synthetic procedures for small molecules, solid phase foldamer synthesis (SPFS), and solution-phase oligomer synthetic steps including characterization are described in detail in the Supplementary Information.

Nuclear magnetic resonance (NMR) spectroscopy

Small molecule NMR spectra were recorded on AVANCE NEO NMR spectrometer 500 MHz (Bruker BioSpin) with CryoProbe™ Prodigy and a BCU II. $CDCl_3$ (δ_H: 7.26, δ_C: 77.16) and DMSO-d_6 (δ_H: 2.50, δ_C: 39.52) were used as solvents and their residual solvent signals were used as internal standard.[33] The raw data was evaluated using *Mnova version 14.0.0* by *Mestrelab Research*. The derived data signals are stated with chemical shift in ppm, their multiplicity (s = singlet, bs = broad singlet, d = doublet, t = triplet, q = quadruplet, m = multiplet, or a combination of these), their coupling constant in Hz and their integrated values.

Reversed-phase (RP) high performance liquid chromatography (HPLC)

Analytical and semi-preparative RP-HPLC were performed on a Thermo Fisher Scientific Ultimate 3000. Preparative RP-HPLC was performed on a Waters system equipped with a 2545 Quaternary Gradient Module. The mobile phase was composed of ultra-pure water (A) and acetonitrile (B) with different additives (trifluoroacetic acid, formic acid, NH_4OAc or NEt_3HOAc) depending on the application. More detail is given in the Supplementary Information.

Mass spectrometry (MS)

Small molecule mass spectra were recorded by direct infusion on a Thermo Finnigan LTQ FT Ultra or Thermo Finnigan LTQ Orbitrap XL mass spectrometer with a resolution of 100.000 at m/z 400. Depending on the sample, mass ranges from 40 to 2000 u were measured. The spray capillary voltage was 3 kV, the heater capillary temperature 250 °C, the nitrogen gas flow 25 units and the auxiliary gas flow 5 units. Foldamer mass spectra were obtained from liquid chromatography (LC)-MS on a micrOTOF II mass spectrometer by Bruker Daltonics and molecules were ionized by Electrospray Ionization $(ESI)^+$ or ESI^- depending on the analyte.

Dynamic light scattering (DLS)

DLS measurements were performed using a DynaPro™ Nanostar™ DLS system (Wyatt Technology, Santa Barbara, CA) and analysis was done by using the DYNAMICS® software (version 8.0.0.77) provided by the manufacturer. Results from 20 consecutive measurements at 25 °C (laser wavelength 661.3 nm) each with an acquisition time of 3 s and a read interval of 1 s were averaged for every sample. The samples were centrifuged before the measurements at 15000 *g* for 10 min and the measurement was performed with the supernatant solution.

Circular dichroism (CD) spectroscopy

CD-spectra were recorded on a Jasco J-1500 CD spectrometer. CD-spectra were recorded from 450 to 250 nm with a continuous scanning rate of 50 nm/min, a Digital Integration Time (D.I.T.) of 0.5 seconds and a bandwidth of 1.00 nm. Spectra are plotted as molar Δε to correct for concentration. Concentrations were measured at the spectrophotometer in parallel to the CD measurements. MQ-dimer concentration (conceptually comparable to base pair concentration) was kept between 150-200 µM. The data shown are the average of two measurements and were smoothed using a Savitzky-Golay filter with a polynomial order of 3.

Expression and purification of the INO80ΔN complex from *C. thermophilum*

The purification of C. thermophilum INO80 followed the protocol as described previously.[31] All subunits of the INO80 complex were cloned and expressed using MultiBac technology (MultiBac™). Gene encoding $Ino80^{718-1848}$ (CtIno80ΔN)–2xFLAG® or $Ino80^{545-850}$ (INO80 A-module constructs) –2xFLAG® was cloned into the pACEBac1 vector; genes for Ies2, Ies6 and Arp5 were cloned into the pIDK vector and Rvb1 and Rvb2 were cloned in pIDC. Together, these subunits were assembled into one single bacmid (produced in *E. Coli* DH10 MultiBac cells). Another bacmid was assembled for the A-module, thereby genes for Ies4 and Taf14 were cloned in pACEBac1. Actin, Arp4 and Arp8 were cloned into a pIDK vector. All recombination steps, using Cre recombinase (NEB), were carried out in PirHC (Geneva Biotech) and *E. coli* XL1-Blue (Stratagene) cells. From each bacmid, baculoviruses were generated in Spodoptera frugiperda (Sf21) insect cells (Thermo Fisher Scientific, #11497013). One liter of Trichoplusia ni High Five cells (Invitrogen, #B85502) was infected with both baculoviruses (1:100) and cultured for 60 hours at 27°C. After harvesting by centrifugation at 4°C for 15 minutes at 4,500 rpm, the cells were lysed in lysis buffer [30 mM HEPES (pH 7.8), 400 mM NaCl, 10% glycerol, 0.25 mM DTT, 0.28 µg/ml pepstatin A, 0.17 mg/ml PMSF, 0.33 mg/ml benzamidine, and 2 mM $MgCl_2$] and gently sonicated for 2 minutes (output control 5 and 50% duty cycle). The lysate was cleared by centrifugation at 20,500g and 4°C for 30 minutes. Followed by this, the supernatant was incubated with 2 ml of ANTI-FLAG® M2 Affinity Gel (Sigma-Aldrich) for 1 hour. After sequential washes with 50 ml of lysis buffer and 75 ml of wash buffer [30 mM HEPES (pH 7.8), 150 mM NaCl, 5% glycerol, and 0.25 mM DTT] the protein was eluted by

three consecutive incubations, each with 1.5 ml of wash buffer supplemented with FLAG® peptide (0.2 mg/ml), for 20 minutes each. The pooled fractions were purified using a Mono Q 5/50 GL 1 mL column (GE Healthcare), yielding highly purified INO80.

Expression and purification of RvLEAM$_{short}$

The RvLEAM$_{short}$ expression plasmid, pET15b-RvLEAM$_{short}$, was a gift from Ci Ji Lim (Addgene plasmid #218122) and expressed and purified as described.[34] Briefly, recombinant RvLEAM$_{short}$ was expressed in *E. coli* BL21 Rosetta (DE3) cells. A 1 mL preculture of Lysogeny Broth (LB) media, supplemented with 34 µg/mL chloramphenicol and 100 µg/mL ampicillin, was grown overnight at 37°C and used to inoculate 1 L terrific broth (TB) media, supplemented with chloramphenicol and ampicillin. Cells were grown to an absorbance (A_{600}) of 0.6 for induction with 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and then grown for another 16 hours at 12°C.Cells were harvested by centrifugation and resuspended in lysis buffer (50 mM HEPES pH 7.5, 300 mM NaCl, 10 mM imidazole, 1 mM DTT, 1 mM PMSF), lyzed by sonication, and the clarified lysate was incubated with pre-equilibrated nickel-NTA resin (Qiagen, 50 mg/mL binding capacity) for 1 hour at 4°C. The resin was washed thrice with 50 mL of lysis buffer before elution with 50 mM HEPES pH 7.5, 300 mM NaCl, 250 mM imidazole, and 1 mM DTT. The protein was concentrated and purified by size exclusion chromatography on a Superdex 75 10/300 column (25 mM HEPES pH 7.5, 100 mM NaCl, 1 mM DTT). Protein-containing fractions were analyzed by SDS-PAGE, pooled, concentrated and stored at –80 °C until usage.

Bio-layer interferometry

BLI measurements were carried out at 25 °C on an Octet R8 BLI Sartorius instrument, using streptavidin (SA) biosensors. The biotinylated DMFO (0.5 µM) or biotinylated double-stranded DNA (1 µM) was loaded on SA sensors for 60 s after a 60 s baseline, to reach 0.4-0.5 nm of wavelength shift. In the case of DNA, one of the two strands was biotinylated on the 5' end (5'-TCGTAGCAAGCTCTAGCACCGCTTAAACGCACGTAC-3'), and DNA was annealed prior to loading on the sensors. The biosensors were washed for 120 s in kinetic buffer: either PBST (PBS with 0.05 % Tween 20) or "low salt phosphate buffer (LSPBT)": Phosphate buffer with 50 mM NaCl and 0.05% Tween 20. Tween 20 was added as a blocking reagent, to reduce non-specific interactions. The association was recorded for 120 s in increasing concentrations of freshly thawed INO80 A-module (from 0 to 45 or 91 nM). The dissociation was recorded in the kinetic buffer. A global curve fitting was next performed using the reprocessing software embedded with the BLI Octet R8 instrument. A global curve fitting was performed to calculate the K_D values. Non-specific interactions of INO80 module A with the SA biosensor were tested and considered negligible (< 0.04 nm shift). All kinetics (curves and values) are accessible in the Supplementary Information.

Cryo-EM grid preparation foldamer clusters

CtINO80$^{545-850}$ A-module and foldamer **40** were mixed in a ratio of 1:1 and dialyzed against cryo-EM reaction buffer I [20 mM HEPES (pH 7.5), 50 mM KCl, 0.25 mM CaCl2, 2 mM MgCl2, and 0.5 mM DTT] for one hour at 25°C. The complex was vitrified at a concentration of 1 mg/ml onto a glow discharged QUANTIFOIL® R2/1 Cu200 grid in the presence of 0.045% octyl-β-glucoside using a Leica EM GP (Leica; 10 °C and 90% humidity).

Cryo-EM grid preparation A-module foldamer

CtINO80ΔN and **40** were mixed in a ratio of 1:2 and dialyzed against cryo-EM reaction buffer II [25 mM HEPES (pH 8), 5 mM EDTA, 80 mM KCl, 0.5 mM DTT], for one hour at 25°C with and without 20x-fold excess LEA-protein. The sample was vitrified at a concentration of 1 mg/ml onto a glow discharged QUANTIFOIL® R2/1 Cu200 grid in the presence of 0.045% octyl-β-glucoside using a Leica EM GP (Leica; 10 °C and 90% humidity).

Cryo-EM data collection

Movies of A-module and foldamer **40** particles embedded in vitreous ice were collected at liquid nitrogen temperature using an FEI Titan Krios G3 transmission electron microscope (300 kV) a GIF quantum energy filter (slit width 20 eV) and a Gatan K2 Summit direct electron detector (software used: EPU 3.5.1, TEM User interface Titan 3.15.1, Digital Micrograph 3.22.1461.0) The movies were recorded in counting mode using EPU acquisition software (Thermo Fisher Scientific) at ×165,000 magnification with a pixel size of 1.049 Å/pixel and nominal defocus range of -1.1 to -2.9µm. The total electron dosage of each movie was 41.709 e/Å2, fractionated into 40 movie frames.

Cryo-EM data processing

Movies of CtINO80ΔN and foldamer **40** particles embedded in vitreous ice were collected at liquid nitrogen temperature using an FEI Titan Krios G3 transmission electron microscope (300 kV) equipped with a Selectris X imaging filter (slit with 10eV) and a Falcon4 direct electron detector. The movies were recorded in counting mode using EPU acquisition software (Thermo Fisher Scientific) at ×165,000 magnification with a pixel size of 0.727 Å/pixel and nominal defocus range of -1.1 to -2.9µm. The total electron dosage of each movie was 40 e/Å2, fractionated into 40 movie frames.

Cryo-EM image processing of foldamer clusters of 40

The movie frames were motion-corrected using MotionCor2.[35] All subsequent processing steps were performed in cryoSPARC (v4.5.3),[36] and the resolutions reported here are calculated based on the gold-standard Fourier shell correlation criterion (FSC = 0.143).[37] The CTF parameters of the dataset (2518 micrographs) were determined using patch CTF estimation

(multi) in cryoSPARC (v4.5.3).[38] The exact processing scheme is depicted in Figure S1. Initial particle picking was done using Blob picker (3400 k particles). Particles were subjected to 2D classification. After manual picking (234 particles), one round of Topaz was performed.[39] 4.8k particles were extracted with a box size of 280 (3x binned) px and a pixel size of 1.049. 4.8 k particles were used for *ab initio* reconstruction and heterogeneous refinement into two classes. One class with 3.9 k particles was used for another round of 3D classification and homogenous refinement (3.9 k particles). The final resolution of the foldamer cluster reconstruction was 13.29 Å. Figures were generated in UCSF ChimeraX 1.6.145.[40]

Cryo-EM image processing of CtINO80ΔN and 40

The movie frames were motion-corrected using MotionCor2.[35] All subsequent processing steps were performed in cryoSPARC (v4.5.3),[36] and the resolutions reported here are calculated based on the gold-standard Fourier shell correlation criterion (FSC = 0.143).[37] The CTF parameters of the dataset (28096 micrographs) were determined using patch CTF estimation (multi) in cryoSPARC (v4.5.3). The exact processing scheme is depicted in Figure S2. Initial particle picking was done using Blob picker (22000 k particles). Particles were subjected to 2D classification, *ab initio* reconstruction (194k particles). After several rounds of heterogeneous refinements and 3D classification, one class with 22 k particles was used for homogenous and non-uniform refinement. The final resolution of the A-module-foldamer complex was 3.59 Å. Figures were generated in UCSF ChimeraX 1.6.145.[40]

Molecular modelling & docking

The molecular models of conjugates and DNA mimics were built in Maestro (Version 11.5) [41] and energy-minimized with parameters in table S1. Foldamer coordinates for the models of conjugates were derived from earlier crystal structures from diethylester-protected polyphosphonate foldamers,[18] with their side-chain protecting groups removed and subsequently energy-minimized according to parameters given in the Supplementary Information. For figure generation of cryo-EM reconstructions, PDB: 8A5P and a model of the foldamer (see above) were rigid body docked into the densities using UCSF ChimeraX 1.6.145.

RESULTS & DISCUSSION

Dimeric building blocks as macro-monomers

Previous DNA mimic foldamer synthesis. Initially, DMFOs have been synthesized exclusively by solution-phase fragment condensation (doubling their length) to give molecules of defined 2^n length with a maximum reported foldamer length of 32 units (mimicking 16 bp DNA).[18] This strategy has two key advantages: first, the coupling reaction rapidly increases chain length (n = 7 in a 2^n scheme yields 128 units); second, it enables scalable production of substantial quantities of material. For these reasons, it has been previously employed for DMFOs and other AOFs.[18, 42, 43] While both seem key advantages, this synthetic strategy is severely hampered by an inflexibility to rapidly change single units in a designed molecule, such as the incorporation of a chiral B^{RMe} (Fig. 1A) monomer **1** for handedness control of sequences (see below).[44] Inspired by peptides, the synthesis of DMFOs has been mostly shifted towards solid-phase chemistry utilizing Fmoc-acid derivatives of Q and M **2-6** (Fig. S3).[14, 17, 19, 45-47] Although some sequence alternation is required, most DMFO sequences consist of extended $(MQ)_n$ segments, enabling the use of dimers as macro-monomers.

Dimers over monomers. The primary advantage of using dimers rather than monomers for chain elongation is that it halves the number of required coupling steps, thereby saving material, reagents, synthesizer time, and solvent. In addition, the respective amines of Q and M that represent the nucleophiles for the chain elongation do vary greatly in reactivity, e.g., benzylamine outcompetes aniline by four orders of magnitude.[48] Coupling on Q thus requires much more reactive electrophiles, typically acyl chlorides that must be generated by oxalyl chloride or Ghosez's reagent if the side chain protecting group (PG) is acid-labile,[14] both practically unfeasible to automatize. Automation was achieved by a complex activation cocktail based on trichloro acetonitrile and triphenylphosphine in a variant of the Appel reaction. This rapidly generates numerous reactive intermediates.[45] For coupling on M, this is unnecessary, so common peptide chemistry coupling reagents such as benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate (PyBOP),[49] 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU),[50] or 1-cyano-2-ethoxy-2-oxoethylidenaminooxy)dimethylamino-morpholino-carbenium hexafluorophosphate (COMU)[51] can be utilized, and excess monomer recovery has been reported.[52] By utilizing dimers **7**-**9**, chain elongation can exclusively performed by coupling on the benzylic amine of M. With this strategy, DMFO synthesis can be fully automated with all common peptide synthesizers under standard peptide-SPS conditions.

Synthesis of Fmoc-dimers. Thus, we have synthesized dimeric macro-monomers **7** and **8** for Q^{4Pho} and Q^{5Pho} polyphosphonates (Fig. 2A, B) as well as **9** (Fig. S1) for polycarboxylate DMFOs. The synthesis of **7** and **9** (Fig. 2A, Scheme S3) is based on quinolinone **10**, which can be produced on large scales based on published procedures.[53] The synthetic route to **7** (Fig. 2A) serves as a representative example for both. The side chain was installed under Mitsunobu-conditions to yield **11**, which was hydrogenated to give amine **12**. Base-mediated transesterification from the C-terminal methyl to allyl ester yielded **13**. Then, the Fmoc-protected monomer **4** was amidated by **13** to give the dimer **14** with PyBOP as a dehydrating reagent. The C-terminal carboxylic acid allyl ester was removed with Pd(0)-mediated cleavage to give the final SPS suitable dimer **7** in an overall yield of 33% over 5 steps. The synthesis of Q^{5Pho}-based dimer **8** (Fig 2B) differs from **7** and **9**, particularly in the strategy used for side chain installation. Starting from compound **15**, that is synthesized in two steps from commercial 2-nitro,5-bromo aniline.[13] As the side chain is installed under Heck conditions, an allyl ester protection of the C-terminal acid is likely to react with itself intermolecularly and was thus deemed unviable. Therefore, the C-terminal acid of **15** was protected as a *tert*-butyl (*t*-Bu) ester by EDC·HCl to give **16**. The resulting Heck reaction with commercial vinyl diethyl phosphonate yielded **17** and the resulting double bond was hydrogenated together with the nitro group to give the amine **18**. Amidation of **4** with **18** (similarly to the synthesis of **14**) yielded the dimer **19** which was deprotected using TFA to give **8** in an overall yield of 49% over 5 steps.

Overall, each of the three dimers is synthesized in 13 steps (including 6 steps for M derived monomers **4** and **6**) from commercial compounds in a convergent route. Synthesizing Fmoc-monomers separately culminates in 12 overall steps, which results in comparable synthetic effort for monomer production. The synthesis of **9** via monomer **5** and intermediates **20-26** is described in the supplementary information.

Foldamer synthesis

Solid-phase foldamer synthesis (SPFS) For foldamer synthesis, a low-loading polystyrene-polyethylene glycol hybrid solid support was utilized for its high swelling potential. The loading was further decreased, as target sequences are both long (72-100 units) and heavy, with an average weight of 373 Dalton per unit, more than three times heavier than the average natural amino acid. As a result of this, **27** with a molecular weight of ca. 34.5 kDa compares to a peptide with over 300 units. Foldamer chain elongation was achieved fully automatically on a PurePep Chorus synthesizer using piperidine as a Fmoc deblocking reagent, PyBOP for coupling, and acetic anhydride as capping reagent (Fig 2C). The coupling solution containing excess dimer was collected and treated with citric acid to facilitate building block recovery.

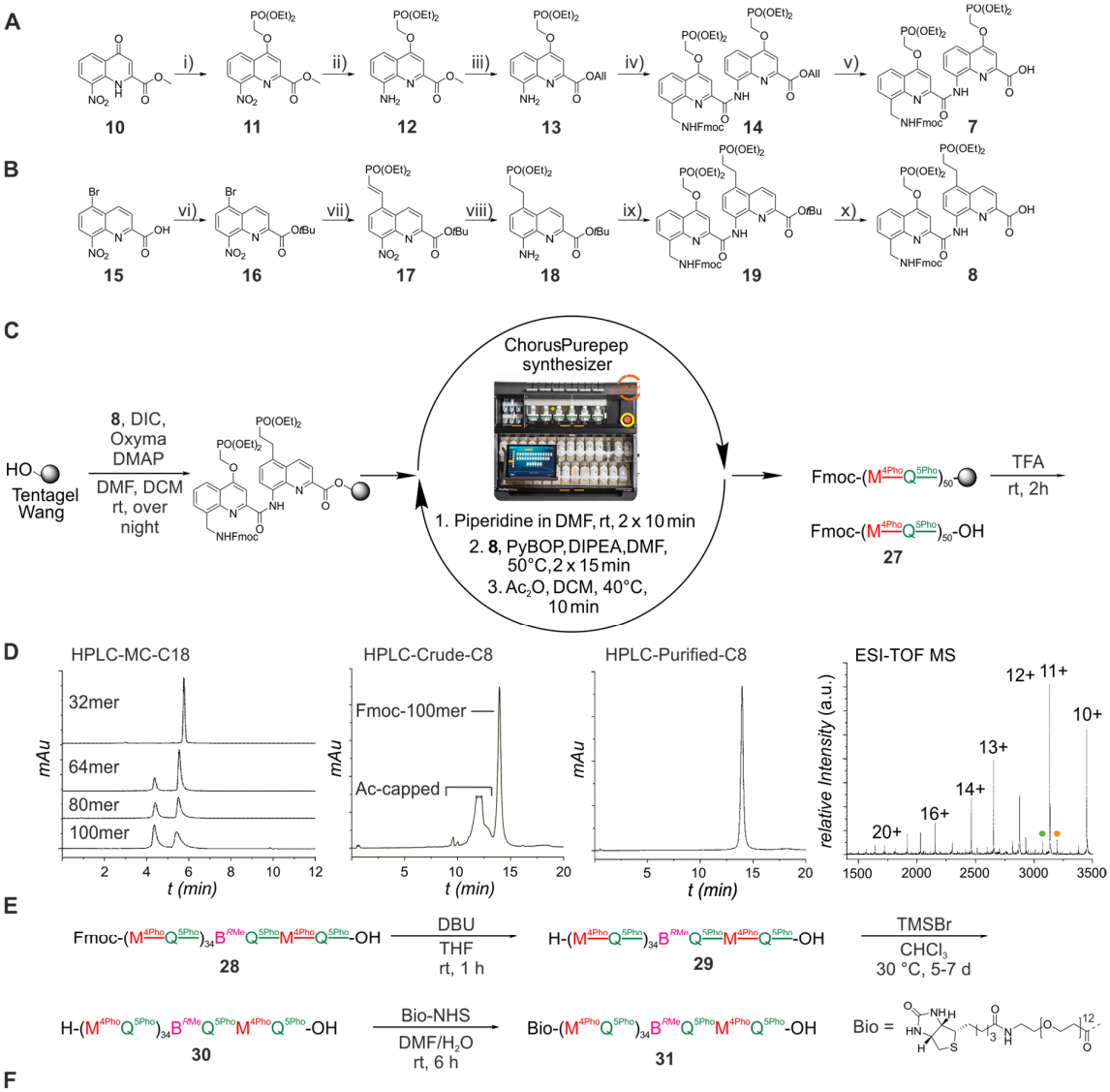


Figure 2 A) Synthesis of dimeric building block **7**: i) $HO(CH_2)PO(OEt_2)$, DIAD, PPh_3, THF, 0 °C to 50 °C, 57%. ii) H_2, Pd/C, EtOAc, rt, 18 h, 97%. iii) AllOH, TEA, 70 °C, 2 d, 76%. iv) **3**, PyBOP, DIPEA, $CHCl_3$, rt, over night, 86%. v) $PhSiH_3$, cat. $Pd(Ph_3)_4$, CH_2Cl_2, rt, 1 h, 91%. B) Synthesis of dimeric building block **8**. vi) *tert*-Butanol, EDC·HCl, DMAP, $CHCl_3$, 0 °C to rt, over night, 82%. vii) $(CH_2)(CH)PO(OEt_2)$, $Pd(PPh_3)_2Cl_2$, K_2CO_3, *o*-xylene, 125 °C, 18 h, 71%. viii) H_2, Pd/C, EtOAc, rt, 16 h, quant. ix) **3**, PyBOP, DIPEA, $CHCl_3$, rt, over night, 98%. x) TFA/DCM, rt, 8 h, 86%. DIAD = diisopropyl azodicarboxylate, TEA = triethyl amine, DIPEA = diisopropyl ethylamine, PyBOP = benzotriazol-1-yloxytripyrrolidinophosphonium hexafluoro-phosphate. C) Resin loading and chain-elongation of **8** to 100-unit globally protected foldamer **27**. D) C18 (left) and C8 (middle, right)-RP HPLC chromatograms of synthesis crude, purified compound **27** and mini-cleavages of intermediate compounds of different length. ESI-TOF mass spectrum of purified compound **27**. Green and orange circles represent impurities of a dimer uncapped deleted species and a dimer over-inserted species respectively. E) Further synthetic steps of **28** to yield final water-soluble compounds **30** and **31** of 72-unit length unlabeled and biotin-labeled chiral foldamers. F) Comprehensive list of all additional synthesized foldamer sequences and intermediates **32-43**.

Resin cleavage was conducted in TFA to yield Fmoc and globally protected foldamers such as **27**. Upon elongation, chromatograms on RP-HPLC with C18 stationary phases broaden upon reaching certain lengths (>64 units) (Fig 2D) but stay significantly sharper on C8 stationary phases. The terminal Fmoc group shifts the target sequence to higher retention times, and serves as a hydrophobic purification tag, a concept that has been explored previously.[54] The mass spectrum of **27** does reveal that deletions (green circle, Fig. 2D) and over insertions (orange circle) cannot be separated with this method, as their retention time overlap due to their high sequence homology. Deletions or over insertions in a monomer-by-monomer synthesis causes consecutive Q or M units that result in a frameshift of the foldamer's grooves (Fig. 1C). In contrast, deleting an entire dimer only shortens the overall sequence length, as all accumulated deletions lead to the same product. In this case, the sequence homology is preserved. In the case of polycarboxylate sequences, which were synthesized with dimer **9**, side chains are protected as *t*Bu-esters. These are removed during resin cleavage to directly yield water-soluble compounds **42** and **43**.

Overall, there is potential to optimize the solid-phase synthesis of DMFOs with this approach, but since every synthetic attempt requires a modest amount of Fmoc-protected dimer (ca. 2.7 g of dimer for 15 μmol scale of **27**), this was not pursued at this stage.

Post SPFS Isolated Fmoc and polydiethyl phosphonate protected foldamers such as **28** may be used directly in global deprotection reactions to give **40**. Otherwise, Fmoc can be removed prior by 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) to yield **29** (also partly cleaving diethyl protecting groups, see oligomer mass spectra). After global deprotection, foldamers **30** and **38** can be labeled with a biotin moiety to give **31** and **39** respectively.

Whereas ESI mass spectra of globally protected foldamers display excellent signal-to-noise ratio, polyanionic compounds display noisier spectra even though their phosphonic acid residues are at least monoanionic at neutral pH.[18]

Foldamer analysis

Controlling foldamer handedness $(MQ)_n$ foldamers are achiral, and in the absence of other chiral information, exist as racemic mixtures of right-handed (*P*)- and left-handed (*M*)-helical conformers. The helix-handedness of foldamers with the length of 16 units such as **41** can be biased up to a diastereomeric ratio (DR) of ca. 99% toward the *M*-helical conformer upon insertion of a single chiral B^{RMe} unit (Figure 1A) in the sequence. For example, when in thermodynamic equilibrium, the ^{1}H nuclear magnetic resonance spectrum of **41** shows two conformers with a ratio of 98 to 2, reflecting this diastereomeric ratio, and its CD spectrum shows a negative band with a maximum at 350 nm (for Q^{5Pho}, 365 nm for Q^{4Pho}-based sequences) because of its predominant *M* helicity. Importantly, it is the *M* helical conformation of $(MQ)_n$ oligomers that mimics the shape and charge distribution of B-DNA because it projects

its phosphonate side chains so that they form two *P* exo-helices that match with the *P* double helix of the phosphates of DNA (Figure 1B). Chiral foldamers have a B^{RMe} unit incorporated close to the C-terminus, because the monomer by monomer incorporation requires a different activation approach (acyl chloride) and thus were synthesized manually up until hexamers using Appel conditions.[45] The analysis of the globally deprotected foldamer **40** by ^{1}H-NMR spectroscopy shows signals that belong to the foldamer *P*-helix (Fig. 3A), which was confirmed by a weak CD intensity (Fig 3B). Remeasuring after heating the sample for 24 h to 55 °C in water, an enrichment of the desired *M*-helical species by both techniques was confirmed. The CD intensity at 350 nm, normalized per residue, can be compared with **41**, a compound of known DR. From this comparison, we deduct that the incorporation of one B^{RMe} unit at the foldamer C-terminal extremity does result in a close to quantitative handedness bias, even for a foldamer of this length. Similar analyses for Q[4]-based sequences with derivatives **36** and **37** as well as **42** and **43** (Figs. S4-S7) confirmed these results.

Foldamer aggregates We attempted to visualize DMFO **40** in complex with the A-module of INO80 (see below) using cryogenic electron microscopy (cryo-EM). However, using the original cryo-EM reaction buffer I (20 mM HEPES (pH 7.5), 60 mM KCl, 0.25 mM $CaCl_2$, 20 µM $ZnCl_2$, and 0.25 mM DTT) resulted in highly clustered foldamers (Figs. 3 D-F, Fig. S1). Interestingly, the foldamers aggregate orderly with regular distance and space between themselves. The model of the foldamer could be fitted into the obtained density, verifying the length of the synthesized object (Fig. 3F). Later, analysis by dynamic light scattering (DLS) revealed that foldamers aggregate in the presence of divalent cations (especially Ca^{2+}, Fig. 3C)

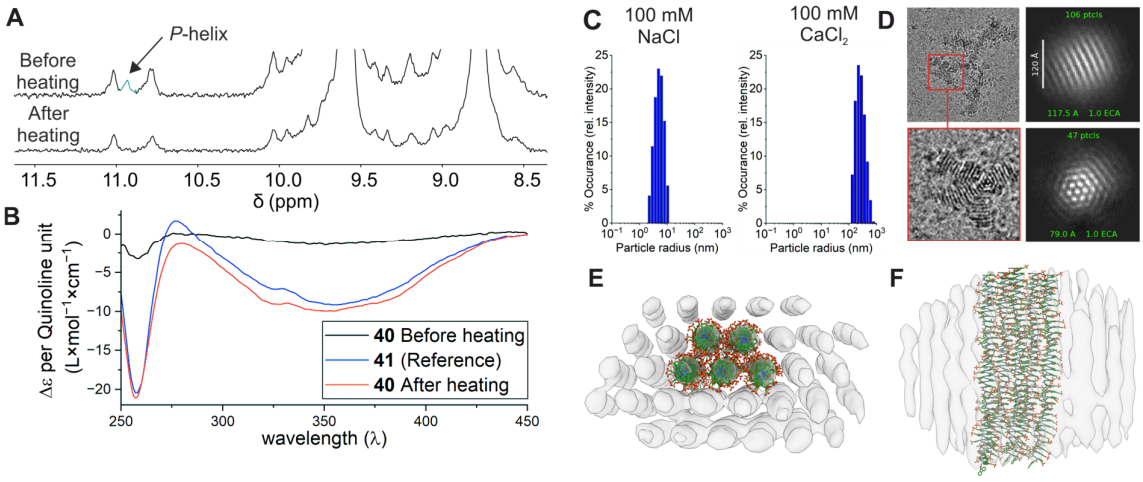


Figure 3 A) Excerpts of the ^{1}H NMR spectra of compound **40** before and after heating the sample to 55 °C for 20 h. The signal attributed to the foldamer *P*-helix highlighted in teal loses intensity upon heating. B) CD-spectra of compound **40** before and after heating and compound **41** with DR ca. 99%. C) Percentage occurrence of particle radii measured by DLS (c = 1mg/mL) with NaCl or $CaCl_2$. D) Representative micrographs and 2D class averages of foldamer aggregates visualized by cryo-EM.3D reconstruction of foldamers. An energy-minimized molecular model of **40** was fitted into the density in topview E) and side view F).

and cause precipitation which is only visible at higher concentrations. Subsequently, buffers containing various additives were tested, and even in the absence of divalent cations, some larger particles were observed (Fig. S3 C), which necessitated the use of EDTA to chelate any residual divalent metal ions.

Analysis of INO80 Arp module – foldamer binding

A 72-unit foldamer corresponds to approximately 36 base pairs of DNA, which matches the length of the modelled DNA segment observed in previous cryo-EM structures of A-module/DNA complexes.[31] We therefore measured the binding affinities of 72-unit foldamers and dsDNA for the A-module of INO80 with Biolayer Interferometry (BLI). Biotinylated DMFOs **31** and **39** were immobilized on streptavidin (SA) sensors. For DNA, one of the two strands was biotinylated on the 5' end, and annealed to its complementary sequence prior to immobilization on SA sensors. The biotinylated molecule was loaded in phosphate buffer saline (PBS) for 60 seconds at a concentration that allowed a 0.4 – 0.6 nm wavelength shift: 500 nM for dsDNA, 1 µM for **31** and **39**.

The A-module of INO80 was stored at −80 °C and freshly thawed before concentration determination and BLI experiments. We worked in two different kinetic buffers, to both of which 0.05 % Tween 20 was added, to reduce non-specific interactions: low salt phosphate buffer (LSPBT) and phosphate-buffered saline (PBST). Non-specific interactions of the A-module with the non-loaded SA biosensor were measured in both buffers, to ensure they were negligible.

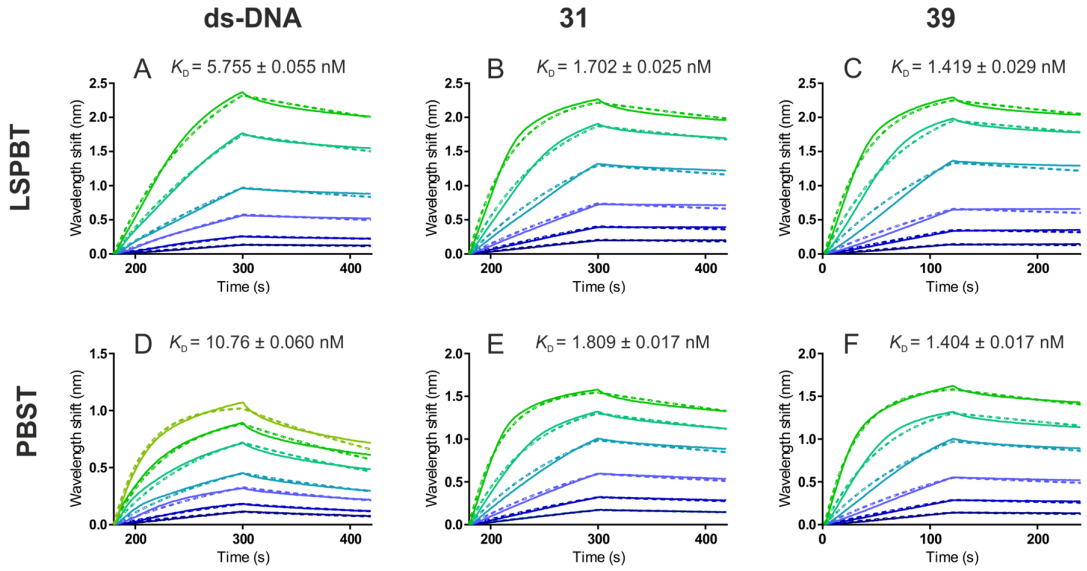


Figure 4 Bio-Layer Interferometry (BLI) sensorgrams of INO80 module A against dsDNA (5'-TCGTAGCAAGCTCTAGCACCGCTTAAACGCACGTAC-3') (A, D), foldamer **31** (B, E) and foldamer **39** (C, F) immobilized on streptavidin sensors. Top figures are in low salt phosphate buffer (LSPBT) and bottom ones in phosphate buffer saline (PBST), both supplemented with 0.05 % Tween 20. Solid lines represent experimental data, and dashed lines are curves fitted to a 1:1 binding model. Calculated K_D values are indicated on the graphs (K_D ± K_D error, in nM). All binding kinetics are listed in the Supplementary Information.

After a baseline, we measured association and dissociation in both buffers. (Figs. 4, S10-S15, Tables S2-S8). The curves were fitted with a 1:1 global curve fitting, that allowed the calculation of the K_D. We chose a range of A-module concentrations up to approximately 10 times the K_D: 45.5 nM in most cases and 91 nM for the binding of dsDNA in PBST, as binding is weaker in this case. Binding to both **31** and **39** is comparable, and stronger (K_D = 1.4–1.8 nM) than for dsDNA: five times stronger in LSPBT and one order of magnitude stronger in PBST. A higher concentration of NaCl affects dsDNA binding. The binding of both DMFOs, on the other hand, does not appear to be affected by salt.

Cryo-EM of INO80 A-module-DMFO complexes

The validation of the binding of foldamer to the INO80 A-module complex, prompted us to revisit cryo-EM visualization attempts. With a revised buffer system (25 mM HEPES (pH 8), 5 mM EDTA, 80 mM KCl, 0.5 mM DTT), foldamer clustering was drastically reduced, and the first 2D classes of INO80 A-module: foldamer **40**-bound classes were obtained (Fig. 5B), but orientational bias prevented 3D reconstruction. Further orientations were derived by the usage of an INO80ΔN construct that includes the INO80 C-module as opposed to the isolated A-module. These could be expanded upon additional 2D classes by the addition of LEA protein as a surfactant.[34] 3D reconstructions were carried out, achieving an overall resolution of 3.59 Å (Fig. S2). The protein architecture is indistinguishable from previously obtained structures of the A-module (PDB: 8A5D). The HSA1 and HSA2 domains serve as a scaffold where Arp8, actin, Ies4 and Arp4 assemble onto (Fig. 5A). Comparing foldamer (this study) and DNA (PDB 8A5P/Q) binding to the A-module, binding modes differ, but both overall bind similar protein domains (Fig. S9). The foldamer prefers binding to the Arp8 subunit and HSA2, while DNA further interacts with the HSA1 domain and the Arp8 N-terminus (Fig. S9), which might be due to the length discrepancy between the two substrates. Nevertheless, we did not obtain further classes in which the foldamer interacts with the HSA1 domain and the Arp8 N-terminus. Overall, the obtained foldamer density is less defined than the protein complex and DNA in DNA/A-module structures.[31] The binding mode of the foldamer may be more heterogeneous, as it can bind in either N-to-C or C-to-N orientation. Further variations that can cause distinct orientational states are lateral displacements or differences in foldamer helical roll. As such, foldamer grooves cannot be allocated in the density, and the overall density does not fully match the predicted foldamer length from molecular modeling, as is also common for unbound overhanging DNA stretches.[31]

When comparing A-module electron density from this study with DNA in curved or straight state (PDB: 8A5P/Q), previously unmodeled electron density of the INO80 motor domain, which leads to the INO80 post-HSA domain (and thus towards the C-module) becomes visible in the presence of foldamer.

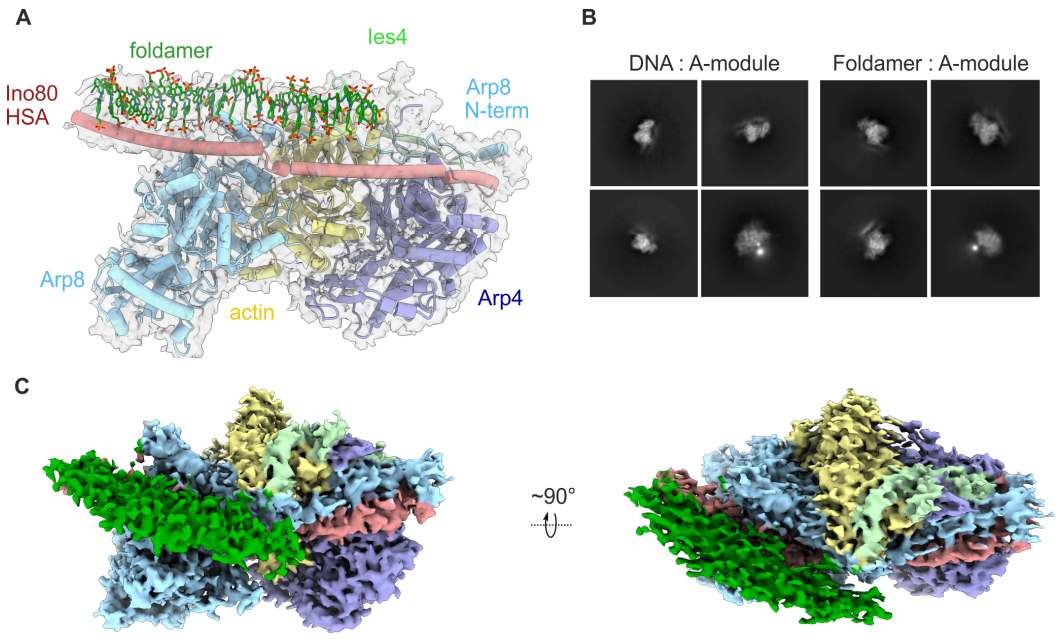


Figure 5 A) Cryo-EM reconstruction (density and structural model) of the *C. Thermophilium* A-module in complex with foldamer **40**. Subunits (HSA1 and HSA2 domains, Arp8, actin, Ies4, Arp4) are color-coded and annotated. Foldamer **40** is depicted in dark green. B) Representative two-dimensional classifications of the A-module in complex with DNA (left) and foldamer (right). The bottom right representations show top views of foldamer or dsDNA. C) Local resolution of the 3D map of the A-module in complex with foldamer **40**, showing all subunits color-coded as in A).

CONCLUSION

We demonstrate that long, water-soluble DMFOs can be synthesized with a dimer-by-dimer strategy reaching dimensions previously inaccessible. The strategy was validated using three distinct Fmoc-protected dimers as macromonomers for SPFS. Coupling efficiencies were consistently high; nevertheless, over-insertion and deletion products were detectable. These side products are negligible for the present applications, as the chosen synthetic strategy ensures sequence homology. Yet there is margin for further optimization. The foldamers DR between *P*- and *M*-helices can be effectively controlled by the incorporation of one CH_3 vs. H chiral differentiation on the foldamers C-terminal extremity. The presence of Ca^{2+} fortuitously enabled direct visualization of the foldamers by cryo-EM due to the formation of clusters albeit at low resolution, nevertheless verifying the overall dimensions of synthesized objects.

Using the A-module of the INO80 chromatin remodeler, we show a single-particle cryo-EM structure of a protein in complex with a large abiotic molecule displaying that DMFOs may represent unique tools for structural biology. Future developments in DMFO design like lowering bending persistence or programmed defined kinking areas may enable reconstruction of nucleosome-like particles without DNA, promote direct engagement with nucleosomes, histones or other large DNA interacting entities, opening additional opportunities for mechanistic and structural studies.

ACKNOWLEDGEMENTS

We thank Leon Stanislawski for the support in small molecule synthesis. We further thank Dr. Lizeth Bodero for operating the Chorus peptide synthesizer and Dr. Lars Allmendinger, for his support with NMR experiments. We further thank Dr. Felix Metzner for his initial work on foldamer aggregates in cryo-EM.

AUTHOR CONTRIBUTIONS

ML conducted the synthesis of small molecules and foldamers, including their analysis. FD and AL expressed the protein and performed biochemical studies. FD and AL prepared samples for cryo-EM measurements. FD processed the cryo-EM data. CD operated the Chorus peptide synthesizer and conceptualized the SPS protocol. LGG performed BLI measurements. ML, FD and LGG wrote the manuscript. KPH and IH supervised the project.

CONFLICT OF INTEREST

There are no conflicts of interest to declare.

REFERENCES

1. E. Nogales, *Nat. Struct. Mol. Biol.*, 2024, **31**, 847-848.
2. K. Luger, A. W. Mader, R. K. Richmond, D. F. Sargent and T. J. Richmond, *Nature (London)*, 1997, **389**, 251-260.
3. A. A. Reyes, R. D. Marcum and Y. He, *J. Mol. Biol.*, 2021, **433**, 166929.
4. C. R. Clapier, J. Iwasa, B. R. Cairns and C. L. Peterson, *Nat. Rev. Mol. Cell Biol.*, 2017, **18**, 407-422.
5. H. F. Noller, *Cell (Cambridge, MA, U. S.)*, 2024, **187**, 6486-6500.
6. L.-F. Song, Z.-H. Deng, Z.-Y. Gong, L.-L. Li and B.-Z. Li, *Front Bioeng Biotechnol*, 2021, **9**, 689797.
7. Y. Yin, R. Arneson, Y. Yuan and S. Fang, *Chem. Sci.*, 2025, **16**, 1966-1973.
8. A. Saebi, J. S. Brown, V. M. Marando, N. Hartrampf, N. M. Chumbler, S. Hanna, M. Poskus, A. Loas, L. L. Kiessling, D. T. Hung*, et al.*, *ACS Chem. Biol.*, 2023, **18**, 518-527.
9. N. Hartrampf, A. Saebi, M. Poskus, Z. P. Gates, A. J. Callahan, A. E. Cowfer, S. Hanna, S. Antilla, C. K. Schissel, A. J. Quartararo*, et al.*, *Science (Washington, DC, U. S.)*, 2020, **368**, 980-987.
10. M. O. Frederick, R. A. Boyse, T. M. Braden, J. R. Calvin, B. M. Campbell, S. M. Changi, S. R. Coffin, C. Condon, O. Gowran, J. McClary Groh*, et al.*, *Org. Process Res. Dev.*, 2021, **25**, 1628-1636.
11. Y. Tan, H. Wu, T. Wei and X. Li, *J. Am. Chem. Soc.*, 2020, **142**, 20288-20298.
12. S. Kwon, V. Morozov, L. Wang, P. K. Mandal, S. Chaignepain, C. Douat and I. Huc, *Org. Biomol. Chem.*, 2024, **22**, 9342-9347.
13. X. Hu, S. J. Dawson, P. K. Mandal, X. de Hatten, B. Baptiste and I. Huc, *Chem. Sci.*, 2017, **8**, 3741-3749.
14. B. Baptiste, C. Douat-Casassus, K. Laxmi-Reddy, F. Godde and I. Huc, *J. Org. Chem.*, 2010, **75**, 7175-7185.
15. M. Zwillinger, P. Soregi, F. Sanchez, C. Douat, M. Csekei, I. Huc and A. Kotschy, *J. Org. Chem.*, 2025, **90**, 3043-3052.
16. J. Buratto, C. Colombo, M. Stupfel, S. J. Dawson, C. Dolain, B. Langlois d'Estaintot, L. Fischer, T. Granier, M. Laguerre, B. Gallois*, et al.*, *Angew. Chem., Int. Ed.*, 2014, **53**, 883-887.
17. V. Corvaglia, D. Carbajo, P. Prabhakaran, K. Ziach, P. K. Mandal, V. Dos Santos, C. Legeay, R. Vogel, V. Parissi, P. Pourquier*, et al.*, *Nucleic Acids Res.*, 2019, **47**, 5511-5521.
18. K. Ziach, C. Chollet, V. Parissi, P. Prabhakaran, M. Marchivie, V. Corvaglia, P. P. Bose, K. Laxmi-Reddy, F. Godde, J.-M. Schmitter*, et al.*, *Nat. Chem.*, 2018, **10**, 511-518.
19. V. Corvaglia, J. Wu, D. Deepak, M. Loos and I. Huc, *Chem. - Eur. J.*, 2024, **30**, e202303650.
20. D. Deepak, J. Wu, V. Corvaglia, L. Allmendinger, M. Scheckenbach, P. Tinnefeld and I. Huc, *Angew. Chem., Int. Ed.*, 2025, **64**, e202422958.
21. M. Loos, F. Xu, P. K. Mandal, T. Chakrabortty, C. Douat, D. B. Konrad, M. Cabbar, J. Singer, V. Corvaglia, T. Carell*, et al.*, *Angew. Chem., Int. Ed.*, 2025, **64**, e202505273.
22. V. Kleene, V. Corvaglia, E. Chacin, I. Forne, D. B. Konrad, P. Khosravani, C. Douat, C. F. Kurat, I. Huc and A. Imhof, *Nucleic Acids Res*, 2023, **51**, 9629-9642.
23. V. Corvaglia, I. A. M. Amar, V. Garambois, S. Letast, A. Garcin, C. Gongora, M. Del Rio, C. Denevault-Sabourin, N. Joubert, I. Huc*, et al.*, *Pharmaceuticals*, 2021, **14**, 624.
24. N. Narayana and M. A. Weiss, *J. Mol. Biol.*, 2009, **385**, 469-490.
25. R. Ebbert, A. Birkmann and H.-J. Schuller, *Mol. Microbiol.*, 1999, **32**, 741-751.
26. J. Jin, Y. Cai, T. Yao, A. J. Gottschalk, L. Florens, S. K. Swanson, J. L. Gutierrez, M. K. Coleman, J. L. Workman, A. Mushegian*, et al.*, *J. Biol. Chem.*, 2005, **280**, 41207-41212.
27. X. Shen, G. Mizuguchi, A. Hamiche and C. Wu, *Nature (London)*, 2000, **406**, 541-544.
28. K. Yen, V. Vinayachandran and B. F. Pugh, *Cell (Cambridge, MA, U. S.)*, 2013, **154**, 1246-1256.
29. S. Eustermann, K. Schall, D. Kostrewa, K. Lakomek, M. Strauss, M. Moldt and K.-P. Hopfner, *Nature (London, U. K.)*, 2018, **556**, 386-390.
30. R. Ayala, O. Willhoft, R. J. Aramayo, M. Wilkinson, E. A. McCormack, L. Ocloo, D. B.

Wigley and X. Zhang, *Nature (London, U. K.)*, 2018, **556**, 391-395.
31. F. Kunert, F. J. Metzner, J. Jung, M. Hopfler, S. Woike, K. Schall, D. Kostrewa, M. Moldt, J.-X. Chen, S. Bantele*, et al.*, *Sci. Adv.*, 2022, **8**, eadd3189.
32. K. R. Knoll, S. Eustermann, V. Niebauer, E. Oberbeckmann, G. Stoehr, K. Schall, A. Tosi, M. Schwarz, A. Buchfellner, P. Korber*, et al.*, *Nat. Struct. Mol. Biol.*, 2018, **25**, 823-832.
33. G. R. Fulmer, A. J. M. Miller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw and K. I. Goldberg, *Organometallics*, 2010, **29**, 2176-2179.
34. K. M. Abe, G. Li, Q. He, T. Grant and C. J. Lim, *Nat. Commun.*, 2024, **15**, 7705.
35. S. Q. Zheng, E. Palovcak, J.-P. Armache, K. A. Verba, Y. Cheng and D. A. Agard, *Nat. Methods*, 2017, **14**, 331-332.
36. A. Punjani, J. L. Rubinstein, D. J. Fleet and M. A. Brubaker, *Nat. Methods*, 2017, **14**, 290-296.
37. S. H. W. Scheres and S. Chen, *Nat. Methods*, 2012, **9**, 853-854.
38. A. Punjani, J. L. Rubinstein, D. J. Fleet and M. A. Brubaker, *Nat Methods*, 2017, **14**, 290-296.
39. T. Bepler, A. Morin, A. J. Noble, J. Brasch, L. Shapiro and B. Berger, *Res Comput Mol Biol*, 2018, **10812**, 245-247.
40. E. F. Pettersen, T. D. Goddard, C. C. Huang, E. C. Meng, G. S. Couch, T. I. Croll, J. H. Morris and T. E. Ferrin, *Protein Sci.*, 2021, **30**, 70-82.
41. Maestro, *Journal*, 2021.
42. Z. Xu, J. Tian, W. Wang, M. Lu, C. Zhang, H. Yu, M. Wang and Z. Dong, *ACS Macro Lett.*, 2025, **14**, 853-857.
43. X. Li, T. Qi, K. Srinivas, S. Massip, V. Maurizot and I. Huc, *Org. Lett.*, 2016, **18**, 1044-1047.
44. D. Bindl, E. Heinemann, P. K. Mandal and I. Huc, *Chem. Commun. (Cambridge, U. K.)*, 2021, **57**, 5662-5665.
45. V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat and I. Huc, *Chem. - Eur. J.*, 2023, **29**, e202300898.
46. M. Loos, *PCCP*, 2025, **1**.
47. M. Loos, L. Thurecht, J. Wu, V. Corvaglia, Z. Liu, V. Pophristic, M. Zacharias and I. Huc, *Chemical Science*, 2026, DOI: 10.1039/D5SC08567E.
48. G. W. Cline and S. B. Hanna, *J. Am. Chem. Soc.*, 1987, **109**, 3087.
49. J. Coste, D. Le-Nguyen and B. Castro, *Tetrahedron Lett.*, 1990, **31**, 205.
50. V. Dourtoglou, J.-C. Ziegler and B. Gross, *Tetrahedron Letters*, 1978, **19**, 1269-1272.
51. A. El-Faham and F. Albericio, *J. Pept. Sci.*, 2010, **16**, 6-9.
52. F. Shamout, L. Fischer, N. L. Snyder and L. Hartmann, *Macromol. Rapid Commun.*, 2020, **41**, 1900473.
53. T. Qi, T. Deschrijver and I. Huc, *Nat. Protoc.*, 2013, **8**, 693.
54. H. L. Ball and P. Mascagni, *Int. J. Pept. Protein Res.*, 1992, **40**, 370.

6.2 Supporting information

6.2.1 Supplementary figures and tables

6.2.1.1 Supplementary figures

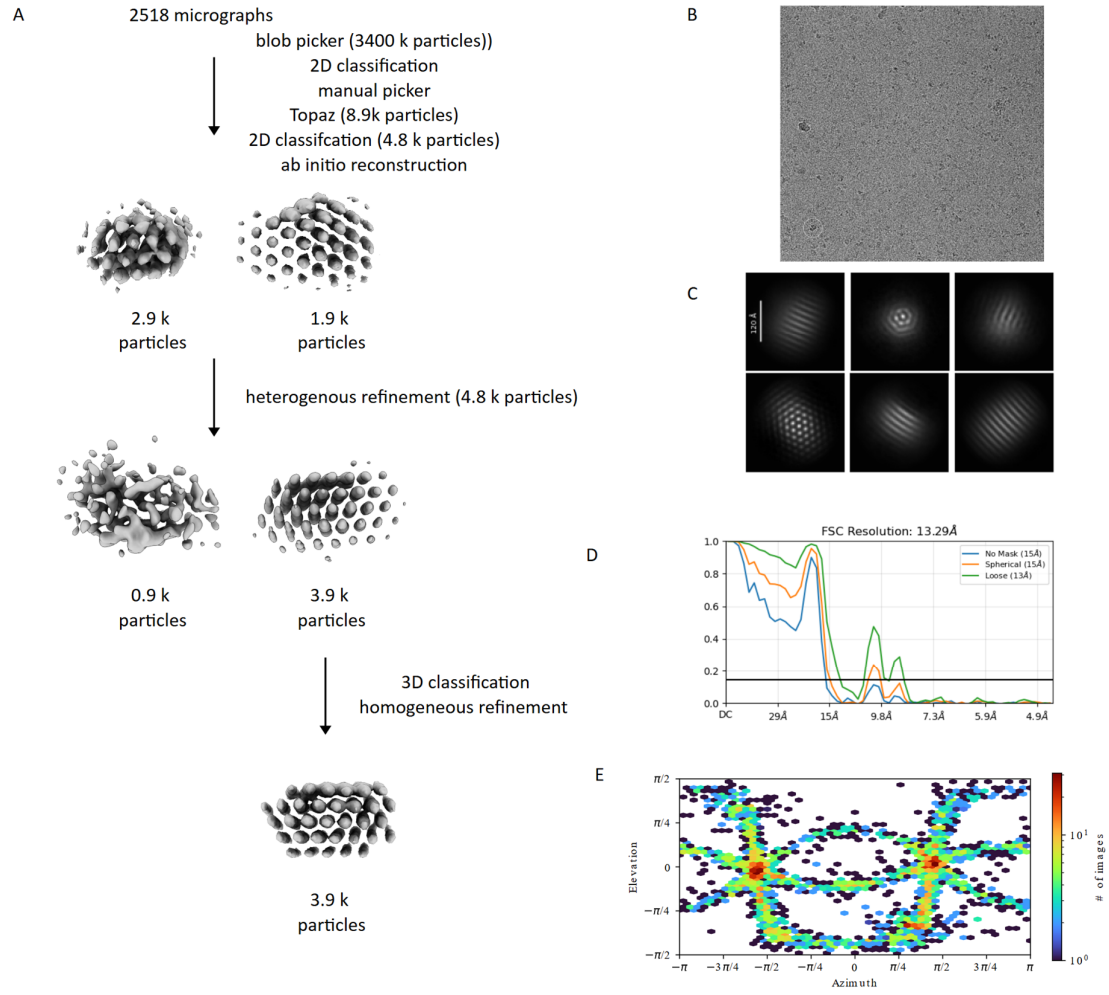


Figure S1 Cryo-EM data analysis of foldamer clusters of **40**. Workflow of cryo-EM data processing in cryoSPARC2. B) Representative micrograph of the foldamer clusters. The displayed micrograph is representative of 2518 movies collected. C) Representative classes of a 2D classification of the particles used for the final reconstruction. D) Histogram of the final reconstruction. The black line indicates the 0.143 cutoff criterion, indicating a nominal resolution of 13.29 Å (cryoSPARC2). F) Angular distribution of the particles used for the final foldamer cluster reconstruction.

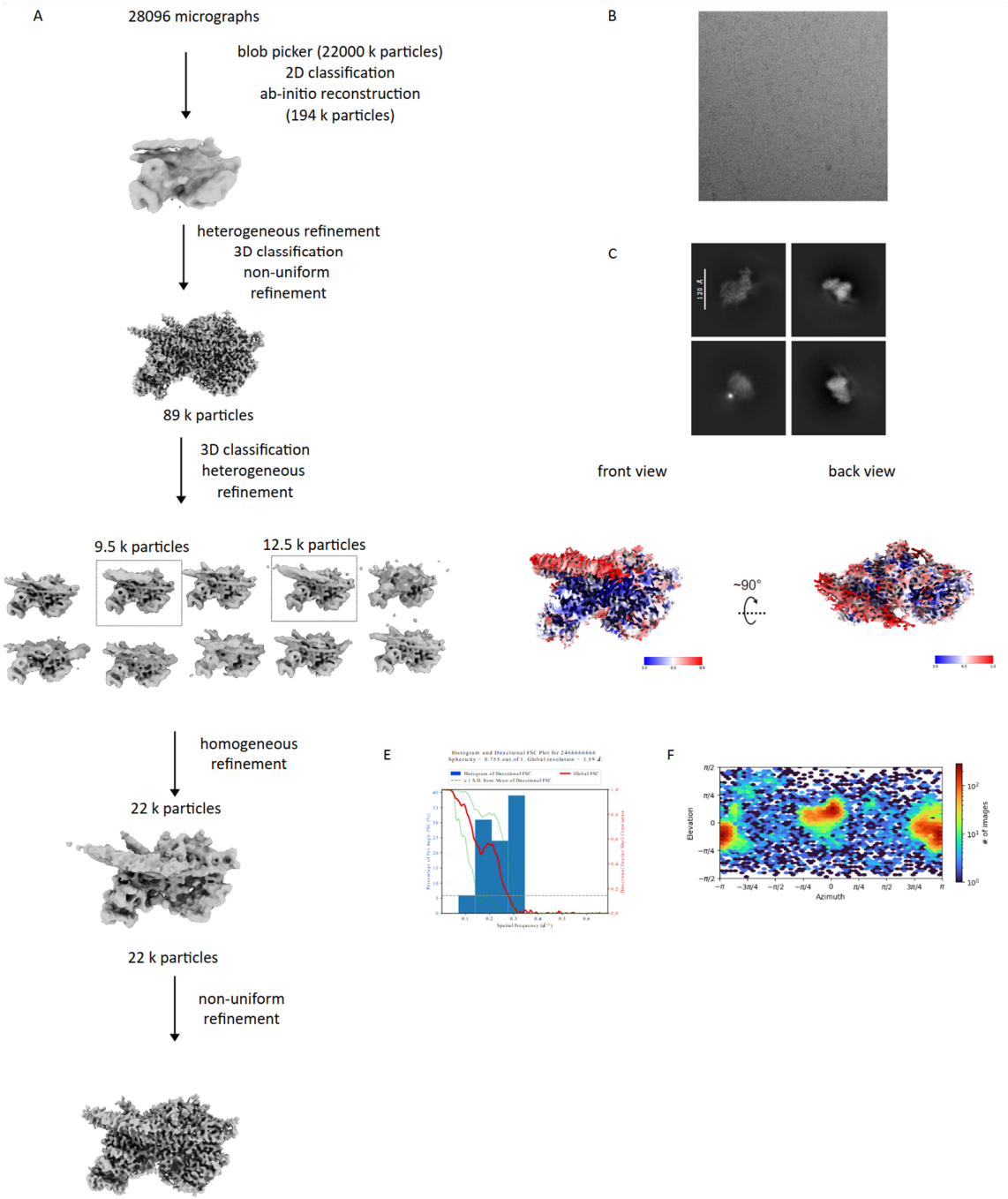


Figure S2 Cryo-EM data analysis of the INO80 A-module interacting with foldamer **40**. A) Workflow of the cryo-EM data processing in cryoSPARC2. B) Representative micrograph of the A-module:foldamer. The shown micrograph is representative of 28096 recorded movies. C) Representative classes of a 2D classification of the particles used for the final A-module:foldamer reconstruction. D) Local resolution map of the Amodule:foldamer calculated in cryoSPARC. Blue indicates higher resolution and red indicates lower resolution. E) Histogram of directional FSC3 (blue) and global FSC curve (red) of the final reconstruction. The spread of directional resolution values is defined as ±1σ (dashed green lines). The grey dashed line indicates the 0.143 cutoff criterion, indicating a nominal resolution of 3.59 Å. F) Angular distribution of particles included in the final INO80 A-module: foldamer reconstruction.

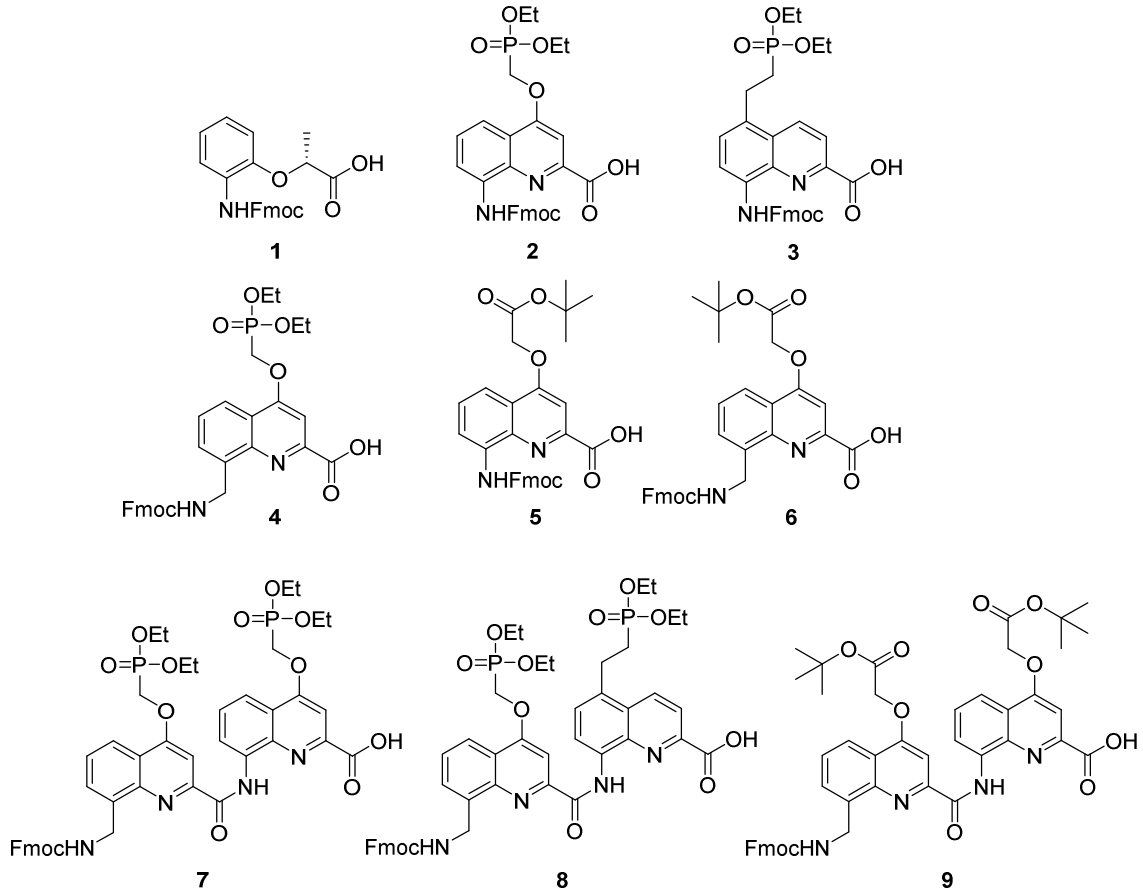


Figure S3 Fmoc-protected amino acids monomers and dimers. Their synthesis has been performed according to published procedures: **1**[1], **2&4**[2], **3**[3] and **5**[3]. The synthesis of **6–9** can be found below.

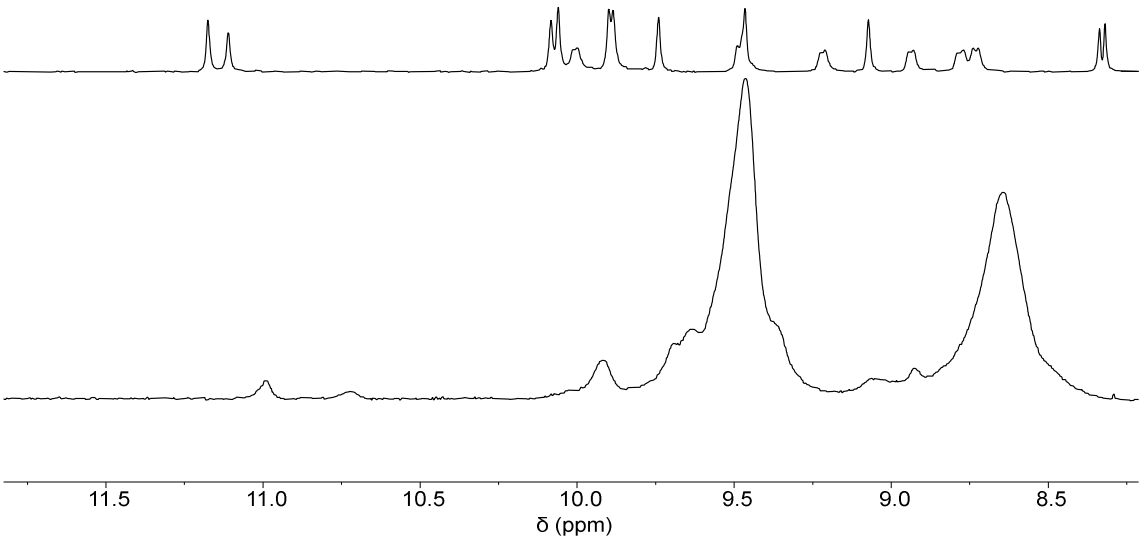


Figure S4 ^{1}H-NMR of the amide region of **37** (top) and **36** (bottom) to compare handedness induction.

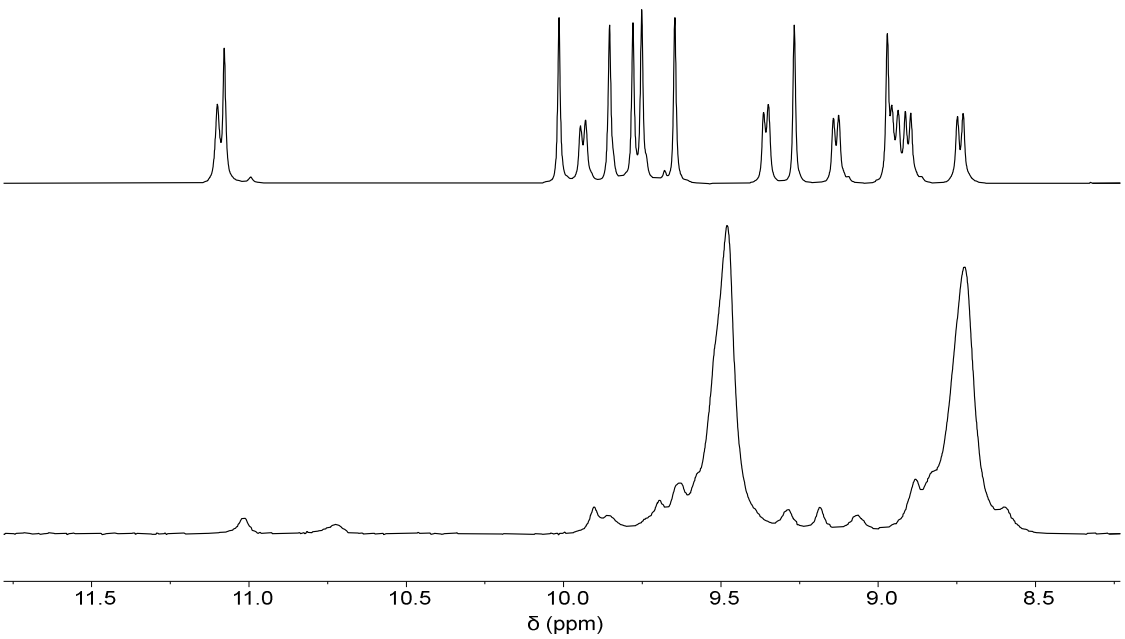


Figure S5 ^{1}H-NMR of the amide region of **43** (top)and **42** (bottom) to compare handedness induction.

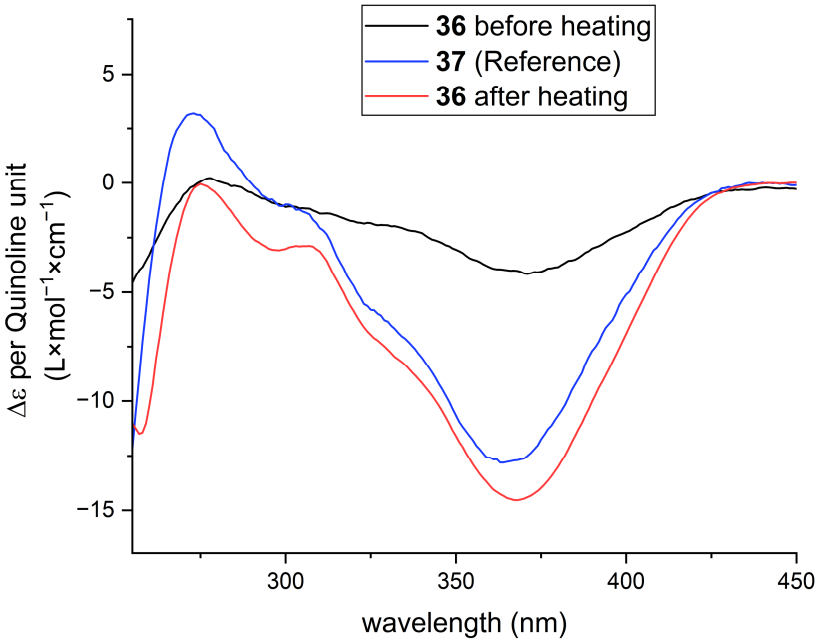


Figure S6 Per quinoline normalized CD-spectra of **37** as a reference (blue) and **36** before and after heating (black and red respectivly).

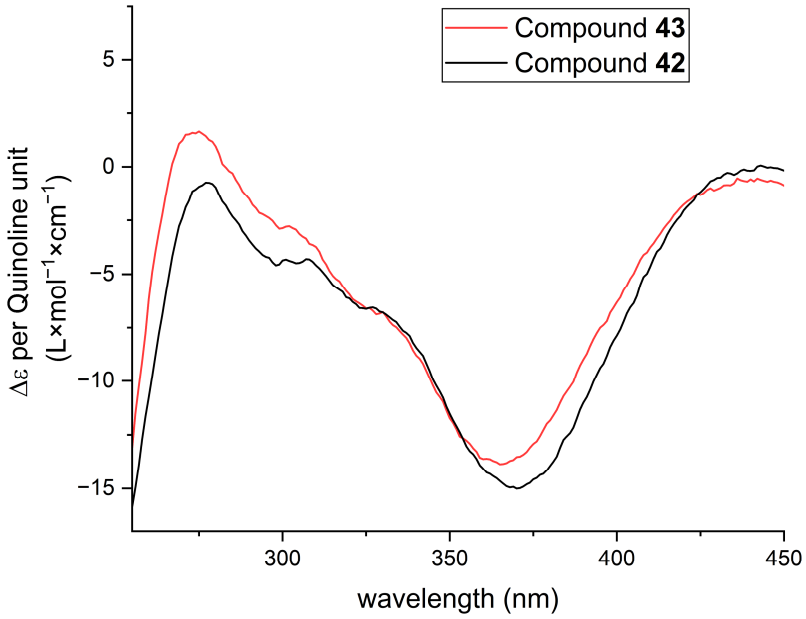


Figure S7 Per quinoline normalized CD-spectra of **43** (red) and **42** (black) before and after heating.

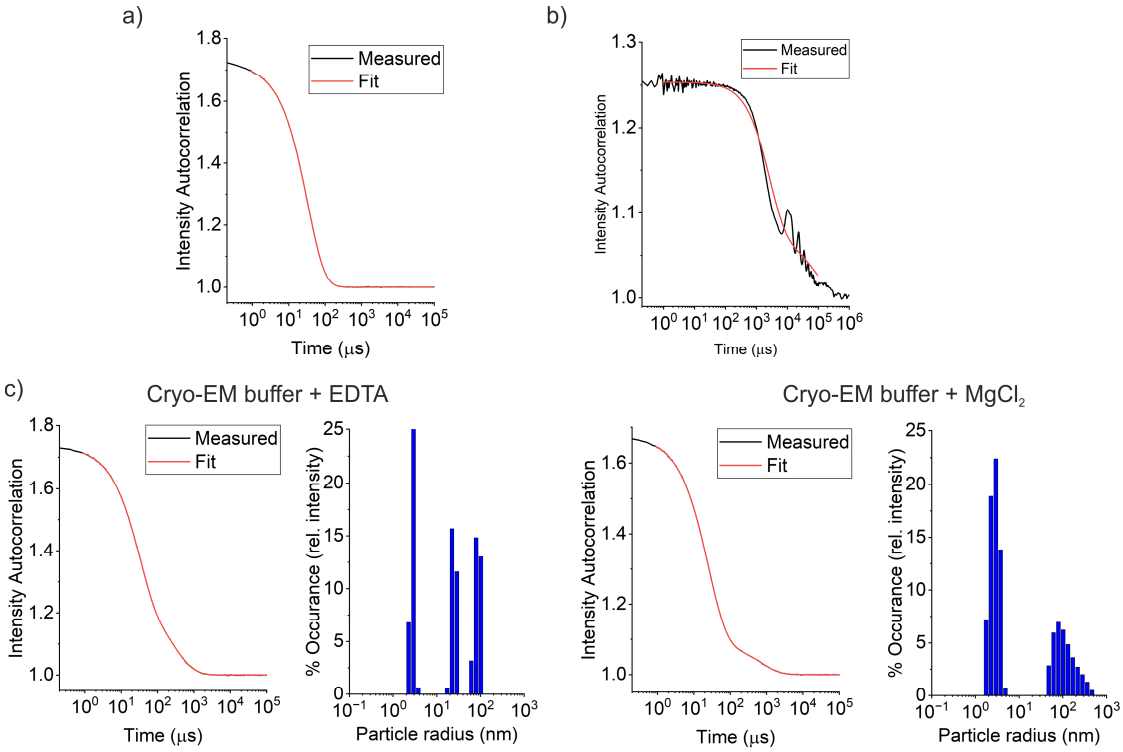


Figure S8 DLS-Autocorrelation functions (black) and their fit (red) of DMFO **40** in 100 mM NaCl (a) and 100 mM Ca_2Cl (b). The noisy curve in (b) is an indication of heavy aggregation in presence of Ca^{2+} c) Autocorrelation functions (black) and respective fit (red) and relative occurrence plots of particle sizes of DMFO **40** in Cryo-EM buffer II + EDTA (left) and Cryo-EM buffer II + $MgCl_2$ (right).

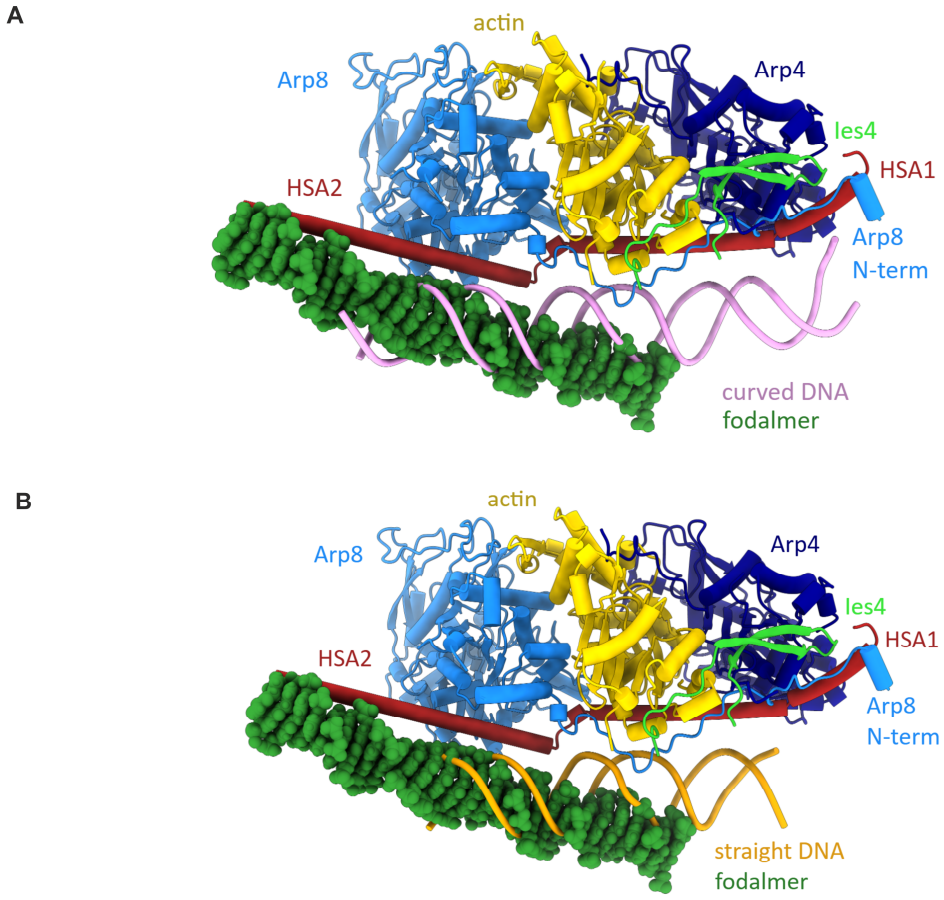


Figure S9 Structural comparison of DNA and foldamer bound A-modules. (A) Structural comparison of the foldamer and A-module bound to curved DNA (PDB: 8A5P).[4] (B) Different binding preference of the foldamer and A-module bound to straight DNA (PDB: 8A5Q).[4]

6.3 Supplementary tables

Table S1 Minimization parameters that were used to build molecular models.

Forcefield	OPLS3
Solvent	Water
Charges from	Force Field
Cutoff	Extended
Method	PRCG
Converge on	Gradient
Convergence threshold	0.001
Minimization mode	Minimization of non-conformers
Maximum iterations	25000

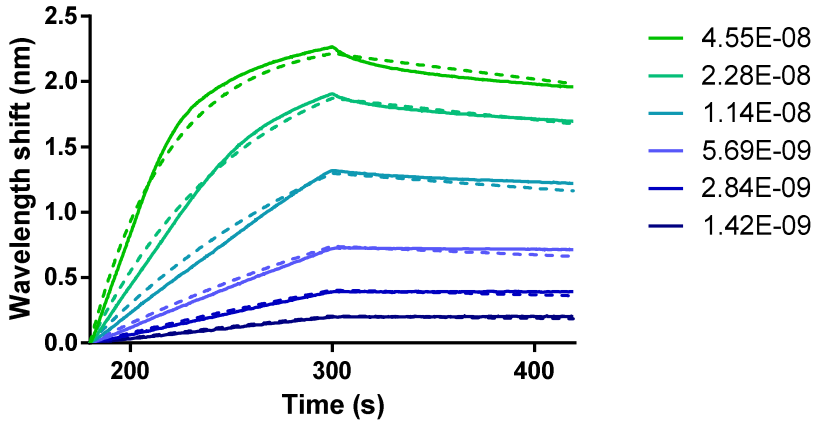
Table S2 Calculated K_D values (in nM) in low salt and high salt buffer.

Buffer	31	39	dsDNA
LSPB	1.702 ± 0.025	1.419 ± 0.029	5.755 ± 0.055
PBST	1.809 ± 0.017	1.404 ± 0.017	10.76 ± 0.060

Table S3 Kinetics of INO80 A-module against foldamer **31** in LSPBT.

K_D (M)	K_D Error	k_a (1/Ms)	k_a Error	k_{dis} (1/s)	k_{dis} Error	RSS	Full R^2
1.702E-09	2.464E-11	5.407E05	2.689E03	9.204E-04	1.251E-05	12.967	0.9964

Associated sensorgram for Table S3.



FigureS10 Kinetics of INO80 A-module against foldamer **31** in LSPBT.

Table S4 Kinetics of INO80 A-module against foldamer **39** in LSPBT.

K_D (M)	K_D Error	k_a (1/Ms)	k_a Error	k_{dis} (1/s)	k_{dis} Error	RSS	Full R^2
1.419E-09	2.892E-11	5.473E05	3.272E03	7.765E-04	1.513E-05	19.8611	0.9952

Associated sensorgram for Table S4.

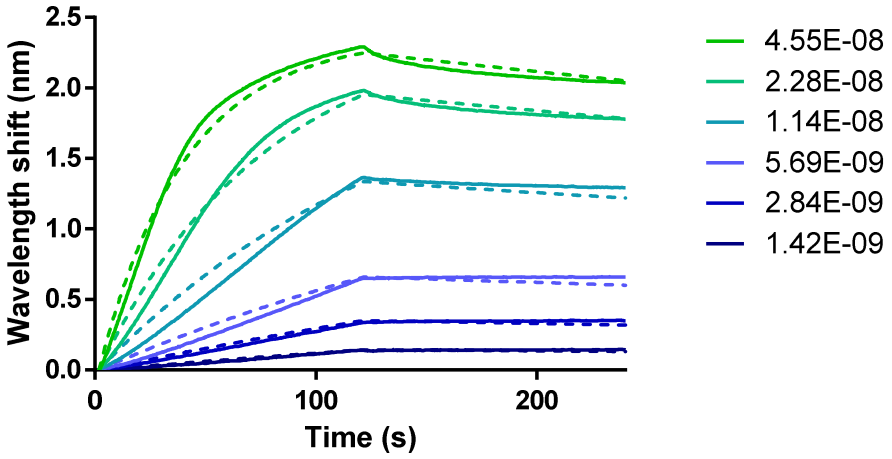


Figure S11 Kinetics of INO80 A-module against foldamer **39** in LSPBT.

Table S5 Kinetics of INO80 A-module against dsDNA in LSPBT.

K_D (M)	K_D Error	k_a (1/Ms)	k_a Error	k_{dis} (1/s)	k_{dis} Error	RSS	Full R^2
5.755E-09	5.509E-11	2.173E05	1.362E03	1.250E-03	9.046E-06	5.5317	0.9984

Associated sensorgram for Table S5

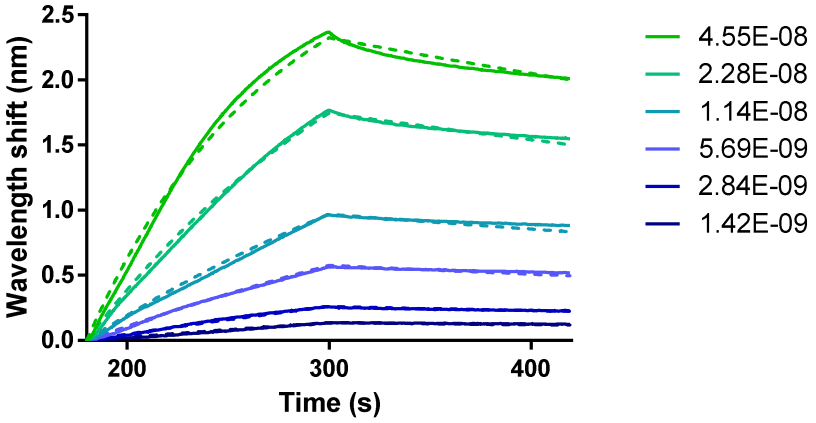


Figure S12 Kinetics of INO80 A-module against dsDNA in LSPBT.

Table S6 Kinetics of INO80 A-module against foldamer **31** in PBST.

K_D (M)	K_D Error	k_a (1/Ms)	k_a Error	k_{dis} (1/s)	k_{dis} Error	RSS	Full R^2
1.809E-09	1.699E-11	6.942E05	2.708E03	1.256E-03	1.073E-05	4.8151	0.9972

Associated sensorgram for Table S6.

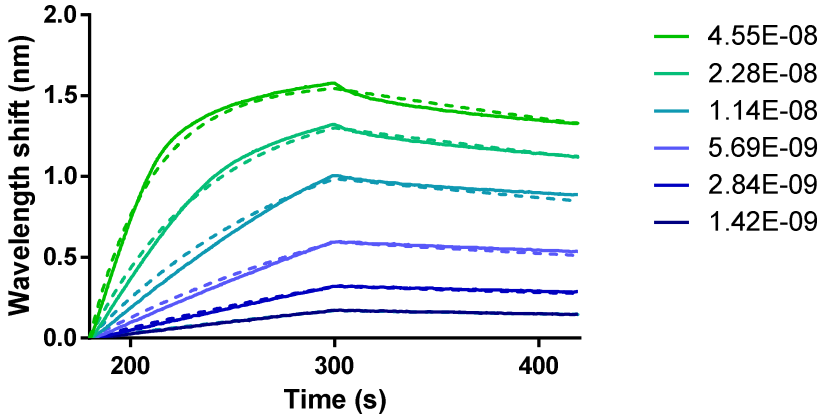


Figure S13 Kinetics of INO80 A-module against foldamer **31** in PBST.

Table S7 Kinetics of INO80 A-module against foldamer **39** in PBST.

K_D (M)	K_D Error	k_a (1/Ms)	k_a Error	k_{dis} (1/s)	k_{dis} Error	RSS	Full R^2
1.404E-09	1.732E-11	7.097E05	2.942E03	9.966E-04	1.158E-05	5.7797	0.9969

Associated sensorgram for Table S7.

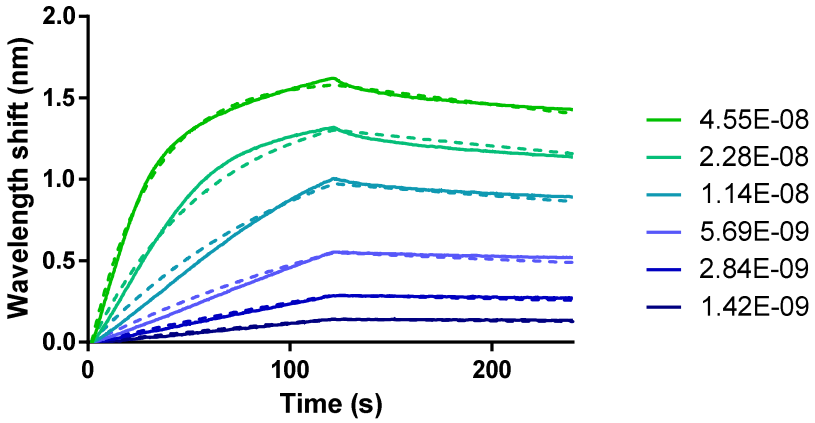


Figure S14 Kinetics of INO80 A-module against foldamer **39** in PBST.

Table S8 Kinetics of INO80 A-module against dsDNA in PBST.

K_D (M)	K_D Error	k_a (1/Ms)	k_a Error	k_{dis} (1/s)	k_{dis} Error	RSS	Full R^2
1.076E-08	6.041E-11	3.372E05	1.504E03	3.628E-03	1.238E-05	2.823	0.996

Associated sensorgram for Table S8.

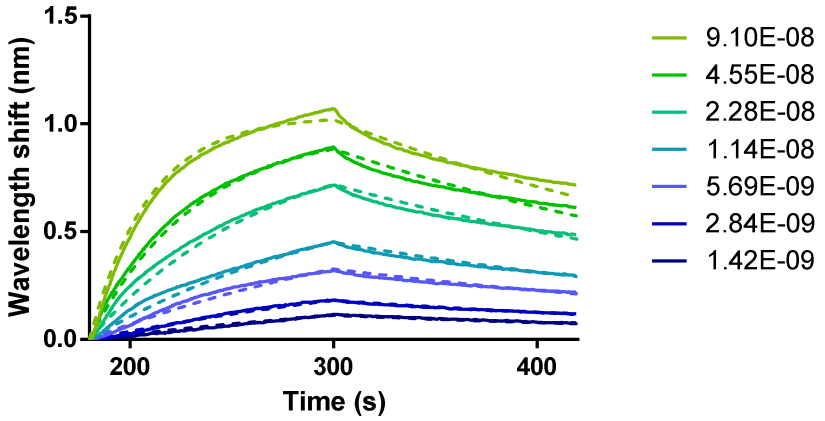


Figure S15 Kinetics of A-INO80 module against dsDNA in PBST.

6.3.1 Materials and Methods

6.3.1.1.1 Materials

Reagents were used as purchased from commercial sources without further purification. Column chromatography purifications were performed on silica gel (230-400 mesh, 40-63 µm, Merck). Thin-layer chromatography was performed on silica gel plates (60-F254, Merck). Reactions requiring anhydrous conditions were performed under nitrogen (dried flowing through $CaCl_2$) with commercial anhydrous solvents except THF and CH_2Cl_2, that were dispensed from a *MBRAUN SPS-800* solvent purification system using alumina columns for drying. $CHCl_3$, TEA and DIPEA were freshly distilled over CaH_2 under N_2-atmosphere. Peptide synthesis grade DMF was purchased from Carlo Erba. Ultrapure water was obtained via a Stakpure Omnia Pure-T UV-TOC ultrapure water system. PyBOP was recrystallized from aceton/Et_2O (1:1, v/v; 10 mL per g) until no presence of pyrrolidine by ^1H NMR spectroscopy was detected.[5]

6.3.1.1.2 Small molecule analysis

Small molecule NMR spectra were recorded on AVANCE NEO NMR spectrometer 500 MHz (Bruker BioSpin) with CryoProbe™ Prodigy and a BCU II. $CDCl_3$ (δ_H: 7.26, δ_C: 77.0), DMSO-d_6 (δ_H: 2.50, δ_C: 39.4) were used as solvents and their residual solvent signals were used as internal standard.[6] The raw data was evaluated using *Mnova version 14.0.0* by *Mestrelab Research*.The derived data signals are stated with chemical shift in ppm, their multiplicity (s = singlet, bs = broad singlet, d = doublet, t = triplet, q = quadruplet, m = multiplet, or a combination of these), their coupling constant in Hz and their integrated values. Small molcule mass spectra were recorded on a Thermo Finnigan LTQ FT Ultra or Thermo Finnigan LTQ Orbitrap XL mass spectrometer with a resolution of 100.000 a m/z of 400. Depending on the sample, mass ranges from 40 to 2000 u were measured. The spray capillary voltage was 3 kV, the heater capillary temperature 250 °C, the nitrogen gas flow 25 units and the aux gas flow 5 units.

6.3.1.1.3 Automated flash chromatography

Automated flash chromatography was performed on a *Puriflash XS 520 Plus* by *Interchim* using a reversed-phase C_{18}-column (*INTERCHIM PURIFLASH C18-HP 15UM F0025 FLASH COLUMN*) with analytical grade MeCN and ultra-pure water mixtures as eluent. Elution was monitored by UV detection at 214, 254 and 300 nm.

6.3.1.1.4 RP-HPLC chromatography

6.3.1.1.4.1 RP-HPLC TFA additive (building blocks & poly diethyl phosphonate foldamers)

Analytical RP-HPLC analysis and semi-preparative purifications (eg. compound) were performed with a TFA buffered system 0.1% TFA in ultra-pure water (referred to as mobile phase A) and 0.1% TFA in HPLC-grade acetonitrile (referred to as mobile phase B) were used on a Thermo Fisher Scientific Ultimate 3000 HPLC system equipped with Macherey-Nagel Nucleodur C18ec (4.6 × 150 mm, 5 µm) or Macherey-Nagel Nucleodur C8 gravity columns (4.6

× 150 mm, 5 µm) and a flowrate of 1 mL/min for analysis and a Waters system equipped with a 2545 Quaternary Gradient Module and a XBridge® Prep C8 OBD™ column (19 × 150 mm, 5 µm) and a flowrate of 25 mL/min for purifications. Elution was monitored by UV detection at 254 and 300 nm with a diode array detector.

6.3.1.1.4.2 RP-HPLC TEAA buffered (polyphosphonic acid foldamers)

Analytical RP-HPLC analysis and semi-preparative purifications in TEAA conditions were performed on a Thermo Fisher Scientific Ultimate 3000 HPLC system using Kinetex C18 EVO (100 Å, 2.1 x 50 mm, 1.8 µm) or Aeris™ WIDEPORE C4 (200 Å, 150 x 2.1 mm, 3.6 µm) with a triethylammonium acetate buffer system with a flowrate of 0.33 mL/min. The mobile phase was composed of 12.5 mM TEAA in ultra-pure water at pH 8.5 (A) and 12.5 mM TEAA in water: acetonitrile mixture (1:2, v/v) at pH = 8.5 (B). Purifications in basic conditions were done on a Thermo Fisher Scientific Ultimate 3000 HPLC system equipped with a a Kinetex C18 EVO column (100 Å, 150 x 10 mm, 5 µm) and a flowrate of 5 mL/min. Elution was monitored by UV detection at 254 and 300 nm with a diode array detector.

6.3.1.1.4.3 RP-HPLC NH_4OAc buffered (polyphosphonic acid foldamers containing N-terminal Fmoc or Biotin)

Analytical RP-HPLC analysis and semi-preparative purifications in TEAA conditions were performed on a Thermo Fisher Scientific Ultimate 3000 HPLC system using Kinetex C18 EVO (100 Å, 2.1 x 50 mm, 1.8 µm) or Aeris™ WIDEPORE C4 (200 Å, 150 x 2.1 mm, 3.6 µm) with a NH_4OAc buffer system with a flowrate of 0.33 mL/min. The mobile phase was composed of 12.5 mM aqueous NH_4OAc at pH 8.5 (A) and acetonitrile pH = 8.5. Purifications in this conditions were done on a Thermo Fisher Scientific Ultimate 3000 HPLC system equipped with a a Kinetex C18 EVO column (100 Å, 150 x 10 mm, 5 µm) and a flowrate of 5 mL/min. Elution was monitored by UV detection at 254 and 300 nm with a diode array detector.

6.3.1.1.5 LCMS analysis

6.3.1.1.5.1 LCMS formic acid/TFA buffered

LCMS analysis of these compounds were performed on a a Thermo Fisher Scientific Ultimate 3000 HPLC system equipped with Macherey-Nagel Nucleodur C18 gravity (2× 50 mm, 1.8 µm) and 0.18% formic acid and 0.02% TFA in ultra-pure water (referred to as mobile phase A) and 0.18% formic acid and 0.02% TFA in LCMS-grade acetonitrile (referred to as mobile phase B) and a flowrate of 0.33 mL/min. For mass spectrometry analysis, the LC system was coupled to a micrOTOF II mass spectrometer by Bruker Daltonics and molecules were ionized by ESI^+.

6.3.1.1.5.2 LCMS NH_4OAc buffered

LCMS analysis of polyanionic compounds were conducted on a Thermo Fisher Scientific Ultimate 3000 HPLC system with an NH_4OAc buffer system consisting of 12.5 mM NH_4OAc dissolved in ultra-pure water and adjusted to pH = 8.5 with aqueous ammonia (referred to as mobile phase A) and LCMS-grade acetonitrile (referred to as mobile phase B) on a Kinetex C18

EVO column (2.1 x 50 mm, 1.8 µm) column and a flowrate of 0.33 mL/min. Elution was monitored by UV detection at 254 and 300 nm with a diode array detector. For mass spectrometry analysis, the LC system was coupled to a micrOTOF II mass spectrometer by Bruker Daltonics and molecules were ionized by ESI^-.

6.3.1.1.6 Oligomer NMR

Spectra were recorded on a Avance III HD 500 MHz Bruker BioSpin spectrometer equipped with a broad band observe 5-mm BB-H&FD CryProbe™ Prodigy.

6.3.1.1.7 Molecular modeling

The molecular models of conjugates and DNA mimics were built in Maestro (Version 11.5)[7] and energy-minimized with parameters in table S1. Foldamer coordinates for the models of conjugates were derived from earlier crystal structures from diethylester-protected polyphosphonate foldamers,[8] stripped of their side-chain protecting groups and energy-minimized according to parameters below.

6.3.2 Synthetic procedures

6.3.2.1 Foldamer synthesis

6.3.2.1.1 Foldamer-solid-phase synthesis

Solid phase synthesis (SPS) was performed on Tentagel® S PHB (90 µm, Initial loading: 0.21 mmol/g, Iris Biotech). manually under MW-irradiation on a CEM Discover (Liberty Bio) microwave oven using a reaction vessel and an internal fiber optic probe for temperature control as described below for manual couplings, or with a fully automated PurePep® Chorus synthesizer for chain elongations.

Loading

Tentagel® S PHB resin was (500 mg, 95.0 µmol, 1 eq.) swollen in DCM (ca. 10 mL) for 5 min, then the resin was filtered off. Respective Fmoc-dimers (1 eq.) were dissolved in DCM/DMF (v/v, 1/1, 5 mL) and added to the resin. OxymaPure (13.5 ,g 95.0 µmol, 1 eq.) was dissolved in DMF (2 mL) and added to the resin. 4-Dimethylaminopyridine (580 µg, 4.75 µmol, 0.05 eq.) dissolved in (DCM, minimal amount) and added to the resin, followed by *N,N'*-diisopropylcarbodiimide (14.9 µL, 95.0 µmol, 1 eq.). The resin was shaken over night, then filtered off, washed with DMF (3 x 5 mL) and DCM (3 x 5 mL). The resin was capped by incubation with Ac_2O/DCM (v/v, 1:1, 10 mL) over 1 h. The resin was filtered off and washed (3 x 5 mL) and subsequently dried by flowing N_2-stream through the resin.

Loading determination

For resin loading estimation, a solution of 20% piperidine in DMF (v/v, 3 mL) was added to a known mass of the previously dried resin (2-3 mg) and agitated for 5 min. After agitation, the resin was allowed to settle at the bottom and the absorbance of the supernatant solution (a prior baseline measurement was conducted) that contained the resin was measured at 290 nm. The resin loading was determined by the following equation.

$$Resin\ loading\ \left[\frac{mmol}{g}\right] = \frac{\left[Abs_{final} - Abs_{initial}\right]}{[2.00 \times m_{resin}]}$$

Fmoc deprotection

To the pre-swollen loaded resin, a 20% solution of piperidine in DMF (3 mL, v/v) was added and the resin was mixed by bubbling N_2-gas through the solution for 7 min. The resin was filtered and washed with DMF (2 x 3 mL) and the deprotection was repeated once for 7 min to give the respective aminoresin. The resin was filtered, washed with DMF (3 x 3 mL), then DCM (3 x 3 mL) and then washed either with anhydrous THF (3 x 3 mL) if the next coupling step was conducted by *in situ* acid chloride base coupling or DMF (2 x 3 mL) if the next coupling step was conducted by PyBOP activation.

Coupling *in situ* acid chloride activation (Manual, 20 µmol scale

The NH_2-resin was suspended in anhydrous THF (1 mL) and *2-4-6*-collidine (23.8 µL, 180 µmol, 9.0 eq.) was added. Concurrently, in a glass vial, monomer (3.0 eq.) and PPh_3 (42 mg, 160 µmol, 8 eq.) were mixed and dissolved in anhydrous $CHCl_3$ (1 mL). Subsequently, trichloroacetonitrile (18.1 µL, 180 µmol, 9 eq.) was added to the vial, which was quickly mixed and then the mixture was added to the pre-swollen resin. After mixing, the reaction vessel was placed under microwave irradiation (25 W, ramp to 50 °C over 5 min, hold at 50 °C for 15 min). The resin was filtered off and washed with anhydrous THF (3 x 3 mL). The coupling step was repeated once. The resin was filtered off and washed with anhydrous THF (3 x 3 mL) and DCM (3 x 3 mL) prior to the capping step.

Coupling PyBOP acitivation (PurePep® Chorus synthesizer, 15 µmol scale)

The NH_2-resin was suspended in DMF (1 mL) and then subsequently monomer (2 eq.) dissolved in DMF (0.5 mL), followed by benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate (2 eq., 15.6 µmol, 30.0 µmol) dissolved in DMF (0.5 mL) followed by DIPEA (10.5 µL, 60.0 µmol, 4 eq.) was added the mixture was heated to 50 °C (PurePep® Chorus induction heating system) for 15 min under consistent shaking and N_2-bubbling. The coupling solution was filtered off and treated for building block recovery (see below). The resin was washed with DMF (3 x 3 mL) and the coupling was repeated once. The Resin was washed with DMF (3 x 3 mL) and DCM (3 x 3 mL) prior to the capping step.

Capping

Ac_2O/DCM (v/v, 1:1, 4 mL) was added to the resin and the mixture was heated to 40 °C for 10 min (either MW-assisted or induction based). The resin was filtered off and washed with DCM (3 x 3 mL) and then DMF (3 x 3 mL) prior to the next deprotection cycle.

Resin cleavage

The resin-bound foldamer was placed in a syringe equipped with a filter and then suspended in TFA (3 mL). The resin was shaken for 1 h at rt. The resin was then filtered off and washed twice with TFA. TFA was removed *in vacuo* and the resulting solid was precipitated in cold Et_2O. The precipitate was centrifuged, and the solvent was decanted to give a yellow solid. The decanted Et_2O was concentrated by rotary evaporation and the precipitation was repeated. The combined precipitates were dissolved in water/MeCN and then lyophilized to give the crude protected foldamer as a yellow solid that was subsequently purified.

For polycarboxylate foldamers, the resin cleavage was performed in 5% H_2O in TFA for 24 h. Precipitation was conducted similarly, but the crude product was dissolved in 200 mM NH_4OAc (pH = 8.5) buffer at prior to lyophilization.

6.3.2.1.2 Solution-phase foldamer synthesis

6.3.2.1.2.1 Removal of N-terminal Fmoc after purification

The purified polydiethyl phosphonate foldamer (10-20 mg) was dissolved in 0.2% 1,8-diazabicyclo[5.4.0]undec-7-ene/THF (v/v, 1.5 mL) and stirred at rt for 1 h. The mixture was subjected to high vacuum for 30 min to obtain a yellow crude oil. The oil was redissolved in Et_2O and left in the freezer (−24 °C) for 2 h to allow for full precipitation. The precipitate was centrifuged and the supernatant decanted. The crude was freeze-dried from MeCN/H_2O (v/v, 1:1) to obtain the amino-foldamer that was directly used in global diethyl-phosphonate removal without further purification.

6.3.2.1.2.2 Removal of the diethyl-phosphonate protecting groups

Purified protected foldamers were dissolved in anhydrous chloroform (1 mL per 10 mg compound) and cooled to 0 °C. TMSBr (0.4 mL per 10 mg compound) was diluted in anhydrous chloroform (0.8 mL per 10 mg compound). Diluted TMSBr-solution was added dropwise over 10 min to the reaction mixture. The reaction mixture was heated to 30 °C and stirred under N_2-atmosphere for 5-7 d until an HPLC- and LCMS aliquot of the reaction mixture showed full cleavage of all phosphonate-diethyl esters. The reaction mixture was evaporated *in vacuo* (40 °C water bath) to give a yellow oil, then co-evaporated with DCM (2x) to give a yellow solid. The solid was suspended in water, basified to pH > 12 with triethylamine and stirred for 30 min (in the case of Fmoc-oligomers, 200 mM NH_4OAc (pH = 8.5) was used). The suspension was filtered through nylon syringe filters to give a pale-yellow solution that was freeze-dried to give the crude deprotected foldamers as yellow solids. The crude deprotected foldamers were then purified by RP-HPLC with varying buffers and gradients listed below.

6.3.2.1.2.3 Biotinylation

Purified poly triethyl ammonium-salts of poly phosphonic foldamers were dissolved in degassed triethyl ammonium acetate buffer (37.5 mM, pH = 8.5, 40 µL per 1mg foldamer). α-Biotin-omega-carboxy succinimidyl ester dodeca(ethylene glycol) (8 eq.) dissolved in degassed DMF (40 µL) was added and the mixture was incubated for 2 h at rt. Then further α-biotin-omega-carboxy succinimidyl ester dodeca(ethylene glycol) (24 eq.) in dissolved in degassed DMF (40 µL) were added and the reaction was left to incubate for further 4 h. The crude mixture was directly purified by semi-preparative RP-HPLC.

6.3.2.1.2.4 Cation exchange chromatography

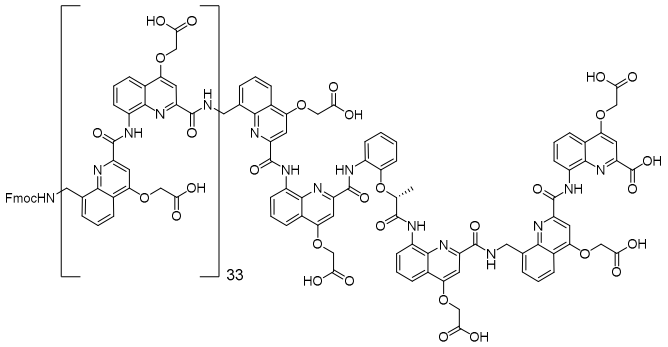
Cation exchange chromatography was performed on Dowex© 50WX4 200-400 (H) resin. The resin was swollen in H_2O the orange solution was decanted. The resin was transferred into a column and washed with H_2O (all with gravity flow) followed by two column volumes (CV) of

2 M HCl solution, washed with H_2O until pH = 6-7 (ca. 5-10 CV), washed with two CV of 2 M NH_4OAc solution, washed again with 5 CV of H_2O. Purified triethylammonium salts of foldamers were dissolved in water and loaded on the column, the column was closed without flow for 2 h, the column was eluted with water (20 mL) and the compound was lyophilized to give the purified foldamer as NH_4^+-salt as yellow solid.

6.3.2.1.3 Synthesized foldamers

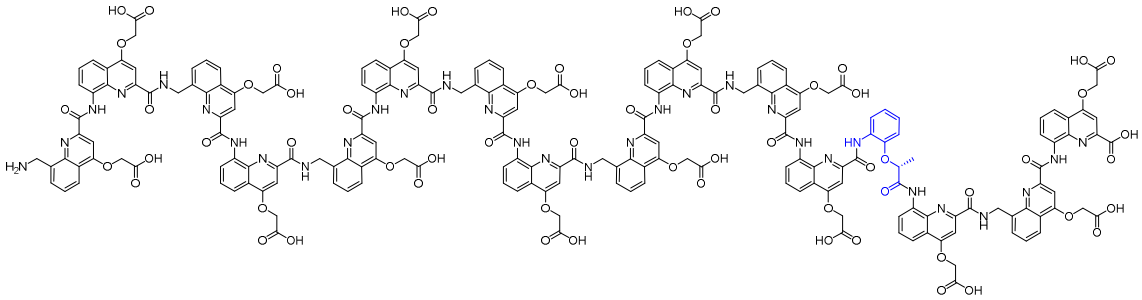
The synthesis of foldamers **34**, **35**, **37** and **41** are described within chapter 5 of this thesis.[3]

6.3.2.1.3.1 **42** Fmoc-polycarboxylate 72mer



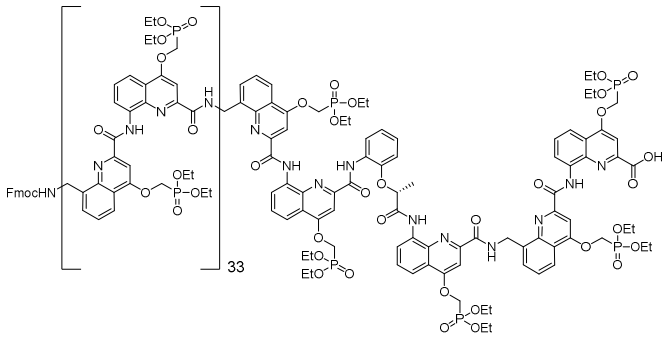
Compound **42** was synthesized on Tentagel Wang resin (15 µmol scale, initial loading: 0.124 mmol/g, 121 mg resin). From the recovered 255 mg resin, 50 mg resin were subjected to resin cleavage to obtain 27.1 mg of crude foldamer that was purified by semi-preparative RP-HPLC (C18 Kinetex Evo, NH_4OAc system, Gradient: 0-20 B in A) and subsequently lyophilized to obtain 14.3 mg (extrapolated yield: 27%) of compound **42** as poly ammoniumsalt. For HPLC and MS data see below.

6.3.2.1.3.2 **43** NH_2-polycarboxylate 16mer



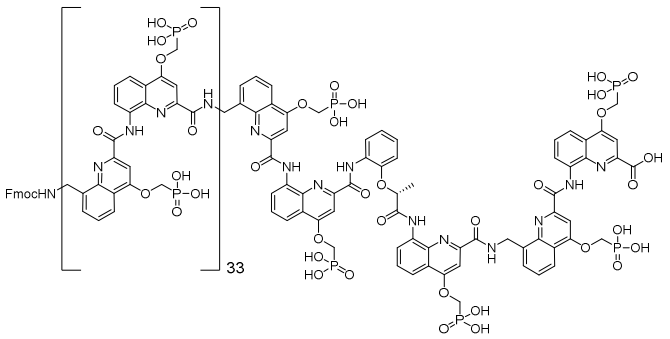
Compound **43** was synthesized on Tentagel Wang resin (15 µmol scale, initial loading: 0.124 mmol/g, 121 mg resin). From the resin cleavage of 198 mg of resin was obtained 61.0 mg of crude foldamer that was purified by semi-preparative RP-HPLC (C18 Kinetex Evo, buffers: TEAA system, Gradient: 0-20 B in A) and subsequently lyophilized to obtain 30.4 mg (yield: 51%) of compound **43** as poly tritheylammonium-salt that was salt exchanged to a poly NH_4^+-salt according to section 6.3.2.1.2.4. For HPLC and MS data see below.

6.3.2.1.3.3 **32** Fmoc-mQQ⁴-diethylprotected 72mer



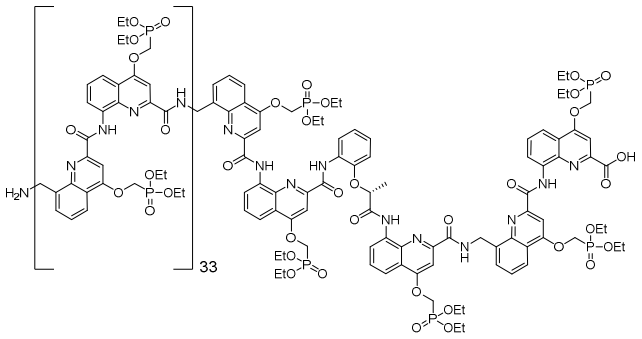
Compound **32** was synthesized on Tentagel Wang resin (15 µmol scale, initial loading: 0.105 mmol/g, 135 mg resin). From the recovered resin was obtained 248 mg of crude foldamer that was purified by semi-preparative RP-HPLC (X-Bridge C8 column, buffers: A = H_2O + 0.1% TFA, B = MeCN + 0.1% TFA, Gradient: 57-61 B in A) and subsequently lyophilized to obtain 49 mg (13%) of compound **32** as a yellow solid. For HPLC and MS data see below.

6.3.2.1.3.4 **36** Fmoc-mQQ⁴-polyphosphonic 72mer



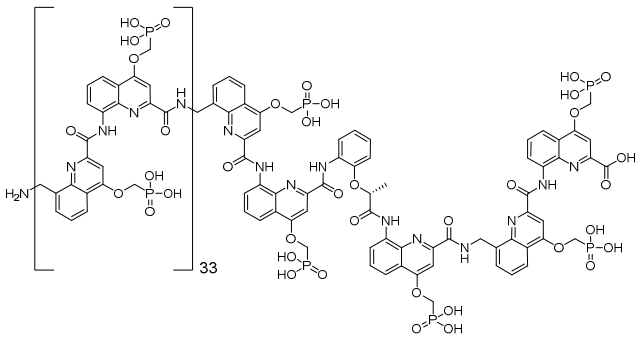
Compound **32** (16.0 mg, 652 nmol) was globally deprotected according to section 3.1.2.2 (stirring time: 6d) and then purified by semi-preparative RP-HPLC (Kinetex C18 Evo, NH_4OAc system, 0-12 % B in A) to give the title compound (5.71 mg, 278 nmol, 43%) as a poly ammonium salt as a yellow solid. For HPLC and MS data see below.

6.3.2.1.3.5 **33** NH_2-mQQ^4-diethylprotected 72mer



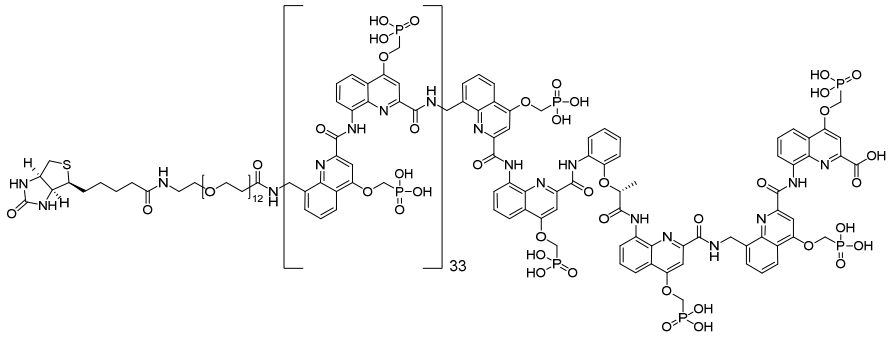
The Fmoc-group of compound **32** (11.8 mg, 485 nmol) was removed according to section 6.3.2.1.2.1 to obtain the title compound and used for global deprotection without any further purification. For HPLC and MS data see below.

6.3.2.1.3.6 **38** NH_2-mQQ^4-polyphosphonic 72mer



Compound **33** (11.8 mg, 485 nmol) was globally deprotected according to section 6.3.2.1.2.2 (stirring time: 4d) and then purified by semi-preparative RP-HPLC (Kinetex C18 Evo, TEAA system, 0-30 % B in A) to give the title compound as a poly triethylammonium salt as a yellow solid (5.11 mg, 251 nmol, 52%) over two steps. For HPLC and MS data see below.

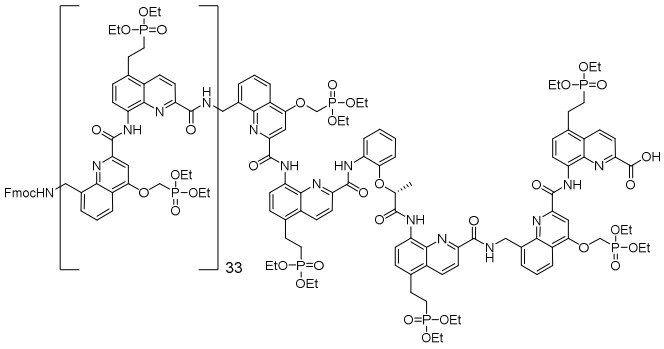
6.3.2.1.3.7 **39** Biotin-mQQ^4-polyphosphonic 72mer



Compound **38** (0.90 mg, 44 nmol) was biotinylated according to section 6.3.2.1.2.3 and then purified by semi-preparative RP-HPLC (Kinetex C18 Evo, TEAA, 0-30 % B in A). Combined fractions were degassed by N_2-bubbling through the solution and then freeze-dried to give the title compound as a poly triethylammonium salt that was ion-exchanged according to section. 6.3.2.1.2.4 to obtain the title compound as a yellow solid, then (0.53 mg, 25 nmol, 57%). For

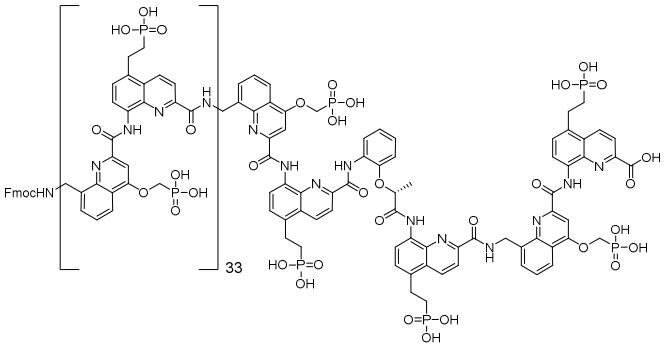
HPLC and MS data see below.

6.3.2.1.3.8 **28** Fmoc-mQQ[5]-diethylprotected 72mer



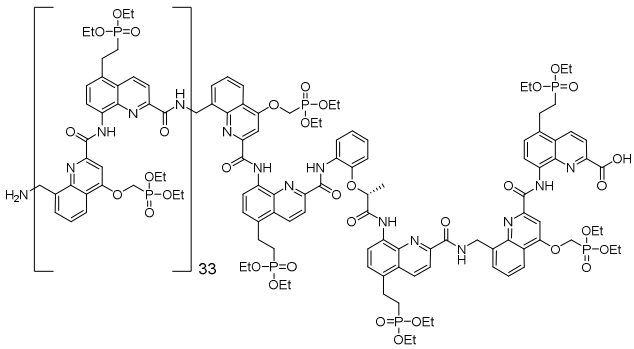
Compound **28** was synthesized on Tentagel Wang resin (15 µmol scale, initial loading: 0.077 mmol/g, 188 mg resin). From the recovered resin was obtained 208 mg of crude foldamer that was purified by semi-preparative RP-HPLC (X-Bridge C8 column, buffers: A = H_2O + 0.1% TFA, B = MeCN + 0.1% TFA, Gradient: 50-54 B in A) and subsequently lyophilized to obtain 84.6 mg (23%) of compound **28** as a yellow solid. For HPLC and MS data see below.

6.3.2.1.3.9 **40** Fmoc-mQQ[5] polyphosphonic 72mer



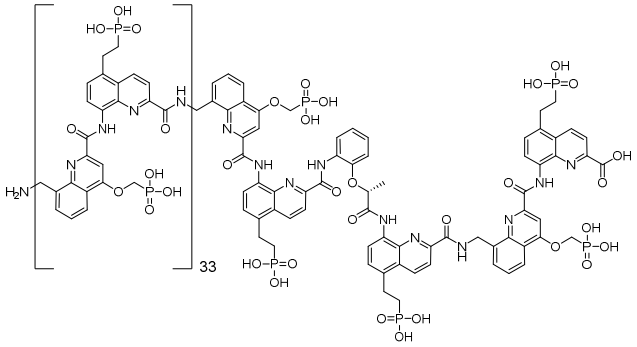
Compound **28** (20.0 mg, 810 nmol) was globally deprotected according to section 6.3.2.1.2.2 (stirring time: 8d) and then purified by semi-preparative RP-HPLC (Kinetex C18 Evo, NH_4OAc system, 0-12 % B in A) to give the title compound (7.17 mg, 346 nmol, 43%) as a poly ammonium salt as a yellow solid. For HPLC and MS data see below.

6.3.2.1.3.10 **29** NH_2-mQQ[5]-diethylprotected 72mer



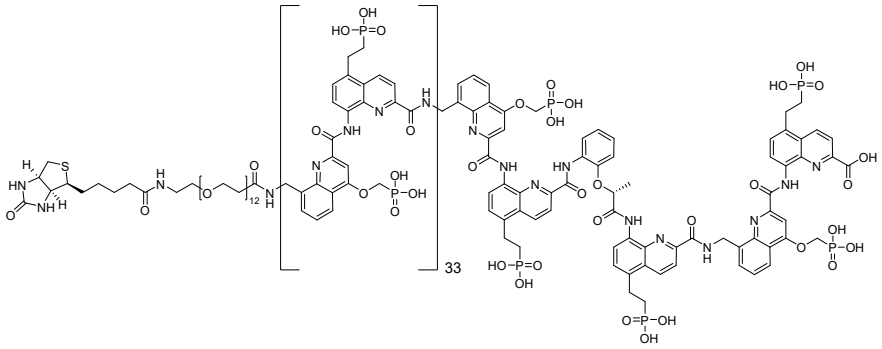
The Fmoc-group of compound **28** (18.9 mg, 772 nmol) was removed according to section 6.3.2.1.2.1 of the SI and used for global deprotection without any further purification. For HPLC and MS data see below.

6.3.2.1.3.11 **30** NH_2-mQQ^5 polyphosphonic 72mer



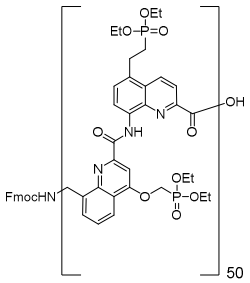
Compound **29** (18.9 mg, 772 nmol) was globally deprotected according to section 6.3.2.1.2.2 (stirring time: 4d) and then purified by semi-preparative RP-HPLC (Kinetex C18 Evo, TEAA, 0-30 % B in A) to give the title compound as a poly triethylammonium salt as a yellow solid (4.63 mg, 226 nmol, 29%) over two steps. For HPLC and MS data see below.

6.3.2.1.3.12 **31** Biotin-mQQ^5 polyphosphonic 72mer



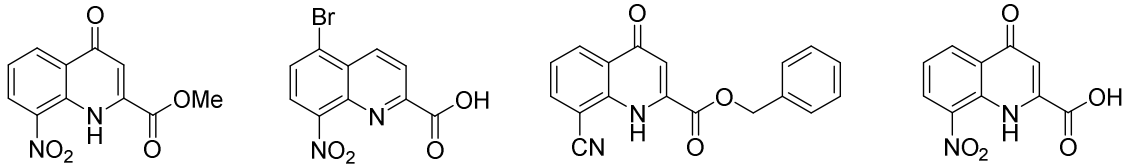
Compound **30** (1.80 mg, 88 nmol) was biotinylated according to section 6.3.2.1.2.3 and then purified by semi-preparative RP-HPLC (Kinetex C18 Evo, TEAA system, 0-30 % B in A). Combined fractions were degassed by N_2-bubbling through the solution and then freeze-dried to give the title compound as a poly triethylammonium salt that was ion-exchanged according to section. 6.3.2.1.2.4 of the SI to obtain the title compound as a yellow solid, then (0.83 mg, 39 nmol, 44%). Purified HPLC shows an additional, more polar peak that has formed after sample purification, likely reflecting biotin oxidation. For HPLC and MS data see below.

6.3.2.1.3.13 mQQ^5-diethylprotected 100mer **27**



Compound **27** was synthesized on Cl-MPA Protide resin (15 µmol scale, initial loading: 0.136 mmol/g, 111 mg resin). From the recovered 450 mg resin, 150 mg resin were subjected to resin cleavage to obtain 80 mg of crude foldamer that was purified by semi-preparative RP-HPLC (X-Bridge C8 column, buffers: A = H_2O + 0.1% TFA, B = MeCN + 0.1% TFA, Gradient: 50-54 B in A) and subsequently lyophilized to obtain 30.1 mg (extrapolated yield: 17%) of compound **27** as as a yellow solid. For HPLC and MS data see below.

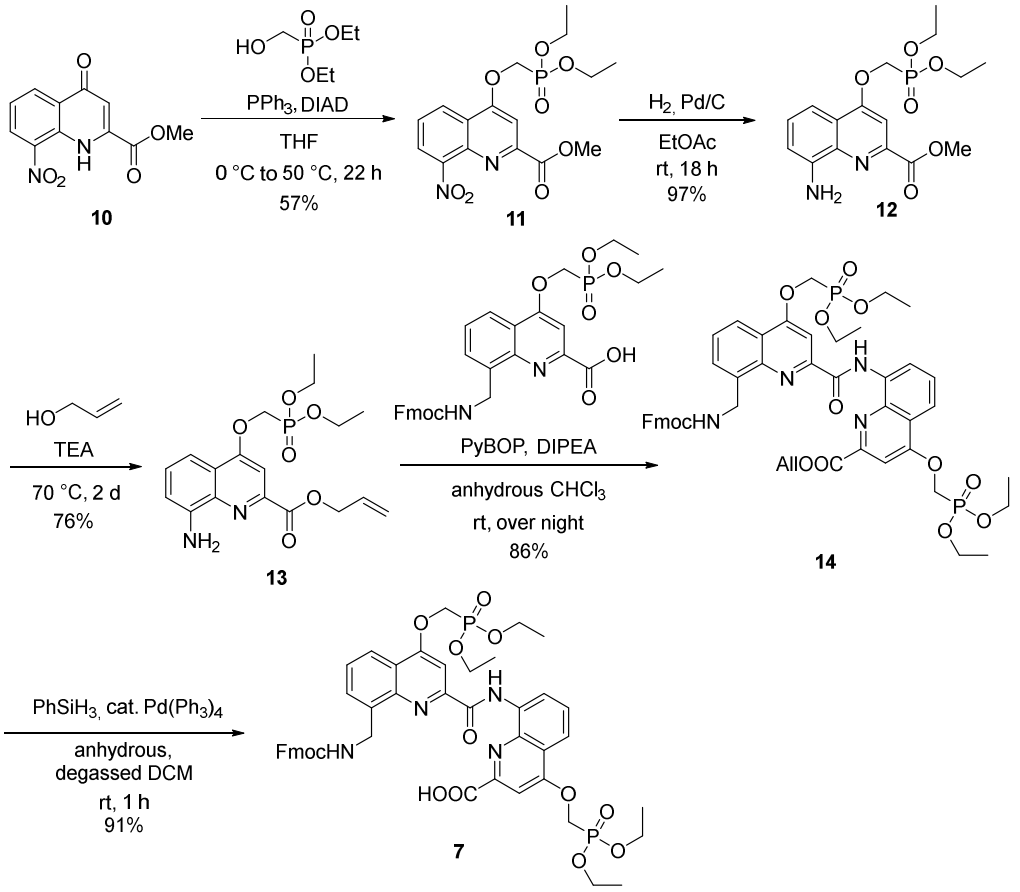
6.3.2.2 Monomer synthesis



10 **15** **20** **22**

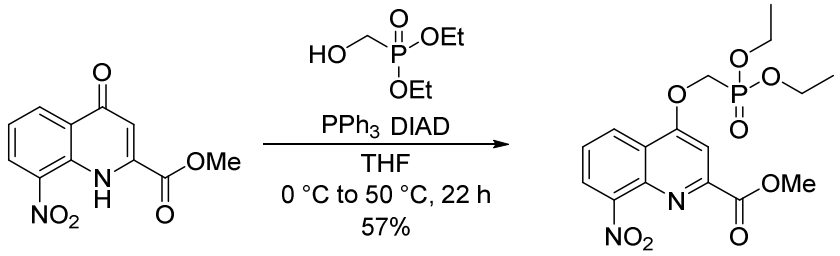
Figure S16 a) Non-commercial precursors for building block synthesis or Fmoc-monomers for solid-phase synthesis were synthesized according to published procedures: **10**[9] , **15**[10], **20**[8], **22**[8].

6.3.2.2.1 $^{m}QQ^{4}$ – dimeric building block



Scheme S1 Synthesis of **7**. Non-commercial precursor **10** was synthesized according to reported procedures.[9]

Synthesis of **11**

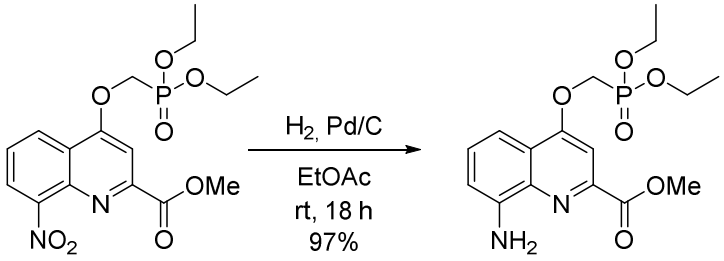


Freshly dried compound **10** (10.0 g, 40.3 mmol, 1 eq.), triphenylphosphine (13.7 g, 52.4 mmol, 1.3 eq.), and diethyl hydroxymethyl phosphonate (6.54 mL, 44.3 mmol, 1.1 eq.) were suspended in anhydrous THF under N_2-atmosphere and cooled to 0 °C. Diisopropyl azodicarboxylate (10.3 mL, 52.4 mmol, 1.3 eq.) was added dropwise over 30 min at 0 °C. The resulting solution was stirred at 0 °C for 1 h, rt for 1.5 h, and 50 °C for 16 h. THF was removed *in vacuo* and co-evaporated with DCM (2x). The brown oil was solubilized in DCM (50 mL) and Et_2O (100 mL) was layered above and let to crystallize (at -24 °C) over night. The resulting solid was dried overnight at the high-vacuum line to remove residual THF. The crude product was purified by flash column chromatography (silica, 80% EtOAc 20% cHex to 100% EtOAc. After evaporation of the solvent, the purified product was dissolved in DCM (10 mL), Et_2O

(50 mL) was layered on top and the product was crystallized at 5 °C overnight, filtered and washed with cold Et_2O (100 mL) to yield the title compound (9.10 g, 22.9 mmol, 57%) as yellow crystalline flaskes.

^{1}H NMR (400 MHz, $CDCl_3$): δ (ppm) = 8.49 (dd, *J* = 8.5, 1.4 Hz, 1H), 8.13 (dd, *J* = 7.5, 1.4 Hz, 1H), 7.75 (s, 1H), 7.70 (dd, *J* = 8.5, 7.5 Hz, 1H), 4.59 (d, *J* = 10.2 Hz, 2H), 4.29 (dq, *J* = 8.5, 7.1 Hz, 4H), 4.04 (s, 3H), 1.39 (t, *J* = 7.1 Hz, 6H).
^{13}C NMR (101 MHz, $CDCl_3$): δ (ppm) = 165.50, 162.42 (d, *J* = 13.2 Hz), 162.35, 151.37, 148.57, 140.25, 126.70, 126.31, 125.48, 123.02, 102.37, 62.98 (d, *J* = 171.8 Hz), 63.39 (d, *J* = 6.5 Hz) 53.63, 16.70 (d, *J* = 5.6 Hz).
HRMS: (ESI$^+$) m/z calcd. for $C_{16}H_{20}N_2O_8P$: 399.0952 $(M+H)^+$; found: 399.0952.

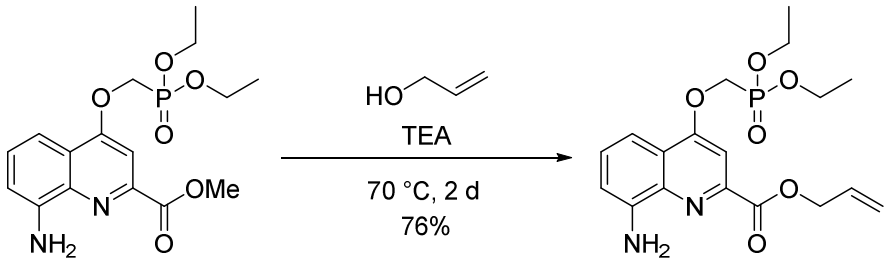
Synthesis of **12**



11 (10.2 g, 25.6 mmol, 1 eq.) was dissolved in EtOAc (425 mL) and the clear solution was degassed by N_2-flow for 15 min. Pd/C (1.2 g) was added and the black suspension was further degassed for 5 min. The solution was put under H_2-atmosphere and the reaction mixture was vigorously stirred for 24 h at rt. The reaction mixture was filtered through celite and washed with EtOAc. The solvent was removed *in vacuo* and the resulting solid was triturated in MeCN for 10 min and freeze-dried to give the title compound (9.18 g, 24.9 mmol, 97%) as a yellow solid.

^{1}H NMR (500 MHz, DMSO-d_6): δ (ppm) = δ 7.65 (s, 1H), 7.40 (t, *J* = 7.9 Hz, 1H), 7.22 (dd, *J* = 8.2, 1.3 Hz, 1H), 6.94 (dd, *J* = 7.7, 1.3 Hz, 1H), 6.06 (s, 2H), 4.82 (d, *J* = 9.9 Hz, 2H), 4.17 (dq, *J* = 8.3, 7.0 Hz, 4H), 3.95 (s, 3H), 1.27 (t, *J* = 7.0 Hz, 6H).
^{13}C NMR (126 MHz, DMSO-d_6): δ (ppm) = 165.39, 161.44 (d, *J* = 13.5 Hz), 146.13, 144.97, 137.14, 129.49, 122.00, 109.88, 106.52, 101.37, 62.39 (d, *J* = 6.1 Hz), 61.73 (d, *J* = 164.5 Hz), 52.58, 16.32 (d, *J* = 5.6 Hz).
HRMS: (ESI$^+$) m/z calcd. for $C_{16}H_{22}N_2O_6P$: 369.1210 $(M+H)^+$; found: 369.1201.

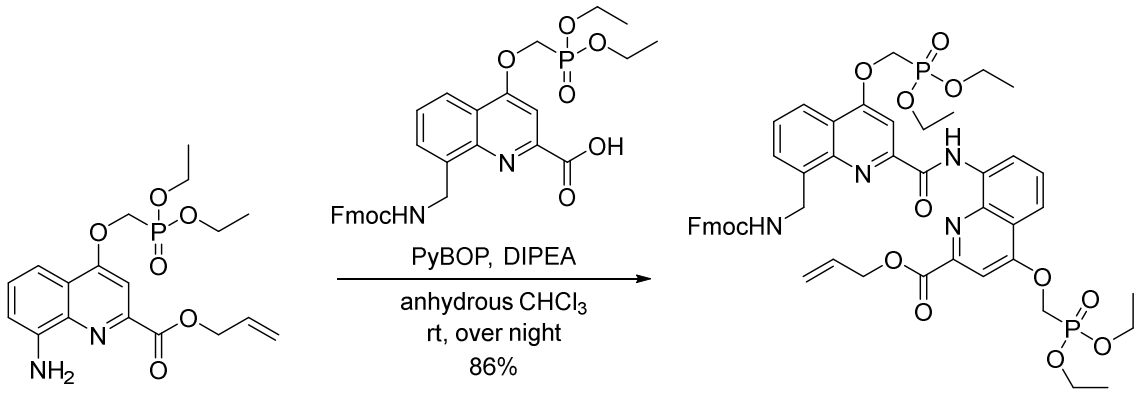
Synthesis of **13**



In a dry flask, **12**(1.00 g, 2.71 mmol, 1 eq.) was dissolved in allyl alcohol (10.2 mL, distilled over K_2CO_3) under N_2-atmosphere and triethylamine (1.14 mL, 8.14 mmol, 3 eq., freshly distilled over CaH_2) was added. The dark yellow solution was stirred for 24 h at 70 °C. After 24 h, triethylamine (1.14 mL. 8.14 mmol, 3 eq., freshly distilled over CaH_2) was added and the solution was further stirred for 26 h at 70 °C. Volatiles were removed *in vacuo* and the crude product was purified by automated reversed-phase chromatography (C18, 10 → 100% MeCN in H_2O) and freeze-dried to give the title compound (817 mg, 2.07 mmol, 76%) as a yellow solid.

^{1}H NMR (500 MHz, DMSO-d_6): δ (ppm) = 7.65 (s, 1H), 7.41 (t, *J* = 8.0 Hz, 1H), 7.23 (dd, *J* = 8.2, 1.3 Hz, 1H), 6.95 (dd, *J* = 7.7, 1.3 Hz, 1H), 6.11 (ddt, *J* = 17.3, 10.6, 5.4 Hz, 1H), 6.06 (s, 2H), 5.48 (dq, *J* = 17.3, 1.7 Hz, 1H), 5.32 (dq, *J* = 10.5, 1.5 Hz, 1H), 4.90 (dt, *J* = 5.5, 1.5 Hz, 2H), 4.83 (d, *J* = 9.9 Hz, 2H), 4.17 (dq, *J* = 8.4, 7.1 Hz, 4H), 1.27 (t, *J* = 7.0 Hz, 6H).
^{13}C NMR (126 MHz, DMSO-d_6): δ (ppm) = 164.52, 161.46 (d, *J* = 13.6 Hz), 146.09, 144.95, 137.15, 132.54, 129.53, 122.03, 118.20, 109.90, 106.57, 101.39, 65.67, 62.39 (d, *J* = 6.0 Hz), 61.72 (d, *J* = 164.2 Hz), 16.33 (d, *J* = 5.7 Hz).
HRMS: (ESI$^+$) m/z calcd. for $C_{18}H_{24}N_2O_6P$: 395.1366 $(M+H)^+$; found: 395.1361.

Synthesis of **14**



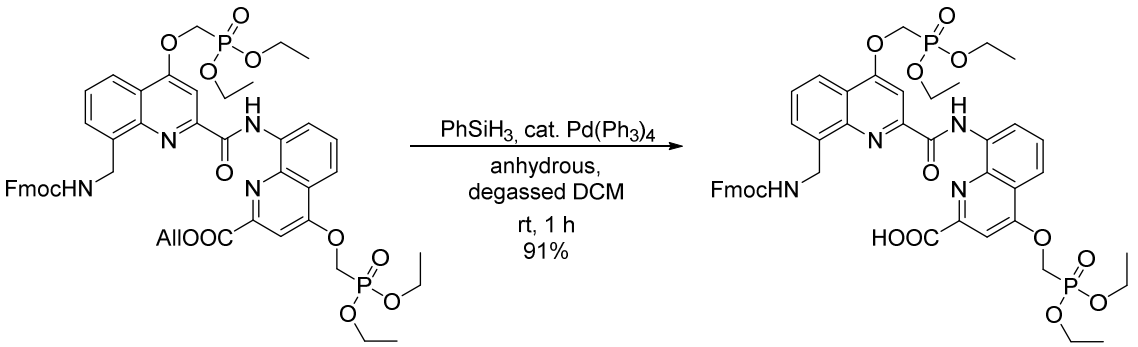
In a dry flask, **4** (4.00 g, 6.77 mmol, 1 eq.), **13** (2.67 g, 6.77 mmol, 1 eq.) and PyBOP (7.05 g, 13.6 mmol, 2 eq.) were dissolved in $CHCl_3$ (130 mL, freshly distilled over CaH_2) and the mixture was cooled to 0 °C. Diisopropylethylamine (1.30 mL, 7.45 mmol, 1.1 eq.) was added at 0 °C and the reaction mixture was stirred at 0 °C for 1 h and at rt for 24 h. Volatiles were removed *in vacuo* and the crude oil was dissolved in DCM (150 mL). The organic layer was washed with $NaHCO_3$ (5%, 2x) and NH_4Cl (sat., 2x) and the aqueous layer was back-extracted with DCM. The combined organic layers were dried over $MgSO_4$, solids were filtered and the solvent was removed *in vacuo*. The mixture was dissolved in MeCN (120 mL) and triturated for 20 min until full precipitation. The solid was isolated by filtration and washed with MeCN to give the title compound (5.6 g, 5.79 mmol, 86%) as a colorless solid.

NMR shows 2 conformers. Major 80%, minor 20%. Resolved signals are given with their respective integration. For unresolved signals, total integration is given. ^{13}C shifts of the minor species are not listed. HPLC trace as additional proof of purity, see section 6.3.3.2.1.

^{1}H NMR (500 MHz, $CDCl_3$) δ (ppm) = 12.43 (s, 0.8H), 12.36 (s, 0.2H), 9.18 (d, *J* = 7.8 Hz, 0.8H), 9.13 (d, *J* = 7.5 Hz, 0.2H), 8.25 (d, *J* = 8.3 Hz, 0.8H), 8.20 (d, *J* = 8.4 Hz, 0.2H), 8.03 – 7.95 (m, 1H), 7.93 – 7.88 (m, 1.6H), 7.76 – 7.71 (m, 1.2H), 7.65 (d, *J* = 7.6 Hz, 1.6H), 7.63 – 7.58 (m, 1.6H), 7.47 – 7.42 (m, 0.4H), 7.39 – 7.32 (m, 2H), 7.28 (t, *J* = 7.7 Hz, 1.6H), 7.17 (t, *J* = 7.5 Hz, 0.4H), 7.04 (t, *J* = 7.2 Hz, 1.6H), 6.21 (t, *J* = 6.2 Hz, 0.8H), 6.11 – 5.95 (m, 1H), 5.63 (t, *J* = 6.3 Hz, 0.2H), 5.40 (d, *J* = 17.3 Hz, 2H), 5.30 – 5.23 (m, 2.4H), 5.06 (d, *J* = 6.6 Hz, 0.4H), 4.96 (d, *J* = 5.8 Hz, 1.60H), 4.86 (d, *J* = 5.0 Hz, 0.4H), 4.64 (d, *J* = 10.3 Hz, 2H), 4.58 (d, *J* = 10.2 Hz, 0.4H), 4.54 (d, *J* = 10.3 Hz, 1.6H), 4.49 (d, *J* = 5.7 Hz, 0.4H), 4.37 – 4.23 (m, 9.6H), 4.18 (bs, 0.2H), 4.06 (t, *J* = 7.3 Hz, 0.8H), 1.40 (td, *J* = 7.0, 4.0 Hz, 12H).

^{13}C NMR (125 MHz, $CDCl_3$): δ (ppm) = 164.40, 163.36, 163.25, 162.98, 162.66, 162.55, 156.97, 150.60, 147.64, 146.17, 144.03, 141.24, 139.89, 137.35, 135.29, 131.82, 131.79, 130.90, 129.14, 127.70, 127.61, 126.92, 125.11, 122.20, 122.03, 121.49, 119.91, 119.38, 118.28, 116.07, 101.56, 98.47, 77.41, 77.16, 76.91, 66.99, 66.64, 63.41, 63.35, 63.30, 61.95, 61.93, 47.35, 42.87, 16.74, 16.69. **HRMS:** (ESI^+) m/z calcd. for $C_{49}H_{53}N_{24}O_{13}P_2$: 967.3079 $(M+H)^+$; found: 967.3054. **HPLC** purity at 300 nm: 98.9%.

Synthesis of **7**



14 (1.00 g, 1.03 mmol, 1 eq.) and tetrakis(triphenylphosphine)-palladium(0) (59.8 mg, 51.7 mmol, 0.05 eq.) were dissolved in degassed, anhydrous DCM (30.0 mL) and put under Ar-atmosphere. Phenylsilane (382 µL, 1.03 mmol, 3 eq.) was added and the flask was covered from light and stirred at rt for 1 h. Volatiles were removed *in vacuo* and the crude product was redissolved in MeCN/DCM (1:1, 30 mL). The product was precipitated from the solution by slow evaporation of DCM from the solvent mixture to give the title compound (873 mg, 941 µmol, 91%) as a grey solid. To remove leftover traces of palladium contamination, the compound can be further purified by flash column chromatography (5→10% MeOH in DCM + 0.1% AcOH).

NMR shows 2 conformers. Major 85%, minor 15%. Resolved signals are given with their respective integration. For unresolved signals, total integration is given. ^{13}C shifts of the minor species are not listed. HPLC trace as additional proof of purity, see section 6.3.3.2.1.

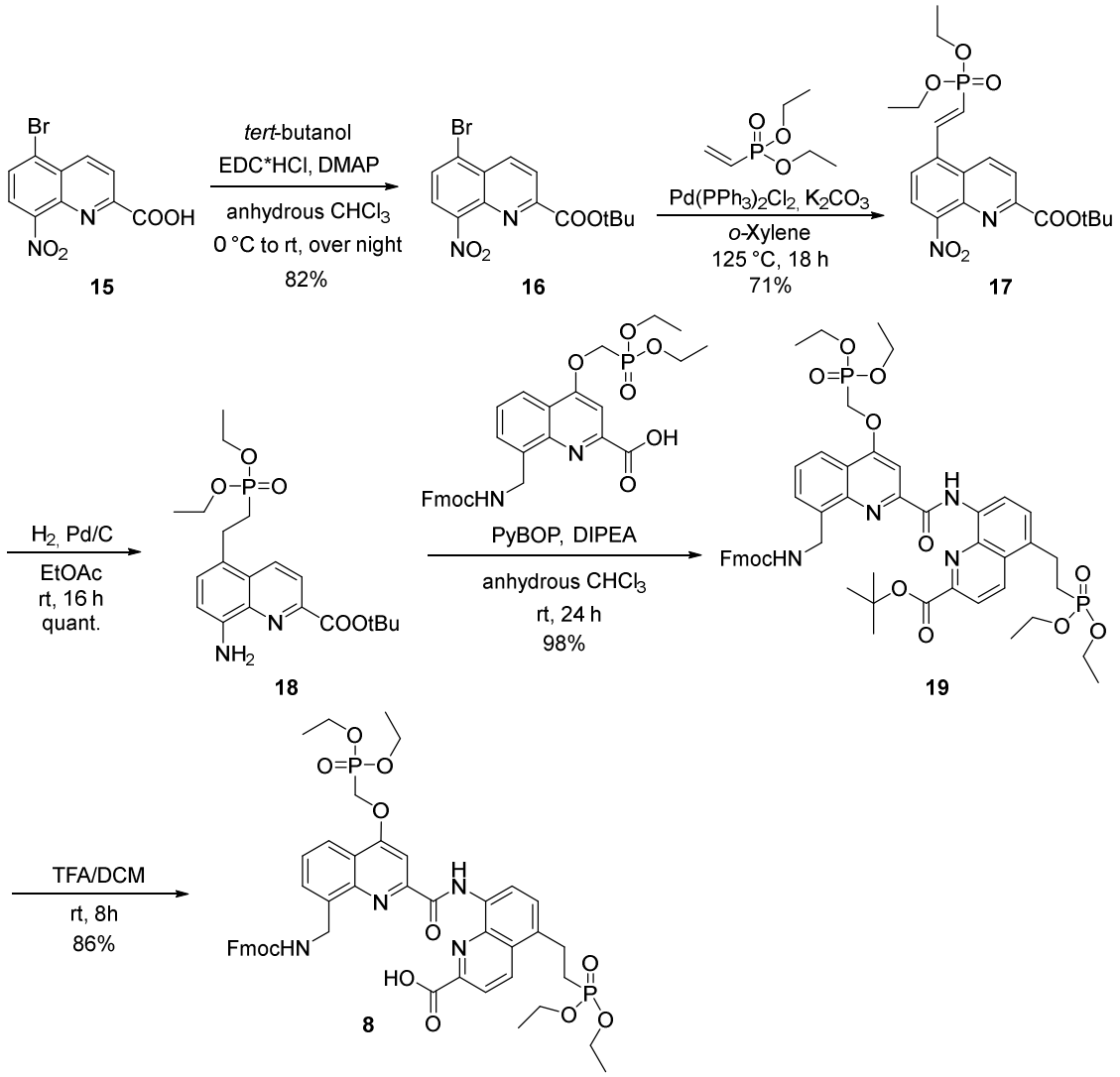
^{1}H NMR (500 MHz, DMSO-d_6): δ (ppm) = 13.65 (bs, 1H), 12.31 (s, 0.15H), 12.28 (s, 0.85H), 9.10 – 9.03 (m, 1H), 8.11 (d, *J* = 8.4 Hz, 1H), 8.04 (s, 0.15H), 8.01 (s, 0.85H), 7.92 – 7.85 (m, 3.7H), 7.81 (t, *J* = 8.1 Hz, 1H), 7.75 – 7.62 (m, 5H), 7.41 (t, *J* = 7.5 Hz, 1.7H), 7.37 (t, *J* = 6.6 Hz, 0.3H), 7.29 (t, *J* = 7.4 Hz, 1.7H), 7.25 (d, *J* = 7.6 Hz, 0.3H), 7.16 (t, *J* = 7.5 Hz, 0.3H), 6.86 (t, *J* = 7.5 Hz, 0.3H), 5.21 – 5.11 (m, 2H), 5.04 – 4.97 (m, 2H), 4.94 (d, *J* = 9.8 Hz, 2H), 4.37 (d, *J* = 6.8 Hz, 1.7H), 4.27 – 4.15 (m, 9.15H), 4.00 (t, *J* = 6.8 Hz, 0.15H), 1.32 – 1.26 (m, 12H).

^{13}C NMR (126 MHz, DMSO-d_6): δ (ppm) = 166.13, 162.74, 162.64, 162.23, 162.09, 161.98, 156.55, 150.02, 144.77, 143.91, 140.78, 138.62, 138.10, 134.69, 128.59, 128.36, 127.78, 127.61, 127.05, 125.22, 121.29, 121.26, 120.14, 119.93, 117.04, 115.47, 102.61, 99.31, 65.25, 62.70, 62.62, 62.53, 62.48, 61.40, 61.32, 46.89, 40.49, 16.43, 16.41, 16.38, 16.37.

HRMS: (ESI$^+$) m/z calcd. for $C_{45}H_{49}N_4O_{13}P_2$: 927.2766 $(M+H)^+$; found: 927.2745.

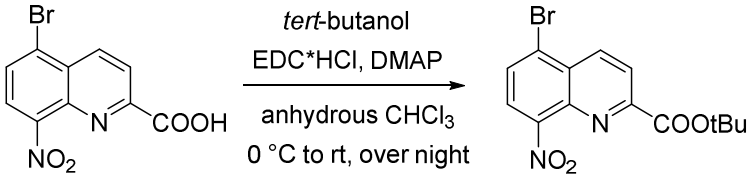
HPLC purity at 300 nm: 99.8%.

6.3.2.2.2 $^{m}QQ^{5}$ – dimeric building block



Scheme S2 Synthesis of **8**. Non-commercial precursor **15** was synthesized according to reported procedures.[10]

Synthesis of **16**



In a dry flask, **15** (1.00 g, 3.37 mmol, 1 eq.) and *N*,*N*-dimethyl-4-aminopyridine (411 mg, 3.37 mmol, 1 eq.) were dissolved in anhydrous $CHCl_3$ (20 mL) under N_2-atmosphere and cooled to 0 °C. EDC-hydrochloride (2.39 g, 13.5 mmol, 4 eq.) and *tert*-butanol (1.28 mL, 13.5 mmol, 4 eq.) were added and the solution was stirred at 0 °C for 2 h and at rt for 18h. The solvent was removed *in vacuo* and the resulting solid was dissolved in EtOAc (30 mL). The

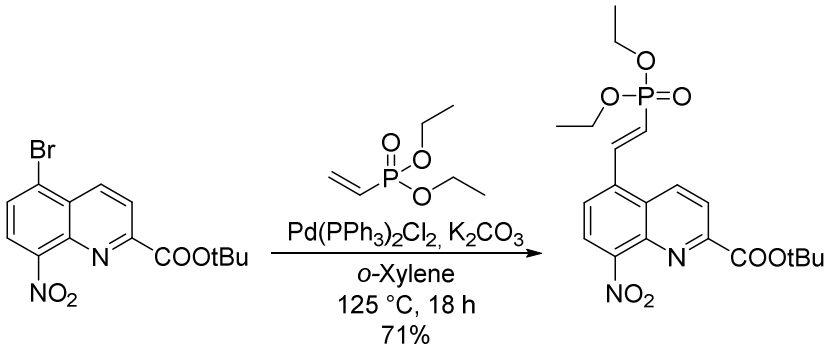
organic phase was washed with aq. NH_4Cl-solution (sat., 2x), aq. $NaHCO_3$-solution (5 % w/v, 2x). The combined aqeous phases were back-extracted with EtOAc, the combined organic phases were washed with brine (1x), dried over $MgSO_4$, filtered and the solvent was removed *in vacuo*. The crude mixture was purified by column chromatography (silica, 20% EtOAc c-Hex, R_F = 0.4) to give the title compound (980 mg, 2.77 mmol, 82%) as an orange solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ (ppm) = 8.74 (d, *J* = 8.8 Hz, 1H), 8.32 (d, *J* = 8.8 Hz, 1H), 7.99 (s, 2H), 1.67 (s, 9H).

^{13}C NMR (126 MHz, $CDCl_3$): δ (ppm) = 163.07, 152.28, 148.26, 139.63, 137.28, 130.66, 129.16, 126.23, 124.82, 123.67, 83.57, 28.17.

HRMS: (ESI^+) m/z calcd. for $C_{14}H_{14}N_2O_4Br$: 374.9956 $(M+Na)^+$; found: 374.9952.

Synthesis of **17**



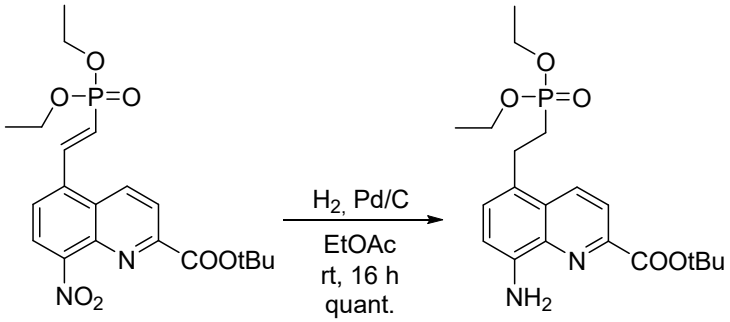
In a dry flask, **16** (802 mg, 2.27 mmol, 1 eq.) and oven dried K_2CO_3 (314 mg, 2.27 mmol, 1 eq.) were suspended in anhydrous *o*-xylene (8 mL) under N_2-atmosphere. Diethyl vinyl phosphonate (418 µL, 2.73 mmol, 1.2 eq.) and bis(triphenylphosphine)palladium(II) dichloride (47.8 mg, 47.8 µmol, 0.03 eq.) were added and the black solution was stirred at 125 °C for 18 h. After cooling, the reaction mixture was diluted with EtOAc (80 mL) and the organic phase was washed with citric acid (5% w/v, 3x). The combined aqueous phases were back-extracted with EtOAc and the combined organic phases were dried over Na_2SO_4, filtered and the solvent was removed *in vacuo*. The crude mixture was purified by column chromatography (silica, EtOAc, R_F = 0.5). Fractions containing the product were combined, the solvent was removed *in vacuo*. To remove traces of catalyst (may disturb following reaction), the mixture was dissolved in DCM (3 mL) and Et_2O was layered on top and the product was left to crystallize at -14 °C. The title compound was isolated by filtration and washed with cold Et_2O to give a grey solid (705 mg, 1.62 mmol, 71%).

^{1}H NMR (500 MHz, $CDCl_3$) δ (ppm) = 8.69 (d, *J* = 8.9 Hz, 1H), 8.29 (d, *J* = 8.9 Hz, 1H), 8.18 (dd, *J* = 22.1, 17.3 Hz, 1H), 8.11 (d, *J* = 7.9 Hz, 1H), 7.89 (d, *J* = 7.9 Hz, 1H), 6.52 (t, *J* = 17.0 Hz, 1H), 4.27 – 4.16 (m, 4H), 1.67 (s, 9H), 1.40 (t, *J* = 7.1 Hz, 6H).

^{13}C NMR (126 MHz, $CDCl_3$): δ (ppm) = 163.22, 151.74, 149.24, 141.95 (d, *J* = 7.2 Hz), 139.25, 136.98 (d, *J* = 23.5 Hz),, 133.38, 127.71, 125.24 (d, *J* = 1.6 Hz), 124.16, 122.96 (d, *J* =

189.0 Hz), 122.96, 83.45, 62.52 (d, *J* = 5.6 Hz), 28.17, 16.63 (d, *J* = 6.2 Hz).
HRMS: (ESI$^+$) m/z calcd. for $C_{20}H_{26}N_2O_7P$: 437.1472 $(M+H)^+$; found: 437.1473.

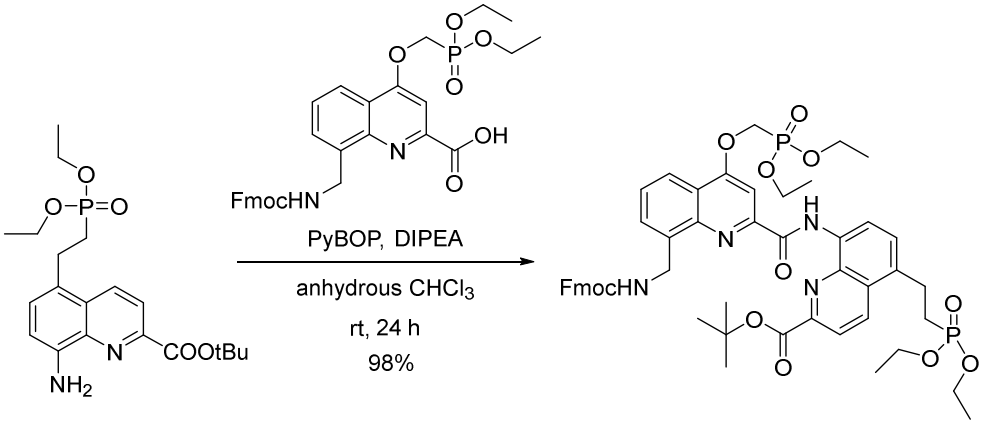
Synthesis of **18**



17 (3.19 g, 7.31 mmol, 1 eq.) was dissolved in EtOAc (135 mL) and the solution was degassed by N_2-bubbling for 15 min. Pd/C (350 mg) was added and the solution further degassed for 5 min. The mixture was put under H_2-atmosphere and stirred at rt for 24 h under vigorous stirring. The crude mixture was filtered through a pad of celite and washed with EtOAc. The solvent was removed *in vacuo* to give the title compound (3.00 g, 7.31 mmol, quant.) as an orange oil.

^{1}H NMR (500 MHz, $CDCl_3$) δ (ppm) = 8.33 (d, *J* = 8.8 Hz, 1H), 8.09 (d, *J* = 8.8 Hz, 1H), 7.27 (d, *J* = 7.7 Hz, 2H), 6.86 (d, *J* = 7.7 Hz, 1H), 5.12 (bs, 2H), 4.20 – 4.03 (m, 4H), 3.28 – 3.17 (m, 2H), 2.14 – 2.01 (m, 2H), 1.67 (s, 9H), 1.33 (t, *J* = 7.1 Hz, 6H).
^{13}C NMR (126 MHz, $CDCl_3$): δ (ppm) = 164.44, 145.94, 144.22, 138.23, 132.78, 129.22, 127.98, 124.65 (d, *J* = 17.3 Hz), 121.17, 109.75, 82.28, 77.48, 77.16, 76.84, 61.81 (d, *J* = 6.5 Hz), 28.32, 27.39 (d, *J* = 138.3 Hz), 24.72 (d, *J* = 4.2 Hz), 16.65 (d, *J* = 6.0 Hz).
^{13}C NMR (101 MHz, $CDCl_3$) δ 124.65 (d, *J* = 17.3 Hz), 82.28, 61.81 (d, *J* = 6.5 Hz), 27.39 (d, *J* = 138.3 Hz), 24.72 (d, *J* = 4.2 Hz), 16.65 (d, *J* = 6.0 Hz).
HRMS: (ESI$^+$) m/z calcd. for $C_{20}H_{30}N_2O_5P$: 409.1887 $(M+H)^+$; found: 409.1887.

Synthesis of **19**



In a dry flask, **4** (4.31 g, 7.30 mmol, 1 eq.), PyBOP (7.59 g, 14.6 mmol, 2 eq.) and diisopropyl ethyl amine (1.40 mL, 8.03 mmol, 1.1 eq.) were dissolved in $CHCl_3$ (140 mL, freshly distilled over CaH_2) under N_2-atmosphere and the mixture was stirred. **18** (2.98 g, 7.30 mmol, 1 eq.) was dissolved in $CHCl_3$ (15 mL, freshly distilled over CaH_2) under N_2-atmosphere, added to the reaction mixture and the orange solution was stirred at rt for 24 h. Volatiles were removed *in vacuo* and the crude oil was dissolved in EtOAc. The organic layer was washed with $NaHCO_3$ (5%, 3x) and NH_4Cl (sat.) and the combined aqueous layers were back-extracted with EtOAc. The combined organic layers were washed with brine, dried over $MgSO_4$, the solids were filtered off and the solvent was removed *in vacuo*. To the residual oil, acetonitrile (50 mL) was added and the mixture was left to precipitate while sonication. The title compound was isolated by filtration and washed with cold (4 °C) acetonitrile (7.01 g, 1.72 mmol, 98%) as yellow solid. NMR shows 2 conformers. Major 82%, minor 18%. Resolved signals are given with their respective integration. For unresolved signals, total integration is given. ^{13}C shifts of the minor species are not listed. HPLC trace as additional proof of purity, see section 6.3.3.2.1.

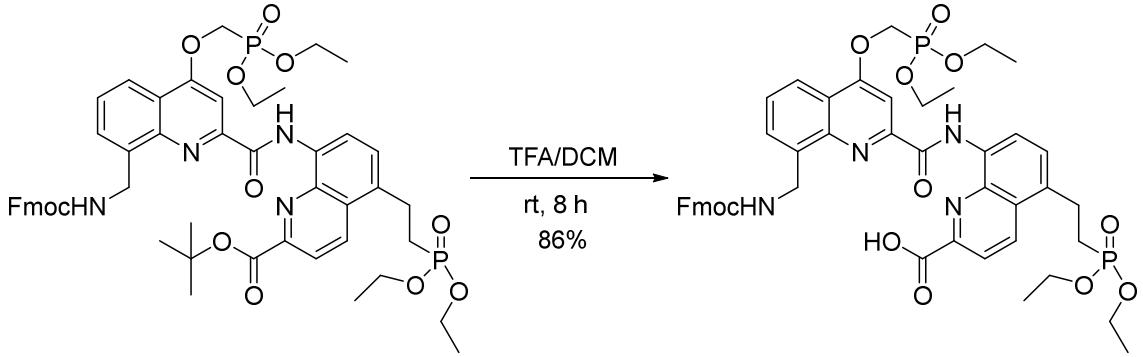
^{1}H NMR (500 MHz, $CDCl_3$): δ (ppm) = 12.54 (s, 0.82H), 12.51 (s, 0.18H), 9.06 (d, *J* = 7.9 Hz, 0.82H), 9.01 (d, *J* = 7.9 Hz, 0.18H), 8.56 – 8.49 (m, 1H), 8.26 – 8.12 (m, 2H), 7.96 – 7.89 (m, 1.8H), 7.72 – 7.64 (m, 2H), 7.64 – 7.56 (m, 1.8H), 7.49 – 7.43 (m, 0.2H), 7.40 (d, *J* = 7.5 Hz, 2.2H), 7.30 (t, *J* = 7.5 Hz, 2H), 7.12 (t, *J* = 7.4 Hz, 0.4H), 7.07 (t, *J* = 7.4 Hz, 1.6H), 6.34 (t, *J* = 6.4 Hz, 0.82H), 5.69 (t, *J* = 5.5 Hz, 0.18H), 5.37 (d, *J* = 6.3 Hz, 1.64H), 5.25 (d, *J* = 6.6 Hz, 0.36H), 4.63 (d, *J* = 10.3 Hz, 2H), 4.43 (d, *J* = 6.4 Hz, 0.36H), 4.31 (dq, *J* = 8.3, 7.1 Hz, 4H), 4.25 (d, *J* = 7.4 Hz, 1.64H), 4.20 – 4.11 (m, 4.2H), 4.08 (t, *J* = 7.4 Hz, 0.82H), 3.38 (td, *J* = 9.9, 5.0 Hz, 2H), 2.21 – 2.11 (m, 2H), 1.64 (s, 9H) 1.40 (t, *J* = 7.1 Hz, 6H), 1.35 (t, *J* = 7.1 Hz, 6H).

^{13}C NMR (126 MHz, $CDCl_3$): δ (ppm) = 163.32, 163.23, 163.21, 162.98, 157.12, 150.56, 146.20, 144.11, 141.28, 139.46, 137.79, 134.40, 133.61, 131.61, 131.48, 130.90, 128.90, 127.67, 127.61, 127.55, 126.97, 125.22, 122.20, 121.65, 121.29, 119.94, 98.39, 82.87, 66.66, 63.34, 63.29, 61.96, 61.91, 47.38, 42.72, 28.25, 27.74, 26.63, 25.07, 25.04, 16.72, 16.68, 16.64.

HRMS: (ESI^+) m/z calcd. for $C_{51}H_{58}N_4O_{12}P_2Na$: 1003.3424 $(M+Na)^+$; found: 1003.3396.

HPLC purity at 300 nm: 99.7%.

Synthesis of **8**



19 (6.20 g, 6.32 mmol, 1 eq.) was dissolved in DCM (43 mL) and TFA (43 mL) was added. The dark orange solution was stirred at rt for 8 h. The solution was poured on ice, then diluted with DCM (50 mL) and the organic phase was separated. The aqueous phase was extracted with DCM (50 mL, 2x) and the combined organic phases were dried over $MgSO_4$, filtered and the solvent was removed *in vacuo*. The crude product was dissolved in DCM/MeOH (1:1, v/v, 60 mL) The product precipitated after 1 h by slow evaporation of DCM at 150-200 mbar (keep outside heating bath). The product was isolated by filtration and washed with -20 °C cold methanol to give the title compound (5.01 g, 5.42 mmol, 86%) as a yellow solid. NMR shows 2 conformers. Major 82%, minor 18%. Resolved signals are given with their respective integration. For unresolved signals, total integration is given. ^{13}C shifts of the minor species are not listed. HPLC trace as additional proof of purity, see section 6.3.3.2.1.

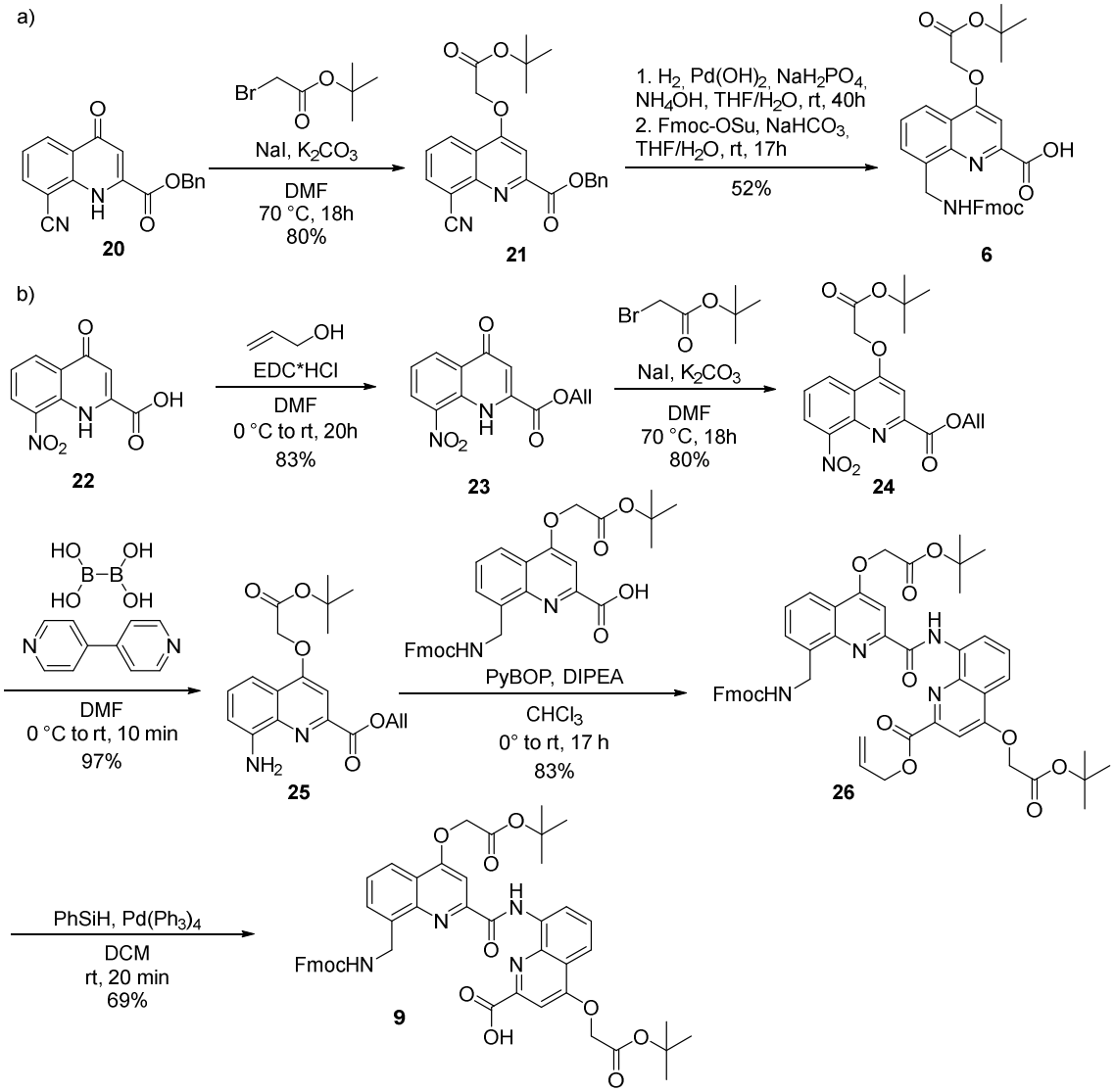
^{1}H NMR (500 MHz, DMSO-d_6): δ (ppm) = 13.61 (s, 1H), 12.26 (s, 0.15H), 12.23 (s, 0.85H), 8.97 (d, *J* = 8.2 Hz, 0.15H), 8.94 (d, *J* = 7.9 Hz, 0.85H), 8.68 (d, *J* = 8.8 Hz, 1H), 8.29 (d, *J* = 8.7 Hz, 1H), 8.11 (dd, *J* = 8.3, 1.4 Hz, 1H), 8.02 (s, 0.15H), 7.98 (s, 0.85H), 7.88 (d, *J* = 7.6 Hz, 1.7H), 7.75 – 7.60 (m, 5.85H), 7.41 (t, *J* = 7.5 Hz, 1.7H), 7.33 (d, *J* = 6.6 Hz, 0.15H), 7.28 (t, *J* = 7.4 Hz, 1.7H), 7.24 (d, *J* = 7.6 Hz, 0.3H), 7.15 (t, *J* = 7.6 Hz, 0.3H), 6.85 (t, *J* = 7.5 Hz, 0.3H), 5.20 – 5.13 (m, 2H), 5.03 – 4.94 (m, 2H), 4.37 (d, *J* = 6.8 Hz, 1.7H), 4.21 (p, *J* = 7.3 Hz, 5.15H), 4.07 – 3.95 (m, 4.15H), 3.31 – 3.22 (m, 2H), 2.23 – 2.12 (m, 2H), 1.30 (t, *J* = 7.0 Hz, 6H), 1.22 (t, *J* = 7.0 Hz, 6H).

^{13}C NMR (126 MHz, DMSO-d_6): δ (ppm) = 165.90, 162.70, 162.59, 162.09, 156.50, 150.07, 146.81, 144.77, 143.88, 140.76, 138.05, 138.02, 134.30, 133.37, 131.97, 131.84, 128.60, 128.30, 127.70, 127.57, 127.19, 127.00, 125.16, 121.69, 121.22, 120.09, 119.91, 116.43, 99.27, 65.22, 62.62, 62.50, 62.45, 61.31, 61.09, 61.04, 40.46, 26.11, 25.02, 24.21, 24.18, 16.38, 16.34, 16.30, 16.25.

HRMS: (ESI$^+$) m/z calcd. for $C_{47}H_{51}N_4O_{12}P_2$: 925.2973 $(M+H)^+$; found: 925.2978.

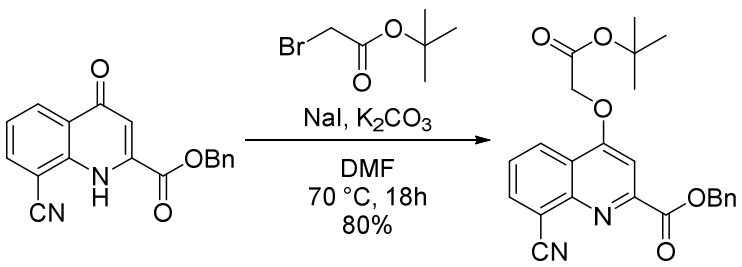
HPLC purity at 300 nm: 98.9%.

6.3.2.2.3 $^{m}QQ^{4Asp}$ – dimeric building block



Scheme S3 a) Synthesis of **6**. Non-commercial precursor **20** was synthesized according to reported procedures.[8] b) Synthesis of **9**. Non-commercial precursor **22** was synthesized according to reported procedures.[8]

Synthesis of **21**



In a dry flask, **20** (5.70 g, 18.7 mmol, 1 eq.), NaI (702 mg, 4.68 mmol, 0.25 eq.) and K_2CO_3 (2.85 g, 20.6 mmol, 1.1 eq., oven-dried at 120 °C) were suspended in anhydrous DMF (105 mL) under N_2-atmosphere. *tert*-butyl bromoacetate (3.27 mL, 22.5 mmol, 1.2 eq.) were added and

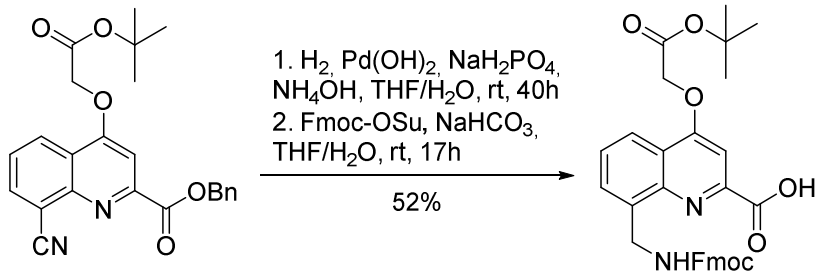
the reaction mixture was stirred at 70 °C for 18 h under N_2-atmosphere. The crude mixture was diluted with EtOAc (500 mL) and washed with H_2O (1x) and aq. LiCl (5%, 5x). The organic phase was dried over $MgSO_4$, filtered and the solvent was removed *in vacuo*. The crude product was triturated in Et_2O (50 mL) for 10 min and left to precipitate at -24 °C for 30 min. The colorless solid was filtered, washed with Et_2O and left to dry at air over night to give the title compound as a colorless solid (5.85 g, 15.0 mmol, 80%).

^{1}H NMR (500 MHz, $CDCl_3$): δ (ppm) = 8.57 (dd, *J* = 8.5, 1.4 Hz, 1H), 8.19 (dd, *J* = 7.2, 1.4 Hz, 1H), 7.67 (dd, *J* = 8.5, 7.2 Hz, 1H), 7.56 (d, *J* = 7.1 Hz, 2H), 7.53 (s, 1H), 7.41 (t, *J* = 7.6 Hz, 2H), 7.36 – 7.32 (m, 1H), 5.52 (s, 2H), 4.84 (s, 2H), 1.49 (s, 9H).

^{13}C NMR (126 MHz, $CDCl_3$): δ (ppm) = 166.07, 164.98, 162.13, 151.16, 148.08, 136.98, 135.59, 128.82, 128.52, 128.34, 127.33, 127.02, 122.36, 116.81, 114.11, 102.52, 83.79, 68.01, 66.08, 28.17.

HRMS: (ESI$^+$) m/z calcd. for $C_{24}H_{23}N_2O_5$: 419.1601 $(M+H)^+$; found: 419.1594.

Synthesis of **6**



21 (5.88 g, 14.1 mmol, 1 eq.) was dissolved in THF (660 mL) and H_2O (240 mL). NaH_2PO_4 x $2H_2O$ (22.0 g, 141 mmol, 10 eq.) were added and the cloudy solution was degassed by bubbling N_2 threw the solution for 30 min in an open flask. The reaction mixture was put under N_2-atmosphere and NH_4OH (20%, 7.45 mL, 42.3 mmol, 3 eq.) and $Pd(OH)_2/C$ (20% w/w, 592 mg) were added. The N_2-atmosphere was replaced by H_2-atmosphere (balloon) and vigorously stirred at rt for 40 h until an HPLC-aliquot did show no residual CN-Q^{Asp}-COOH (note that the debenzylation is rapid while the reduction of the nitrile takes significantly longer, when monitoring by TLC, make sure to co-spot the acid). The black suspension was degassed by N_2-bubbling through the solution to disperse residual dissolved H_2 in solution and the compound was used directly for the next step with no evaporation of solvent or further purification. To the black suspension $NaHCO_3$ (5.90 g, 70 mmol, 5 eq.) was added. Fmoc-OSu (4.74 g, 14.1 mmol, 1 eq.) was dissolved in THF (50 mL) and added dropwise over 10 min. The suspension was stirred at rt for 17 h. The crude mixture was filtered through celite, and washed with THF. The mixture was concentrated in vacuo until ca 300 mL remained. The aq. suspension was acidified with citric acid (5%) to pH = 3, and the aq. phase was extracted with DCM (3 x). The combined organic phases were dried over MgSO4, filtered, and the solvent was removed in vacuo to give a dark-green oil. The oil was quickly taken up in MeCN (50 mL) and

triturated to give a colorless precipitate after 1 minute. MeCN(100 mL) was added, further triturated and left standing at rt over night to allow full precipitation. The product was isolated by filtration and washed with −24 °C MeCN (2x, 50 mL) and the solid was freeze-dried to give the title compound a colorless solid (4.03 g, 7.27 mmol, 52% over 2 steps).NMR shows 2 conformers. Major 87%, minor 13%. Resolved signals are given with their respective integration. For unresolved signals, total integration is given. ^{13}C shifts of the minor species are not listed. HPLC trace as additional proof of purity, see section 6.3.3.2.1.

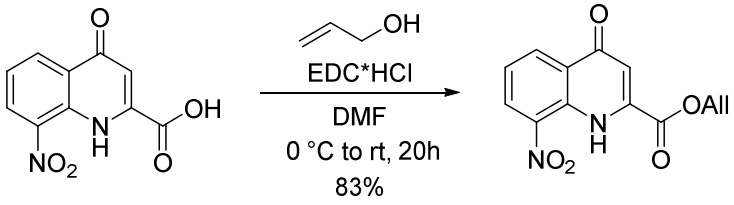
^{1}H NMR (500 MHz, DMSO-*d*$_6$): δ (ppm) = 13.08 (bs, 1H), 8.16 (dd, *J* = 8.2, 1.6 Hz, 1H), 7.92 – 7.86 (m, 2.48H), 7.78 (d, *J* = 7.7 Hz, 0.26H), 7.69 (d, *J* = 7.6 Hz, 2.13H), 7.64 (dd, *J* = 7.1, 1.6 Hz, 1.13H), 7.50 (s, 1H), 7.41 (t, *J* = 7.5 Hz, 1.74H), 7.37 (d, *J* = 7.5 Hz, 0.26H), 7.30 (t, *J* = 7.1 Hz, 1.74H), 7.08 (t, *J* = 7.5 Hz, 0.26H), 5.12 (s, 2H), 4.85 (d, *J* = 6.1 Hz, 1.74H), 4.75 (d, *J* = 6.3 Hz, 0-26H), 4.40 (d, *J* = 6.8 Hz, 1.74H), 4.32 (d, *J* = 6.4 Hz, 0.26H), 4.25 (t, *J* = 6.7 Hz, 0.87H), 4.13 (t, *J* = 6.5 Hz, 0.13H), 1.44 (s, 9H).

^{13}C NMR (126 MHz, DMSO-*d*$_6$): δ (ppm) = 166.88, 165.91, 161.58, 156.54, 148.24, 145.39, 143.87, 140.77, 137.89, 128.41, 127.64, 127.59, 127.03, 125.12, 121.27, 120.38, 120.12, 100.93, 82.10, 65.67, 65.28, 46.84, 40.59, 27.66.

HRMS: (ESI$^-$) m/z calcd. for $C_{32}H_{29}N_2O_7$: 553.1980 (M−H)$^-$; found: 553.1962.

HPLC purity at 300 nm: 97.25%.

Synthesis of **23**



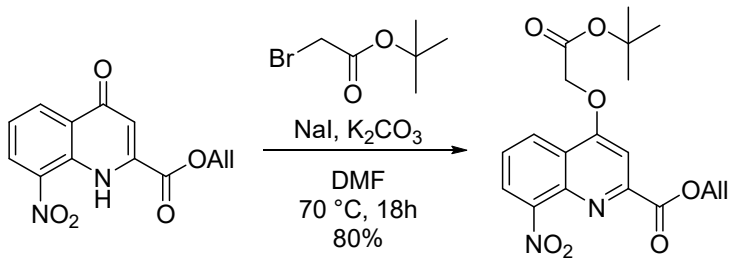
22 (5.27 g, 22.3 mmol, 1.0 eq.) and EDC*HCl (6.42 g, 33.5 mmol, 1.5 eq.) were put in a dry flask under N_2 atmosphere. The solids were suspended in anhydrous DMF (72 mL) and stirred under N_2-atmosphere. Allyl alcohol (1.43 g, 1.68 mL, 24.5 mmol, 1.1 eq.) was dropwise added at 0 °C over 5 min, warmed to rt and stirred for 20 h under N_2-atmosphere. The mixture was poured on ice, and the precipitate was vacuum filtered and washed with cold water. The received solid was dried in a vacuum oven at 60 °C over night to give the title compound as a (5.09 g, 18.6 mmol, 83%) yellow solid.

^{1}H NMR (500 MHz, CDCl$_3$): δ (ppm) = 11.80 (s, 1H), 8.77 – 8.69 (m, 2H), 7.49 (t, *J* = 8.0 Hz, 1H), 7.10 (d, *J* = 1.8 Hz, 1H), 6.04 (ddt, *J* = 16.6, 10.4, 5.9 Hz, 1H), 5.50 (d, *J* = 17.2 Hz, 1H), 5.41 (d, *J* = 10.4 Hz, 1H), 4.96 (d, *J* = 5.8 Hz, 2H).

^{13}C NMR (126 MHz, CDCl$_3$): δ (ppm) = 178.09, 161.63, 137.58, 136.09, 135.36, 134.41, 131.29, 130.49, 128.42, 122.92, 120.46, 113.69, 68.02.

HRMS: (ESI$^-$) m/z calcd. for $C_{13}H_9N_2O_5$: 273.0517 (M−H)$^-$; found: 273.0515.

Synthesis of **24**

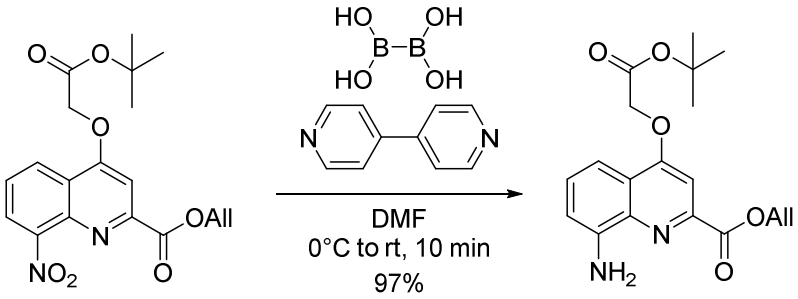


In a dry flask, **23** (8.32 g, 30.3 mmol, 1 eq.), NaI (1.14 g, 7.58 mmol, 0.25 eq.) and oven-dried K_2CO_3 (4.61 g, 33.4 mmol, 1.1 eq.) were put under high-vacuum for 30 min, then put under N_2-atmosphere and suspended in anhydrous DMF (150 mL). *tert*-Butyl bromoacetate (5.30 mL, 36.4 mmol, 1.2 eq.) were added and the suspension was stirred at 70 °C for 18 h. The mixture was poured into water (500 mL) and extracted with EtOAc (500 mL). The organic phase was washed with aq. LiCl (5%, 5x), then dried over $MgSO_4$, filtered and the solvent was removed *in vacuo* to give the title compound as an off-white solid (10.1 g, 26.0 mmol, 86%).

^{1}H NMR (500 MHz, $CDCl_3$): δ (ppm) = 8.56 (dd, *J* = 8.5, 1.4 Hz, 0H), 8.12 (dd, *J* = 7.5, 1.4 Hz, 0H), 7.68 (dd, *J* = 8.5, 7.5 Hz, 0H), 7.55 (s, 0H), 6.08 (ddt, *J* = 17.2, 10.5, 5.6 Hz, 1H), 5.52 (dq, *J* = 17.2, 1.6 Hz, 0H), 5.35 (dq, *J* = 10.4, 1.3 Hz, 1H), 4.92 (dt, *J* = 5.6, 1.4 Hz, 1H), 4.86 (s, 1H), 1.52 (s, 9H).
^{13}C NMR (126 MHz, $CDCl_3$): δ (ppm) = 166.03, 164.70, 161.82, 151.27, 148.58, 140.35, 131.52, 126.63, 126.50, 125.39, 123.13, 119.15, 102.44, 83.84, 67.02, 66.12, 28.19.
HRMS: (ESI^+) m/z calcd. for $C_{19}H_{20}N_2O_7Na$: 411.1168 $(M+Na)^+$; found: 411.1166.

Synthesis of **25**

WARNING: This transformation is very exothermic and can cause a runaway reaction due to excessive heat production. Do not run in a closed system. Do not scale up this reaction before further tests.

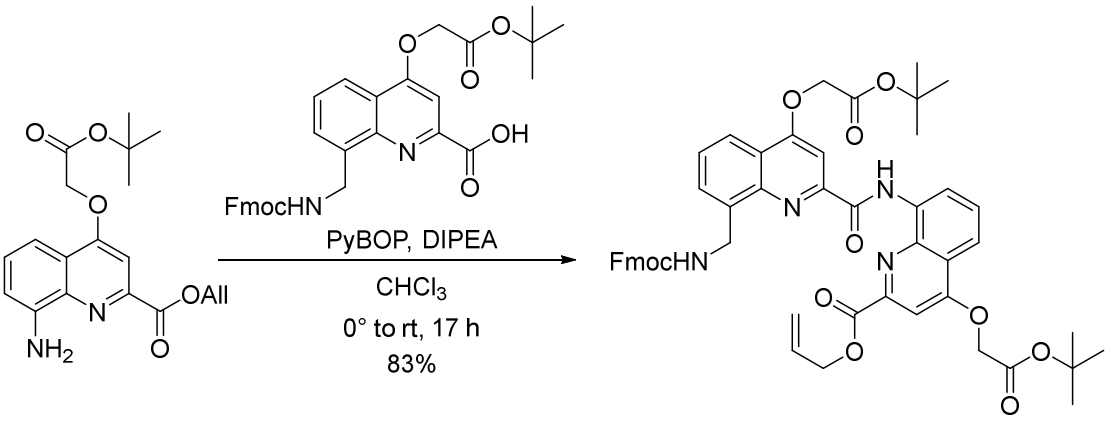


24 (4.00 g, 10.3 mmol, 1 eq.) , 4,4-bipyridine (8.04 mg, 51.5 µmol, 0.005 eq.) were dissolved in DMF (80 mL) and cooled to 0 °C. Hypodiboric acid (2.82 g, 31.4 mmol, 3.05 eq.) were added portion wise over 2 min. The mixture was allowed to warm to rt and stirred for 10 min (this

will heat up higher than rt at this stage, add back the ice bath to slowdown if necessary). A color change from yellow to dark blue (ca. 10 seconds duration) to orange after 5 min indicates completion of the reaction. The mixture was poured in water (500 mL) and extracted with EtOAc (300 mL). The organic phase was washed with aq. LiCl (5%, 5x), dried over $MgSO_4$, filtered and the solvent was removed *in vacuo* to give the title compound (3.60 g, 10.0 mmol, 97%) as a yellow to orange solid.

^{1}H NMR (500 MHz, $CDCl_3$): δ (ppm) = 7.58 (dd, *J* = 8.4, 1.3 Hz, 1H), 7.41 – 7.34 (m, 2H), 6.95 (dd, *J* = 7.5, 1.2 Hz, 1H), 6.10 (ddt, *J* = 17.2, 10.4, 5.7 Hz, 1H), 5.48 (dd, *J* = 17.1, 1.5 Hz, 1H), 5.33 (dq, *J* = 10.3, 1.3 Hz, 1H), 5.12 (s, 2H), 4.92 (dt, *J* = 5.8, 1.5 Hz, 2H), 4.79 (s, 2H), 1.51 (s, 9H).
^{13}C NMR (126 MHz, $CDCl_3$): δ (ppm) = 166.77, 165.35, 161.55, 145.60, 145.03, 138.68, 132.12, 129.22, 122.85, 118.93, 111.19, 109.88, 101.06, 83.22, 66.53, 65.90, 28.19.
HRMS: (ESI^+) m/z calcd. for $C_{19}H_{22}N_2O_7Na$: 381.1426 $(M+Na)^+$; found: 381.1423.

Synthesis of **26**



6 (4.01 g, 7.23 mmol, 1 eq.), **25** (2.54 g, 7.09 mmol, 0.98 eq.) and PyBOP (7.53 g, 14.46 mmol, 2 eq.) were dissolved in anhydrous $CHCl_3$ (120 mL) under N_2-atmosphere and cooled to 0 °C. To the mixture, DIPEA (2.52 mL, 14,46 mmol, 2 eq.) was added and the mixture was allowed to warm to rt and stirred at rt for 17h. $CHCl_3$ was removed *in vacuo* and the crude oil was diluted with EtOAc. The organic phase was washed with $NaHCO_3$ (3x, 5%, w/v), then citric acid (3x, 5%, w/v). The aqueous phases were back extracted with EtOAc (separately) and the combined organic phases were washed with brine, then dried over $MgSO_4$, filtered and the solvent was removed *in vacuo*. The mixture was resolubilized in EtOAc/cHex (1:1, v:v) and filtered trough a pad of silica, washed with EtOAc/cHex (1:1, v:v). The solvent was removed *in vacuo* and the crude oil was precipitated from MeCN. The compound was isolated by filtration and washed with MeCN (−24 °C) and left to dry at air to yield the title compound as orange solid (5.25 g, 5.87 mmol, 83%).

NMR shows 2 conformers. Major 83%, minor 17%. Resolved signals are given with their respective integration. For unresolved signals, total integration is given. ^{13}C shifts of the minor species are not listed. HPLC trace as additional proof of purity, see section 6.3.3.2.1.

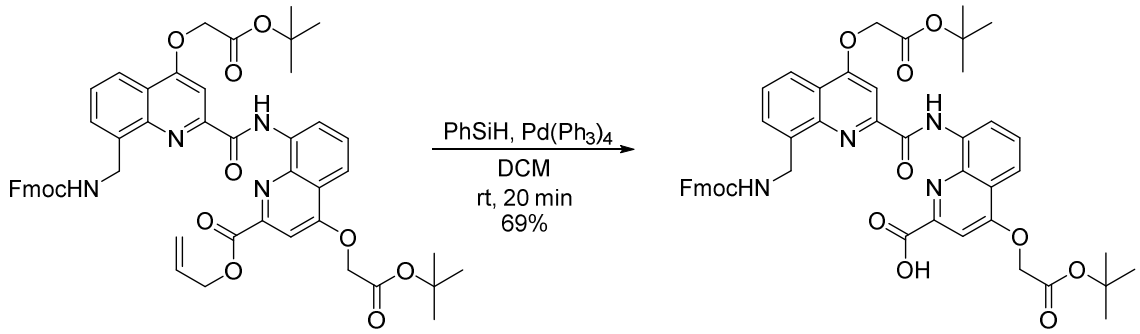
^{1}H NMR (500 MHz, $CDCl_3$): δ (ppm) = 12.41 (s, 0.83H), 12.35 (s, 0.17H), 9.16 (d, *J* = 7.7 Hz, 0.83H), 9.11 (d, *J* = 7.9 Hz, 0.17H), 8.33 (d, *J* = 8.3 Hz, 0.83H), 8.28 (d, *J* = 7.8 Hz, 0.17H), 8.07 (d, *J* = 8.3 Hz, 1H), 7.89 (d, *J* = 6.9 Hz, 0.83H), 7.79 – 7.68 (m, 2.45H), 7.64 (d, *J* = 7.6 Hz, 1.68H), 7.60 (t, *J* = 7.7 Hz, 0.83H), 7.46 – 7.40 (m, 1.49H), 7.37 (d, *J* = 7.5 Hz, 1.68H), 7.32 (d, *J* = 7.7 Hz, 0.34H), 7.28 (d, *J* = 7.5 Hz, 1.68H), 7.16 (t, *J* = 7.3 Hz, 0.34H), 7.04 (t, *J* = 7.4 Hz, 1.67H), 6.19 (t, *J* = 6.4 Hz, 0.83H), 6.01 (ddt, *J* = 16.4, 10.9, 5.8 Hz, 1H), 5.61 (t, *J* = 6.0 Hz, 0.17H), 5.39 (dd, *J* = 17.0, 1.6 Hz, 1H), 5.29 – 5.21 (m, 2.7H), 5.08 (d, *J* = 6.6 Hz, 0.31H), 4.95 (d, *J* = 5.8 Hz, 1.67H), 4.91 (s, 2.13H), 4.84 (s, 0.35H), 4.80 (s, 1.68H), 4.49 (d, *J* = 5.8 Hz, 0.32H), 4.26 (d, *J* = 7.3 Hz, 1.67H), 4.17 (t, *J* = 6.7 Hz, 0.17H), 4.06 (t, *J* = 7.3 Hz, 0.83H), 1.55 (s, 9H), 1.51 (s, 9H).

^{13}C NMR (126 MHz, $CDCl_3$): δ (ppm) = 166.68, 166.47, 164.53, 163.08, 162.69, 162.02, 156.97, 150.62, 147.52, 146.31, 144.10, 141.27, 140.03, 137.24, 135.30, 131.91, 130.91, 128.95, 127.60, 127.52, 126.96, 125.16, 122.40, 122.20, 121.87, 119.88, 119.16, 118.23, 116.42, 101.77, 98.68, 83.46, 83.38, 66.86, 66.65, 66.06, 65.91, 47.40, 42.92, 28.26, 28.20.

HRMS: (ESI^+) m/z calcd. for $C_{51}H_{51}N_4O_{11}$: 895.3549 $(M+H)^+$; found: 895.3552

HPLC purity at 300 nm: 98.85%.

Synthesis of **9**



26 (1.50 g, 1.68 mmol, 1 eq.) and $Pd(PPh_3)_4$ (51.9 mg, 44.9 µmol, 0.03 eq.) were dissolved in DCM (50 mL) under N_2-atmosphere and the flask was protected from light. Phenylsilane (0.62 mL, 544 mg, 5.03 mmol, 3 eq.) was added and the reaction mixture was stirred for 20 min at rt. Following the completion of the reaction the mixture was washed with aq. citric acid solution (3x, 5% w/v) and the combined aqueous phases were extracted with DCM (3x). The combined organic phases were dried over $MgSO_4$ and the solvent was removed *in vacuo*. From the crude oil, the compound was precipitated from acetonitrile, sonicated to allow for full precipitation, filtered and washed with cold (−24°C) acetonitrile. The compound was further purified by column chromatography (silica, EtOAc 50:50 + 0.1% NEt_3 → 80:20 + 0.1% AcOH) to afford the tittle compound (987 mg, 1.15 mmol, 69%) as a colorless solid.

NMR shows 2 conformers. Major 85%, minor 15%. Resolved signals are given with their respective integration. For unresolved signals, total integration is given. ^{13}C shifts of the minor species are not listed. HPLC trace as additional proof of purity, see section 6.3.3.2.1.

^{1}H NMR (500 MHz, $CDCl_3$): δ (ppm) = 13.65 (bs, 1H), 12.31 – 12.24 (m, 1H), 9.07 – 9.00 (m, 1H), 8.18 (d, *J* = 8.0 Hz, 1H), 7.96 (dd, *J* = 8.4, 1.3 Hz, 1H), 7.87 (d, *J* = 7.5 Hz, 1.7H), 7.80 – 7.73 (m, 1.15H), 7.71 (s, 1H), 7.69 – 7.61 (m, 4.5H), 7.59 (s, 1H), 7.40 (t, *J* = 7.5 Hz, 1.7H), 7.37 – 7.30 (m, 0.3H), 7.30 – 7.22 (m, 2.15H), 7.15 (t, *J* = 7.5 Hz, 0.3H), 6.86 (t, *J* = 7.5 Hz, 0.3H), 5.22 – 5.11 (m, 6.05H), 4.34 (d, *J* = 6.9 Hz, 1.7H), 4.21 (t, *J* = 6.9 Hz, 1.15H), 4.00 (t, *J* = 6.7 Hz, 0.15H), 1.48 (s, 9H), 1.45 (s, 9H).

^{13}C NMR (126 MHz, $CDCl_3$): δ (ppm) = 166.90, 166.81, 166.07, 162.16, 162.11, 161.46, 156.51, 149.76, 148.76, 144.88, 143.87, 140.74, 138.73, 137.96, 134.56, 128.39, 128.33, 127.55, 127.00, 126.55, 125.19, 124.72, 121.33, 121.31, 120.27, 120.07, 117.06, 115.81, 102.28, 98.91, 82.15, 82.12, 65.78, 65.26, 46.85, 40.54, 27.70, 27.66.

HRMS: (ESI^+) m/z calcd. for $C_{48}H_{47}N_4O_{11}$: 855.3236 $(M+H)^+$; found: 855.3242

HPLC purity at 300 nm: 99.77%.

6.3.2.3 Building block recovery

During automated solid-phase synthesis with a Chorus multi-channel peptide synthesizer the lines intended for automated resin cleavage and collection of resin-cleaved material were manipulated to directly lead into a bottle with aq. citric acid (5% w/v) as a quenching solution that was constantly stirred. Each coupling solution was subsequently led into the quenching solution. After completion of the synthesis, to the mixture, EtOAc was added and the mixture was stirred until solubilization of precipitates. The aqueous phase was separated and the organic phase was washed with aq. LiCl (5x, 5%, w/v). See example HPLC chromatograms below. Recovered dimers were purified by silica column chromatography. The column was packed in 0.5% DIPEA, 2.5% MeOH, 97% DCM (v:v:v) and equilibrated for 1 column volume, the compound was loaded dissolved in 0.5% DIPEA, 2.5% MeOH, 97% DCM (v:v:v) and eluted for 3 CV 0.5% DIPEA, 2.5% MeOH, 97% DCM (v:v:v), then 3% MeOH, 97% DCM (v:v), and then eluted by 3% to 10% MeOH in DCM + 0.1% AcOH. Combined pure fractions were concentrated and then washed with water (2x). The combined organic phases were dried over $MgSO_4$, filtered, the solvent was removed *in vacuo* and the product was lyophilized from $MeCN/H_2O$ (1:1, v:v) to yield recovered building blocks in yields higher than 50% of excess used.

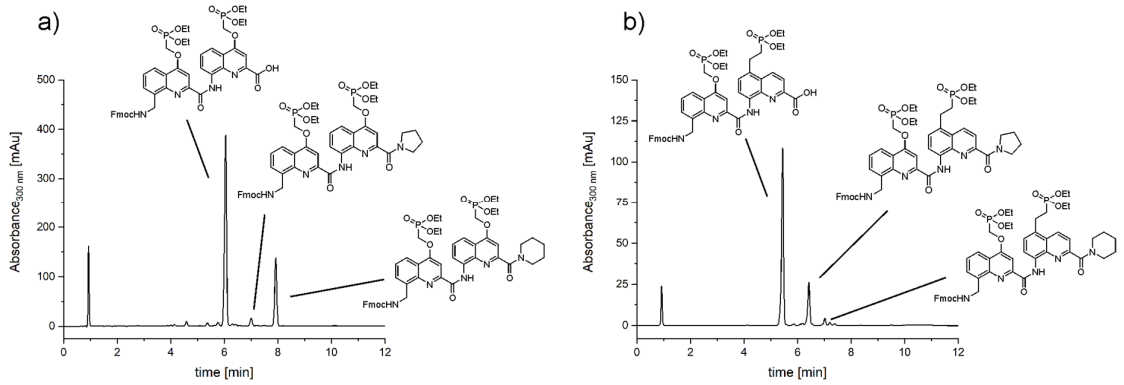


Figure S17 HPLC traces (gradient 50-100, TFA buffered HPLC) of quenched and extracted recovered Fmoc dimer building blocks **7** (a) and **8** (b).

6.3.3 Spectra & Chromatograms

6.3.3.1 NMR spectra

6.3.3.1.1 Small molecules

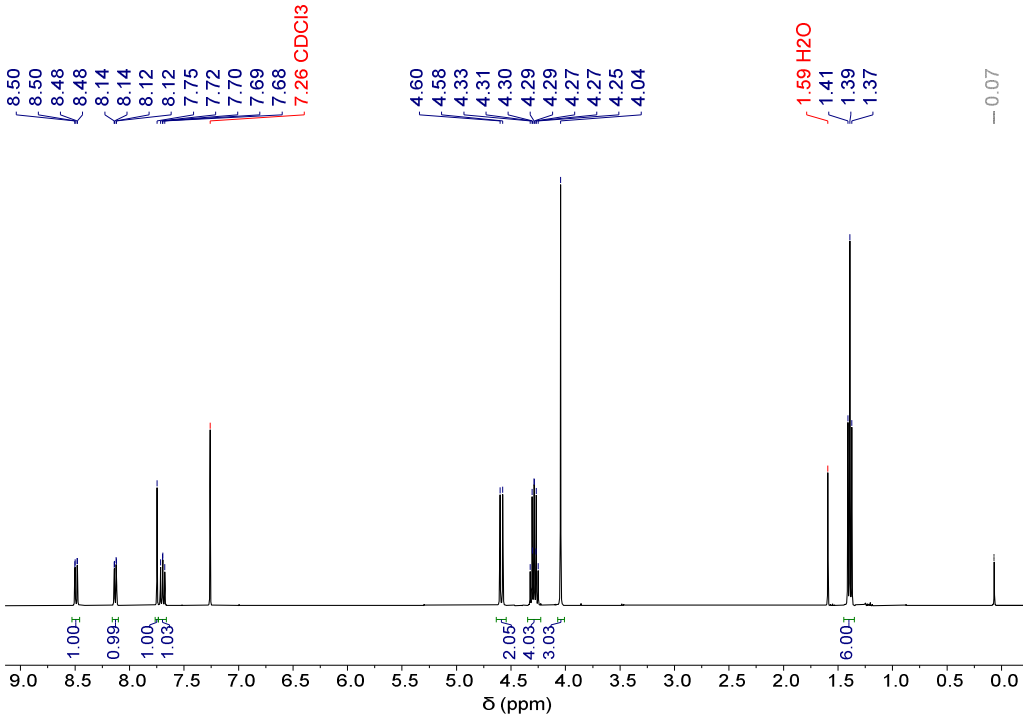


Figure S18 [1]H NMR spectrum of compound **11**.

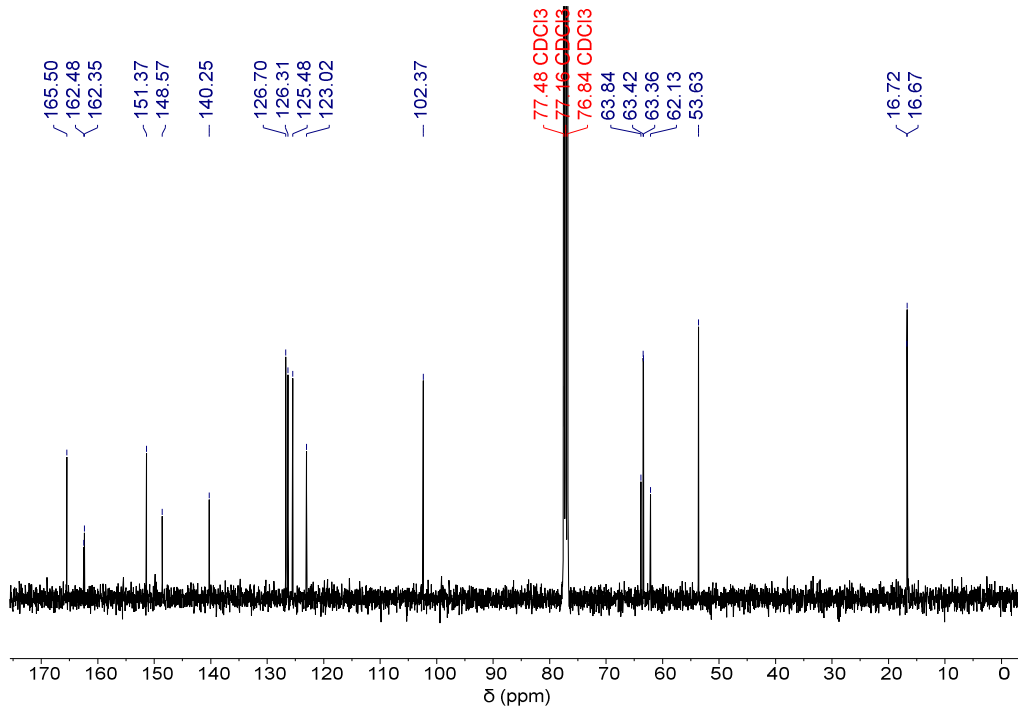


Figure S19 [13]C NMR spectrum of compound **11**.

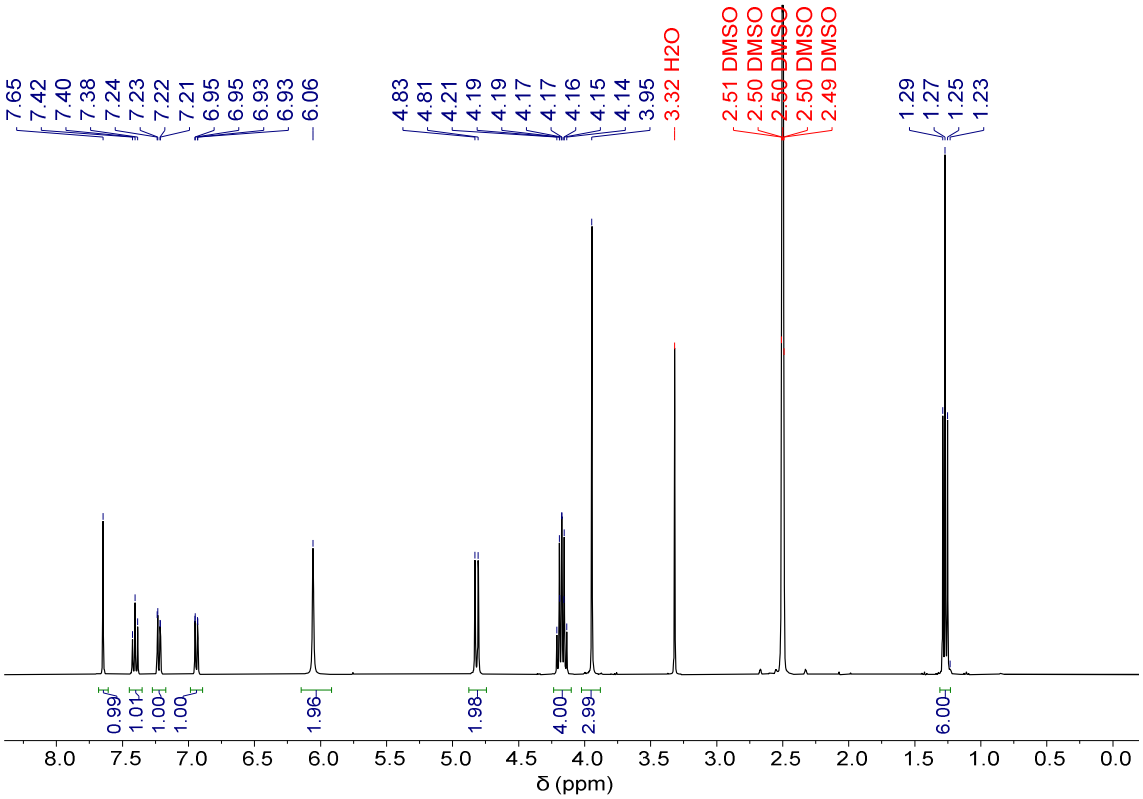


Figure S20 [1]H NMR spectrum of compound **12**.

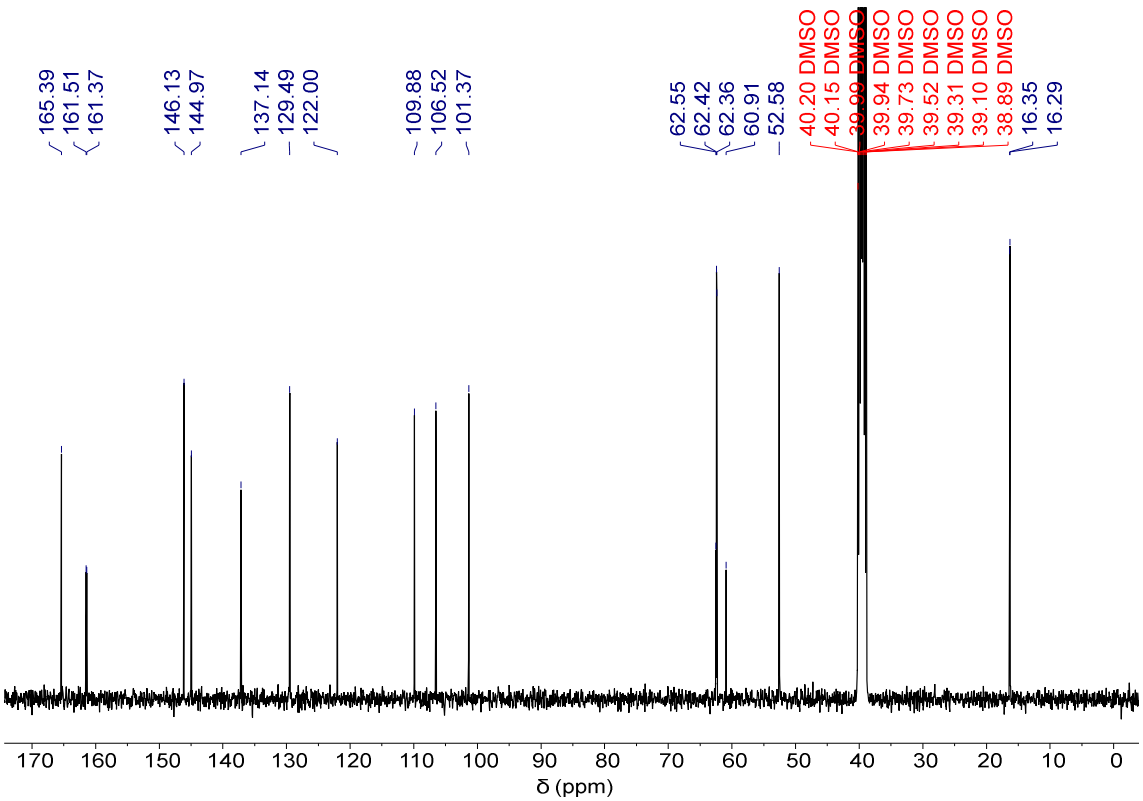


Figure S21 [13]C NMR spectrum of compound **12**.

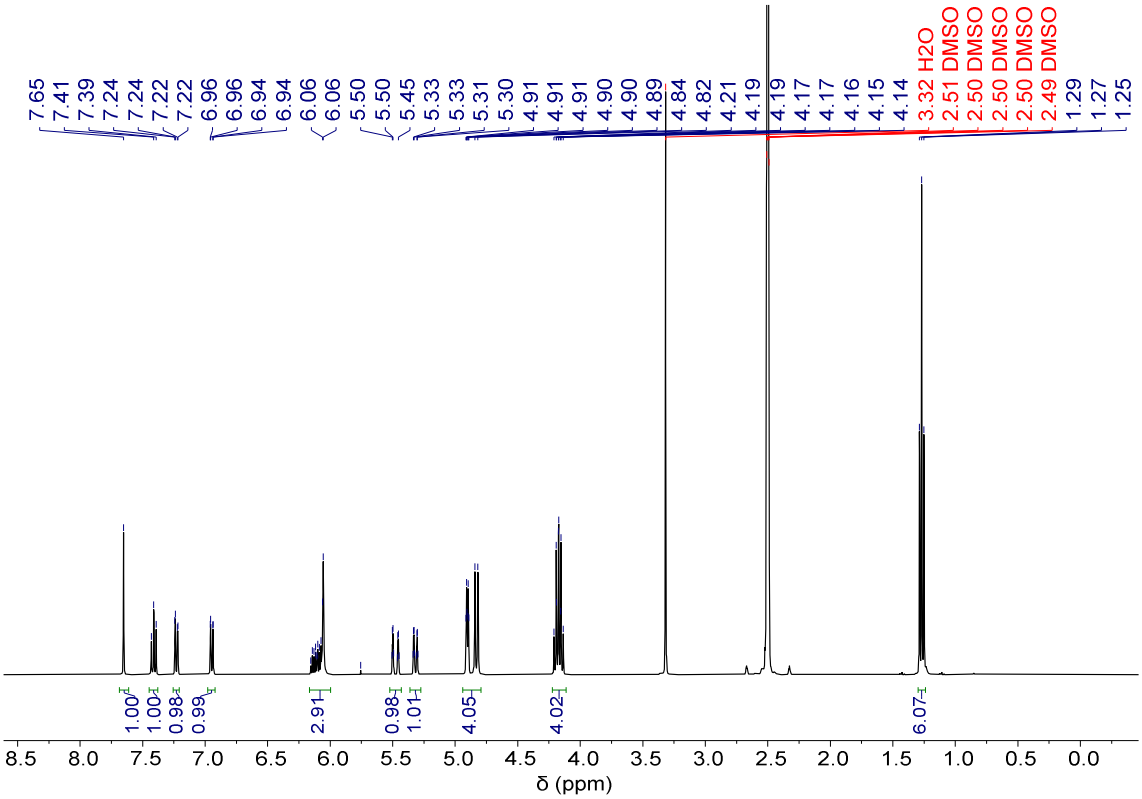


Figure S22 [1]H NMR spectrum of compound **13.**

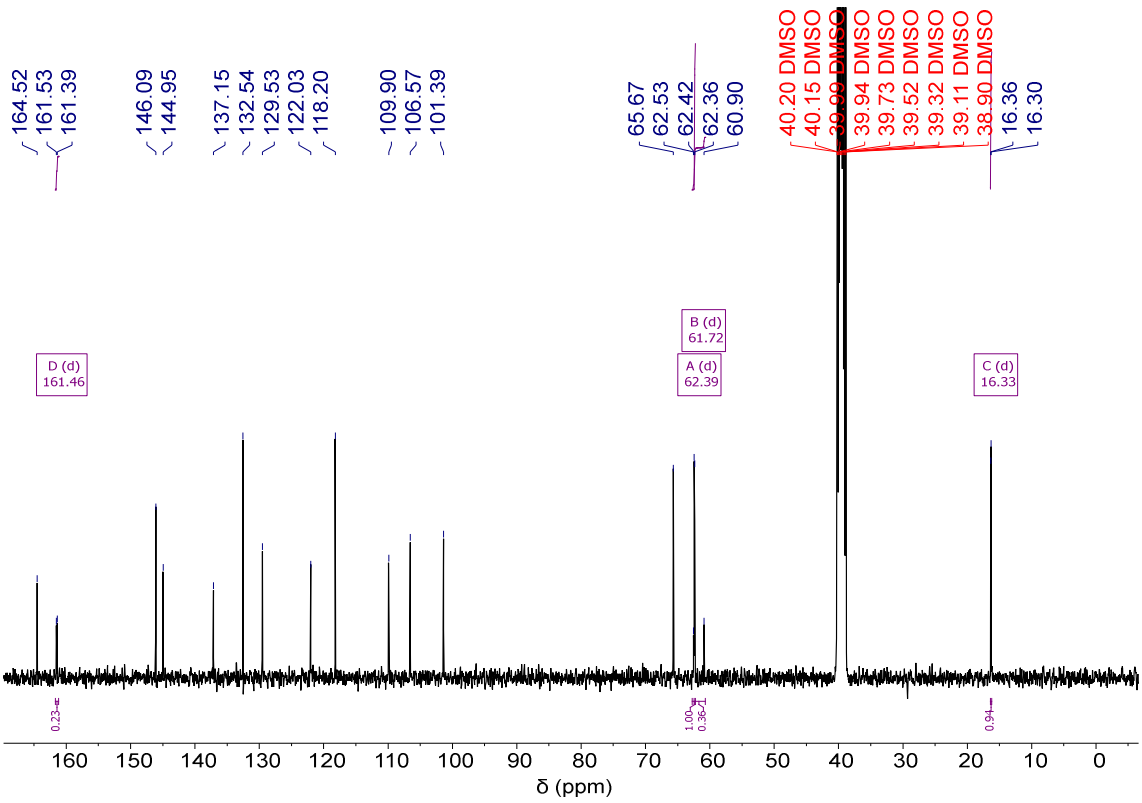


Figure S23 [13]C NMR spectrum of compound **13.**

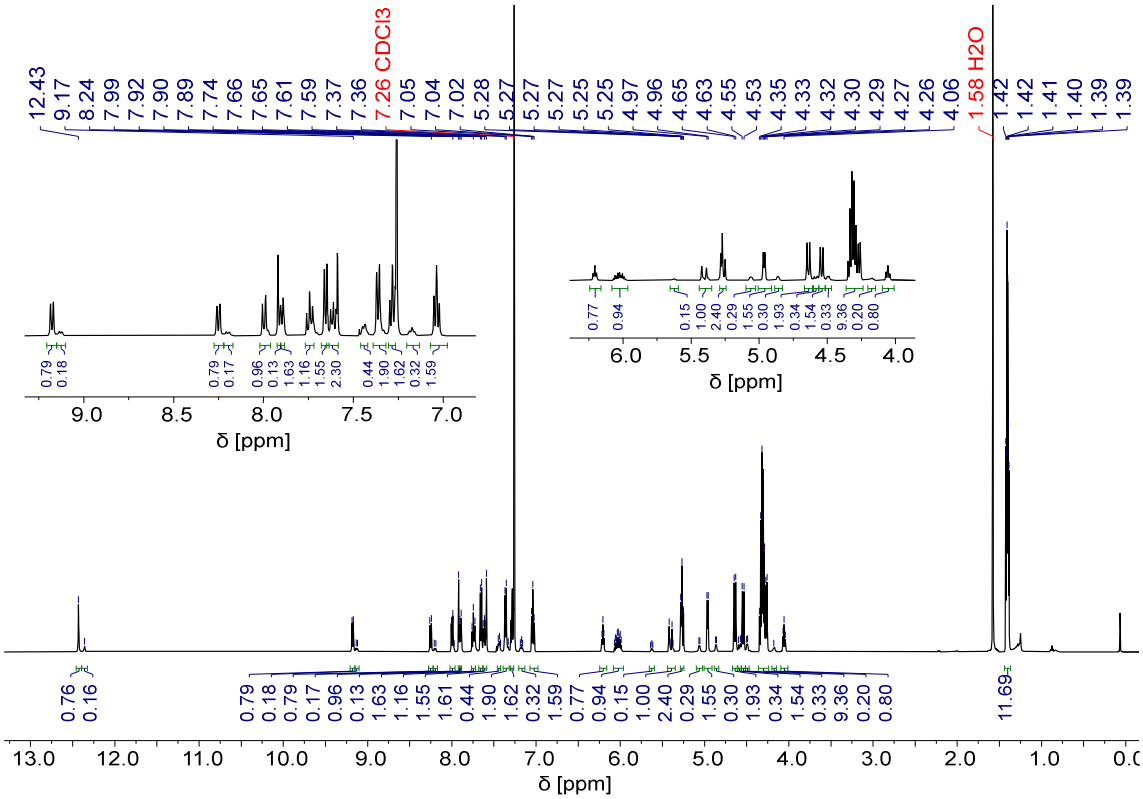


Figure S24 [1]H NMR spectrum of **14.**

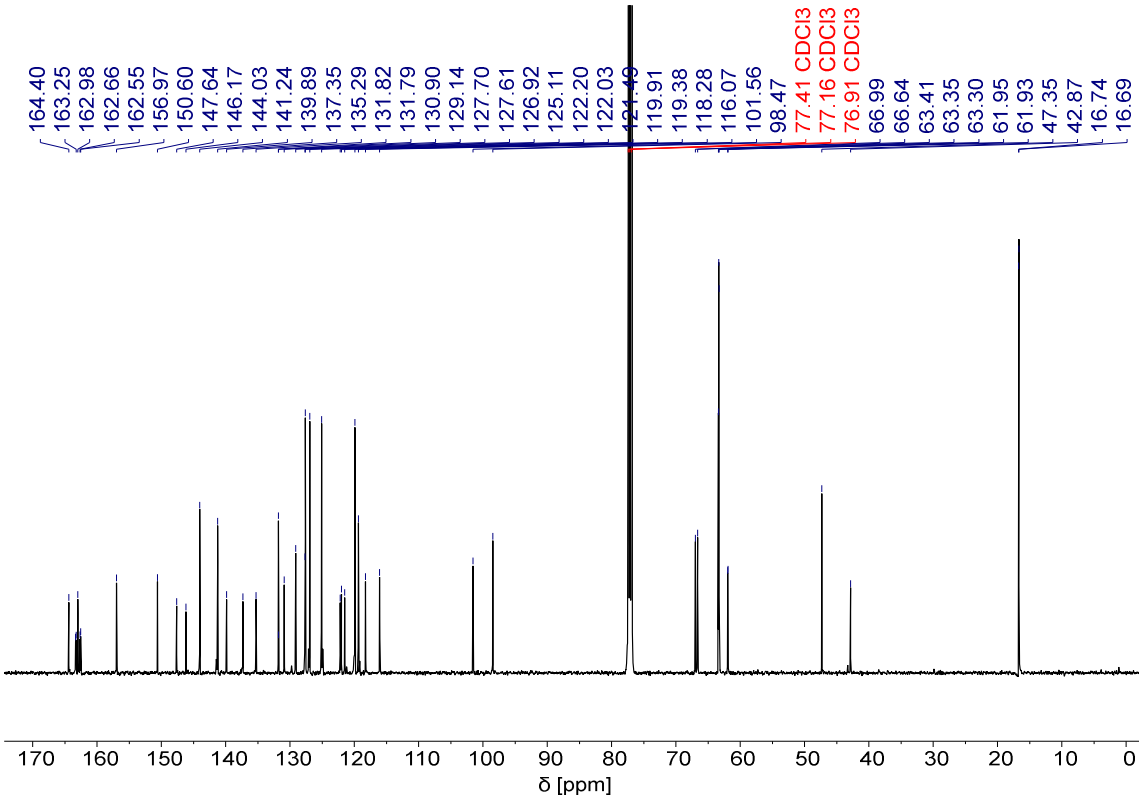


Figure S25 [13]C NMR spectrum of **14.**

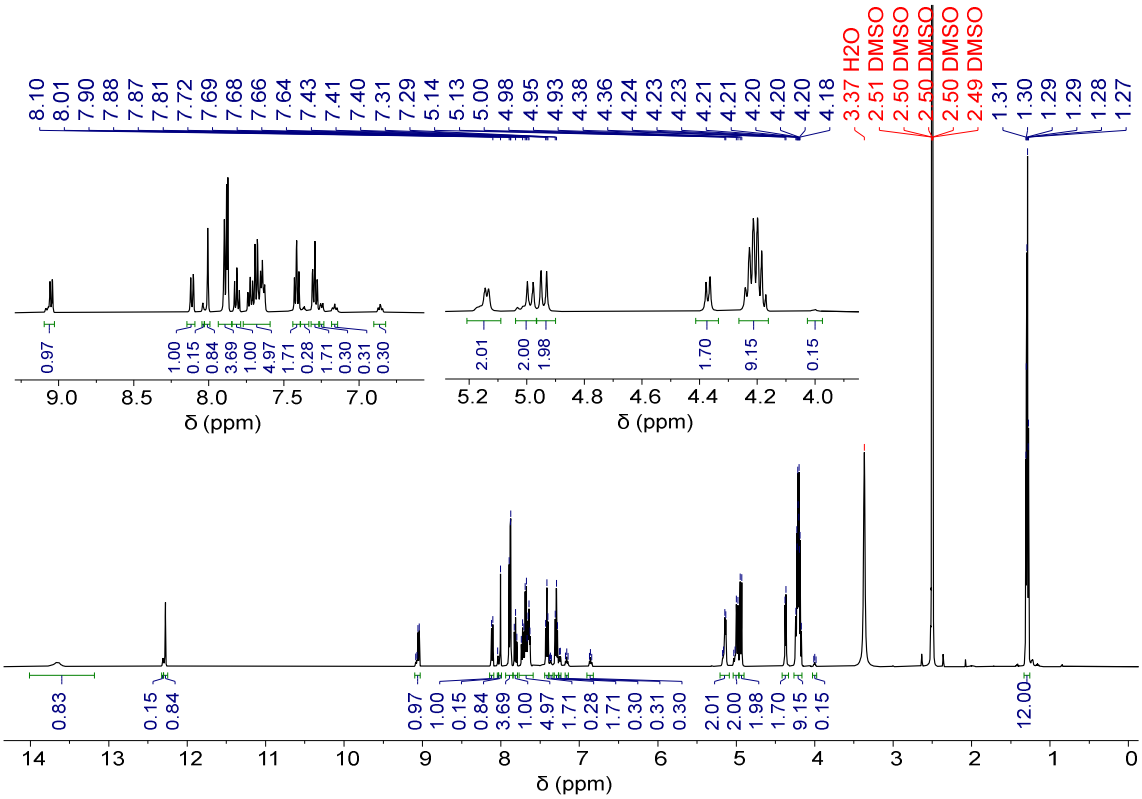


Figure S26 ^{1}H NMR spectrum of compound **7**.

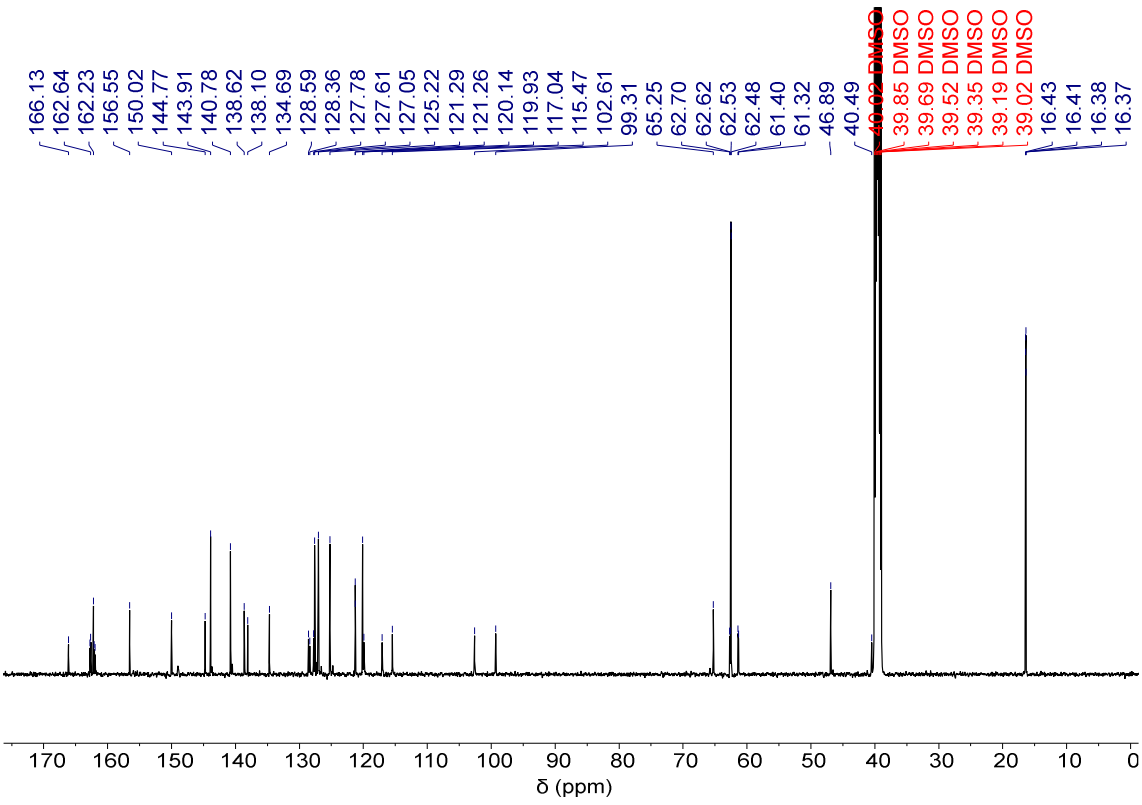


Figure S27 ^{13}C NMR spectrum of compound **7**.

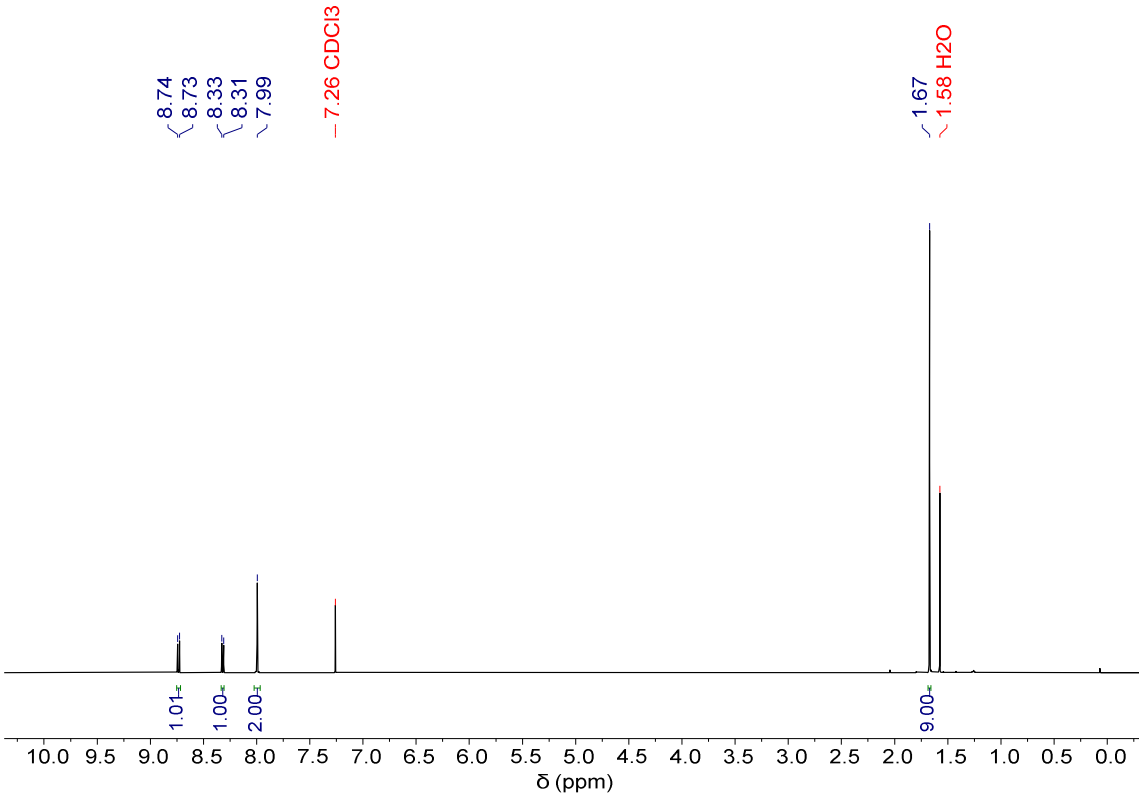


Figure S28 ^{1}H NMR spectrum of compound **16**.

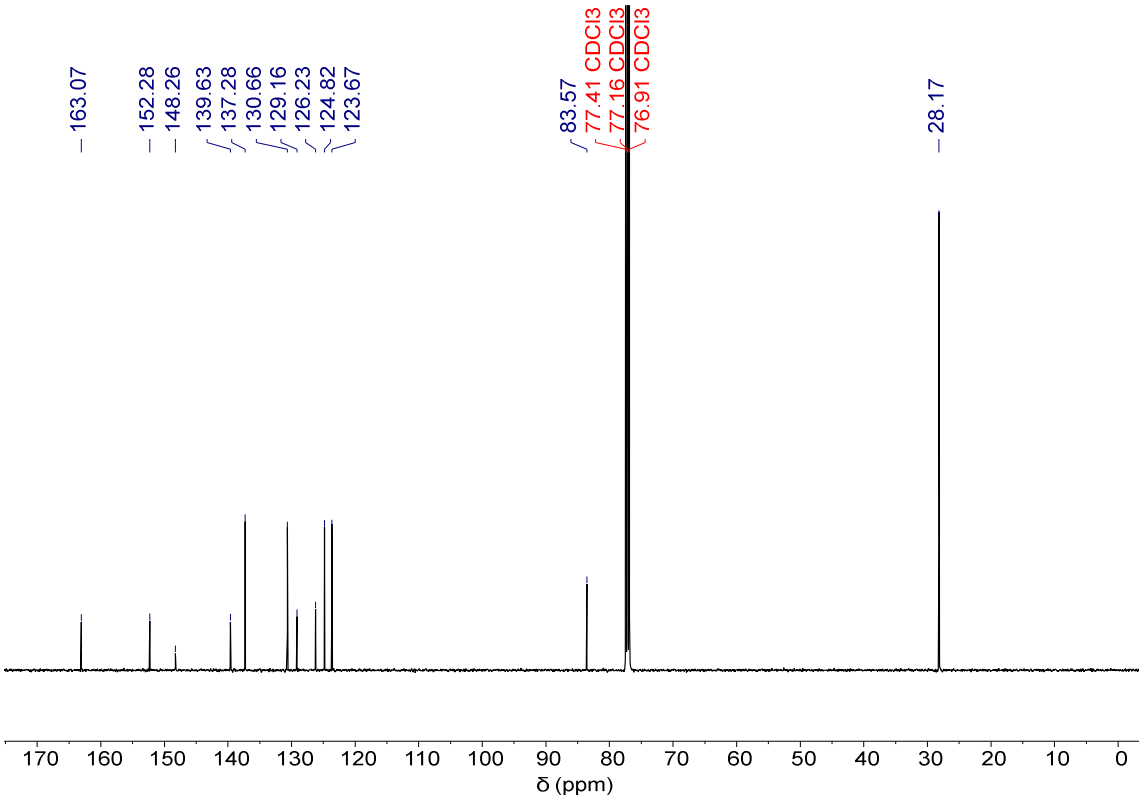


Figure S29 ^{13}C NMR spectrum of compound **16**.

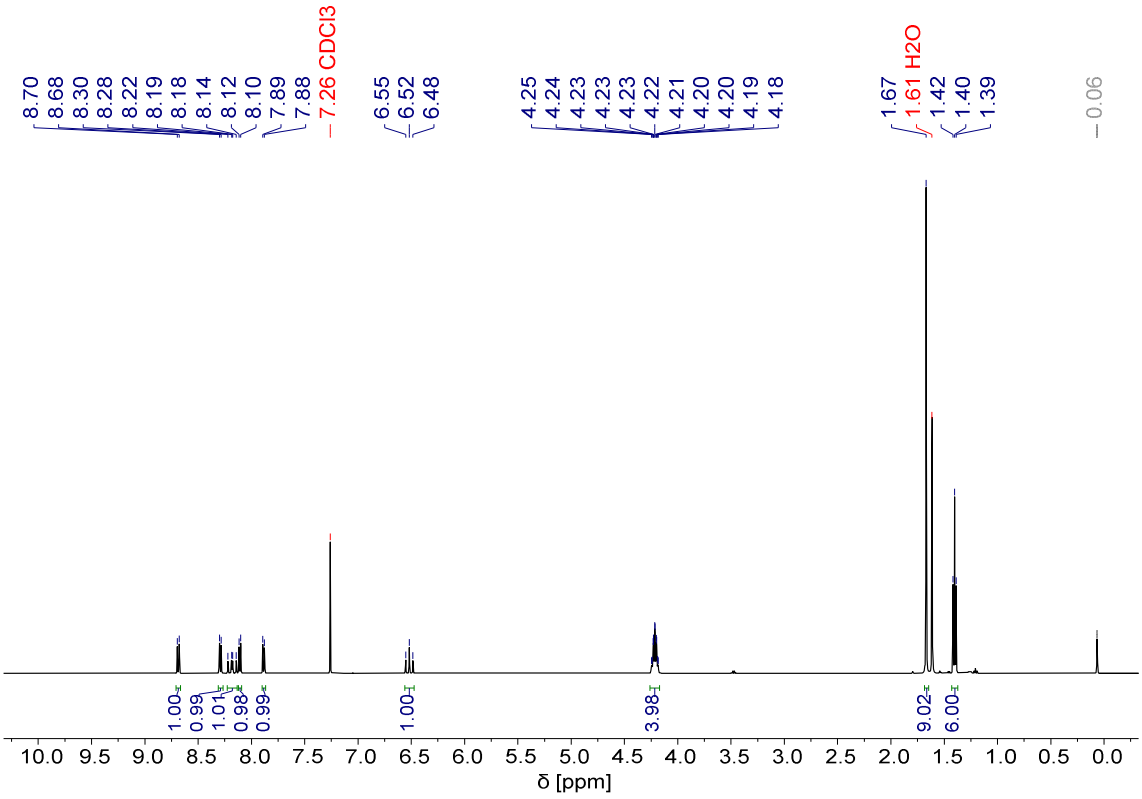


Figure S30 ^{1}H NMR spectrum of compound **17**.

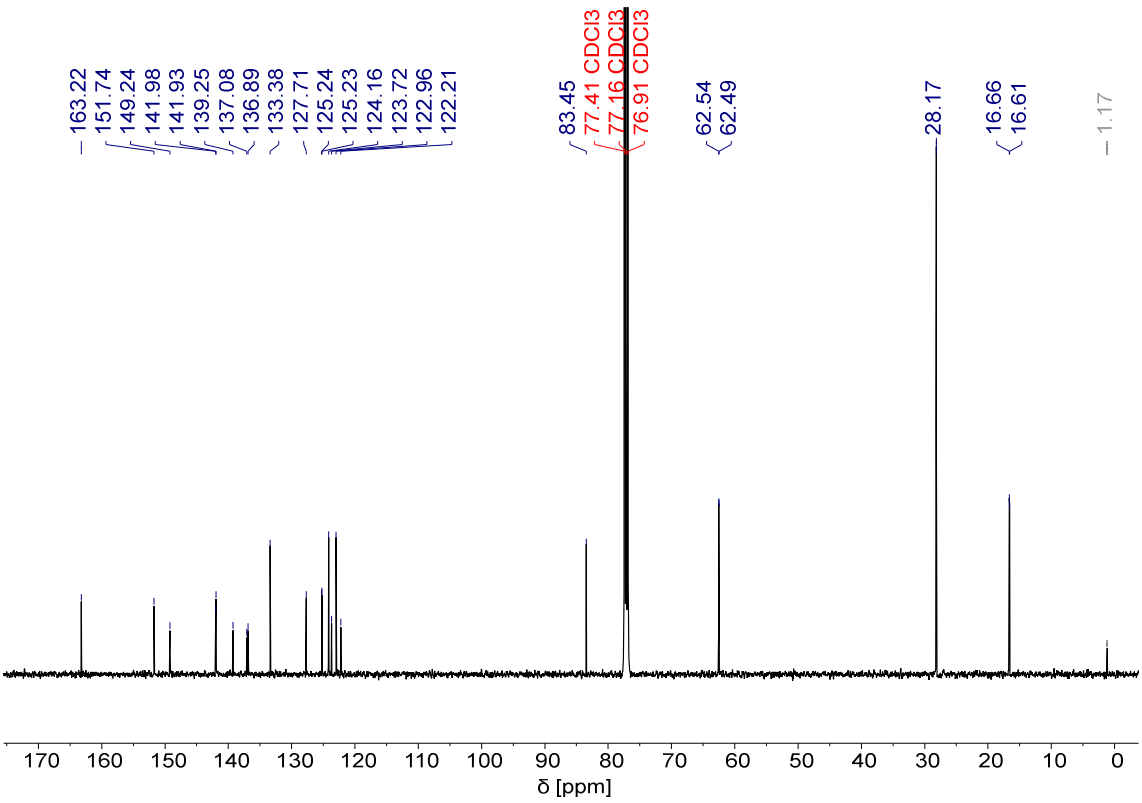


Figure S31 ^{13}C NMR spectrum of compound **17**.

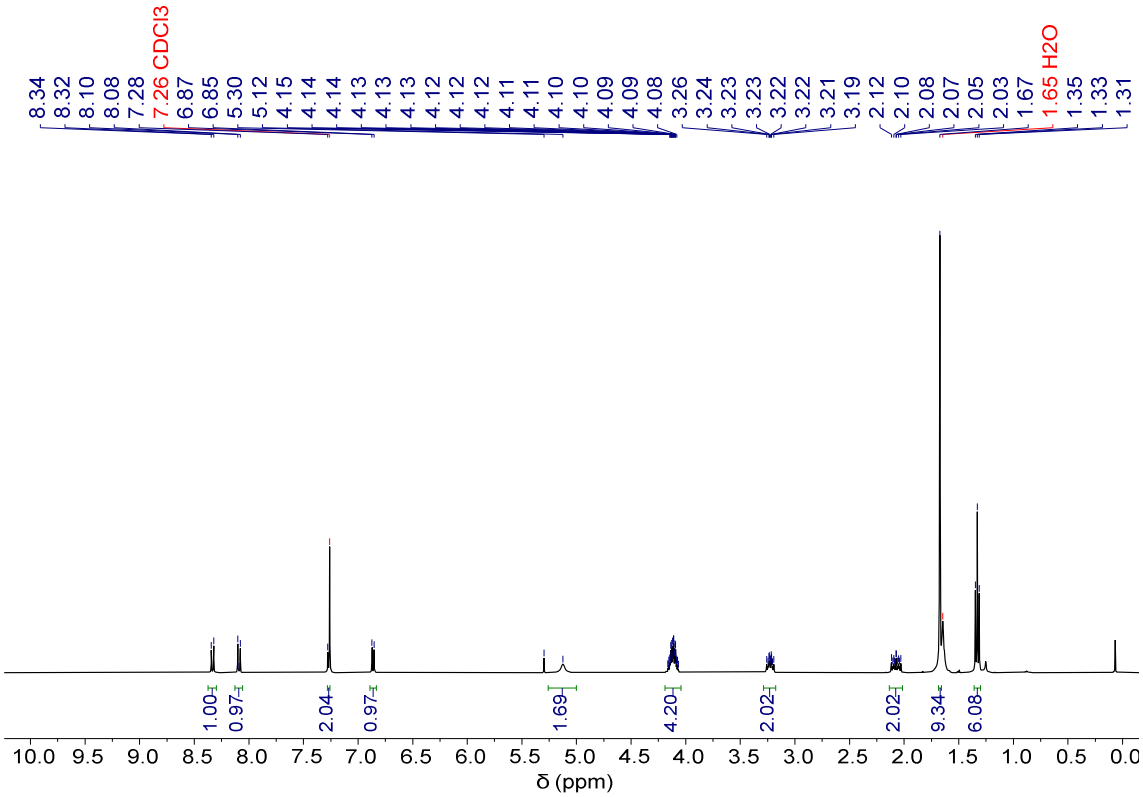


Figure S32 [1]H NMR spectrum of compound **18**.

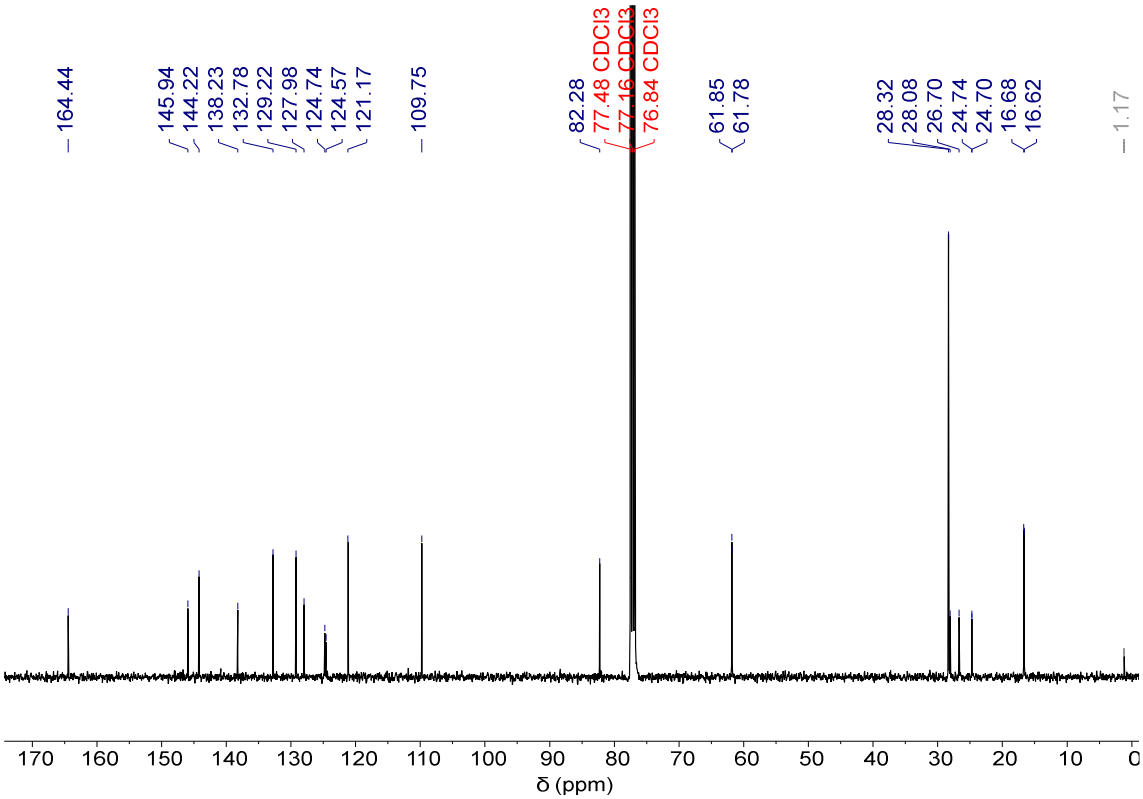


Figure S33 [13]C NMR spectrum of compound **18**.

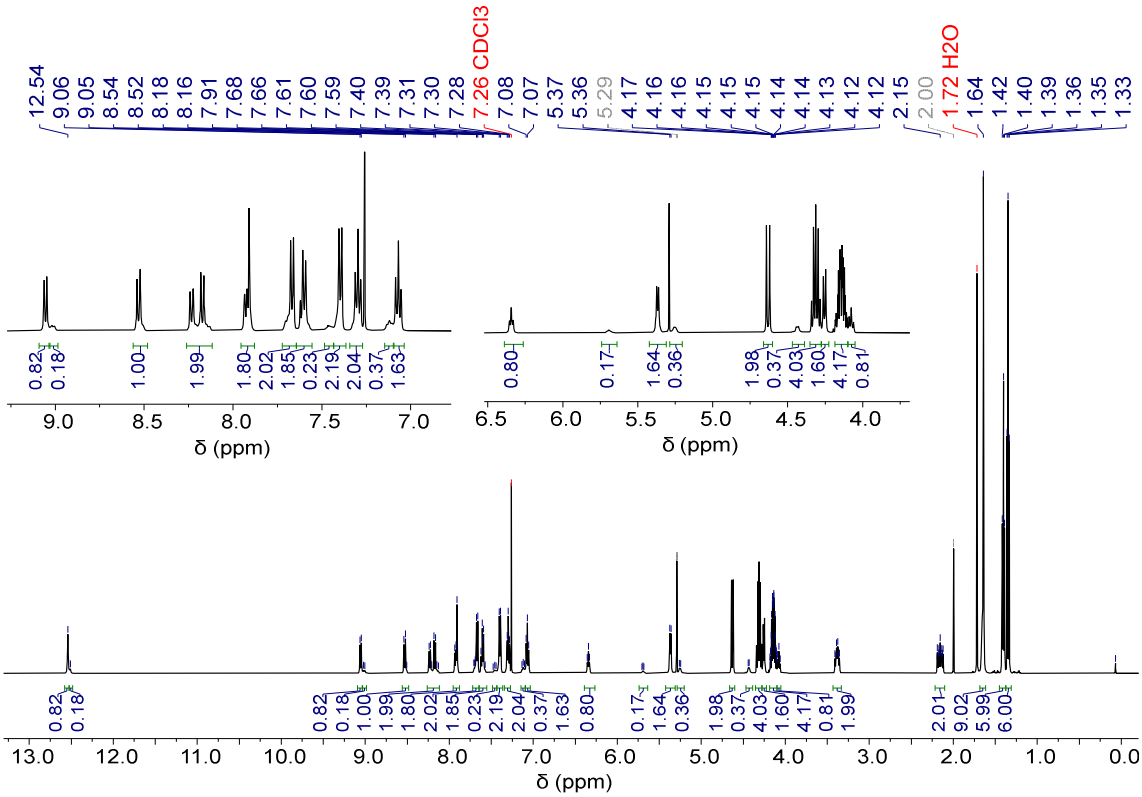


Figure S34 [1]H NMR spectrum of compound **19**.

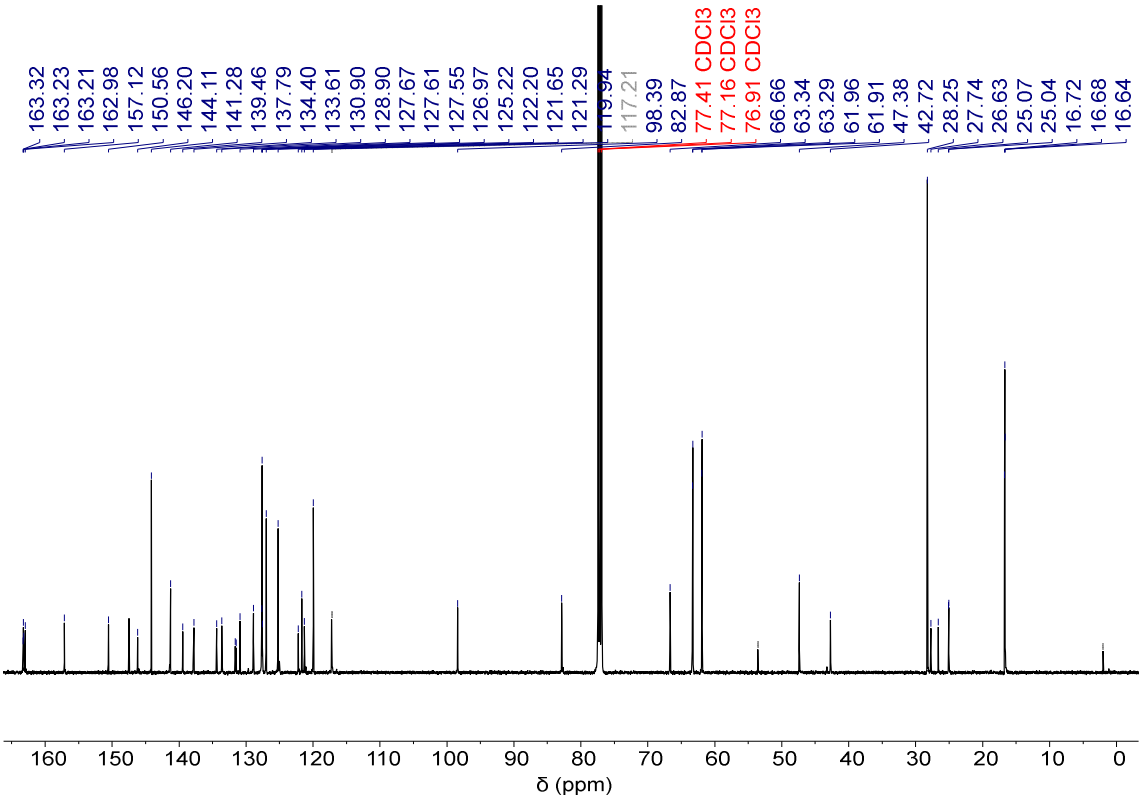


Figure S35 [13]C NMR spectrum of compound **19**.

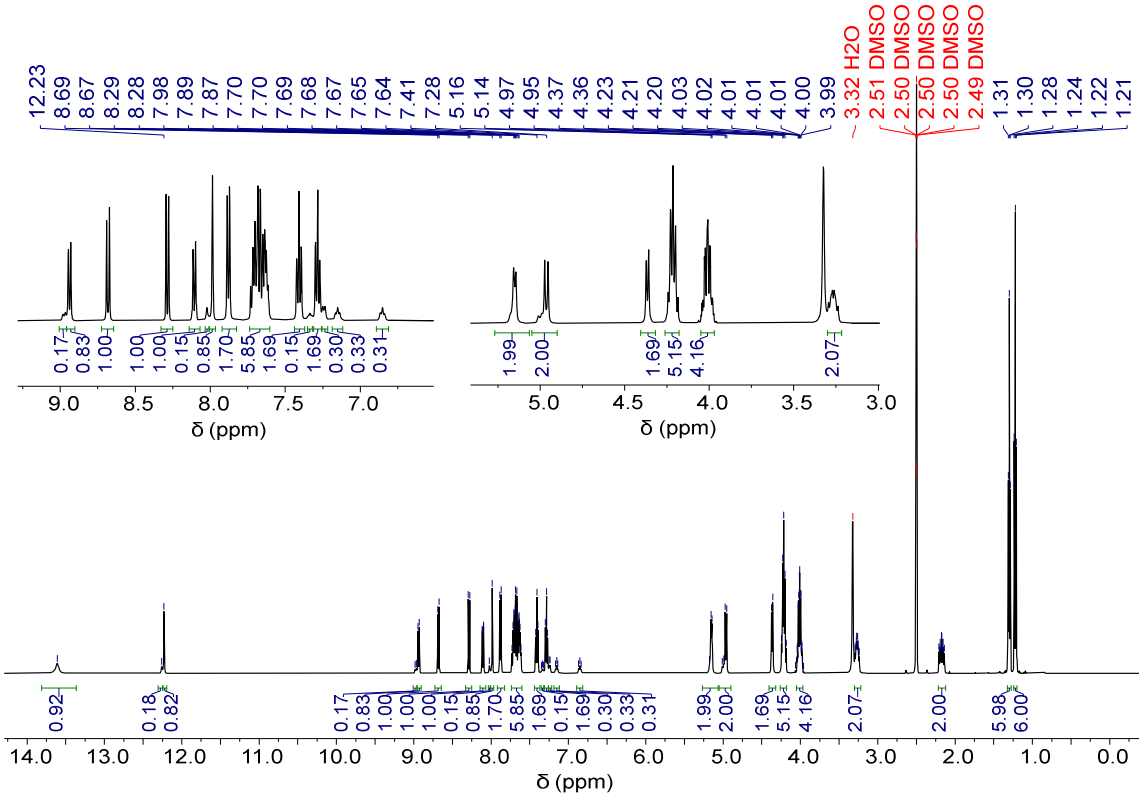


Figure S36 ^{1}H NMR spectrum of compound **8**.

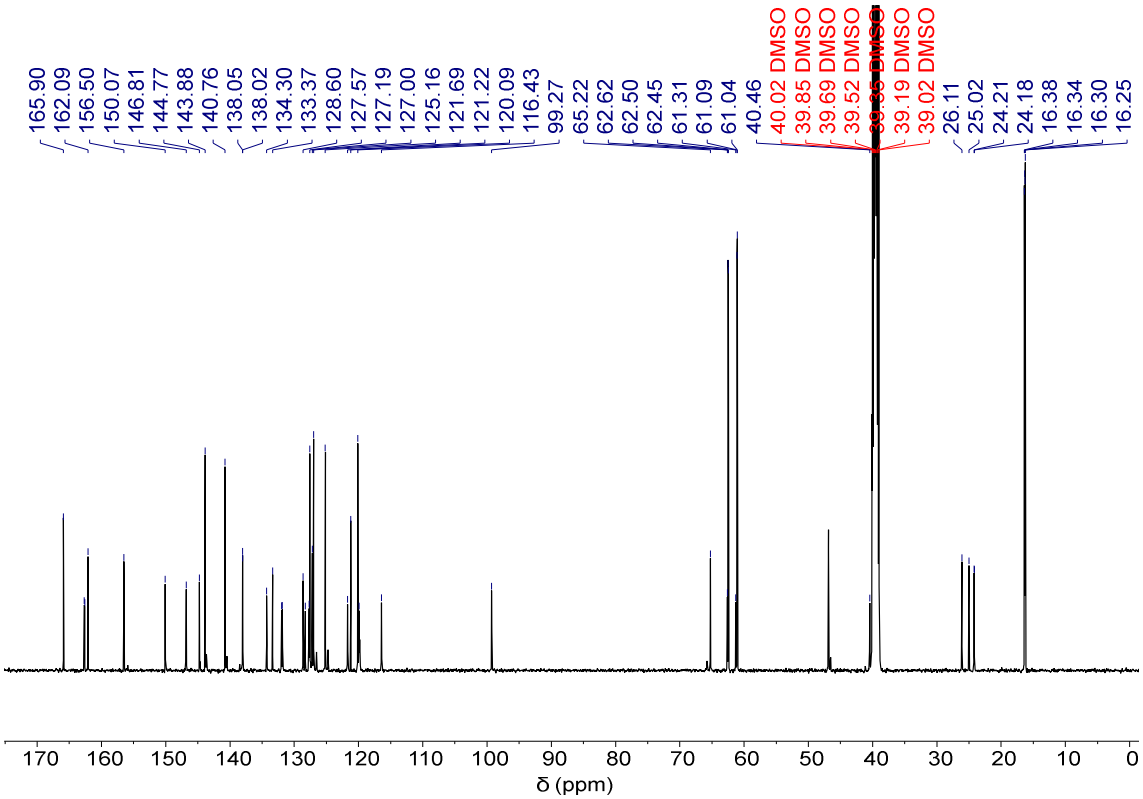
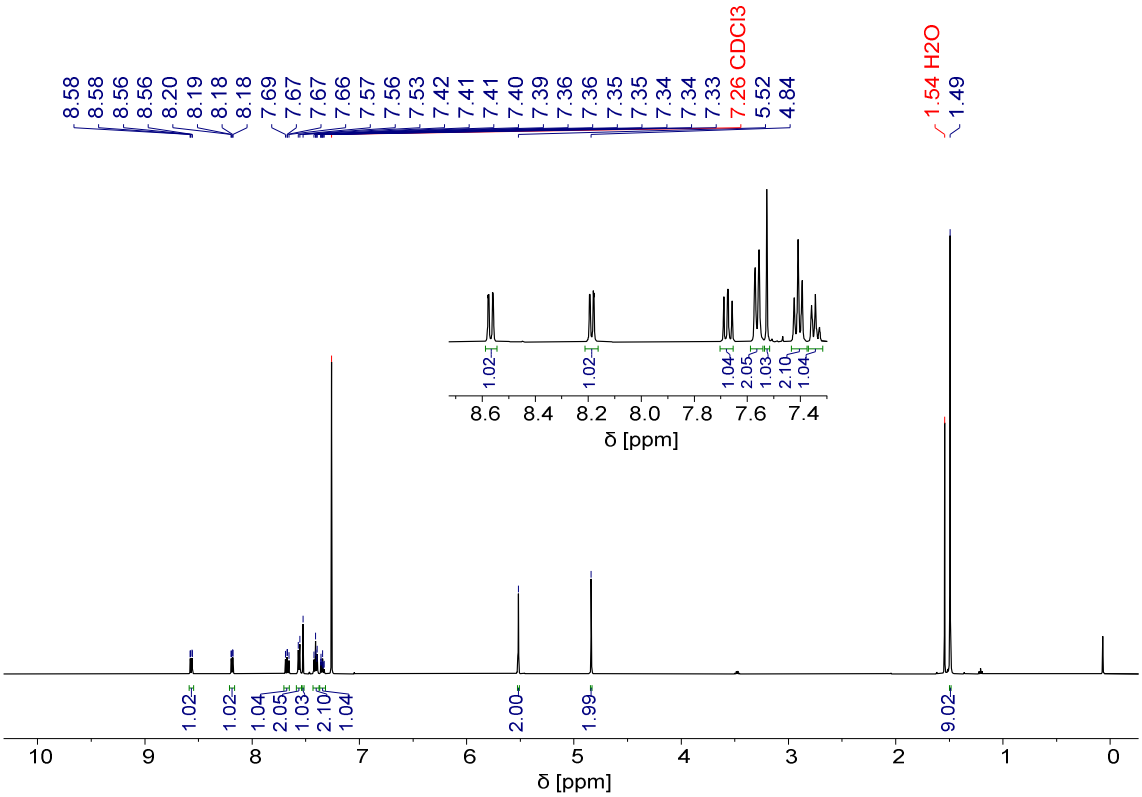


Figure S37 ^{13}C NMR spectrum of compound **8**.



Figures S38 [1]H NMR spectrum of compound **21**.

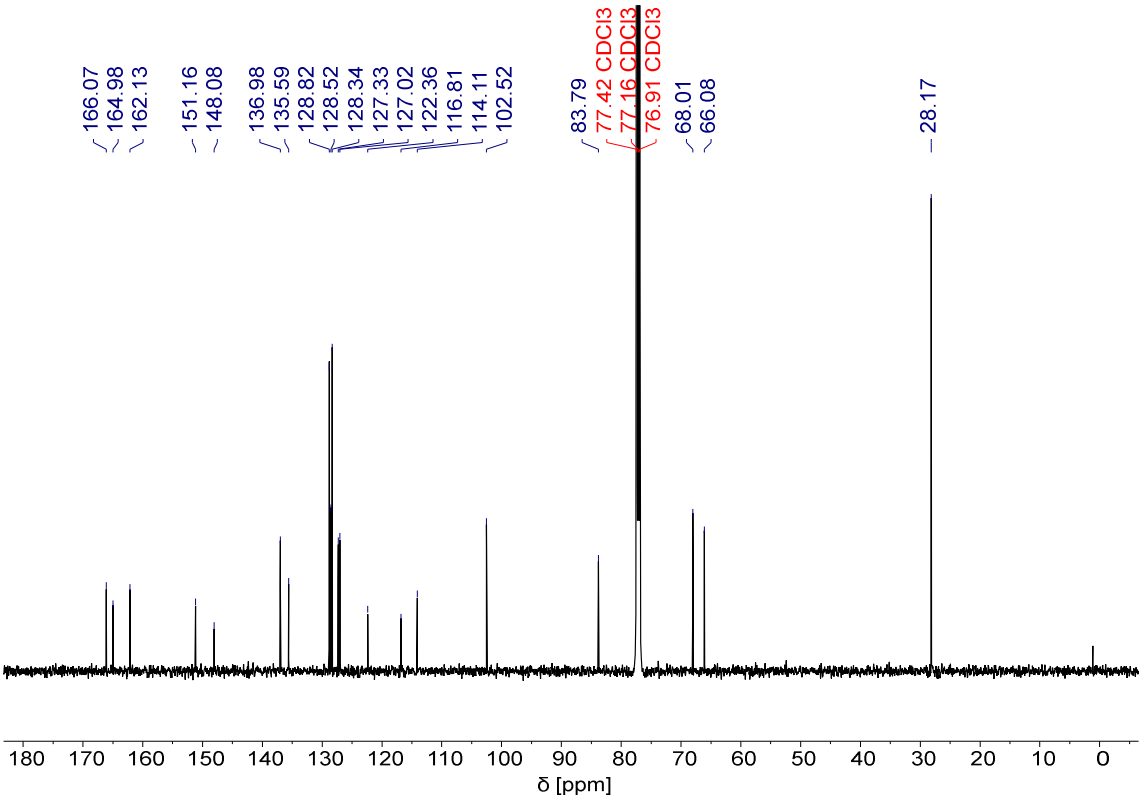


Figure S39 [13]C NMR spectrum of compound **21**.

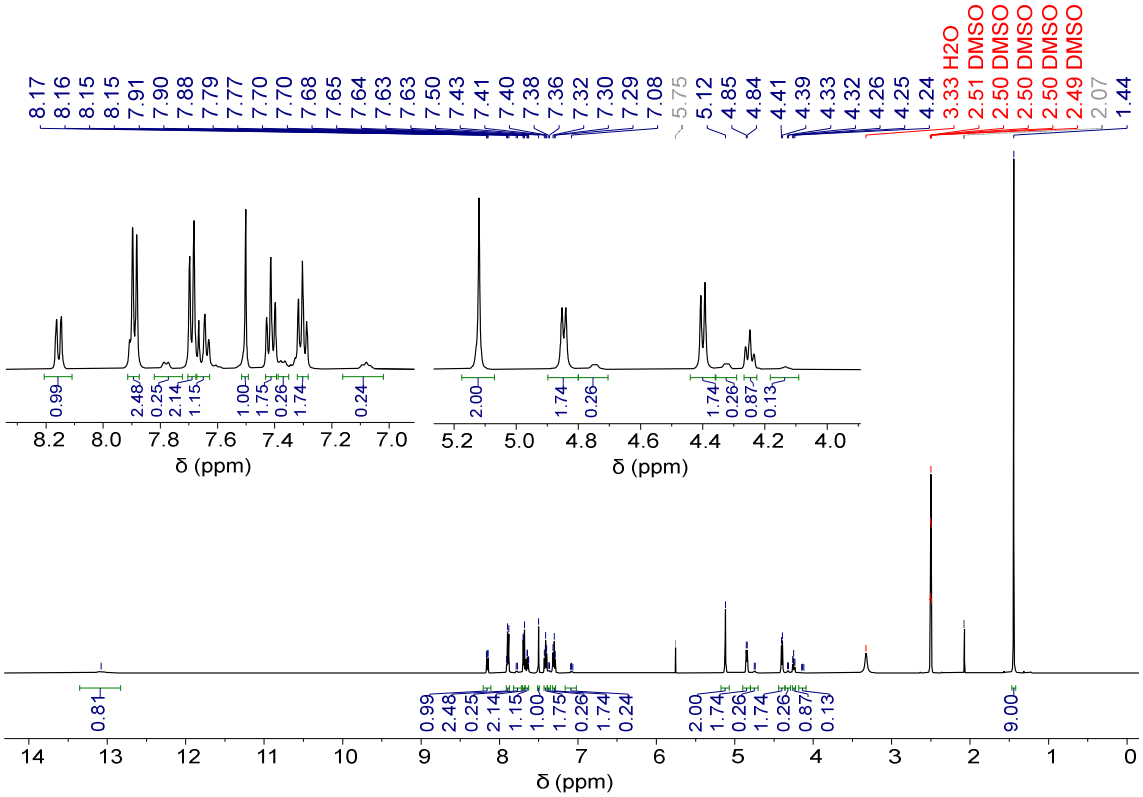


Figure S 40 ^{1}H NMR spectrum of compound **6**.

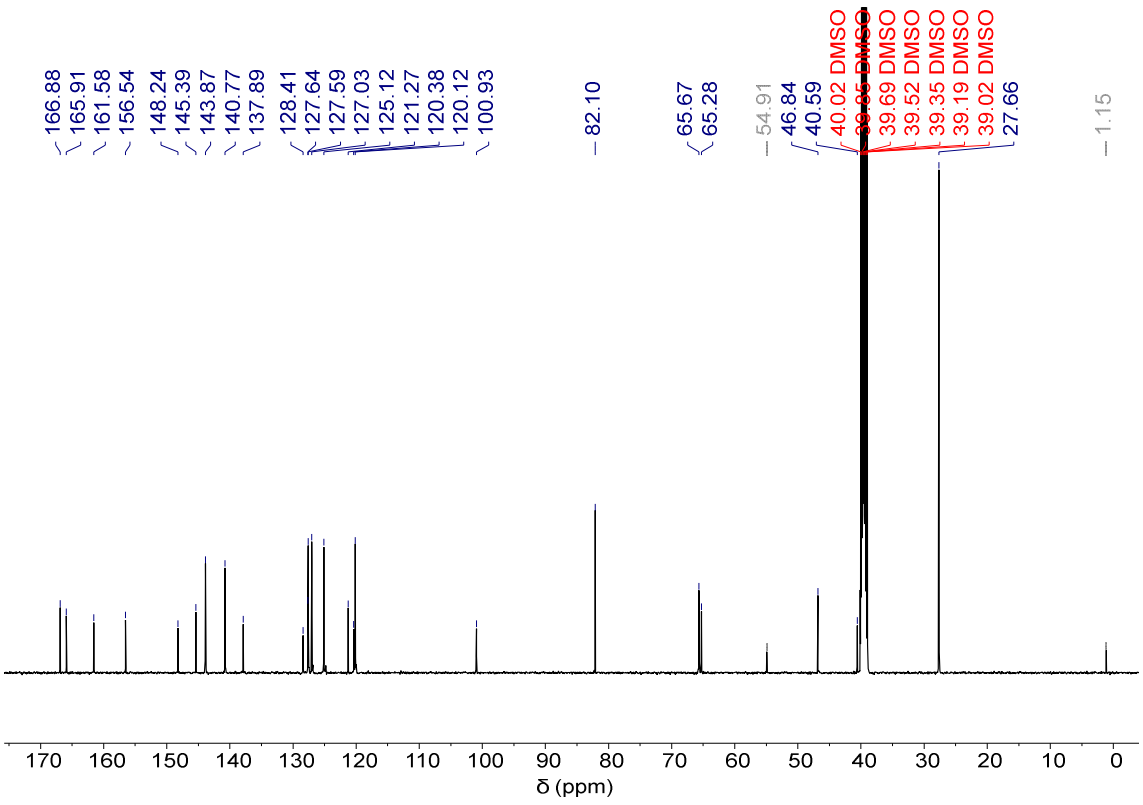


Figure S41 ^{13}C NMR spectrum of compound **6**.

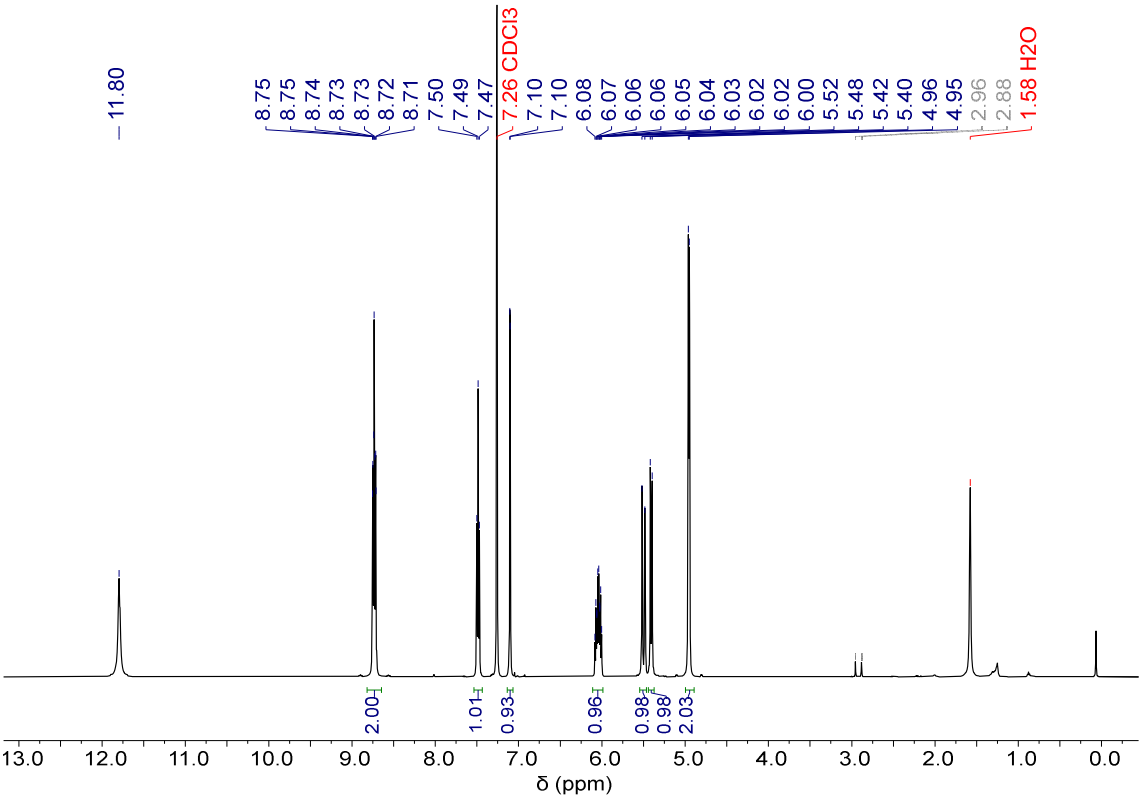


Figure S42 ^{1}H NMR spectrum of compound **23**.

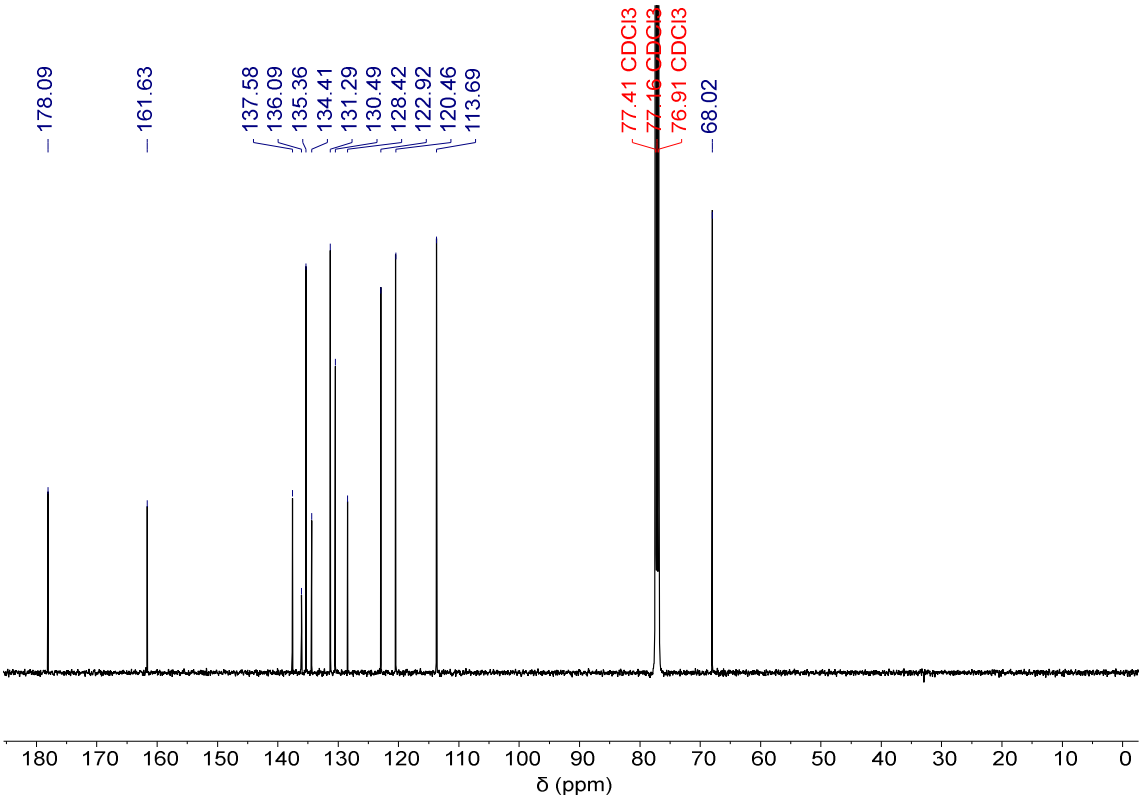


Figure S43 ^{13}C NMR spectrum of compound **23**.

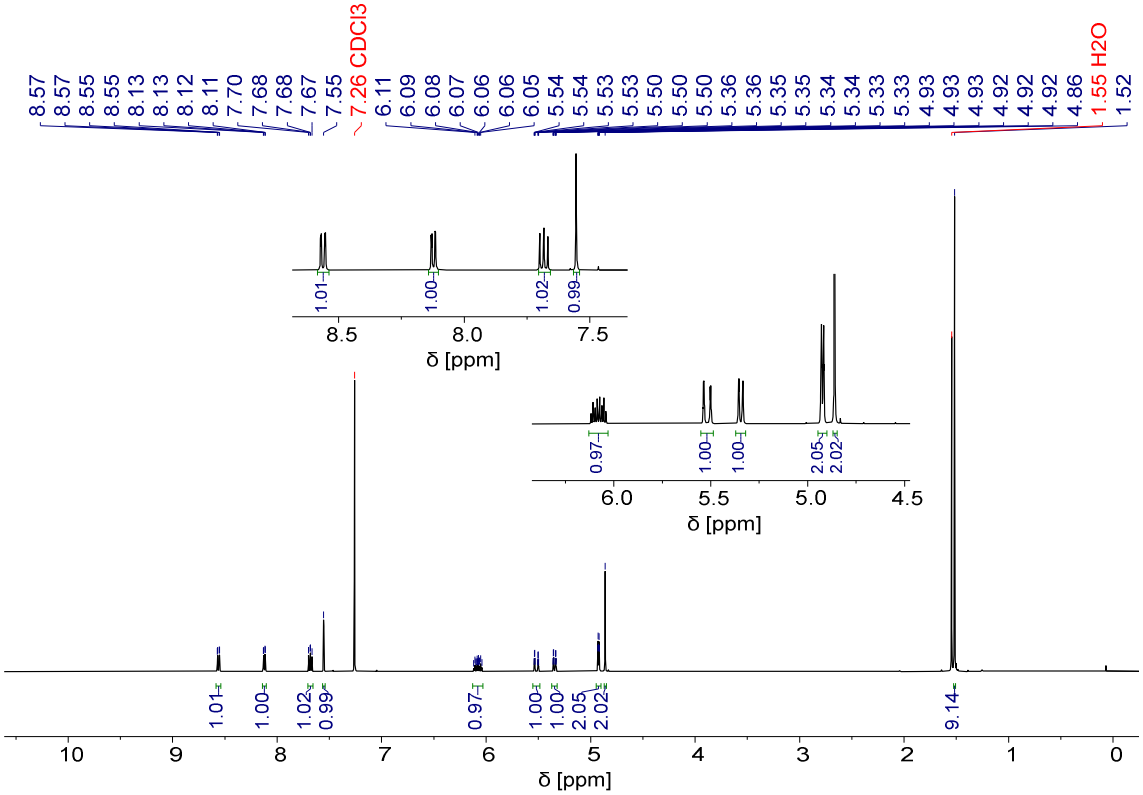


Figure S44 [1]HNMR spectrum of compound **24**.

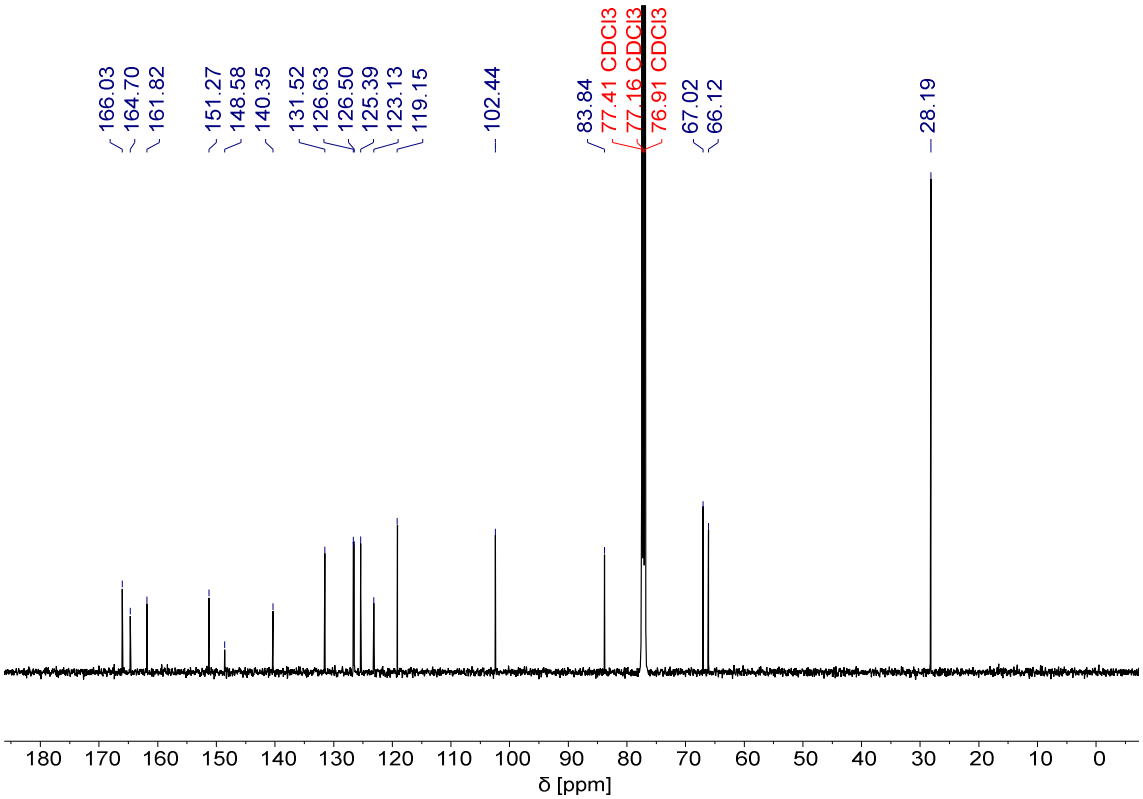


Figure S45 [13]CNMR spectrum of compound **24**.

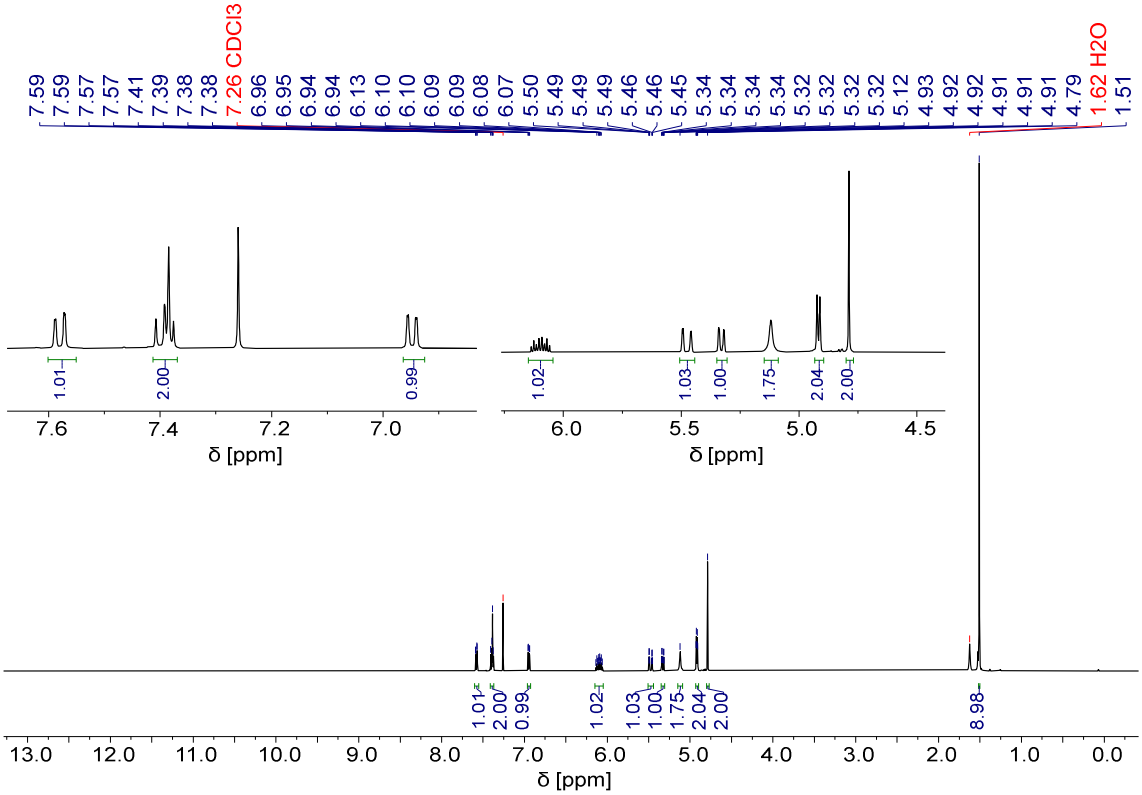


Figure S46 ^{1}HNMR spectrum of compound **25**.

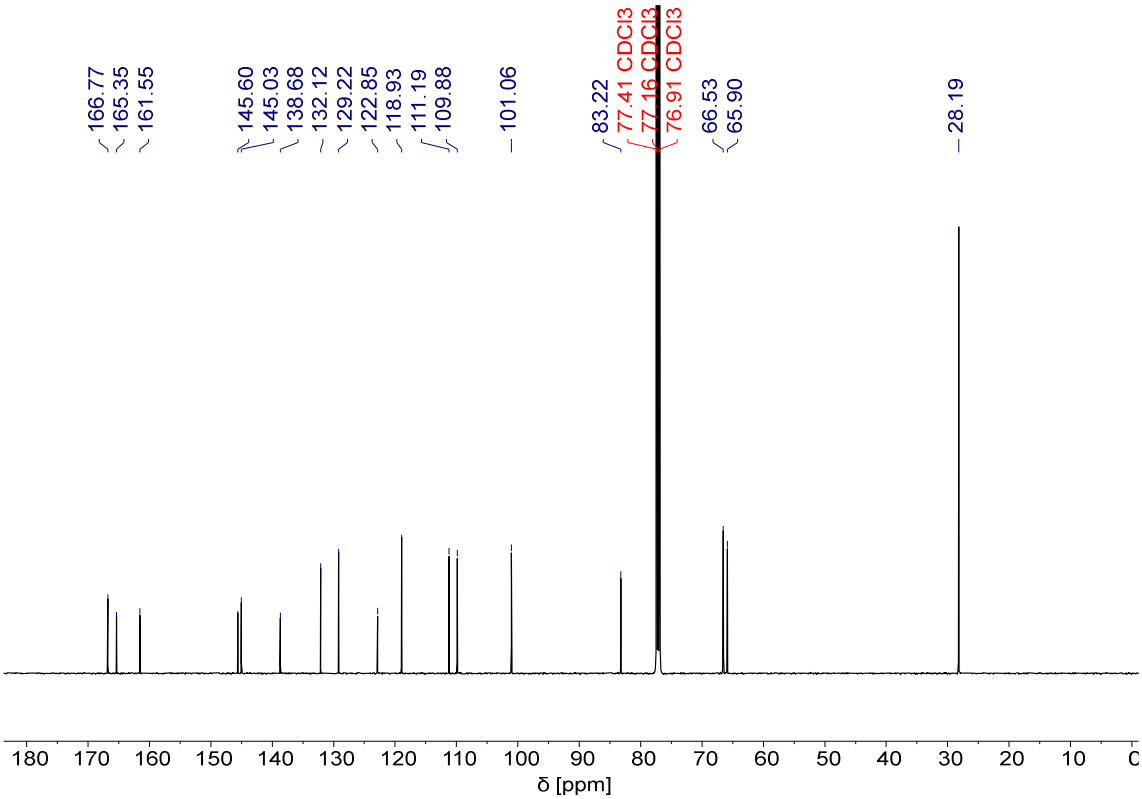


Figure S47 ^{13}CNMR spectrum of compound **25**.

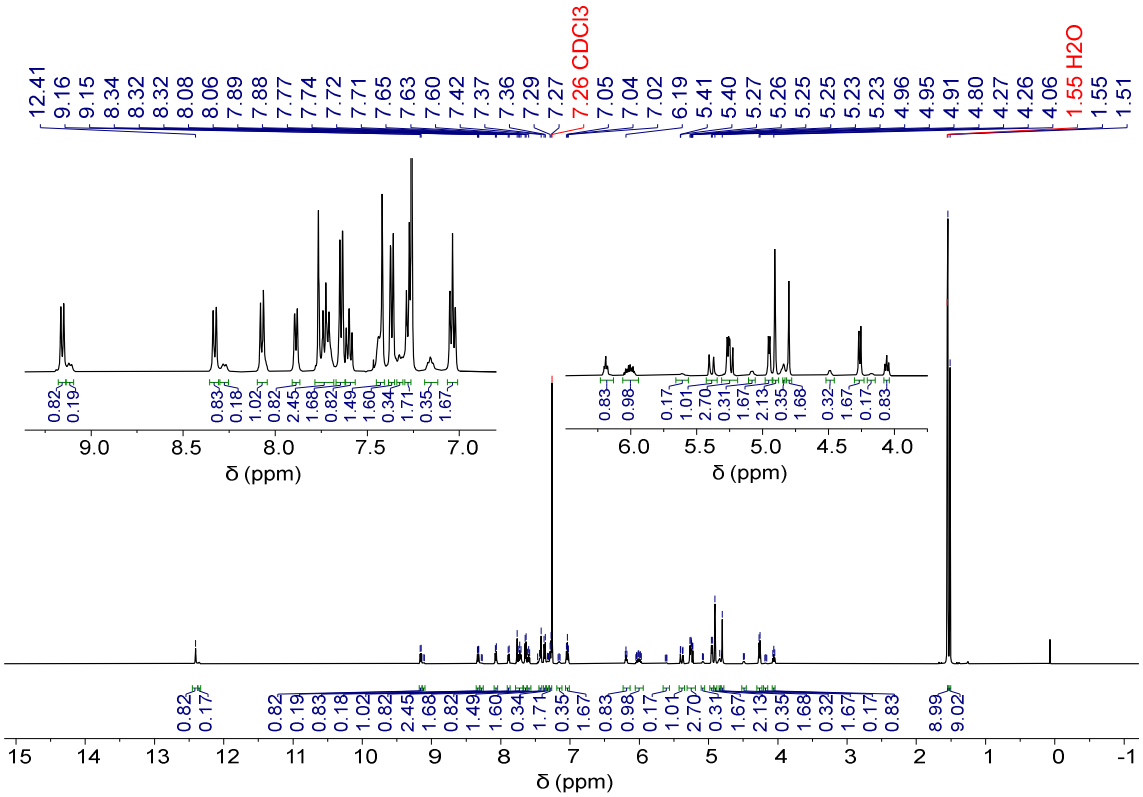


Figure S48 ^{1}HNMR spectrum of compound **26**.

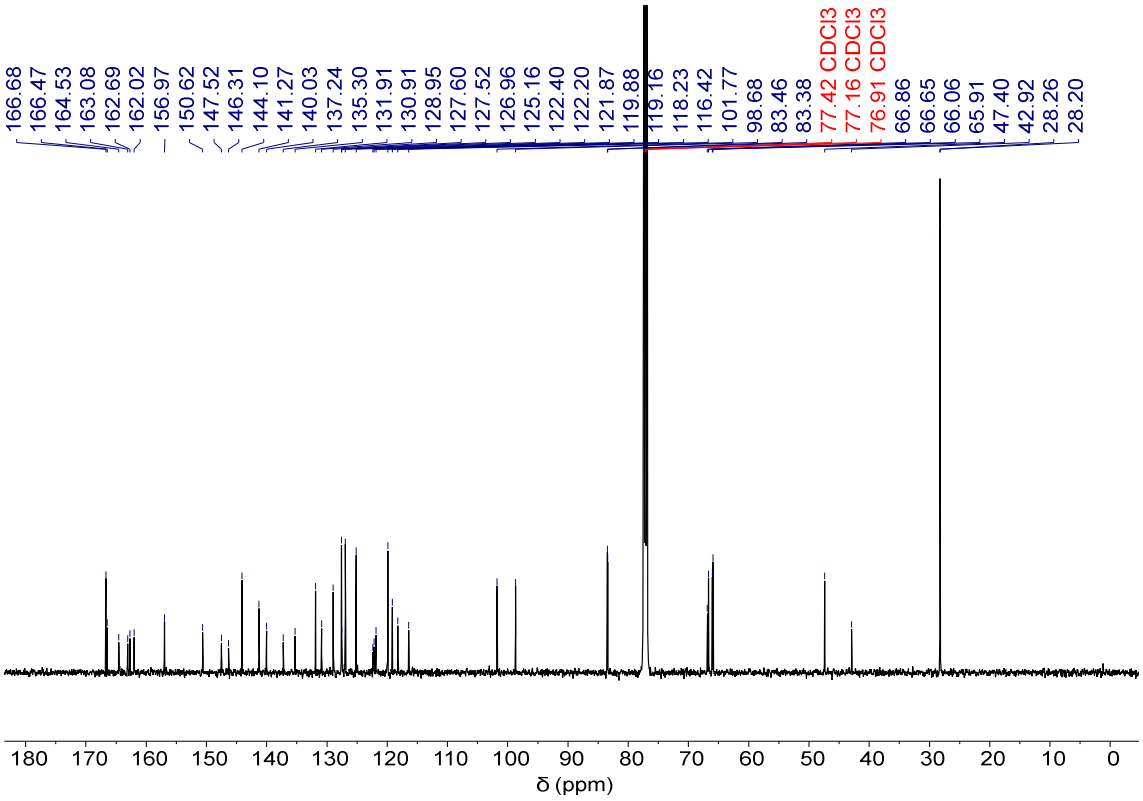


Figure S49 ^{13}CNMR spectrum of compound **26**.

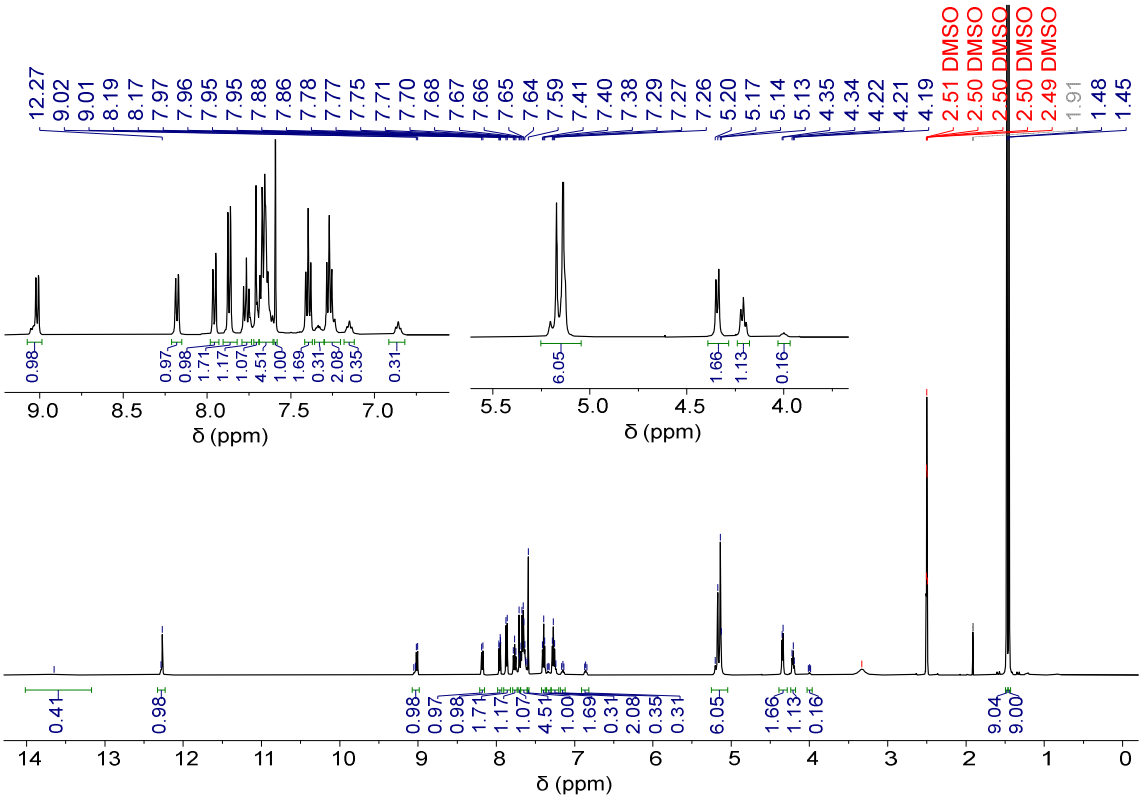


Figure S51 [1]HNMR spectrum of compound **9**.

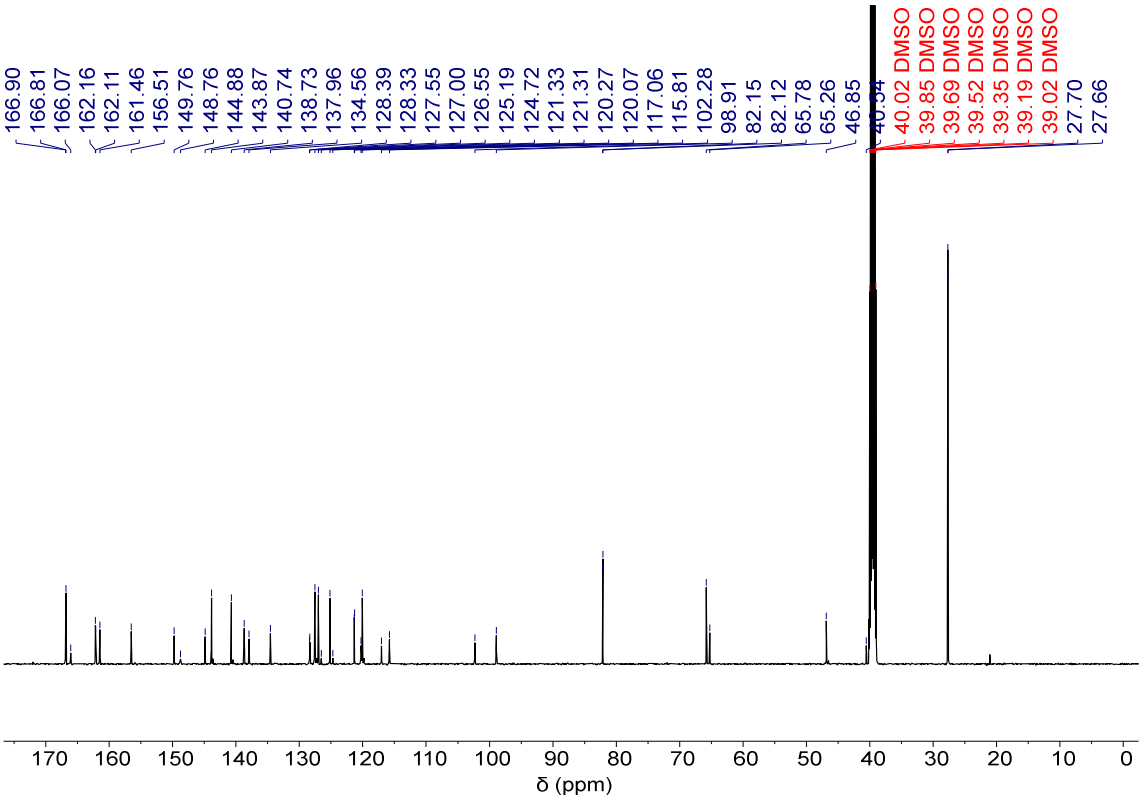


Figure S52 [13]C NMR spectrum of compound **9**.

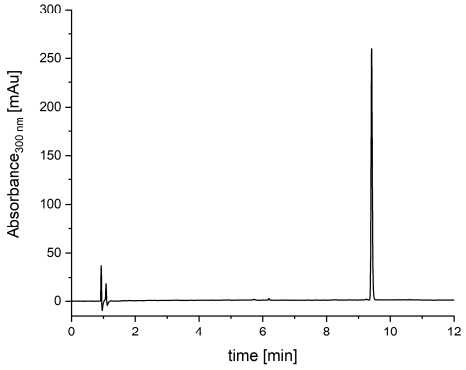
6.3.3.2 HPLC chromatograms

6.3.3.2.1 Small molecule chromatograms

Compound **14**

Column: C18 nucleodur

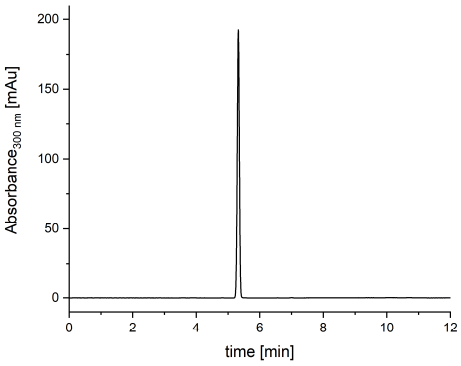
Gradient: 5-100 % B in A, TFA system



Compound **7**

Column: C18 nucleodur

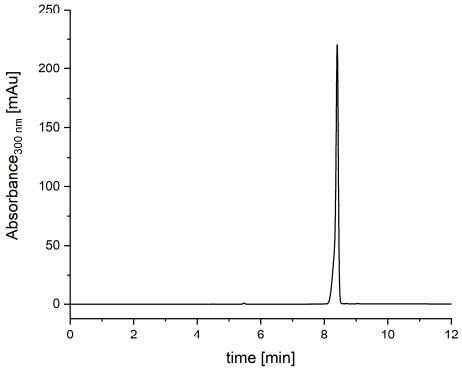
Gradient: 50-100 % B in A, TFA system



Compound **19**

Column: C18 nucleodur

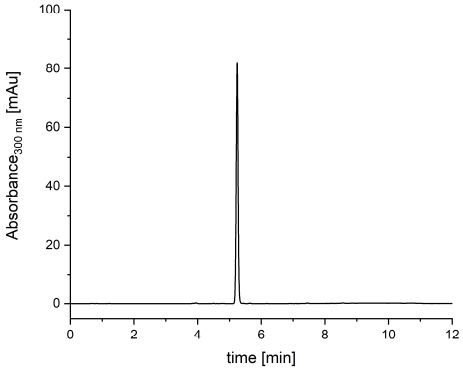
Gradient: 5-100 % B in A, TFA system



Compound **8**

Column: C18 nucleodur

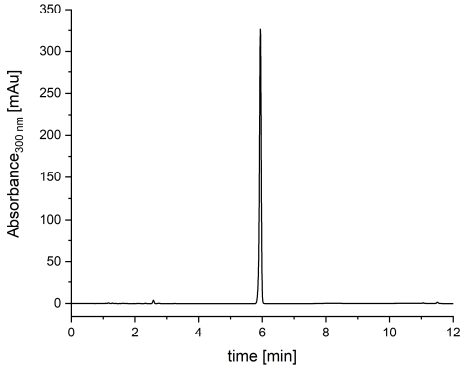
Gradient: 50-100 % B in A, TFA system



Compound **6**

Column: C18 nucleodur

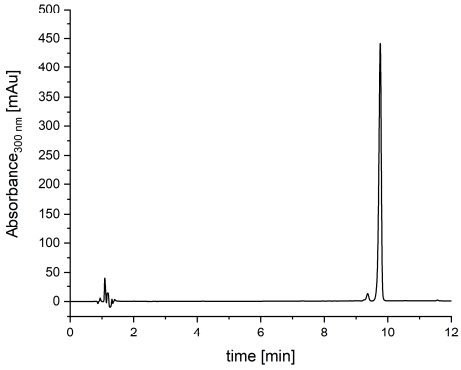
Gradient: 50-100 % B in A, TFA system



Compound **26**

Column: C18 nucleodur

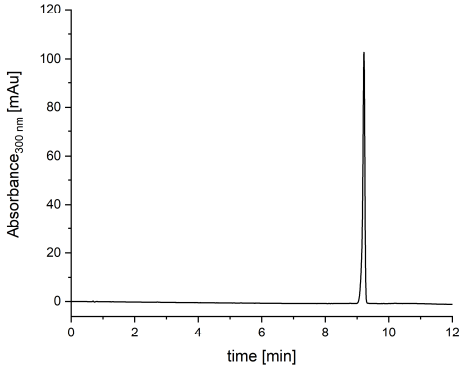
Gradient: 70-100 % B in A, TFA system



Compound **9**

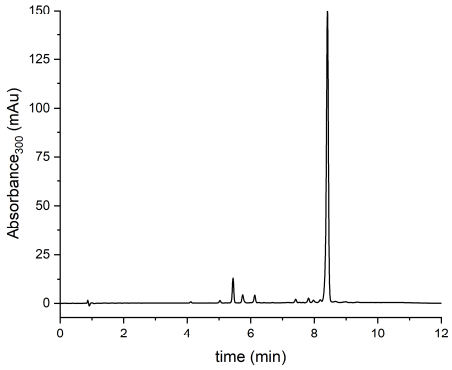
Column: C18 nucleodur

Gradient: 50-100 % B in A, TFA system

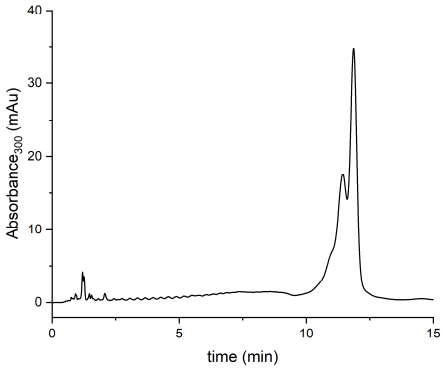


6.3.3.2.2 Foldamer chromatograms (acid-buffered HPLC)

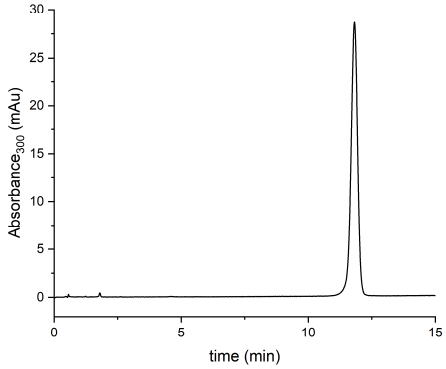
Compound **32** hexamer intermediate
Column: C18 nucleodur
Gradient: 30-100 % B in A, 0.1% TFA system



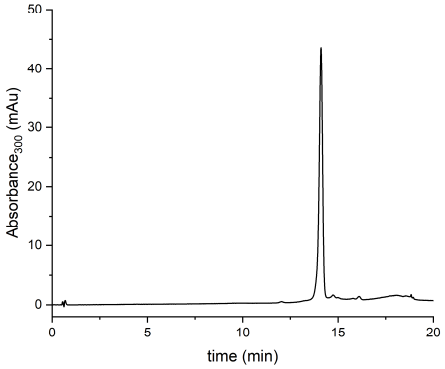
Compound **32** crude product
Column: C8 gravity
Gradient: 55-75 % B in A, TFA system



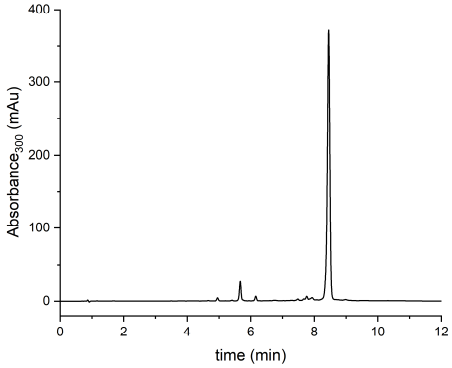
Compound **32** purified product
Column: C8 gravity
Gradient: 55-75 % B in A, TFA system



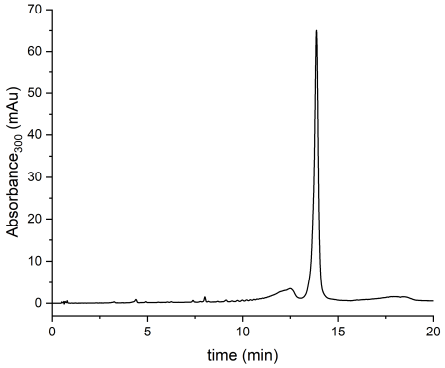
Compound **33**
Column: C8 gravity
Gradient: 30-70 % B in A, TFA system



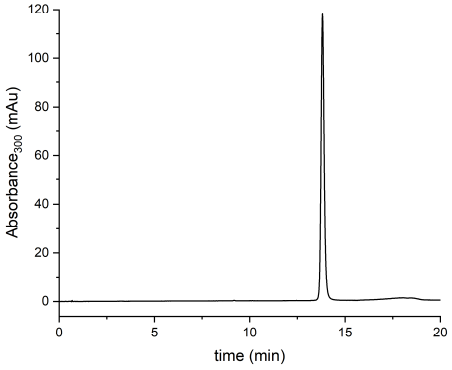
Compound **28** hexamer intermediate
Column: C18 nucleodur
Gradient: 30-100 % B in A, TFA system



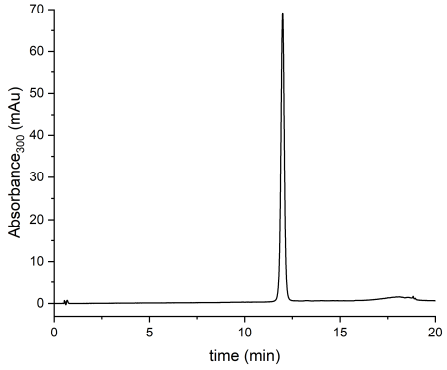
Compound **28** crude product
Column: C8 gravity
Gradient: 30-70 % B in A, TFA system



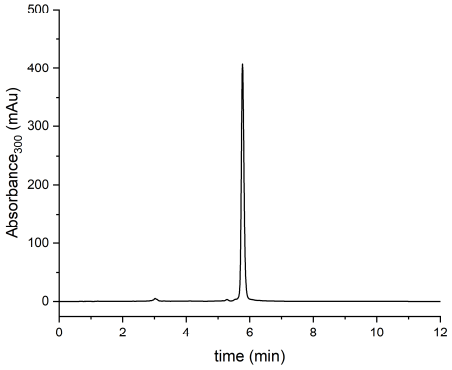
Compound **28** purified product
Column: C8 gravity
Gradient: 30-70 % B in A, TFA system



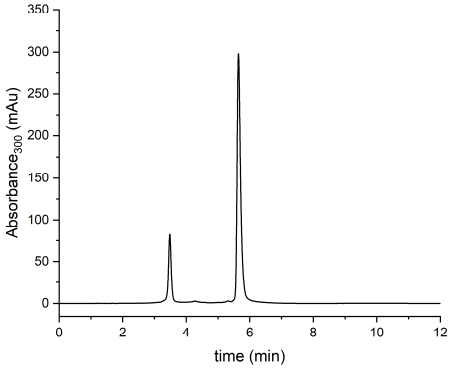
Compound **29** purified product
Column: C8 gravity
Gradient: 30-70 % B in A, TFA system



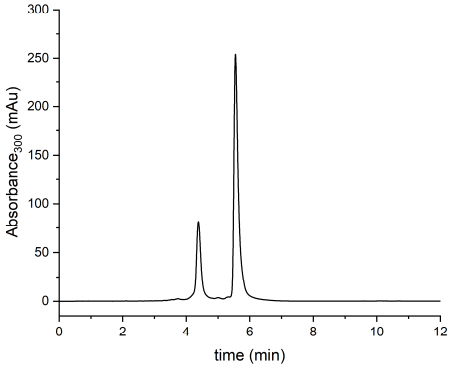
Compound **27** 32-unit intermediate
Column: C18 nucleodur
Gradient: 50-100 % B in A, TFA system



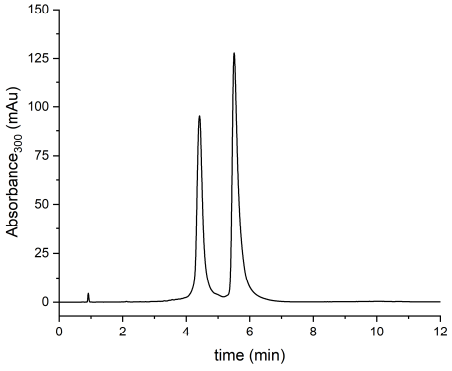
Compound **27** 48-unit intermediate
Column: C18 nucleodur
Gradient: 50-100 % B in A, TFA system



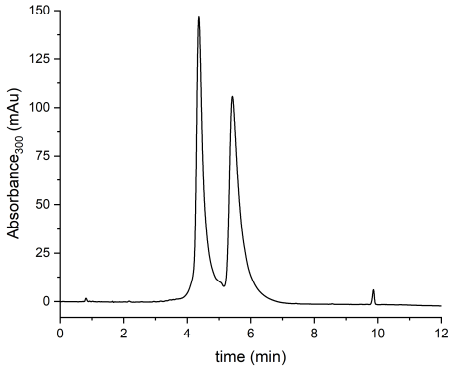
Compound **27** 64-unit intermediate
Column: C18 nucleodur
Gradient: 50-100 % B in A, TFA system



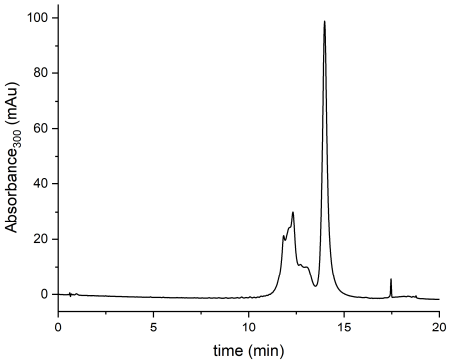
Compound **27** 80-unit intermediate
Column: C18 nucleodur
Gradient: 50-100 % B in A, TFA system



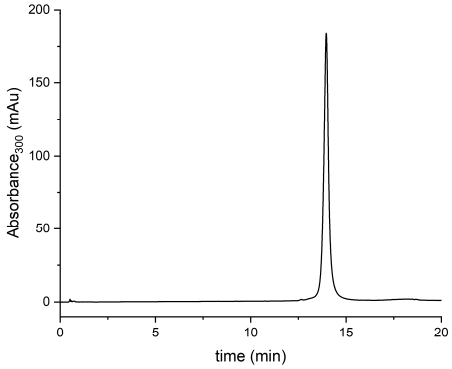
Compound **27** 100-unit crude product
Column: C18 nucleodur
Gradient: 50-100 % B in A, TFA system



Compound **27** 100-unit crude product
Column: C8 gravity
Gradient: 30-70 % B in A, TFA system

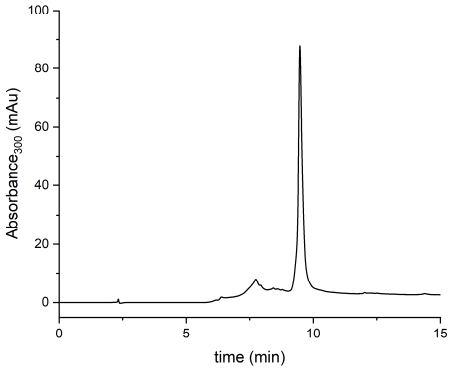


Compound **27** 100-unit purified product
Column: C8 gravity
Gradient: 30-70 % B in A, TFA system

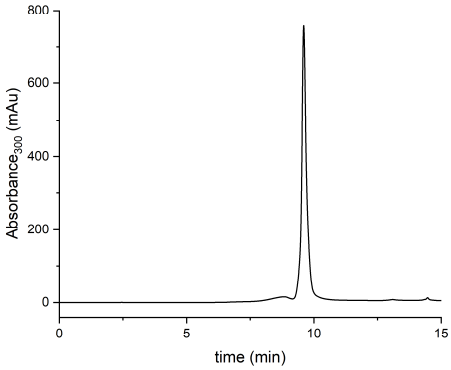


6.3.3.2.3 Foldamer chromatograms (basic-buffered HPLC)

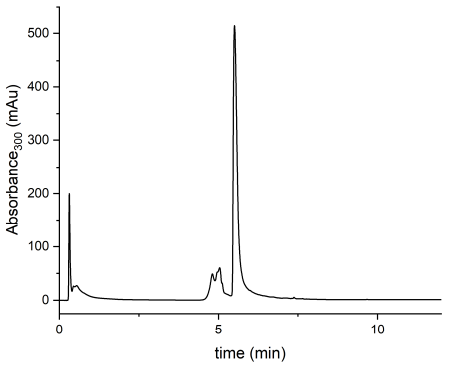
Compound **43** crude product
Column: Nucleodur C18
Gradient: 0-50 % B in A, TEAA system



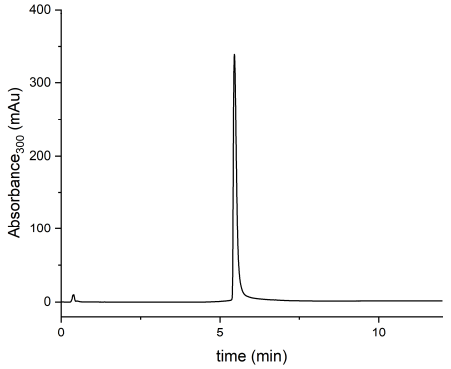
Compound **43** purified product
Column: Nucleodur C18
Gradient: 0-50 % B in A, TEAA system



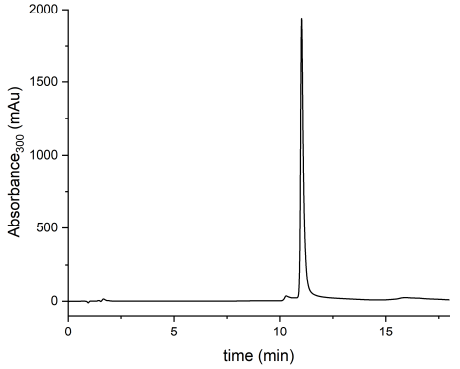
Compound **42** crude product
Column: Kinetex C18 Evo
Gradient: 0-70 % B in A, NH_4OAc system



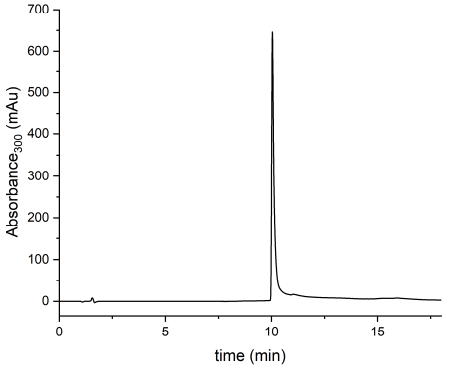
Compound **42** purified product
Column: Kinetex C18 Evo
Gradient: 0-70 % B in A, NH_4OAc system



Compound **36** purified product
Column: Aeris C4
Gradient: 0-100 % B in A, TEAA system



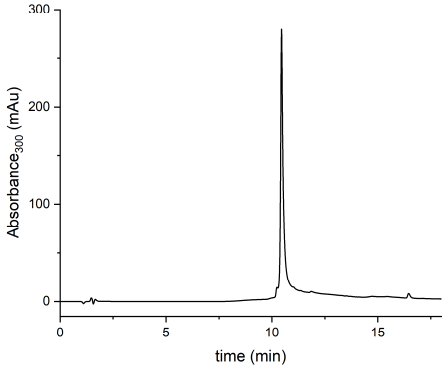
Compound **38** purified product
Column: Aeris C4
Gradient: 0-100 % B in A, TEAA sys



Compound **39** purified product

Column: Aeris C4

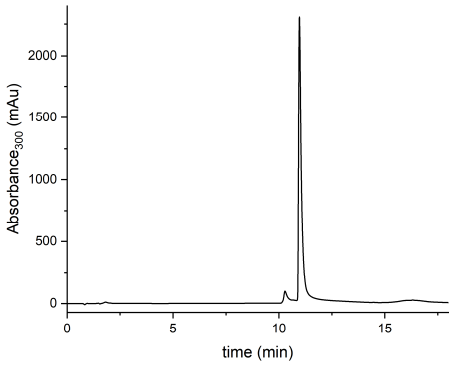
Gradient: 0-100 % B in A, TEAA system



Compound **40** purified product

Column: Aeris C4

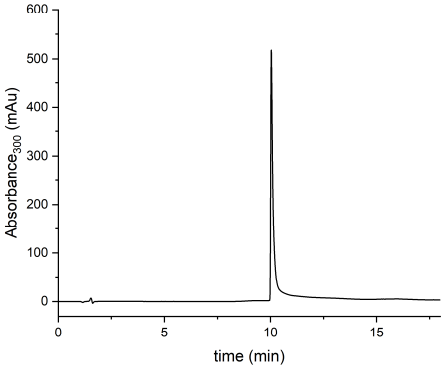
Gradient: 0-100 % B in A, TEAA system



Compound **30** purified product

Column: Aeris C4

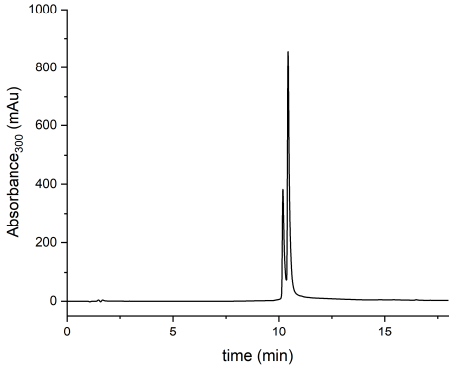
Gradient: 0-100 % B in A, TEAA system



Compound **31** purified product (side product corresponds to biotin oxidation)

Column: Aeris C4

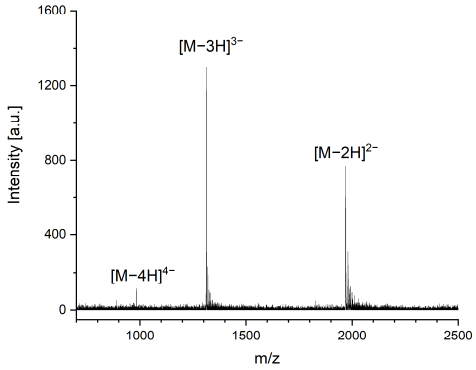
Gradient: 0-100 % B in A, TEAA system



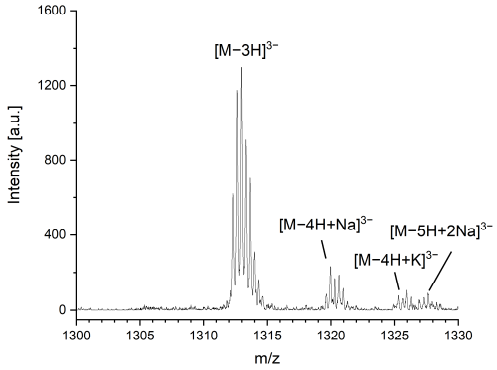
6.3.3.3 Mass spectra

6.3.3.3.1 Foldamer mass spectra

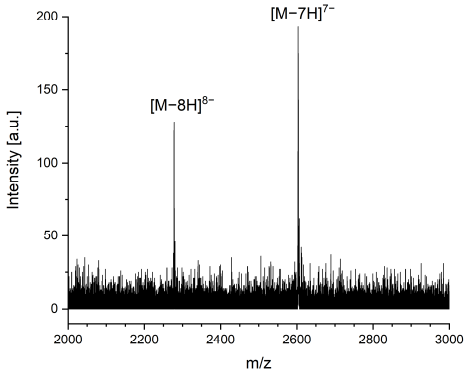
Compound **43**



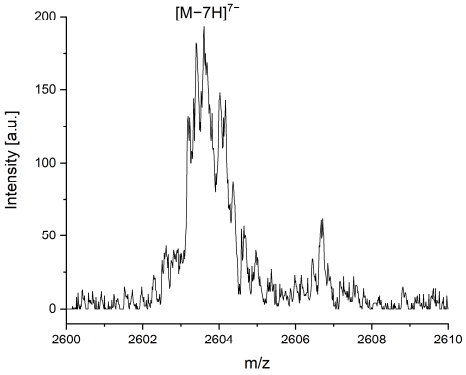
Compound **43** zoom



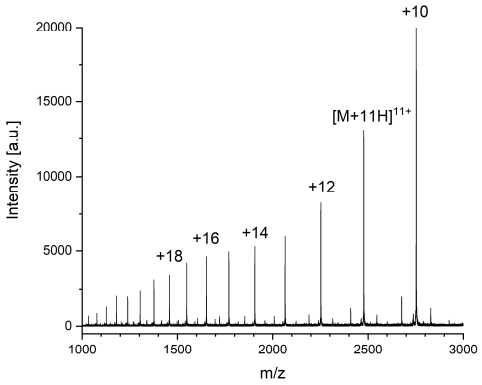
Compound **42**



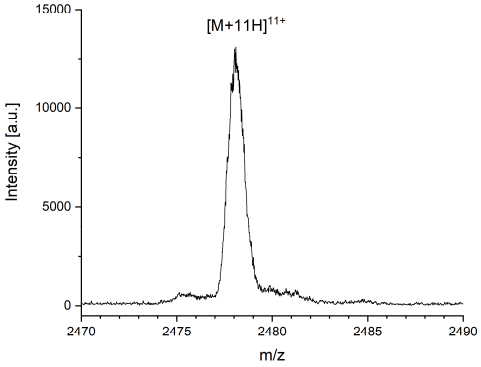
Compound **42** zoom



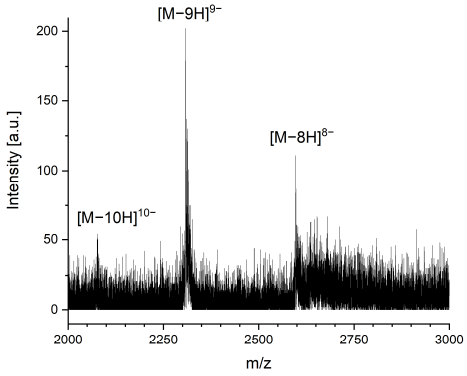
Compound **32**



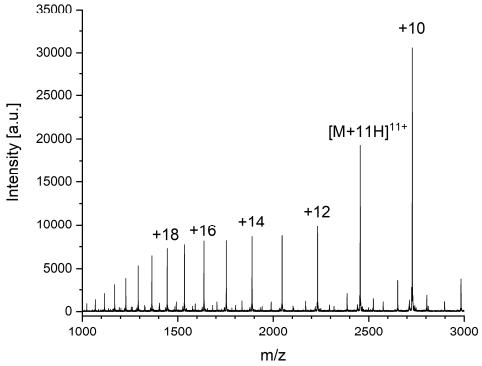
Compound **32** zoom



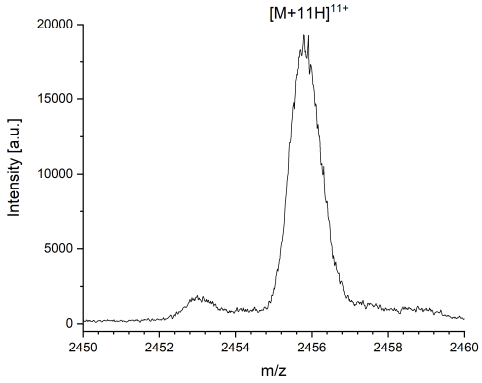
Compound **36**



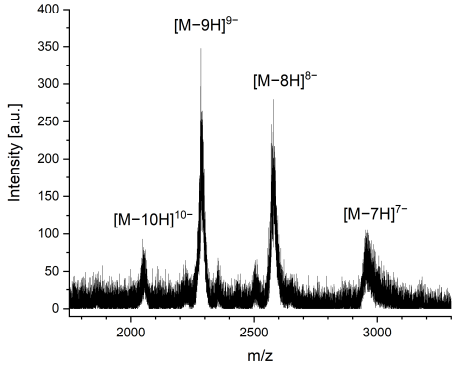
Compound **33**



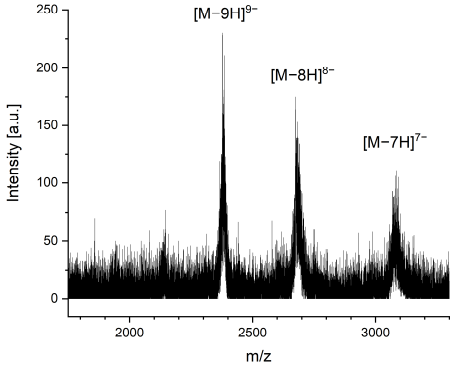
Compound **33** zoom



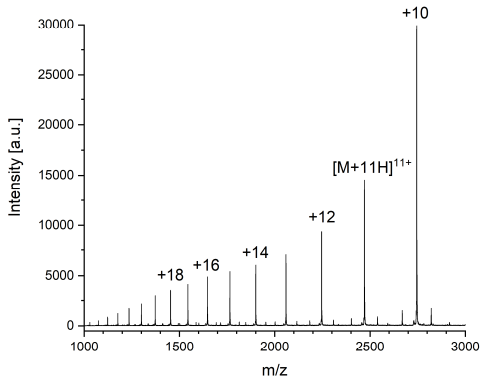
Compound **38**



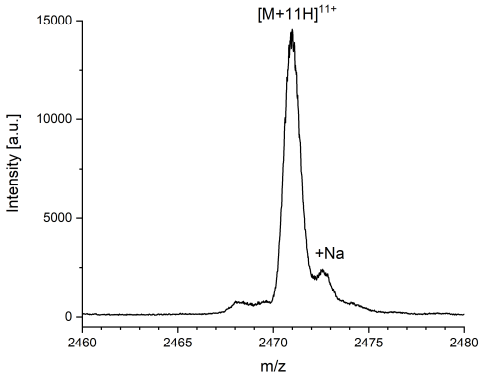
Compound **39**



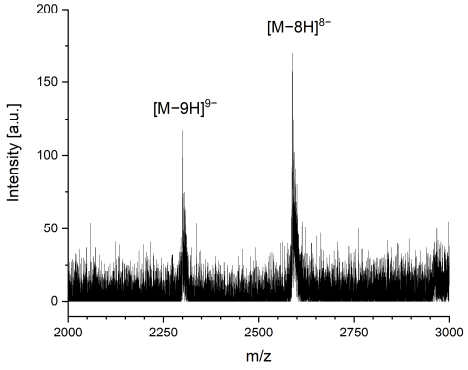
Compound **28**



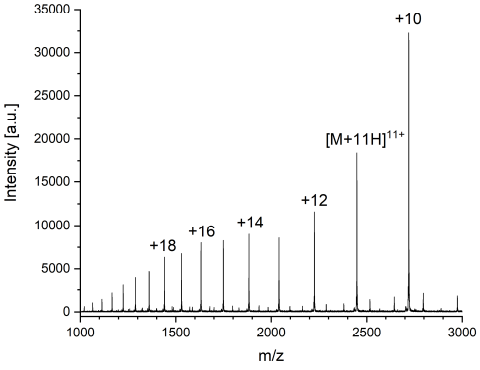
Compound **28** zoom



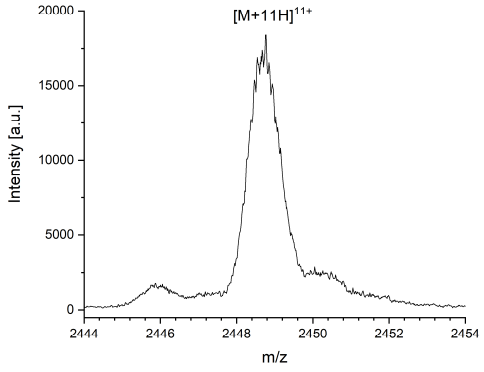
Compound **40**



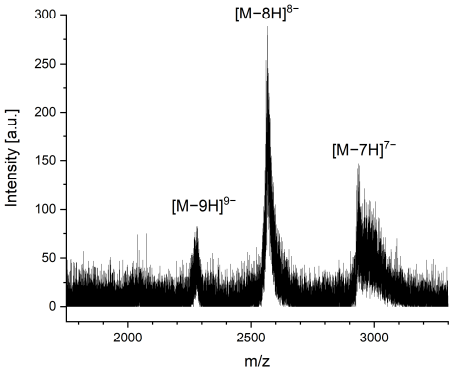
Compound **29**



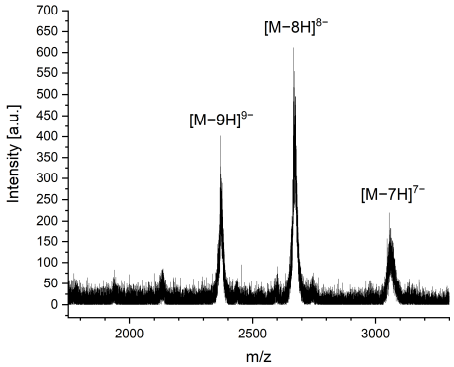
Compound **29** zoom



Compound **30**



Compound **31**



6.3.4 References

1. D. Bindl, E. Heinemann, P. K. Mandal and I. Huc, *Chem. Commun. (Cambridge, U. K.)*, 2021, **57**, 5662-5665.
2. V. Corvaglia, J. Wu, D. Deepak, M. Loos and I. Huc, *Chem. - Eur. J.*, 2024, **30**, e202303650.
3. B. Baptiste, C. Douat-Casassus, K. Laxmi-Reddy, F. Godde and I. Huc, *J. Org. Chem.*, 2010, **75**, 7175-7185.
4. F. Kunert, F. J. Metzner, J. Jung, M. Hopfler, S. Woike, K. Schall, D. Kostrewa, M. Moldt, J.-X. Chen, S. Bantele*, et al.*, *Sci. Adv.*, 2022, **8**, eadd3189.
5. J. R. Dormoy and B. Castro, *Tetrahedron Lett.*, 1979, DOI: 10.1016/s0040-4039(01)95397-3, 3321.
6. G. R. Fulmer, A. J. M. Miller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw and K. I. Goldberg, *Organometallics*, 2010, **29**, 2176-2179.
7. Maestro, *Journal*, 2021.
8. K. Ziach, C. Chollet, V. Parissi, P. Prabhakaran, M. Marchivie, V. Corvaglia, P. P. Bose, K. Laxmi-Reddy, F. Godde, J.-M. Schmitter*, et al.*, *Nat. Chem.*, 2018, **10**, 511-518.
9. H. Jiang, J.-M. Leger and I. Huc, *J. Am. Chem. Soc.*, 2003, **125**, 3448-3449.
10. X. Hu, S. J. Dawson, P. K. Mandal, X. de Hatten, B. Baptiste and I. Huc, *Chem. Sci.*, 2017, **8**, 3741-3749.

7 Chimeric molecules consisting of DNA and a DNA mimic foldamer

Decoy oligonucleotides have proven to be a powerful strategy for modulating gene expression by sequestering transcription factors away from their genomic targets, thus silencing their activity. By presenting short stretches of DNA that mimic natural binding sites, they can effectively block the assembly of transcriptional complexes by competitively inhibiting transcription factors of choice.[1] Their major strengths lie in specificity, modularity, and rapid programmability as they can be tailored to any specific DNA-binding protein simply by altering sequence. Their biggest hurdles lie in their intracellular stability and their efficient delivery.
This paper introduces a bifunctional linker that simultaneously forms a hairpin turn in B-DNA and serves as anchor of a DNA-mimic foldamer helix. This will allow us to build natural/abiotic chimeras where a DNA duplex and a foldamer sit in structural register and each other's grooves are prolonged by the other.

For nucleic-acid decoys, this hybrid strategy is significant for two reasons. First, it opens routes to decoys that combine the natural sequence specificity of DNA with the chemical robustness towards proteolytic degradation, tunable topography and high binding affinity of foldamers, potentially improving stability and off-rate control relative to classic oligo decoys. Second, many TFs do not present a pre-formed DNA-binding surface; instead, folding and ordering are coupled to dimerization, cofactor engagement and or DNA docking. By offering both a native DNA handle and an abiotic DNA-like surface in one construct, these chimeras could better capture or sequester TFs along this coupled folding/assembly pathway.

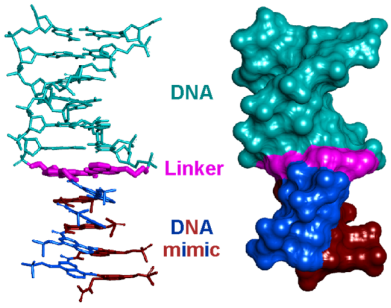


Figure 23 A DNA and DNA mimic foldamer chimeric molecule.

Contributions

The project was planned in collaboration with D.B. Konrad, T. Carell and I. Huc. The synthetic route for the linkage moiety was developed by D.B. Konrad, V. Carvaglia, M. Cabbar and me. Characterization of small-molecule intermediates was conducted by me. Oligonucleotide synthesis and analysis were conducted by F. Xu and J. Singer. P.K. Mandal conducted crystallization experiments, data collection (with assistance of G. Bourenkov and M. Agthe) at beamlines of EMBL Hambur. P.K. Mandal processed, solved and, refined the crystal datasets. Foldamer synthesis and analysis were conducted by me. Foldamer-DNA ligations, purifications and characterizations were performed by me. T. Chakrabortty expressed and

purified the proteins. C. Douat performed BLI measurements. The manuscript was written by me in collaboration with P.K. Mandal and I. Huc.

References

1. M. Hecker and A. H. Wagner, *Biochem. Pharmacol. (Amsterdam, Neth.)*, 2017, **144**, 29-34.

7.1 Publication (accepted): Interfacing B-DNA and DNA Mimic Foldamers

Manuel Loos,[a] Felix Xu,[b] Pradeep K. Mandal,[a]† Tulika Chakrabortty,[a] Céline Douat,[a] David B. Konrad,[a]‡ Melis Cabbar,[a] Johannes Singer,[b] Valentina Corvaglia,[a]‖ Thomas Carell,[b] and Ivan Huc*[a]

[a] M. Loos, Dr. P. K. Mandal, Dr. T. Chakrabortty, Dr. C. Douat, Prof. Dr. D. B. Konrad, M. Cabbar, Dr. V. Corvaglia, Prof. Dr. I. Huc
Department of Pharmacy
Ludwig-Maximilians-Universität München
Butenandtstr. 5–13, 81377 München (Germany)
E-mail: ivan.huc@cup.lmu.de
† Current address: Institute of Science and Technology Austria, Am Campus 1, 3400 Klosterneuburg (Austria)
‡ Current address: Department of Pharmaceutical Sciences, Universität Wien, Josef-Holaubek-Platz 2, 1090 Vienna (Austria)
‖ 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, Milan (Italy)
[b] Dr. F. Xu, J. Singer, Prof. Dr. T. Carell
Department of Chemistry
Ludwig-Maximilians-Universität München
Butenandtstr. 5–13, 81377 München (Germany)

Supporting information for this article is given via a link at the end of the document.

Abstract: A linker unit was designed and synthesized that can serve both as a hairpin turn in a DNA duplex and anchor point for an aromatic helical foldamer mimicking the shape and surface properties of B-DNA. Methods were developed to synthesize natural/non-natural chimeric molecules combining foldamer and DNA segments. The ability of the linker to position the foldamer helix and the duplex DNA so that their rims and grooves are in register, despite their completely different chemical nature, was demonstrated using single crystal X-ray diffraction, circular dichroism and molecular models. Bio-layer interferometry confirmed that artificial hairpin DNA duplexes keep their ability to bind to DNA binding proteins. The chimeric molecules may pave the way to competitive inhibitors of protein-DNA interactions involving sequence-selective DNA-binding proteins.

DNA mimic foldamers are single stranded aromatic helices having an overall shape and charge distribution similar to those of B-DNA duplexes. Oligoamides composed of alternating M and Q quinoline-derived amino acids (Figure 1a,b) constitute a prototypical example.[1,2] Like DNA mimic proteins[3,4] and some anionic polymers, *e.g.* heparin, DNA mimic foldamers may recognize DNA-binding proteins with high affinity and show potential for interfering with some DNA-protein interactions.[1,5-8] Nevertheless, $(MQ)_n$ oligoamides consist of a simple repeat motif and lack sequence features. They are thus less well suited to target proteins that recognize DNA sequence-selectively such as transcription factors or restriction enzymes. A possible approach to overcome this limitation would be to combine an unnatural $(MQ)_n$ oligoamide with natural DNA in a chimeric helical molecule that may benefit from the properties of both worlds, taking inspiration from other hybrid sequences that blend distinct molecular backbones, *e.g.* DNA-peptide nucleic acid chimeras[9-16] and peptides that integrate α-amino acids with some of their homologues[17-20] or analogues.[21,22] However, creating a DNA-$(MQ)_n$ interface that would allow for a structurally consistent arrangement of the DNA

and foldamer subcomponents is challenging because of their completely distinct chemical nature and conformational behavior. Here, we report how we successfully reached this milestone through the design and synthesis of a linker that can serve as both a B-DNA duplex hairpin turn and a single helical $(MQ)_n$ extension, thereby placing the grooves and arrays of negatives charges of the two segments in register.

In $(MQ)_n$ sequences, an MQ dimer brings a structural contribution equivalent to that of a base pair (bp) in a B-DNA duplex.[1] It raises the single helix by ca. 3.5 Å – i.e. the thickness of an aromatic ring – and it spans ca. 0.9 helix turn (equivalent to a twist of –0.1 turn), which results in a twist angle of 36°.with respect to the next MQ dimer in the sequence, similar to the twist angle between bp's in B-DNA. An overlay of an MQ dimer and a dG-dC Watson-Crick-Franklin bp (Figure 1c) highlights their comparable sizes. Specifically, the distance between the nitrogen atoms bonded to the deoxyribose of dA and dT matches with the distance between the carbon atoms in position 4 of M and in position 5 of Q.

This match inspired the design of foldamer-DNA linker X shown in magenta in Figure 1d. X is an MQ analogue that bears the benzylic amine of M and lacks the carboxylic acid of Q. As such, it can be placed at the C terminus of an $(MQ)_n$ helix. Furthermore, X possesses two hydroxy-terminated side chains so that it can be inserted in a DNA sequence via phosphodiester linkages and serve as a hairpin turn within a unimolecular duplex. Thus, the side chain OH groups of the first and second quinoline rings of X are equivalent to deoxyribose 5'- and 3'-OH groups, respectively. The side chains are in position 4 and 5 of the first and second quinoline rings of X, respectively, so as to place X where a DNA bp would be located. The length of the side chains – four and three methylene units – were made to match the number of atoms involved if the connecting units had been deoxyribose (Figure 1d). The torsional flexibility inherent to simple alkyl side chains was intended to give some conformational freedom to the hairpin structure. Notably, the overlap mentioned above between the nitrogen atoms of the bases and the C4/C5 of the quinoline rings does not hold for the atoms immediately adjacent – the ribose C1' carbon atoms and the first exocyclic atoms of the side chains – because of different bond orientations (Figure 1c), and this deviation must be accommodated. Besides, making the structure too rigid from the start may increase chances of failure. Nevertheless, a possible drawback of the flexibility of the side chains of X may be that it allows for two distinct orientations in which either prochiral face of X may stack with the adjacent DNA bp. Molecular models with both orientations suggest that the aromatic overlap between the stacked bp and quinoline rings is larger in the desired orientation (Figure S1 in Supporting information). This matters because the other orientation would be conducive of a foldamer helix handedness opposite to that required when $(MQ)_n$ is linked to an X-DNA hairpin (see below).

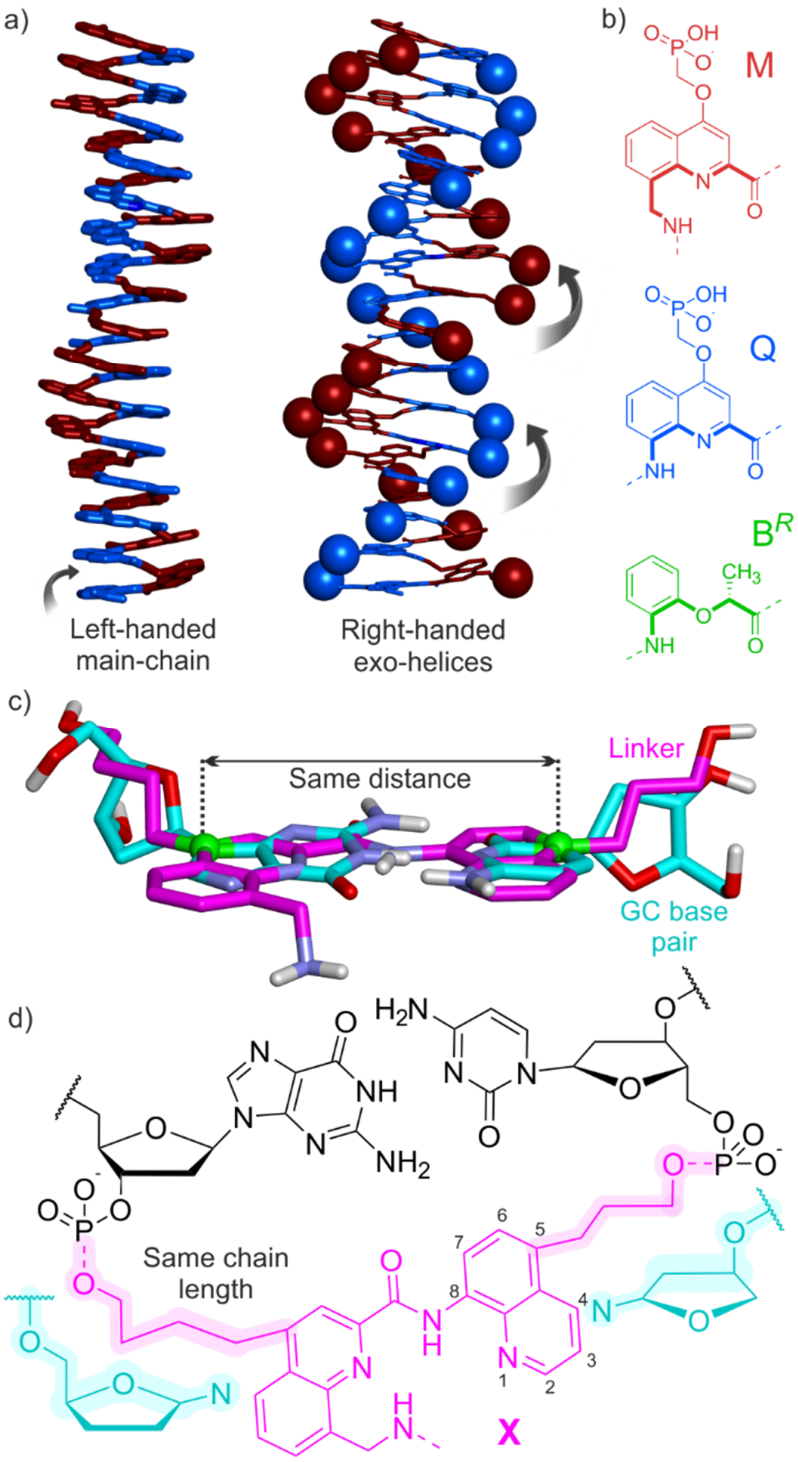


Figure 1 a) Crystal structure of an $(MQ)_{16}$ foldamer.[1] Diethyl ester-protected phosphonic acids side chains have been omitted in the structure at left. The phosphorous atom of the side chain is shown in space-filling representation at right. Atoms are colored in blue and red for Q and M units, respectively. b) Amino acid building blocks of DNA mimic foldamers. The thick bonds belong to the inner rim of the helix. c) Overlay of an MQ dimer (magenta) and an GC base pair (cyan). Overlapping atoms are displayed as green balls (quinoline Q4 carbon and guanine N9 atoms, quinoline Q5 carbon and cytosine N1 atoms). (d) Linker X (magenta) embedded as a turn-unit for a DNA duplex. The highlighted atoms of the side chains of X match in number with the highlighted atoms of the oxolane rings shown in cyan below.

Figure 2 depicts the synthesis of **1**, a precursor of X suitably protected and activated for oligodeoxyribonucleotide (ODN) solid phase synthesis (SPS). The Fmoc group protects the amine during ODN SPS and may be orthogonally removed afterwards for the subsequent conjugation of a pre-synthesized $(MQ)_n$ segment. Compound **1** was prepared from **6**, the 8-cyano-4-(*1H*)-quinolone precursor of M,[1] and from commercially available 5-bromoquinoline **2**. Both side-chains were installed on bromo-quinoline derivatives **3** and **7** via Sonogashira cross-coupling reactions. The alkyne of **4** was then hydrogenated together with the 8-nitro group, whereas the alkyne and nitrile of **8** were hydrogenated in separate steps.

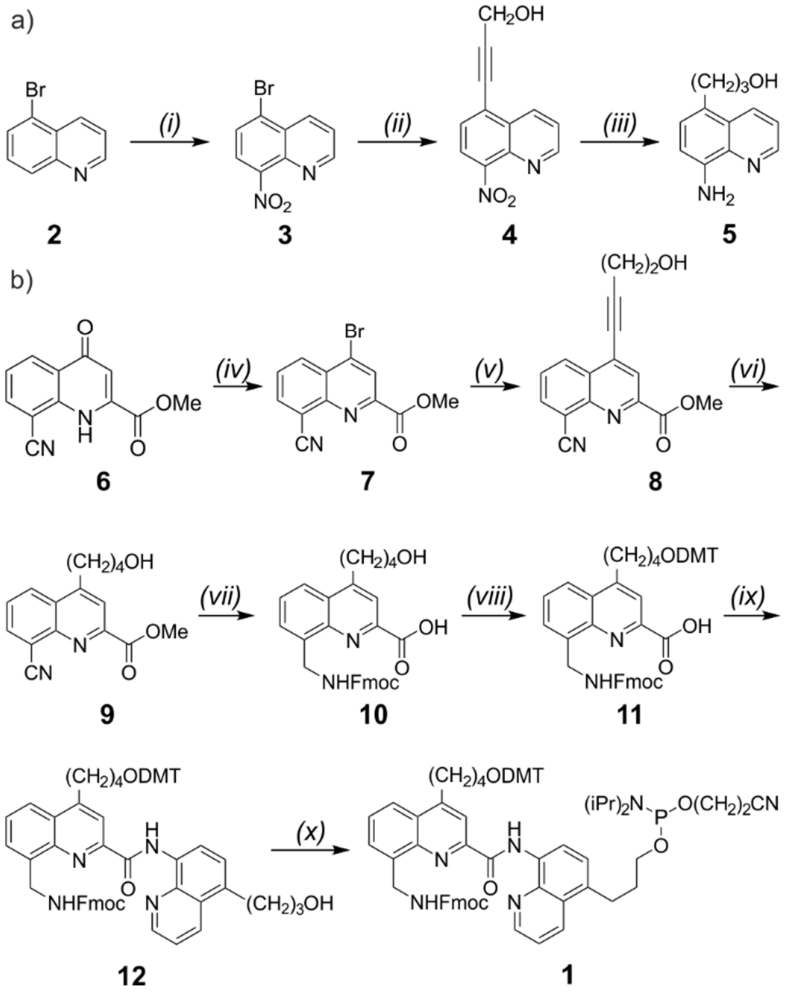


Figure 2 Synthesis of precursor **5**. (*i*) KNO_3, H_2SO_4, 0 °C to rt, 18 h, 85%; (*ii*) propargyl alcohol, CuI, xPhos, $Pd(dppf)Cl_2$, TEA/dioxane, 85 °C, 3 h, 85%; (*iii*) NH_4CO_2H, H_2, Pd/C, $Pd(OH)_2$/C, MeOH, rt, 20 h, 53%. b) Synthesis of DMT-protected phosphoramidite linker **1** for oligonucleotide synthesis. (*iv*) $POBr_3$, DMF, 80 °C, 2 h, 97%; (*v*) 3-butynol, CuI, xPhos, $Pd(dppf)Cl_2$, TEA/dioxane, 85 °C, 3 h, 75%; (*vi*) H_2, Pd/C, THF, rt, 2 h, 97%; (*vii*) 1. LiOH, THF/H_2O, 30 min, rt, 2. H_2, $Pd(OH)_2$/C, NaH_2PO_4; NH_4OH, THF/H_2O, 2 d, rt; 3. Fmoc-OSu, $NaHCO_3$; THF/H_2O, rt, 4 h, 38% over 3 steps; (*viii*) DMTCl, pyridine, rt, 2.5 h, 70%; (*ix*) **5**, PyBOP, DIPEA, $CHCl_3$, 0 °C, 4 h, 91%; (*x*) CEDCl, DIPEA, CH_2Cl_2, 0 °C to rt, 2 h; 80%. Abbreviations: triethylamine (TEA), 1,1'bis(diphenylphosphino)ferrocene (dppf), dimethoxytrityl (DMT), diisopropylethylamine (DIPEA), fluorenylmethyloxycarbonyl (Fmoc); succinimidyl (Su); benzotriazol-1-yl-oxytripyrrolidinophosphonium-hexafluorophosphat (PyBOP); 2-cyanoethyl *N,N*-diisopropyl-chlorophoshporamidite (CEDCl).

To differentiate the two OH groups, one DMT protecting group was installed before amide coupling of the two quinoline precursors. Acylation of **5** by **11** occurred selectively on the 8-amino group to yield **12**, which was activated to give SPS-ready linker **1**. SPS was performed with cyanoethyl phosphoramidite activation using standard building blocks for D-ODNs and with protecting groups allowing UltraMild deprotection for L-ODNs.

A series of ODNs was prepared having up to twenty residues and incorporating one or two linker units to form one or two hairpin turns (**13**-**21** in Figure 3a). A typical scale was 4 μmol, eventually performed four times in parallel. Final yields after purification using high-performance liquid chromatography (HPLC) ranged from 6 to 14% (Table S1). A representative example is highlighted in Figure S2.

Sequences **13**-**15**, **20** and **21**, were all expected to fold as single hairpins due to their complementary DNA segments. The latter two were used to assess binding of transcription factors that recognize the GGA motif (see below). Self-complementary sequences **18** and **19** have two hairpins and could in principle be macrocyclized by a single ligation step. Sequences **16** and **17** are complementary to each other and their duplex could in principle be macrocyclized through two ligation steps. However, efforts to implement ligation within the DNA hairpins were not pursued at this stage. Sequence **13**, where X links complementary d(GT_4G) and d(CA_4C) segments was specifically prepared for structural investigations because related sequences with a stilbenediether linker had been shown to form particularly stable hairpins and to produce crystals suitable for x-ray diffraction (XRD) analysis.[23,24] The melting temperature of **13** was measured at 78 °C, while the duplex d(GT_5G)•d(CA_5C) duplex is not expected to be stable at room temperature, indicating a considerable stabilization (Figure S3). Using the sitting drop vapor diffusion method, crystals of **13** were obtained under multiple crystallization conditions (Figure S4). However, single crystal XRD only showed a weak DNA fiber-like diffraction pattern (Figure S5). Encouraged by our previous success in crystallizing diverse DNA sequences using racemic DNA mixtures,[25-27] we prepared **14**, the L enantiomer of **13**. Crystals of the **13**+**14** racemate diffracting to a resolution of 2.5 Å were obtained readily and the structure could be solved in centrosymmetric space group *P*-1 using an earlier structure[23,24] as molecular replacement model (Figures 3b-d, S6-S10, Table S2). The asymmetric unit consisted of two independent molecules and four Mg^{2+} ions. Both molecules adopted a hairpin B-form DNA conformation albeit with some variations of sugar puckering (Tables S3-S7) However, the structure of the two hairpins differed significantly for both the DNA components and the X linker (Figures 3, S8). Highlighting these differences, the root mean square deviation of the two hairpins superimposed based on calculations performed for all DNA atoms was 1.27 Å.[28] Importantly, in both molecules, the quinoline rings of X linkers are coplanar and stacked to the adjacent dG–dC bp in the predicted preferred orientation (Figure S1).The X units have different twist angles with respect to the dG–dC bp but in both cases, this twist extends the twist between DNA bp's (Figure 3e,f, S10). X is thus equivalent to

an additional bp. To assess the coherence of the position of X units with respect to where a base pair would be instead, each hairpin molecule was overlaid with a copy of itself frame-shifted by one bp along the helix axis (Figure 3e,f, S10). X units were thus made to overlap with adjacent dG-dC bp. For one hairpin, the match was excellent (Figure 3f) whereas for the other the X unit was in part twisted away from an ideal bp position (Figure 3e), reflecting the flexibility of the side chains of X. Altogether the solid-state hairpin structures validate the initial design.

a)

13	D-d(5'-GTT TTG **X**CA AAA C-3')
14	L-d(5'-GTT TTG **X**CA AAA C-3')
15	d(5'-ACA GGA T**X**A TCC TGT-3')
16	d(5'-pC**X**G TCC TA-3')
17	d(5'-pAGG ATG **X**C-3')
18	d(5'-CTG **X**CA GGA TG**X** CAT Cp-3')
19	d(5'-CTG T**X**A CAG GAT GT**X** ACA TCp-3')
20	d(5'-CAC ATC CTG T**X**A CAG GAT GTG-3')
21	d(5'-GAC AGG ATG T**X**A CAT CCT GTC-3')

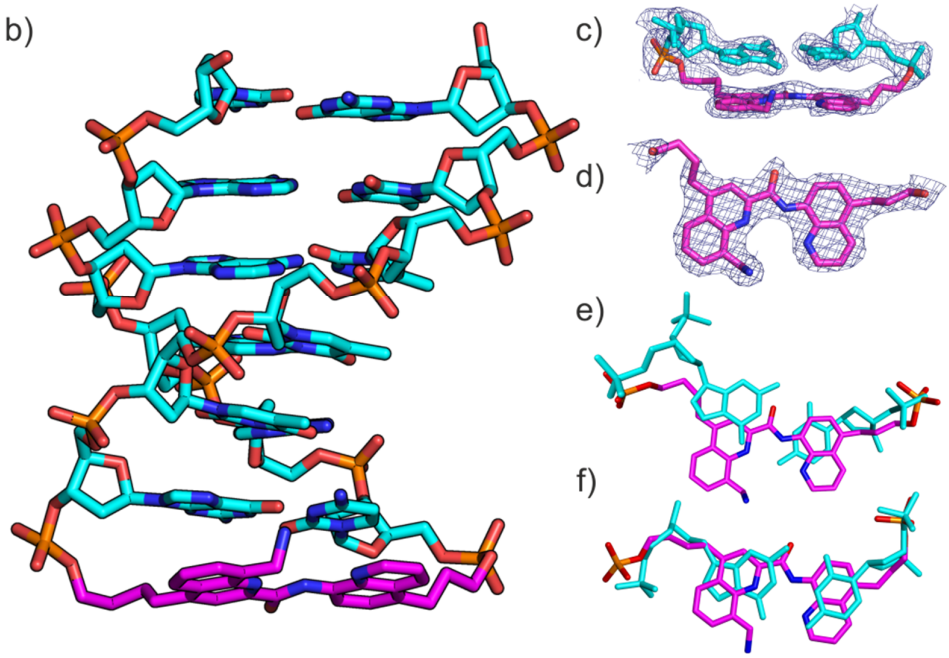


Figure 3 a) Synthesized oligonucleotides with the linker X highlighted in bold. All strands were synthesized from D-deoxynucleoside phosphoramidites except **14**, the enantiomer of **13**, which was made using L-deoxynucleosides. All sequences are terminated by desoxyribose hydroxy groups at 3' and 5' ends, unless they are indicated to be phosphate terminated by a p. b)-f) Solid state structure of **13**. One (hairpin 1) out of the two molecules present in the asymmetric unit is shown in b) with the linker carbon atoms colored in magenta and other carbon atoms and nitrogen, oxygen and phosphorus atoms colored in cyan, blue, red and orange respectively. The final Fourier $2F_o-F_c$ electron density at 2.5 Å resolution and contoured at 1 σ are shown around the linker and adjacent GC bp of hairpin 1 (c) and around the linker in top view (d). The overlay of the linker (carbon atoms shown in purple) and a base pair (in cyan) where it would be expected upon extending the DNA duplex in hairpin 1 (e) and hairpin 2 (f). See Figure S9 for details

We next explored the conjugation of $(MQ)_n$ segments to the benzylic amine of X after its installation within an ODN. Conjugation was first attempted on unpurified ODNs still attached to the SPS glass beads with their protecting groups still on, except for the Fmoc group of X. Using unprotected Ac-$(MQ)_4$-OH **22**, that is, with free phosphonic acid side chains, the terminal carboxylic acid function was successfully activated with EDC (1-ethyl-3-(3-dimethylaminopropyl) carbo-diimide hydrochloride) but no conjugation product could be detected. This route was not pursued but deserves to be revisited because fragment condensation of foldamers on peptides or other foldamers on solid phase is reliable.[29,30] Conjugation was successful when performed in solution using purified deprotected ODNs containing X. However, mass spectrometry revealed the presence of dehydrated impurities.[31] These side reactions could be suppressed by using DMTMM (4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methyl-morpholine hydrochloride) instead of EDC as coupling reagent. These milder conditions required excess foldamer (4.6 equiv.). Final purification was achieved by HPLC using an NH_4OAc buffer system, leading to good separation of the ODN-foldamer conjugates from excess foldamer. Conjugates **24** and **25** (Figure 4a) were prepared and purified on a 50-150 nmol scale. The melting temperature of **24** was measured at 87°C (Figure S3), a slight increase compared to **13**, indicative of a restriction of the flexibility of X, exerted by the foldamer fragment.

Circular dichroism (CD) spectroscopy was used to assess the conformation of the chimeric molecules in solution. $(MQ)_n$ foldamers are achiral and exist as racemic mixtures of right-handed (*P*) and left-handed (*M*) helical conformers. The helix handedness can be biased quantitatively toward the *M* helical conformer upon insertion of a single chiral B^R unit (Figure 1b) in a sequence.[2,32] For example, the ^1H nuclear magnetic resonance spectrum of **23** shows a single set of sharp resonances indicating the presence of a single diastereomeric conformer, and its CD spectrum shows a negative band as a result of its *M* helicity, with a maximum at 365 nm, a wavelength at which DNA does absorb (Figures 4b, S11).

Importantly it is the *M* helical conformation of $(MQ)_n$ that mimics the shape and charge distribution of B-DNA because it projects its phosphonate side chains so that they form two *P* exo-helices that match with the *P* double helix of DNA (Figure 1a).[33,34] Within a DNA-foldamer chimera, the orientation of the X unit with respect to the DNA duplex dictates the handedness of the foldamer. If X is oriented as in the original design and in the solid state structure of **13**, the foldamer helix must adopt the desired *M* handedness whereas a 180° flip of X would induce *P* helicity (Figure S12).[35] Satisfactorily, the CD spectrum of **24**, which possesses a foldamer segment identical to **23** but lacking the B^R unit, shows a negative band at 365 nm indicating it is left-handed. The intensity of the band and its overlap with the CD spectrum of **23** indicate that foldamer helix handedness bias in **24** is essentially quantitative and that its shape is not altered from its conjugation to the ODN (changes in helix shape are typically reflected in the CD spectrum). As expected, the CD spectrum of **25**, the L enantiomer of **24**, shows a band with

opposite sign (Figure 4b). Our attempts to crystallize enantiopure **24** or the **24**+**25** racemic mixture did not yield diffracting crystals. Nevertheless, the solution data clearly establish that $(MQ)_n$ segments connected to a DNA duplex via an X linker adopt the desired *M* helicity, that is, DNA chirality is transferred to the foldamer helix. The energy-minimized structure shown in Figure 4c, with the arrays of anions and the grooves of the DNA and foldamer segment in register, represents a plausible model of the conformation of **24**.

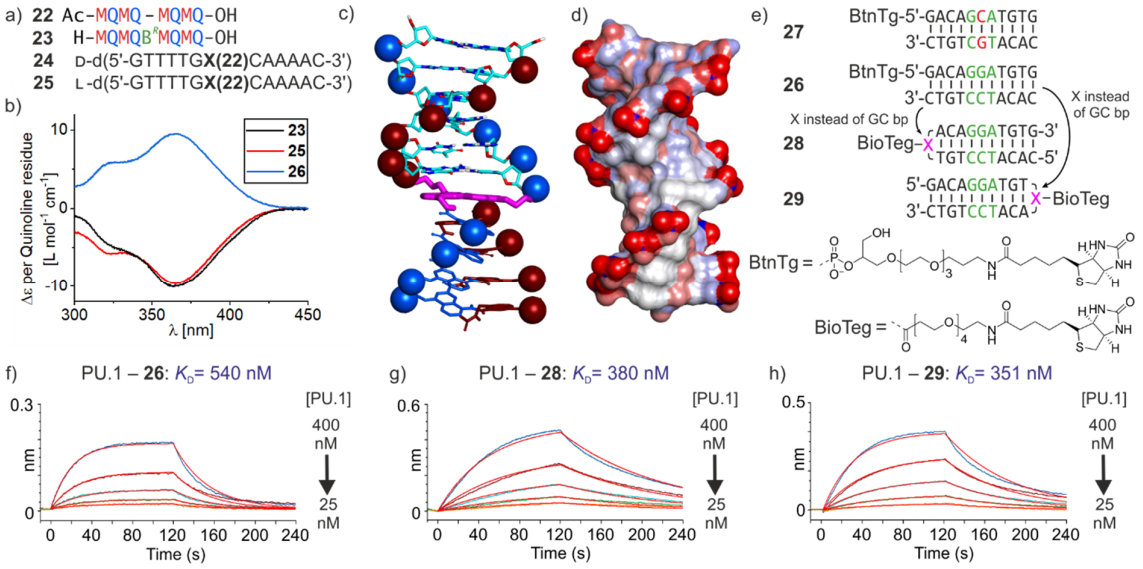


Figure 4 a) Sequences of synthesized foldamers and foldamer-DNA conjugates. **X(22)** indicates an X unit, with its amine acylated with the acid of foldamer **22**. b) CD spectra of conjugates **24** and F comparison with reference foldamer **23**. Spectra have been normalized per quinoline residues. Sample solutions were typically 60 or 80 µM in 2 mm cuvettes. c) Energy-minimized model of foldamer-DNA conjugate **24** with DNA atoms colored in cyan, red, blue and white for carbon, oxygen, nitrogen and hydrogen atoms respectively. Linker atoms are colored in magenta. Foldamer atoms are displayed in blue and red for Q and M respectively. The two phosphorus atoms are shown in space filling representation. DNA phosphorus atoms are colored according to the foldamer *exo*-helix that they extend. d) Solvent accessible surface for 24 colored according to atom charge. e) Biotinylated DNA duplexes with (**26**) and without (**27**) GGA consensus sequence and biotinylated hairpins **28** and **29** containing X. BLI sensorgrams of the binding of PU.1 to immobilized **26** (f), **28** (g) and **29** (h).

Finally, we used bio-layer interferometry (BLI) to assess the binding of two transcription factors, PU.1[36] and SAP1,[37] to biotinylated hairpin ODNs containing X and immobilized on streptavidin biosensors (Figures 4d-h, S13-S16, Table 1). Both proteins bind selectively to DNA sequences that include a d(GGA) motif. We first confirmed binding to **26**, a biotinylated 11-bp DNA duplex including GGA. We also verified the drop in binding affinity when GGA was replaced by GCA using duplex **27**.

We then measured binding of PU.1 and SAP1 to sequences **28** and **29**, the biotinylated analogues of **20** and **21**, respectively which correspond to duplex **26** where either of the terminal bp has been replaced by a biotinylated X. K_D values were in all cases similar to those measured with the original duplex DNA (Table 1), suggesting that the X unit does not perturb the DNA structure and does not hamper protein binding in these two cases. We nevertheless noted that both association and dissociation were faster with hairpin duplexes **28** and **29** than with **26** (Table S8).

Table 1 K_D values of biotinylated DNA duplexes and X-containing hairpins against transcription factors PU.1 and

K_D (nM)[a]	**26**[b]	**27**[b]	**28**[b]	**29**[b]
PU.1	330 ± 9	1600 ± 50	380 ± 0.5	351 ± 0.4
SAP1	1080 ± 6	–[c]	1100 ± 130	880 ± 138

SAP1.

[a] Calculated using a 1:1 model and global curve fitting from BLI sensorgrams measured in 25 mM Na_2HPO_4 (pH 7.5) containing 250 mM NaCl, 1 mM EDTA, 0.05% tween 20. [b] Biotinylated ligands were immobilized on streptavidin sensors and protein analytes were in solution. [c] No binding detected.

The purpose of this work was to create chimeric DNA-foldamer molecules with a structurally consistent arrangement of the DNA and foldamer subcomponents despite their different chemical nature. We have designed and synthesized an artificial linker based on a diamide quinoline unit that can be incorporated into DNA by standard phosphoramidite synthesis and that promotes a DNA hairpin structure. The linker can be elongated by a DNA mimic foldamer that extends the grooves and double helical array of negative charges of DNA. The *M* handedness of the helical foldamer is then controlled by the D stereochemistry of the DNA. The next step will be to explore how DNA-foldamer hybrids can be exploited as a new generation of DNA decoys[38,39] to target sequence-selective DNA binding proteins. While the foldamer component has been shown to enhance affinity for DNA-binding proteins,[1,5-8] the DNA component may help promote selectivity for defined protein targets by including a consensus sequence, in particular for transcription factors many of which have been considered undruggable.

Supporting Information

The authors have cited additional references within the Supporting Information.[40-64]

Acknowledgements

We acknowledge financial support from the Deutsche Forschungsgemeinschaft (DFG) via CRC 1309 (A03, C07, ID: 325871075), CRC 1361 (P2, ID: 393547839) and TRR 237 (A27, ID: 369799452), from European Union's Horizon Europe Framework Program under Grant

Agreement No. ERC-2021-ADG-320892, and from European Union's Horizon 2020 Research and Innovation Program under Grant Agreement No. 741912 (EPiR) and Marie Sklodowska-Curie Grant Agreement No 861381 (Nature-ETN). Further support was obtained by the BMBF Cluster4Future program (Cluster for Nucleic Acid Therapeutics Munich, CNATM, ID: 03ZU1201AA) and by the intl. graduate program RNAmed (Elite network of Bavaria). D.B.K thanks the Fonds der chemischen Industrie for a Liebig Fellowship. We thank T. Landmann for his contribution to ODN synthesis and G. Bourenkov and M. Agthe for assistance during synchrotron measurements in beamlines P13 and P14 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Germany).

Keywords: DNA mimic foldamer • DNA hairpin • Chimeric molecules • Helical molecules • Transcription factor

References

1. 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.
2. V. Corvaglia, J. Wu, D. Deepak, M. Loos, I. Huc, *Chem. - Eur. J.* **2024**, *30*, e202303650.
3. H.-C. Wang, C.-H. Ho, K.-C. Hsu, J.-M. Yang, A. H. J. Wang, *Biochemistry* **2014**, *53*, 2865-2874.
4. D. T. F. Dryden, *Trends Biotechnol.* **2006**, *24*, 378-382.
5. 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.
6. V. Corvaglia, I. A. M. Amar, V. Garambois, S. Letast, A. Garcin, C. Gongora, M. Del Rio, C. Denevault-Sabourin, N. Joubert, I. Huc, P. Pourquier, *Pharmaceuticals* **2021**, *14*, 624.
7. V. Corvaglia, D. Carbajo, P. Prabhakaran, K. Ziach, P. K. Mandal, V. Dos Santos, C. Legeay, R. Vogel, V. Parissi, P. Pourquier, I. Huc, *Nucleic Acids Res.* **2019**, *47*, 5511-5521.
8. D. Deepak, J. Wu, V. Corvaglia, L. Allmendinger, M. Scheckenbach, P. Tinnefeld, I. Huc, *Angew. Chem. Int. Ed.* **2025**, *64*, e202422958.
9. P. J. Finn, N. J. Gibson, R. Fallon, A. Hamilton, T. Brown, *Nucleic Acids Res.* **1996**, *24*, 3357-3363.
10. E. Uhlmann, D. W. Will, G. Breipohl, D. Langner, A. Ryte, *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 2632-2635.
11. L. Moggio, A. Romanelli, R. Gambari, N. Bianchi, M. Borgatti, E. Fabbri, I. Mancini, B. di Blasio, C. Pedone, A. Messere, *Biopolymers* **2007**, *88*, 815-822.
12. Z. Bajor, G. Sagi, Z. Tegyey, L. Oetvoes, *Nucleosides, Nucleotides Nucleic Acids* **2003**, *22*, 1215-1217.
13. M. Inagaki, R. Uematsu, T. Mizutani, D. Unabara, Y. Araki, S. Sakamoto, H. Kashida, M. Nishijima, H. Asanuma, Y. Inoue, T. Wada, *Chem. Lett.* **2019**, *48*, 341-344.
14. L. Petraccone, B. Pagano, V. Esposito, A. Randazzo, G. Piccialli, G. Barone, C. A. Mattia, C. Giancola, *J. Am. Chem. Soc.* **2005**, *127*, 16215-16223.
15. L. Moggio, L. De Napoli, B. Di Blasio, G. Di Fabio, J. D'Onofrio, D. Montesarchio, A. Messere, *Org. Lett.* **2006**, *8*, 2015-2018.
16. G. Barone, L. De Napoli, G. Di Fabio, E. Erra, C. Giancola, A. Messere, D. Montesarchio, L. Petraccone, G. Piccialli, *Nucleosides, Nucleotides Nucleic Acids* **2003**, *22*, 1089-1091.

17. K. L. George, W. S. Horne, *Acc. Chem. Res.* **2018,** *51*, 1220-1228.
18. W. S. Horne, S. H. Gellman, *Acc. Chem. Res.* **2008**, *41*, 1399-1408.
19. I. L. Karle, H. N. Gopi, P. Balaram, *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 3716-3719.
20. I. M. Mandity, E. Weber, T. A. Martinek, G. Olajos, G. K. Toth, E. Vass, F. Fulop, *Angew. Chem., Int. Ed.* **2009**, *48*, 2171-2175.
21. L. Mauran, C. Assailly, S. R. Goudreau, B. Odaert, G. Guichard, M. Pasco, *ChemBioChem* **2024**, *25*, e202400427.
22. L. Cussol, L. Mauran-Ambrosino, J. Buratto, A. Y. Belorusova, M. Neuville, J. Osz, S. Fribourg, J. Fremaux, C. Dolain, S. R. Goudreau, N. Rochel, G. Guichard, *Angew. Chem., Int. Ed.* **2021**, *60*, 2296-2303
23. M. Egli, V. Tereshko, G. N. Mushudov, R. Sanishvili, X. Liu, F. D. Lewis, *J. Am. Chem. Soc.* **2003**, *125*, 10842-10849.
24. F. D. Lewis, X. Liu, Y. Wu, S. E. Miller, M. R. Wasielewski, R. L. Letsinger, R. Sanishvili, A. Joachimiak, V. Tereshko, M. Egli, *J. Am. Chem. Soc.* **1999**, *121*, 9905-9906.
25. P. K. Mandal, G. W. Collie, B. Kauffmann, I. Huc, *Angew. Chem., Int. Ed.* **2014**, *53*, 14424-14427.
26. P. K. Mandal, G. W. Collie, S. C. Srivastava, B. Kauffmann, I. Huc, *Nucleic Acids Res.* **2016**, *44*, 5936-5943.
27. P. K. Mandal, G. W. Collie, B. Kauffmann, I. Huc, *Acta Crystallogr., Sect. D: Struct. Biol.* **2022**, *78*, 709-715.
28. E. Krissinel, K. Henrick, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *D60*, 2256-2268.
29. S. Dengler, C. Douat, I. Huc, *Angew. Chem. Int. Ed.* **2022**, *61*, e202211138.
30. S. Wang, J. Sigl, L. Allmendinger, V. Maurizot, I. Huc, *Chem. Sci.* **2025**, *16*, 1136-1146.
31. For solubilization of $(MQ)_n$ sequences and oligonucleotides, water is necessary. Hence, we work with a large excess of dehydration reagent. This has head to further dehydration products, that neither were observed on oligonucleotides or foldamer monomeric building blocks, but on foldamers alone. We hence theorize the crosslinking of two i+2 phosphonic acid groups on the foldamer exo-helix intramolecularly.
32. D. Bindl, E. Heinemann, P. K. Mandal, I. Huc, *Chem. Commun.* **2021**, *57*, 5662-5665.
33. That the M helix of (MQ)n generates P exohelices of negatively charged side chains may not be intuitive. It stems from the fact that an MQ dimer spans less than a helix turn. The twist angle with the next MQ dimer is negative, and the side chains of the next MQ dimer are shifted backwards with respect to the helix sense. See ref 34 for similar effects in helical peptides.
34. J. M. Martinez-Parra, R. Gomez-Ojea, G. A. Daudey, M. Calvelo, H. Fernandez-Caro, J. Montenegro, J. Bergueiro, *Nat. Commun.* **2024**, *15*, 6987.
35. For related work on handedness communication between two helices connected by a linker, see: a) V. Maurizot, C. Dolain, Y. Leydet, J.-M. Leger, P. Guionneau, I. Huc, *J. Am. Chem. Soc.* **2004**, *126*, 10049-10052. b) C. Dolain, J.-M. Leger, N. Delsuc, H. Gornitzka, I. Huc, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 16146-16151.
36. C. Dolain, J.-M. Léger, N. Delsuc, H. Gornitzka, I. Huc, *Proc. Natl. Acad. Sci. USA* 2005, **102**, 16146–16151.
37. J. R. Terrell, S. J. Taylor, A. L. Schneider, Y. Lu, T. N. Vernon, S. Xhani, R. H. Gumpper, M. Luo, W. D. Wilson, U. Steidl, G. M. K. Poon, *Cell Rep.* **2023**, *42*, 112671.
38. Y. Mo, B. Vaessen, K. Johnston, R. Marmorstein, *Mol. Cell* **1998**, *2*, 201-212.
39. G. Casas, F. Perche, P. Midoux, C. Pichon, J. M. Malinge, *Mol. Ther. Nucleic Acids* **2022**, *29*, 162–175
40. B. Johari, M. Moradi, *Methods Mol. Biol.* **2022**, *2521*, 207–230.
41. R. E. Franklin, R. G. Gosling, *Nature (London, U. K.)* **1953**, *172*, 156.
42. G. R. Fulmer, A. J. M. Miller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw, K. I. Goldberg, *Organometallics* **2010**, *29*, 2176-2179.
43. M. Cianci, G. Bourenkov, G. Pompidor, I. Karpics, J. Kallio, I. Bento, M. Roessle, F. Cipriani, S. Fiedler, T. R. Schneider, *J. Synchrotron Radiat.* **2017**, *24*, 323-332.
44. P. Evans, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2006**, *D62*, 72-82.
45. P. R. Evans, G. N. Murshudov, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *69*, 1204-1214.

46. W. Kabsch, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 125-132.
47. C. Vonrhein, C. Flensburg, P. Keller, A. Sharff, O. Smart, W. Paciorek, T. Womack, G. Bricogne, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 293-302.
48. 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.
49. M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 235-242.
50. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 486-501.
51. M. D. Winn, G. N. Murshudov, M. Z. Papiz, *Methods Enzymol.* **2003**, *374*, 300-321.
52. G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 355-367.
53. A. A. Vagin, R. A. Steiner, A. A. Lebedev, L. Potterton, S. McNicholas, F. Long, G. N. Murshudov, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *D60*, 2184-2195.
54. A. T. Bruenger, *Nature* **1992**, *355*, 472.
55. A. W. Schuettelkopf, D. M. F. van Aalten, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *D60*, 1355-1363.
56. A. Wilson, *Acta Crystallogr.,* **1950**, *3*, 397-398.
57. H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res.* **2000**, *28*, 235-242.
58. W.L. Delano, 2002 The PyMOL Molecular Graphics System, Delano Scientific, San Carlos **2002**.
59. X. -J. Lu, W. K. Olson, *Nat. Protoc.* **2008**, *3*, 1213-1227.
60. Schrödinger, LLC. *Maestro*; Schrödinger, LLC: New York, NY, **2021**.
61. N. Yang, X. Su, V. Tjong, W. Knoll, *Biosens. Bioelectron.* **2007**, *22*, 2700-2706.
62. S. Wang, G. M. K. Poon, W. D. Wilson, *Methods Mol. Biol.* **2015**, *1334*, 313-332.
63. C. Perez, A. M. Barkley-Levenson, B. L. Dick, P. F. Glatt, Y. Martinez, D. Siegel, J. D. Momper, A. A. Palmer, S. M. Cohen, *J. Med. Chem.* **2019**, *62*, 1609-1625.
64. V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat, I. Huc, *Chem. Eur. J.* **2023**, *29*, e202300898.

7.2 Supplementary information

7.2.1 Supplementary figures and tables

7.2.1.1 Supplementary figures

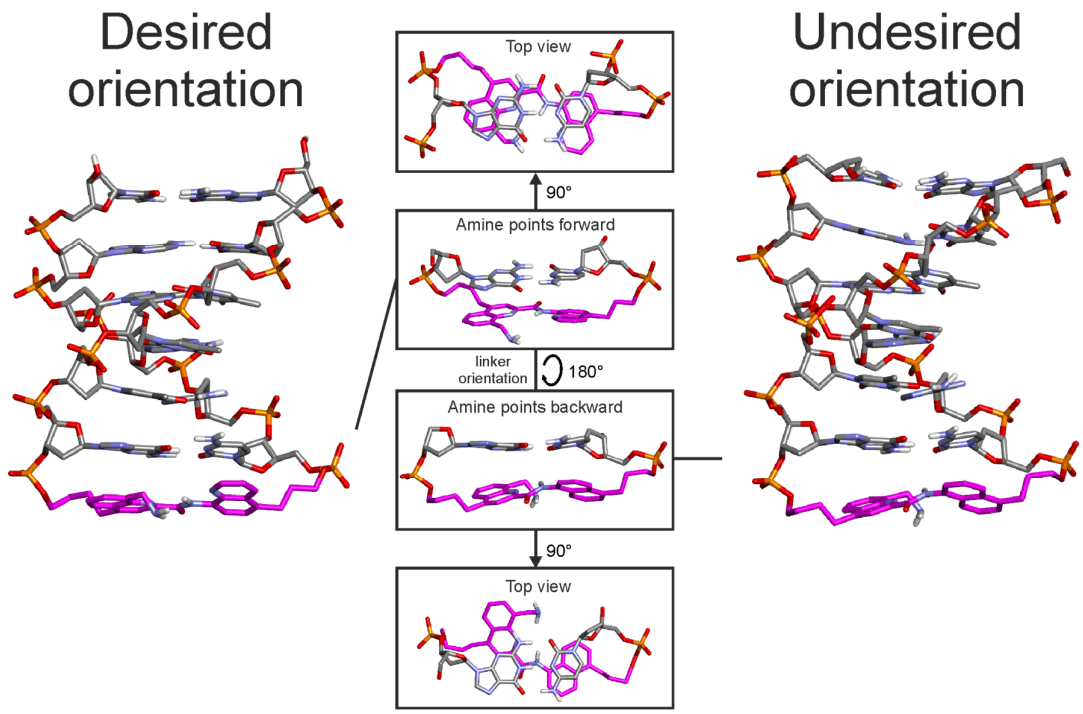


Figure S1. Energy-minimized molecular models of the two orientations that are possible for the diquinoline linker X within a DNA hairpin. The orientation depicted with the amine pointing forward (left) is the preferred orientation for a continuation of the DNA double-helical array of negative charges when a foldamer is appended to the turn unit. Atoms are colored in grey, red, orange and white for carbon, nitrogen, oxygen, phosphorus and hydrogen respectively. Non-polar hydrogen atoms are hidden. Carbon atoms of the linker unit are colored in magenta.

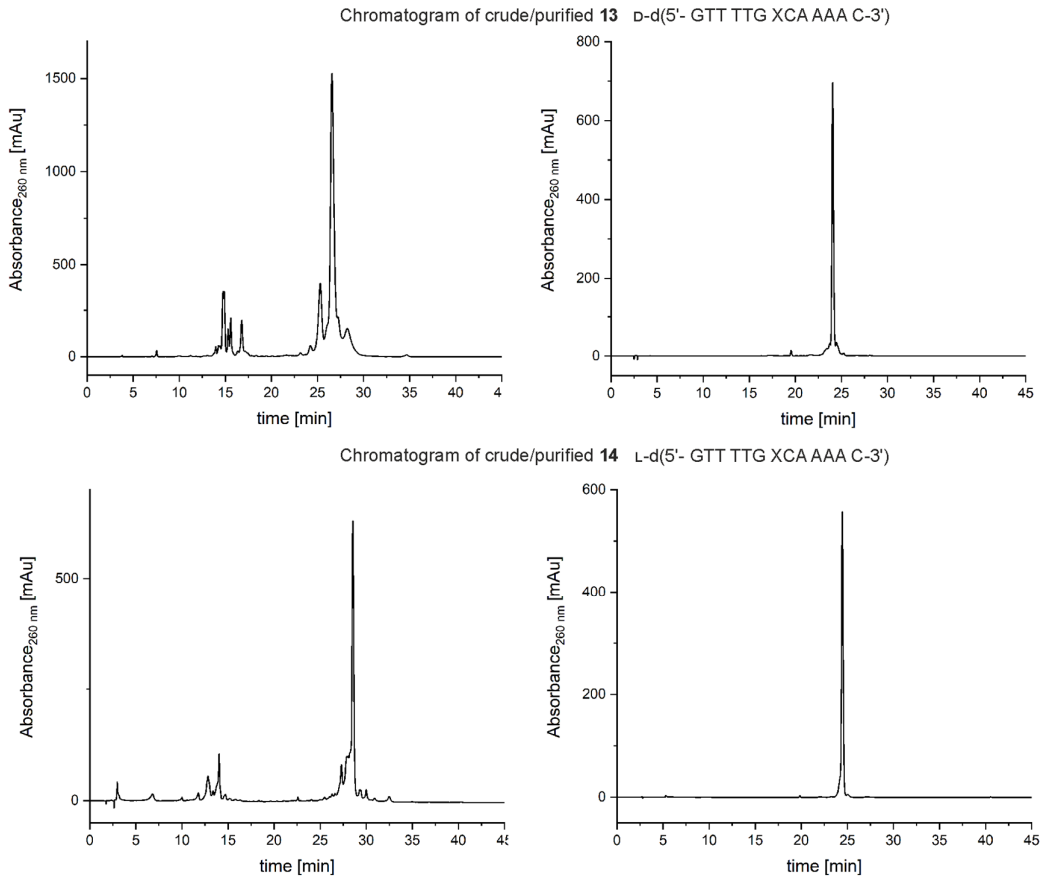


Figure S2. Representive examples of crude and purified chromatograms for ODN synthesis using standard phosphoramidite chemistry (**13**) and ODN synthesis using ultraMild oligonucleotide phosphoramidite chemistry. (**14**).

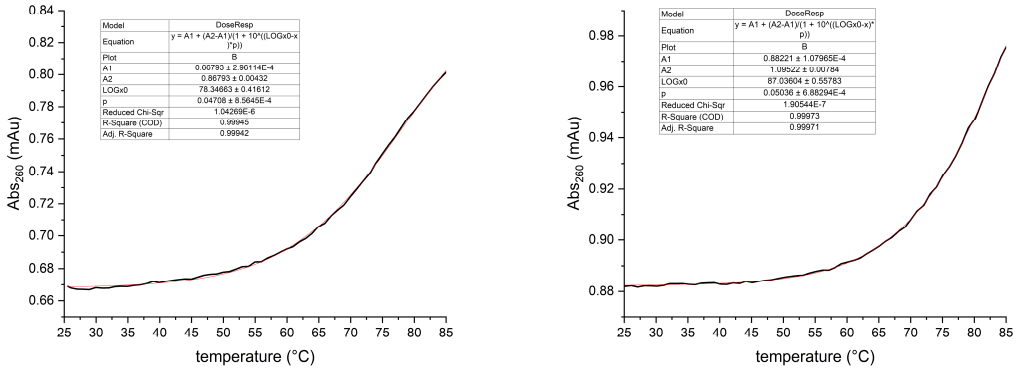


Figure S3. Melting curve of **13** (left) measured in a 1 cm cuvette at a concentration of 3.5 µM in a 1 cm cuvette and **24** (right) at a 25 µM concentration in a 2 mm quartz glass cuvette in sodium phosphate buffer (100 mM, pH = 7.5). Data was fit with a sigmoidal DoseResp function and their melting point was read from their respective inflection points.

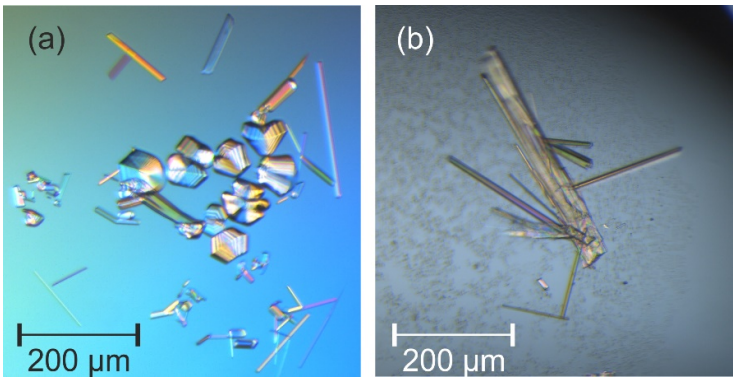


Figure S4. Representative examples of crystals of **13** (D-d(GT_4G)–X(H)–d(CA_4C)) and racemic mixture **13/14** D/L-d(GT_4G)X(H)d(CA_4C) observed under crossed polarizing microscope.

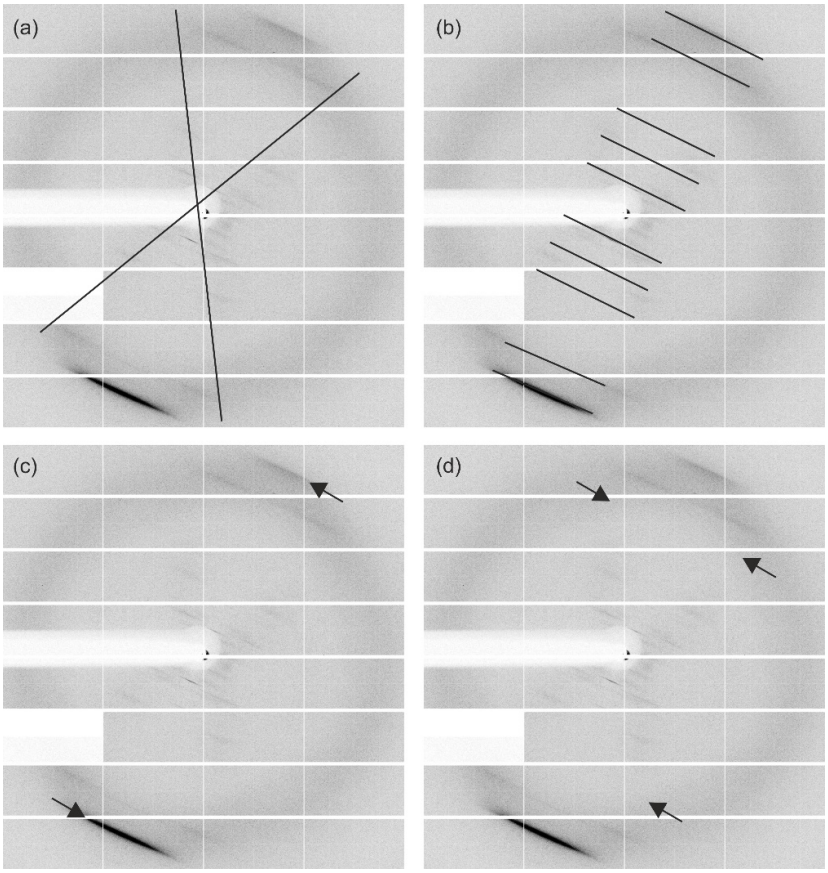


Figure S5. Diffraction of crystals of **13** (D-d(GT_4G)–X(H)–d(CA_4C)). The weak "X"-form distribution of diffraction peaks (a) indicates a helical structure[40] while the distances between the layer lines (b) imply the period of a helical turn. The broad extended lines at the top and bottom (c) correspond to small periodical features- the base pairs or linker. The absence of broad peaks (d) indicates the modulation of the interference pattern *i.e.* interference between the waves diffracted at two helical structures that are relatively shifted to each other along the long axis, implying a double helical structure. The experiment was carried out as a single crystal XRD experiment (unlike X-ray diffraction experiments using DNA fibers) and a limited amount of structural information was obtained. However, the DNA fiber-like pattern implied that the single crystals indeed corresponded to a hairpin with B-DNA like features and were not belonging to any small-molecule components of the crystallization reagents such as salts, buffer and precipitants that tend to diffract to high resolution and possess smaller unit cells.

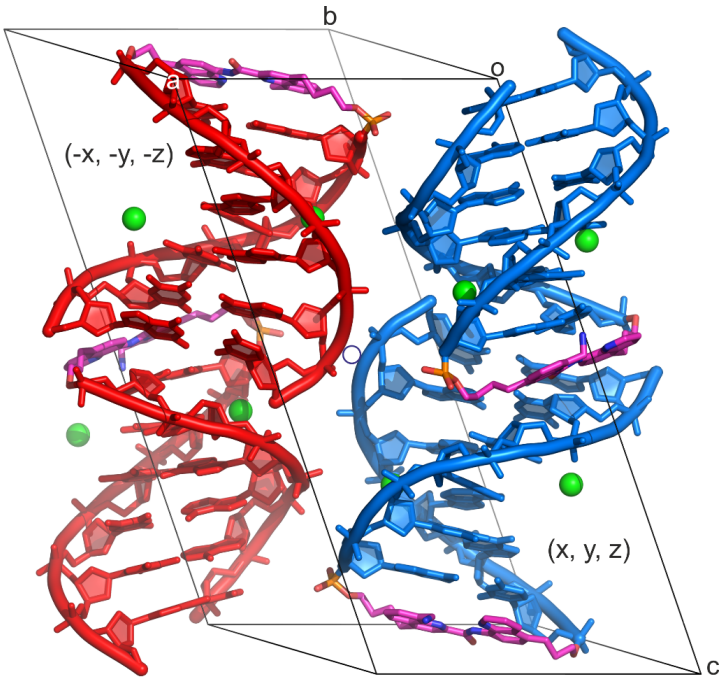


Figure S6. Unit cell arrangement of L/D-d(GT4G)–X–d(CA4C) hairpins **13/14** in P-1 space group. A blue circle represents the point of inversion. The L and D enantiomers are colored in red and blue, respectively. The diquinoline X linker atoms are colored magenta, red, blue and orange for carbon, oxygen, nitrogen and phosphorus respectively. Mg^{2+} ions are shown as green spheres. Bound water molecules have been omitted for clarity. The asymmetric unit consists of two right-handed D hairpins in (x, y, z).

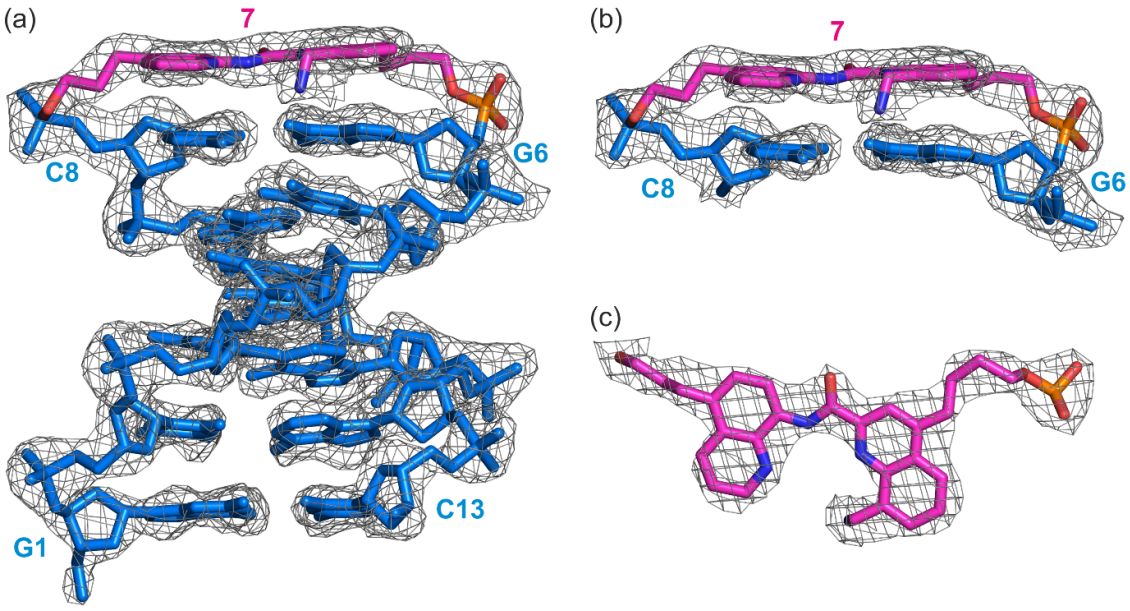


Figure S7. Quality of the structure. The final Fourier 2Fo–Fc electron density at 2.5 Å resolution and contoured at 1 σ around (a) Hairpin 1, (b) its residues G6-X7-C8, and (c) the X linker. DNA residues are colored blue and the linker atoms are colored in magenta, red, blue and orange for carbon, oxygen, nitrogen and phosphorus respectively.

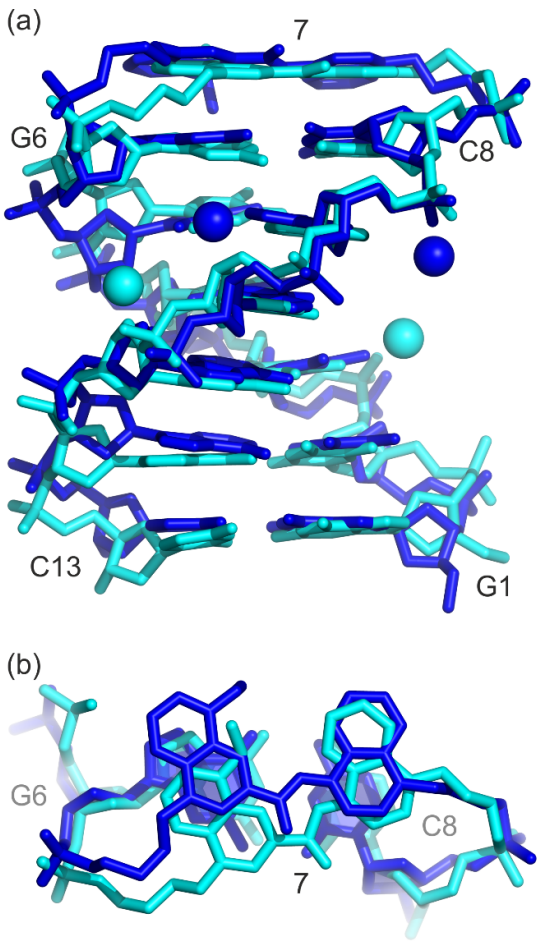


Figure S8. Conformations of D-d(GT_4G)–X–d(CA_4C). a) Overall superposition of Hairpin 1 (blue) and Hairpin 2 (cyan) structures along with the Mg^{2+} ions coordinating their minor groove. b) Geometries and relative orientations of the X linkers in the two hairpins. The view is approximately perpendicular to the adjacent G:C base pair.

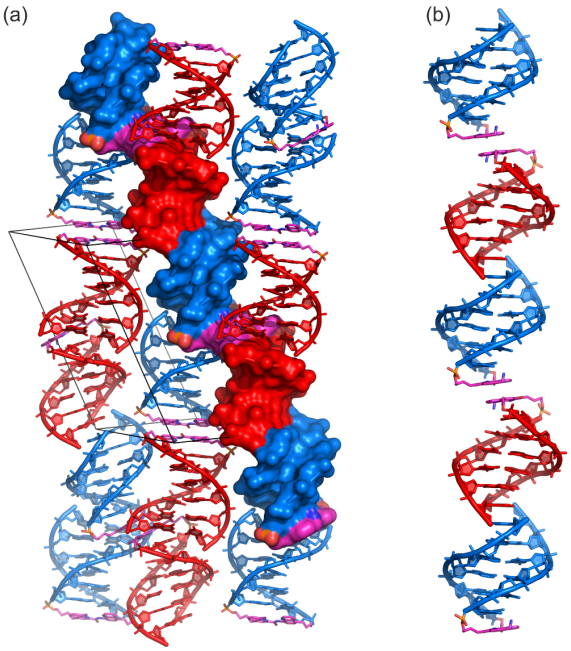


Figure S9. (a) Crystal packing arrangement L/D-d(GT_4G)–X–d(CA_4C) hairpins **13/14**. The L and D enantiomers are colored in red and blue, respectively. The X linker atoms are colored in magenta, red, blue and orange for carbon, oxygen, nitrogen and phosphorus respectively. Mg^{2+} ions and bound water molecules have been omitted for clarity. A heterochiral pseudo-continuous hairpin stacking arrangement is shown in (a) as a surface and in (b) as stick representations, respectively. The presence of a racemic mixture and of the diquinoline linker X gave rise to a novel packing arrangement of hetero-chiral DNA hairpins stacked into pseudo-continuous via the terminal residues of the DNA stem or the diquinoline linkers.

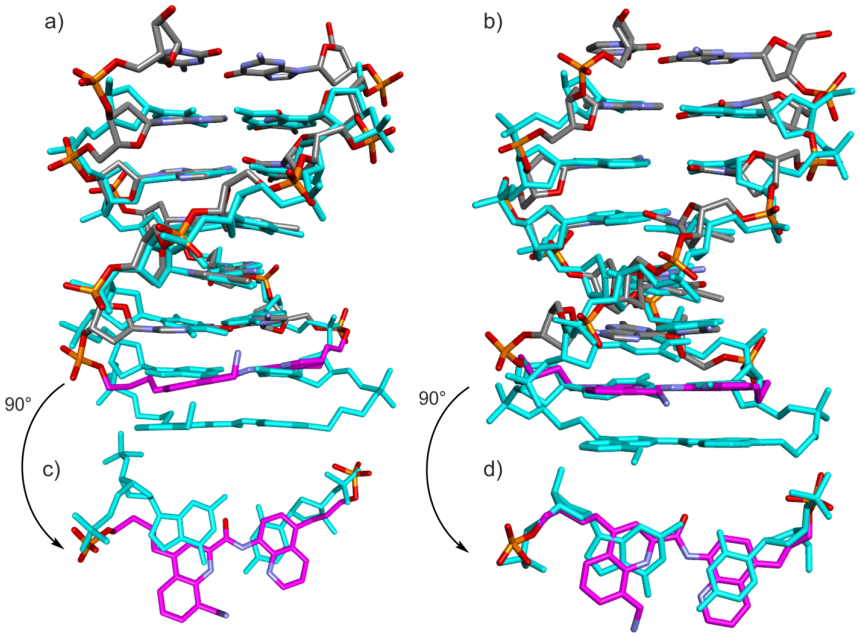


Figure S10. Crystal structure of hairpin 1 (a) and hairpin 2 (b) with linker carbon atoms in colored in magenta and carbon, nitrogen, oxygen and phosphorus colored in grey, blue, red and orange respectively. Additionally, another molecule of the same hairpin, shifted and fitted to match the DNA backbone was overlayed in cyan to highlight the orientation of the linker position in reference to a DNA base pair.

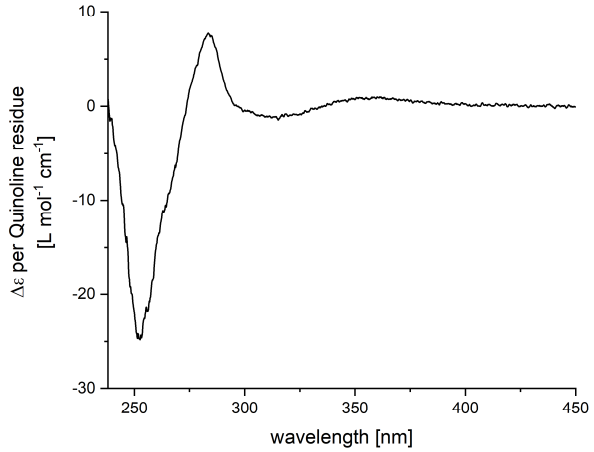


Figure S11. CD spectrum of compound **13** normalized per quinoline residue. Note the weak intensity of the positive band at 350 nm which can be assigned to the diquinoline linker chromophore, stacked on the adjacent base-pair with a right-handed twist as an extension of the B-DNA duplex.

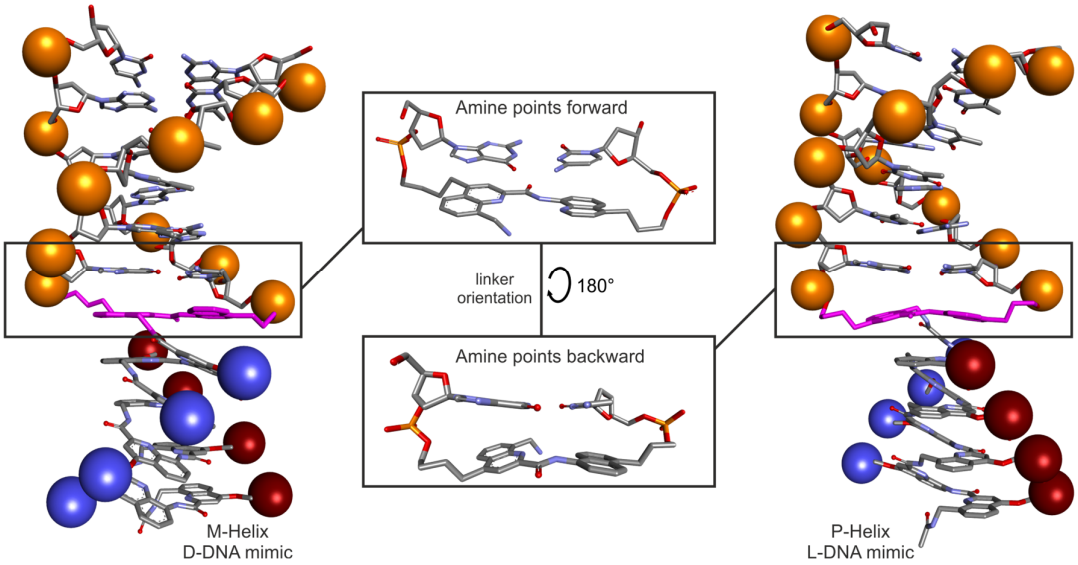


Figure S12. Energy-minimized molecular models of DNA hairpins with an appended foldamer segment in the two possible orientations of the diquinoline hairpin turn X. Only the orientation depicted with the amine pointing forward (left) gives rise to a continuation of the double-helical array of negative charges from the DNA to the foldamer. In the other orientation (right), the arrays of negative charges have opposite handedness in the DNA and foldamer segments. Atoms are color in grey, blue, red, orange and white for carbon, nitrogen, oxygen, phosphorus and hydrogen respectively. Phosphorus atoms are displayed in space-filling representation in respective blue, red and orange for Q units, M units and nucleotides, respectively. Non-polar hydrogen atoms have been omitted for clarity. Carbon atoms of the linker unit X are colored in magenta.

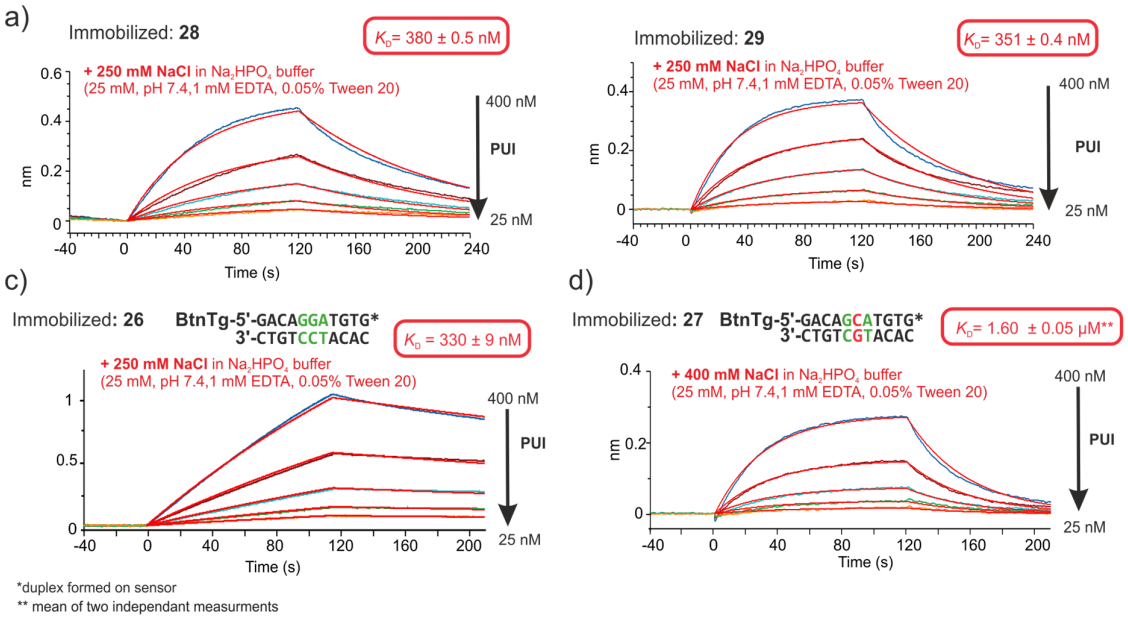


Figure S13. BLI sensorgrams of PU.1 against biotinylated hairpins containing X **28** (a) and **29** (b) with their respective fits and buffer composition. c) Immobilized dsDNA (see Figure S14 and section 7.2.2.6 for general information) containing the cognate GGA binding site with comparable length and sequence as **28** and **29**. For a-c), the K_D values are in the same high nM range. d) Immobilized dsDNA containing a non-cognate GCA motif measurements against PU.1. The GC to CG mutation in the cognate binding site leads to a drop in K_D value.

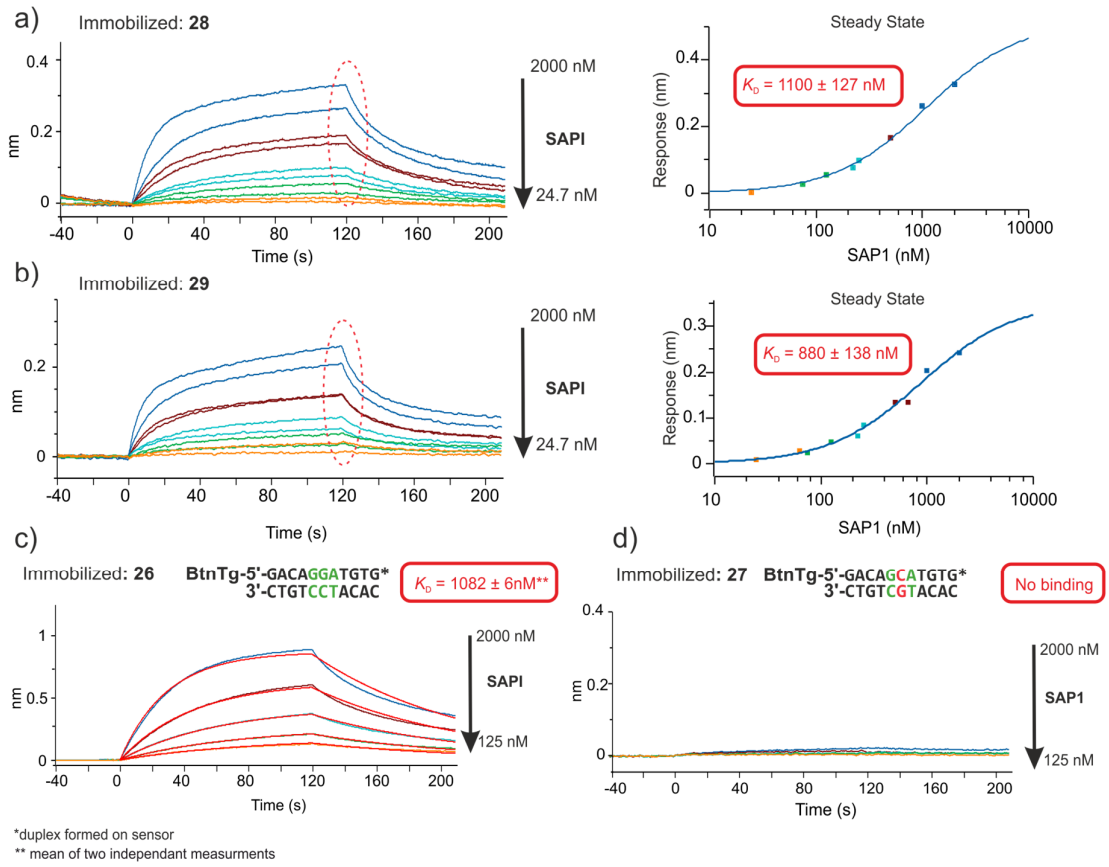


Figure S14. BLI sensorgrams of SAP1 against biotinylated hairpins containing X **28** (a) and **29** (b) with the sensorgrams (left) and a steady-state fit with the values at the end of the association phase at 120 seconds (highlighted by the dashed red ellipse in the sensorgrams). c) Immobilized dsDNA (see Figure S14 and section 2.1.6 for general information) containing the cognate GGA binding site with comparable length and sequence as **28** and **29**. For a-c), the K_D values are in the same low µM to high nM range. d) Immobilized dsDNA containing a non-cognate GCA motif measurements against SAP1. The GC to CG mutation in the cognate binding site leads to a complete loss in binding.

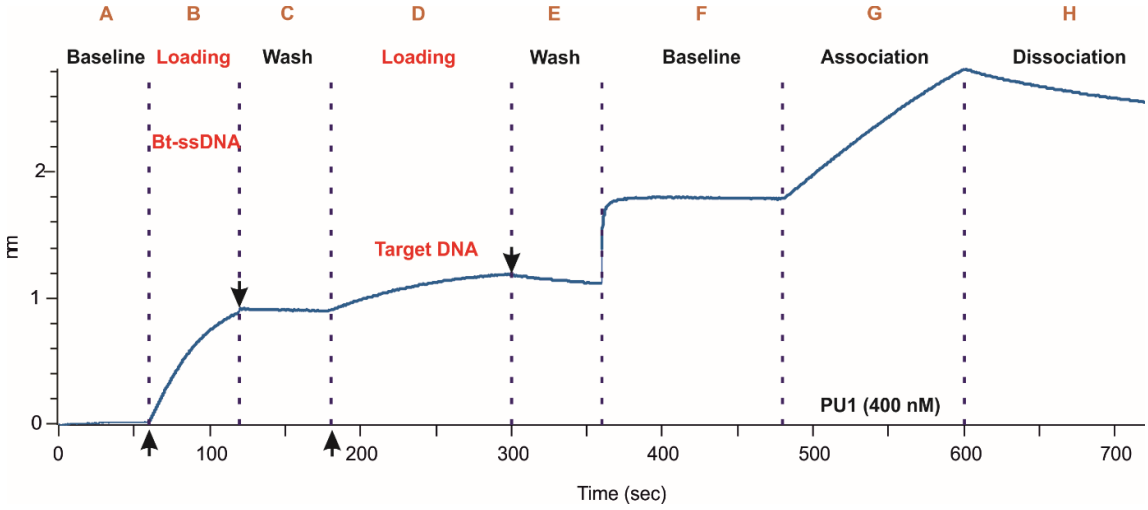


Figure S15. Sensorgrams of ssDNA immobilization, target DNA hybridization (**26**) on biosensors and kinetic curves of association and dissociation exemplified at the highest PU1 concentration (400 nM). The height independent steps (*A*-*H*) are shown, starting from a baseline in PBS of the SA biosensors (*A*) to the dissociation step (*H*). The biotinylated ssDNA (Bt-ssDNA) was loaded at 0.5 µM for 60 seoncds and target dsDNA hybridization (0.5 mM) was performed over 120 seconds in PBS buffer. At the end of each immobilization step, the SA-biosensors were washed (↓ arrows). For the subsequent baseline (*F*), association (*G*) and dissociation (*H*) steps, the kinetic buffer was 25 mM Na_2HPO_4 (pH 7.5), 250 mM NaCl, 0.01 mM EDTA and 0.05% tween 20 buffer.

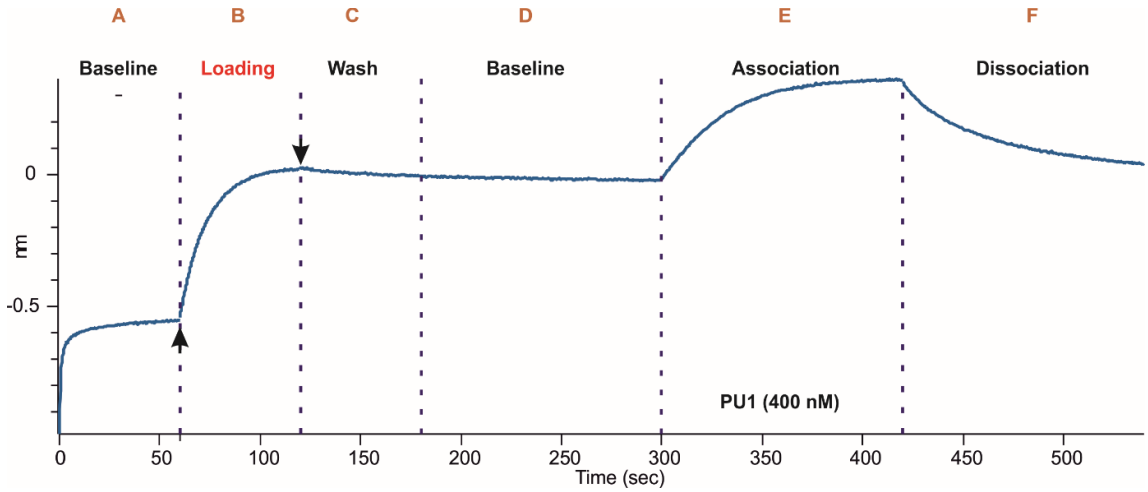


Figure S16. Sensorgrams of hairpin **29** immobilization on SA-biosensors and kinetic curves of association and dissociation exemplified at the highest PU1 concentration (400 nM). The six independent steps (*A*-*F*) are shown. The biotinylated **29** was loaded at 2 mg/mL in 25 mM Na_2HPO_4 (pH = 7.5) buffer containing 250 mM NaCl, 1 mM EDTA and 0.05 % tween. At the end of the loading step, the SA-biosensors were washed with the buffer (↓ arrows). For the subsequent baseline, association and dissociation steps, the kinetic buffer remained identical.

7.2.1.2 Supplementary tables

Table S1 Synthesized oligonucleotides, conjugates and biotinylated oligonucleotides, their calculated and found exact masses and yields. Oligonucleotides were synthesized on four cartridges with 4 µmol initial capacity each. Yields **13-21** are from ODN synthesis. Yields of **24** and **25** are from DNA-foldamer conjugation, **28** and **29** are yields of biotinylation of **20** and **21**.

Sequence name	Sequence (5'→3')	Calculated $[M-nH]^{n-}$	Found (*m/z*)	Yield (Isolated) Mass (Percent)
13	D-d(5'- GTT TTG **X**CA AAA C-3')	n=1; 4161.8	4162.8	13%
14	L-d(5'-GTT TTG **X**CA AAA C-3')	n=1; 4161.8	4161.4	11%
15	d(5'- ACA GGA T**X**A TCC TGT-3')	n=1; 4782.4	4782.5	12%
16	d(5'-pC**X**G TCC TA-3')	n=1; 2654,5	2654.0	14%
17	d(5'-pAGG ATG **X**C-3')	n=1; 2743,5	2742.4	14%
18	d(5'-CTG **X**CA GGA TG**X** CAT Cp-3')	n=1; 5383.8	5383.7	8%
19	d(5'-CTG T**X**A CAG GAT GT**X** ACA TCp-3')	n=1; 6614.3	6614.4	6%
20	d(5'-CAC ATC CTG T**X**A CAG GAT GTG-3')	n=1; 6636.2	6638,1.	10%
21	d(5'-GAC AGG ATG T**X**A CAT CCT GTC-3')	n=1; 6636.237	6636.5	9%
24	D-d(5'- GTT TTG **X(22)**CA AAA C-3')	n=4; 1624.271	1624.277	338 µg (42%)
25	L-d(5'-GTT TTG **X(22)**CA AAA C-3')	n=4; 1624.271	1624.363	904 µg (39%)
28	d(5'-CAC ATC CTG T**X**A CAG GAT GTG-3')	n = 5; 1420.687	n = 5; 1420.515	1.31 mg (67%)
29	d(5'-GAC AGG ATG T**X**A CAT CCT GTC-3')	n = 5; 1420.687	n = 5; 1420.515	182 µg (33%)

Table S2 Summary of X-ray diffraction and refinement data for racemic **13**/**14**. Values in parenthesis refer to data in the highest resolution shell.

Parameters	L/D-d(GT_4G)–X–d(CA_4C)
Data collection	
Wavelength	0.976 Å
Resolution[a] range	28.56 – 2.5 Å (2.59 – 2.5 Å)
Space group	*P*-1
Unit cell	a = 26.517 Å, b = 30.040 Å, c = 45.224 Å α = 105.62°, β = 105.77°, γ = 94.34°
Total reflections	7970 (794)
Unique reflections	3985 (397)
Multiplicity	2.0 (2.0)
Completeness	90.2 % (88.9%)
Mean I/ σ (I)	13.98 (5.53)
Wilson B-factor	49.79 Å^2
R-merge	3.44 % (19.18 %)
R-meas	4.87 % (27.12 %)
CC1/2	0.998 (0.901)
Refinement	
No. of Reflections	3985 (396)
R_{work}	27.82 % (33.89 %)
R_{free}	32.79 % (38.31 %)
No. of non-H atoms DNA hairpin Metal ions, Solvent	739 560 4 Mg^{2+}, 175
R.m.s.d. bond R.m.s.d. angle	0.014 Å 1.92°
Average B factor	34.25 Å^2
PDB ID #	8Q60

Table S3 Definition sugar structural parameters**.**

v0:	C4'-O4'-C1'-C2'
v1:	O4'-C1'-C2'-C3'
v2:	C1'-C2'-C3'-C4'
v3:	C2'-C3'-C4'-O4'
V4:	C3'-C4'-O4'-C1'
tm	the amplitude of pucker
P	the phase angle of pseudorotation

Table S4 Sugar puckering of strand 1 of hairpin 1.

Base	v0	v1	v2	v3	v4	tm	p	Puckering
1G	-17.9	-11.7	35.2	-45.5	40.6	46.2	40.5	C4'-exo
2T	-7.1	-10.8	23.7	-27.8	22.0	28.0	32.3	C3'-endo
3T	-23.6	32.4	-28.6	15.6	4.9	32.1	152.7	C2'-endo
4T	-25.2	32.1	-26.7	12.7	7.7	31.7	147.5	C2'-endo
5T	-40.2	44.8	-33.2	10.5	18.4	44.9	137.8	C1'-exo
6G	-18.9	31.1	-31.1	20.5	-1.2	32.4	163.5	C2'-endo

Table S5 Sugar puckering of strand 2 of hairpin 1.

Base	v0	v1	v2	v3	v4	tm	p	Puckering
1C	-35.5	15.0	9.80	-30.3	42.3	41.2	76.3	O4'-endo
2A	-22.8	35.9	-34.1	21.7	0.20	36.2	160.5	C2'-endo
3A	-15.2	30.7	-34.0	26.4	-7.30	34.2	173.4	C2'-endo
4A	-28.2	36.5	-30.7	15.5	7.70	35.9	149.0	C2'-endo
5A	-11.8	27.6	-31.8	26.0	-9.10	31.8	177.5	C2'-endo
6C	-30.9	36.5	-28.4	11.0	12.3	36.2	141.9	C1'-exo

Table S6 Sugar puckering of strand 1 of hairpin 2.

Base	v0	v1	v2	v3	v4	tm	P	Puckering
1G	2.0	16.7	-27.6	29.3	-20.1	29.8	202.2	C3'-exo
2T	-33.6	41.7	-34.3	15.8	10.9	41.3	146.3	C2'-endo
3T	-33.4	36.8	-26.6	7.8	15.9	36.8	136.3	C1'-exo
4T	-25.3	33.8	-28.7	14.5	6.6	33.2	149.9	C2'-endo
5T	-23.3	34.0	-31.1	18.3	2.8	34.0	156.4	C2'-endo
6G	-27.4	37.8	-35.1	20.1	4.4	38.6	155.3	C2'-endo

Table S7 Sugar puckering of strand 2 of hairpin 2.

Base	v0	v1	v2	v3	v4	tm	P	Puckering
1C	-39.1	37.0	-21.1	-0.9	24.7	39.2	122.5	C1'-exo
2A	-34.7	32.1	-18.2	-1.3	22.4	34.6	121.7	C1'-exo
3A	-15.8	30.6	-33.2	25.1	-6.2	33.5	171.6	C2'-endo
4A	-25.5	32.3	-27.1	13.2	7.3	31.9	148.1	C2'-endo
5A	-9.4	29.2	-37.0	32.7	-14.9	37.1	184.5	C3'-exo
6C	6.2	-20.6	26.5	-23.4	11.1	26.6	5.4	C3'-endo

Table S8 K_D values and related association (k_{on}) and dissociation (k_{off}) constant values calculated after global curve fitting using 1:1 model.

Compound	K_D ($s^{-1}M^{1-}$)	k_{on} ($s^{-1}M^{-1}$)	k_{off} ($s^{-1}M^{-1}$)
26	3.30×10^{-7}	5.43×10^{3}	1.79×10^{-3}
28	3.80×10^{-7}	2.65×10^{4}	1.01×10^{-2}
29	3.51×10^{-7}	4.35×10^{4}	1.53×10^{-2}

7.2.2 Materials and Methods

7.2.2.1 General methods for synthesis and analysis for small molecule intermediates, foldamers and conjugates

Reagents were used as purchased from commercial sources without further purification. Column chromatography purifications were performed on silica gel (230-400 mesh, 40-63 µm, Merck). Thin-layer chromatography was performed on silica gel plates (60-F254, Merck). Reactions requiring anhydrous conditions were performed under nitrogen dried over $CaCl_2$, with commercially anhydrous solvents unless states otherwise. Anhydrous THF and DCM for solid- and solution-phase synthesis was dispensed from a *MBRAUN SPS-800* solvent purification system using alumina columns for drying. $CHCl_3$, DIPEA, and NEt_3 were freshly distilled over CaH_2 under N_2-atmosphere.

NMR spectra were recorded on the spectrometers *Avance III HD 400 MHz Bruker BioSpin* and *Avance III HD 500 MHz Bruker Biospin* for 400 MHz and 500 MHz. $CDCl_3$ (δ_H: 7.26, δ_C: 77.0), DMSO-d_6 (δ_H: 2.50, δ_C: 39.4) and D_2O (δ_H: 4.79) were used as solvents and their residual solvent signals were used as internal standards.[41] NMR spectra of the oligomers were recorded in H_2O/D_2O 9:1 (vol/vol), 50 mM NH_4HCO_3 at 298 K. The raw data was evaluated using *Mnova version 14.0.0* by *Mestrelab Research*. The derived data signals are stated with chemical shift in ppm, their multiplicity (s = singlet, bs = broad singlet, d = doublet, t = triplet, q = quadruplet, m = multiplet, or a combination of these), their coupling constant in Hz and their integrated values. Water suppression was performed with excitation sculpting. Reactions requiring anhydrous conditions were performed under nitrogen. Analytical RP-HPLC analysis and semi-preparative purifications in acidic conditions were performed on a Thermo Fisher Scientific Ultimate 3000 HPLC system using Macherey-Nagel Nucleodur C18 columns (4.6 × 100 mm, 5 µm and 10 × 250 mm, 5 µm). When using acidic conditions (TFA buffered system), 0.1% TFA was added to aqueous mobile phase (referred to as mobile phase A) and to acetonitrile (referred to as mobile phase B). When using basic conditions, the column used for analysis was Macherey-Nagel Nucleodur C18 (4.6 × 100 mm, 5 µm). The mobile phase was composed of 12.5 mM TEAA in water at pH 8.5 (A) and 12.5 mM TEAA in water: acetonitrile mixture (1:2, v/v) at pH = 8.5 (B) (*TEAA buffer system*) or 12.5 mM aqueous NH_4OAc buffer adjusted to pH = 8.5 with aqueous ammonia (A) and acetonitrile (B) *(NH_4OAc-buffer system*). For analytical runs a flow rate of 1.0 mL/min was applied; semi-preparative RP-HPLC purification were performed at a flow rate of 5.0 mL/min. UV absorbance was monitored at 254 and 300 nm with a diode array detector if not stated otherwise. LC-MS chromatograms for acidic conditions were recorded on a Thermo Scientific Dionex UltiMate 3000 equipped with a C18 gravity (2.1 x 50 mm) with a flow of 0.33 mL min-1. 0.1% of formic acid in water (A) and 0.1% of formic acid in acetonitrile (B) were used as mobile phase. For basic conditions (e.g. water soluble oligomers and DNA strands) the column used was a Kinetex C18 EVO column (2.1 x 50 mm, 1.8 µm) with 12.5 mM aqueous NH_4OAc buffer adjusted to pH 8.5 (A) and LC-

MS grade acetonitrile (B). Elution was monitored by UV detection at 254 and 300 nm with a diode array detector. The LC system was coupled to a micrOTOF II mass spectrometer by Bruker Daltonics and molecules were ionized by ESI (for molecules **22**, **22a**, **23**, **24**, **25**, **28**, **29**). Circular dichroism (CD) spectra were measured on Jasco J-810 spectrometer. Spectra were measured at 25 °C, over a wavelength range of 300–500 nm, with a response time of 1 s, a continuous scanning mode, sensitivity set to standard and a scanning speed of 50 nm/min for 2 scans. Both foldamer and conjugate were solubilized in NH_4HCO_3 (50mM, pH = 8.5) and were measured at Δε values were normalized for the number of quinoline units. UV-spectroscopic measurements for determining protein, nucleic acid or foldamer concentrations were performed on a NanoDrop™ OneC (ThermoFisher Scientific). Variable temperature UV-Vis measurements nucleic acid melting were conducted on a JASCO V-650 spectrometer at 260 nm using a 10 mm quartz cuvette with a scanning rate of 1°C/min or on a Jasco J-1500 spectrometer with a 2 mm cuvette.

7.2.2.2 General methods for nucleic acid synthesis and purification

DNA strands not containing X were purchased from Sigma-Aldrich and used without further purification. Phosphoramidites of canonical deoxyribonucleosides (Bz-dA-CE, Ib-dG-CE, Ac-dC-CE and dT-CE) were purchased from Sigma-Aldrich or LinkTech. UltraMild L-DNA phosphoramidites (beta-L-Pac-dA-CE, beta-L-iPr-Pac-dG-CE, beta-L-Ac-dC-CE, beta-L-dT-CE) were purchased from Glen Research. Oligonucleotides were synthesized on a 1 or 4 µmol scale using high Load Glen UnySupport™ or 3'-CPR II CPG (for 3' phosphate modifications) as solid supports using an DNA automated synthesizer (Applied Biosystems 394 DNA/RNA Synthesizer) with a standard phosphoramidite chemistry. For the linker incorporation, the linker was dissolved in MeCN/DCM (v/v; 1:9) to a 0.15 M concentration and the coupling times were extended to 310 s. Oligonucleotides were synthesized in DMT-OFF mode using DCA as a deblocking agent in CH_2Cl_2, Activator 42® or ETT as activator in MeCN, Ac_2O (or Pac_2O for UltraMild phosphoramidites) as capping reagent in pyridine/THF and I_2 as oxidizer in pyridine/H_2O. The cleavage and deprotection of the CPG bound oligonucleotides were performed in an aqueous solution of 30% NH_4OH and heating at 50°C for 5 h. Subsequently, the supernatant was collected, and the beads were washed with water (3 × 0.3 mL). The combined aqueous solutions were concentrated under reduced pressure using a SpeedVac concentrator and subsequently lyophilized. The oligonucleotides were further purified by semi-preparative RP-HPLC using a 1260 Infinity II Manual Preparative LC System from Agilent (G7114A detector) equipped with the column VP 250/10 Nucleodur 100-5 C18ec from Macherey Nagel. A flow rate of 5 mL/min with varying gradients between 0% to 30-60% of buffer B in 45 min was applied for the purifications. The following buffer system was used: buffer A: 100 mM NEt_3/HOAc (pH 7.0) in H_2O and buffer B: 100 mM NEt_3/HOAc (pH 7.0) in H_2O/MeCN (v/v; 1:4). The purified oligonucleotides were analyzed by analytical RP-HPLC on a 1260 Infinity II LC System from Agilent (G7165A detector) equipped with the column EC

250/4 Nucleodur 100-3 C18ec from Macherey Nagel using a flow of 1 mL/min and gradients of 0% to 30-60% of buffer B in 45 min. Finally, the purified oligonucleotides were desalted using a C18 RP-cartridge from Waters. The absorbance of the synthesized oligonucleotides in H_2O solution were measured using an IMPLEN NanoPhotometer® N60/N50 at 260 nm. The extinction coefficients of the oligonucleotides were calculated using the OligoAnalyzer Version 3.0 from Integrated DNA Technologies. For oligonucleotides containing the foldamer linker an estimated extinction coefficient of 11000 $M^{-1}cm^{-1}$ was used for the calculations. The structural integrity of the synthesized oligonucleotides was analyzed by MALDI-TOF mass measurement. For this purpose, the synthesized oligonucleotides **13-21** (2-3 µL) were desalted on a 0.025 µm VSWP filter (Millipore), co-crystallized in a 3-hydroxypicolinic acid matrix (HPA, 1 µL) and measured on a Bruker Autoflex II. Biotinylated strands were measured on a microTOF II mass spectrometer by Bruker Daltonics and molecules were ionized by ESI.

7.2.2.3 Crystallography

7.2.2.3.1 Crystallization and data collection

The lyophilized powder of **13** D-d(GT_4G)–X–d(CA_4C) was dissolved using ultra-pure water such that the final concentration was 1mM. Crystallization screening was carried out by sitting drop vapor diffusion technique at 293 K. Multiple crystallization hits (> 10) were obtained from the screens and one condition that was reproducible was optimized for XRD studies. Colorless, prisms (Figure S4) were obtained within three days from drops prepared by mixing 1µL of **13** D-d(GT_4G)–X–d(CA_4C) and 1µL of crystallization reagent containing 10% v/v (+/-)-2-Methyl-2,4-pentanediol (MPD), 40 mM Sodium cacodylate trihydrate buffer (pH 6.0), 12 mM Spermine tetrahydrochloride, 80 mM Sodium chloride and equilibrated against a reservoir of 30% v/v MPD (50 µL). Crystals were fished using microloops and plunged into liquid nitrogen directly without further cryoprotection.

The diffraction experiment was performed on the beamline P13[42] operated by EMBL Hamburg, at the Petra III storage ring (DESY, Hamburg). Crystal characterization was carried out using 4 images at T = 100 K and λ = 0.775 Å, exposure time for 0.04 s and 1° oscillation per frame. However, the characterization revealed weak, DNA fiber diffraction-like pattern.

We eventually used the racemic crystallographic approach for DNA[25] to improve the quality of crystal and diffraction. The L-enantiomer of d(GT_4G)–X–d(CA_4C) **14** was thus synthesized and the lyophilized powder was dissolved in ultra-pure water. The racemic mixture of L/D-d(GT_4G)–X–d(CA_4C) was prepared by mixing the enantiopure solutions to a final concentration of 1 mM. Colorless needles (Figure S4) of **13/14** L/D- d(GT_4G)–X–d(CA_4C) were obtained within 7 days by hanging drop vapor diffusion technique at 293 K from drops prepared by mixing 1.25 µL of L/D-d(GT_4G)–X–d(CA_4C) and 1.25 µL of the reservoir solution (400 µL) containing 50% (v/v) Polyethylene glycol 200, 100 mM MES buffer (pH 6.5) and 200 mM Magnesium chloride. Crystals were fished using microloops and plunged into liquid nitrogen directly without further cryoprotection.

The diffraction experiment was performed on the beamline P14 operated by EMBL Hamburg, at the Petra III storage ring (DESY, Hamburg). Diffraction data was measured at T = 100 K and λ = 0.976 Å (12.7 keV) using a DECTRIS EIGER2 X 16M detector. The crystal was exposed for 0.08 s and 0.1° oscillation per frame and a rotation pass of 360°. Diffraction was processed using the *autoPROC* pipeline.[43-46] The crystal belonged to the centrosymmetric space group *P*-1 and selected crystal data and data collection parameters are summarized in Table S2. The crystallographic asymmetric unit consisted of two molecules.

7.2.2.3.2 Structure determination and refinement

The structure was determined by molecular replacement (MR) method using the program *Phaser*[47] within the *CCP4* suite.[48] The crystal structure of D-d(GT_4G)–Sd2–d(CA_4C) (PDB ID: 1PUY)[23] was used to build a search model of **13** such that the stilbene diether linker atoms were replaced by diquinoline linker atoms. During energy-minimization, full positional restraints were put on DNA atoms (for parameters, see modeling section). The stem residues were numbered G1, T2, T3, T4, T5, G6 in strand 1 and C8, A9, A10, A11, A12 and C13 in strand 2, and diquinoline linker X is residue 7.

The best MR solution was obtained with a translation-function Z (TFZ) score of 10.4 and a log-likelihood gain (LLG) score of 274 revealed two D-d(GT_4G)–X–d(CA_4C) hairpins (named hereafter Hairpins 1 and 2). Graphical analyses of the model and the electron density maps were carried out using *Coot*.[49] The initial maps showed clearly resolved electron density for the DNA duplex portions d(GT_4G), d(CA_4C); and the diquinoline linker X. The initial map also allowed unambiguous identification of the chirality of symmetry related D- and L- enantiomers (Figure S6).

Refinement of the coordinates was carried out initially with rigid-body refinement and subsequently in combination of translation libration screw (TLS) refinement[50] using d(GT_4G) and d(CA_4C) as a separate TLS groups and restrained refinement in the program *REFMAC5*[51] with maximum-likelihood targets and the *REFMAC5* dictionary.[52] 10% of the unique reflections were used to calculate R_{free}.[53] The planar conformation of the diquinoline linker moiety was modelled using σ-weighted omit and $2F_o$-F_c electron density maps. *PRODRG* was used for the generation of diquinoline linker ligand restraint dictionary[54] and used for refinement in *REFMAC5*. An example of the final electron density are depicted in Figure S7. Beside the synthetic hairpin DNA, four Mg^{2+} ions and 175 solvent molecules were identified. Two out of four Mg^{2+} hexahydrate coordination were identified. The complete model was subsequently refined to yield final values for the R_{factor} and R_{free} of 27.82 % and 32.79 %, respectively. Further refinement did not lead to a better convergence nor an improvement in the refinement statistics. The refinement *R*-values were high compared to typical macromolecular structure depositions. Poor refinement statistics for centrosymmetric crystal are inherent to centric data.[55]

The coordinates and structure factors have been deposited in the Protein Data Bank[56] (PDB

ID: 8Q60). Selected refinement parameters are given in Table S2. Figures for the structure were prepared using *PyMOL*.[57] Root mean square deviation (r.m.s.d.) values were determined using *Superpose*.[58] Structural parameters of DNA were calculated using *3DNA*.[59]

7.2.2.4 Molecular modeling

The molecular models of conjugates and DNA hairpins were built in Maestro (Version 11.5)[60] and energy-minimized with parameters in table S2. Hairpin models were built based on the crystal structure of stilbene-linked ds-DNA with the same stem d(GT_4G)–Sd2–d(CA_4C) (PDB ID: 1PU)[23] with full positional restraints on DNA atoms. Conjugate models were built based on the crystal structure of **13/14** and subsequently energy-minimized with full positional restraints on DNA atoms. Foldamer coordinates for the models of conjugates were derived from earlier crystal structures from diethylester-protected polyphosphonates, stripped of their side-chain protecting groups and energy-minimized according to parameters below.[1]

Table S 9 Minimization parameters used to build molecular models.

Forcefield	OPLS3
Solvent	Water
Charges from	Force Field
Cutoff	Extended
Method	PRCG or TNCG
Converge on	Gradient
Convergence threshold	0.01
Minimization mode	Minimization of non-conformers
Maximum iterations	20000

7.2.2.5 Recombinant protein expression and purification

The pET28a (+)-SAP1 plasmid for SAP1 (UniProt accession number: P28324, aa 1 - 93) with a N-terminal hexa-histidine tag overexpression was obtained from GeneScript.[37] Similarly, the pET28a (+)-PU.1plasmid for PU.1(UniProt accession number: P17947, aa 165-270) with a N-terminal hexa-histidine tag overexpression was obtained from GeneScript.[36] For overexpression and purification, we followed exactly the same protocol. The full-length proteins were expressed in *E. coli* BL21 plysS cells. Overnight pre-cultures in Luria broth (LB) supplemented with 50 µg/ml kanamycin were diluted 1000- fold with fresh 4L LB media and grown at 37 °C until OD600 reached 0.6. The expression was induced by addition of isopropyl 1-thio-β-D-galactopyranoside to make the final concentration 0.5 mM, and the culture was incubated for 15 hours at 22 °C. Next, the cells were harvested at 4000 rpm (J-LITE® JLA9.1000 Rotor, Beckman Coulter) and stored at −80 °C. The following purification steps were carried out at 4°C. Briefly, the cells were resuspended in 50 mL of 20 mM HEPES buffer pH 7.5, 500 mM NaCl containing 20 mM imidazole, and 1 mM phenylmethylsulfonyl fluoride (PMSF). The cells were lysed by sonication and the lysate was clarified by centrifugation at

18000 rpm (JA-25.5 Rotor, Beckman Coulter) for 40 min. The supernatant was equilibrated with nickel-nitrilotriacetic acid (Ni-NTA) agarose beads by gentle mixing for 1 hour. The mixture was then applied to a gravity flow column, drained out, and washed with 20mM HEPES,500mM NaCl, and 100mM imidazole. The protein was eluted with 20mM HEPES, 500mM NaCl, and 400mM imidazole, and immediately dialysed with 20mM HEPES, 150mM NaCl buffer. For storage, the protein was concentrated, frozen in liquid nitrogen and stored in a -80 °C freezer and thawed prior to BLImeasurements.

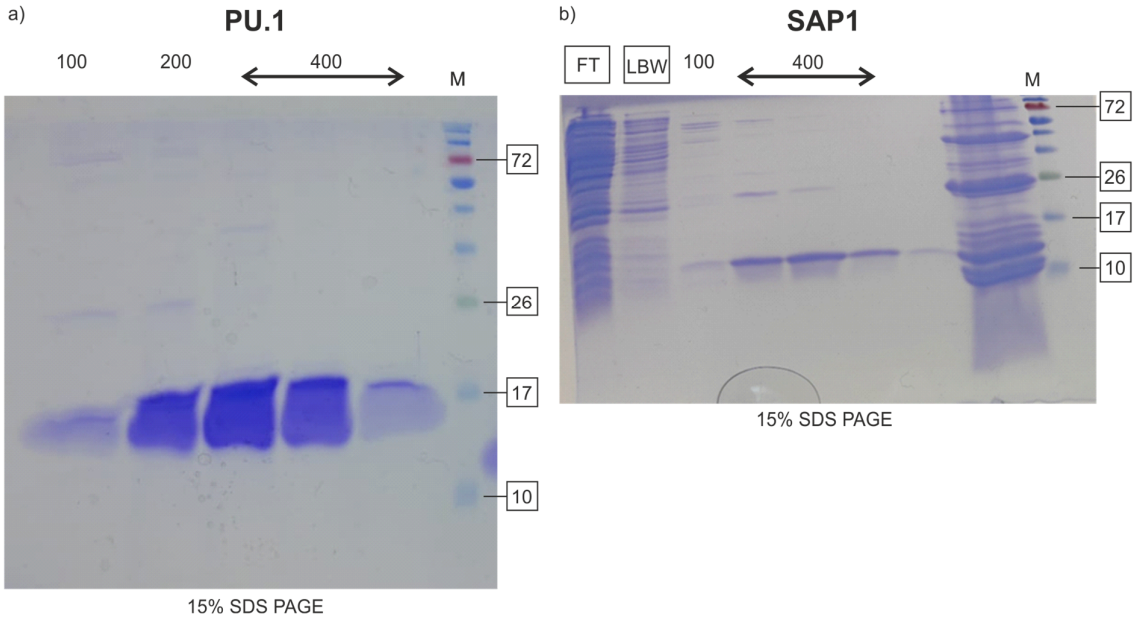
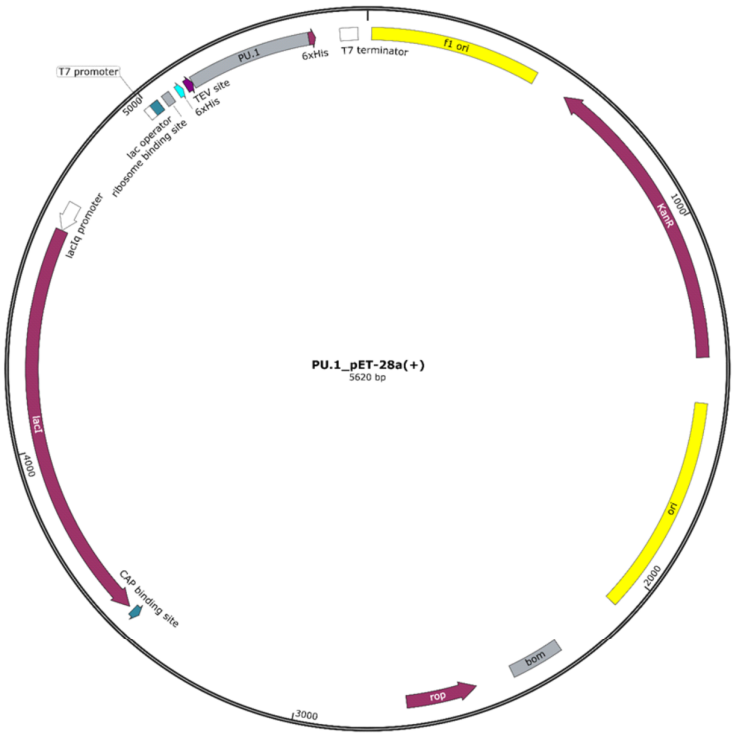


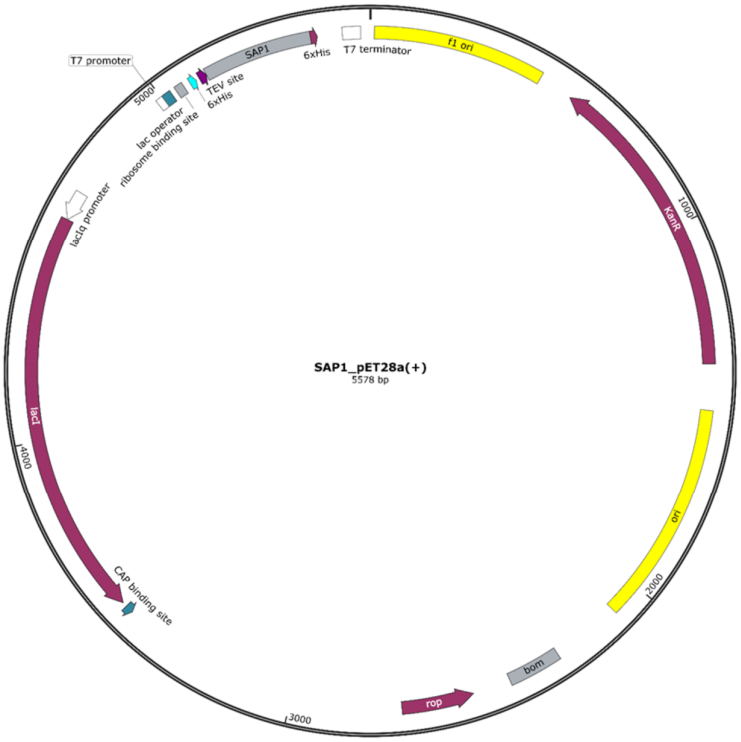
Figure S17. a) SDS-PAGE of PU.1 analysis. lane1: lysis buffer+100mM imidazole wash; lane 2: lysis buffer + 200 mM imidazole wash; lane 3-5: lysis buffer + 400 mM imidazole elution; lane 6: molecular weight marker (kDa). b) SDS-PAGE of SAP1 analysis. lane1: flow through (FT); lane 2: lysis buffer wash (LBW); lane3: lysis buffer + 100mM imidazole wash, lane 4-7: lysis buffer + 400mM Imidazole elution; lane 8: different sample; lane 9: molecular weight marker (kDa).



ATGGGCAGCAGC CATCATCATCATCATCAC AGCAGCGGC GAGAATCTTTATTTTCAGGGC CATATGGGTAGCAAGAAGAAGATTCG
TCTGTATCAGTTCCTGCTGGACCTGCTGCGTAGCGGCGACATGAAAGATAGCATTTGGTGGGTGGACAAAGACAAGGGTACCTTC
CAGTTTAGCAGCAAACACAAAGAGGCGCTGGCGCACCGTTGGGGTATCCAGAAAGGCAACCGTAAGAAAATGACCTACCAAAAG
ATGGCGCGTGCGCTGCGTAACTATGGTAAAACCGGCGAAGTGAAGAAAGTTAAGAAAAAGC TGACCTACCAATTTAGCGGTGAA
GTGCTGGGTCGTGGTGGTCTGGCGGAGCGTCGTCATCCGCCGCAT

MGSSHHHHHHSSGENLYFQ|GHMGSKKKIRLYQFLLDLLRSGDMKDSIWWVDKDKGTFQFSSKHKEALAHRWGIQKGNRKKMTYQK
MARALRNYGKTGEVKKVKKKLTYQFSGEVLGRGGLAERRHPPH

Figure S18. pET28a(+) plasmid map (top), coding sequence (middle), and "PU.1" protein expression product (bottom) used in this study; | indicates the TEV protease cleavage site.



ATGGGCAGCAGC CATCATCATCATCATCAC AGCAGCGGC GAGAATCTTTATTTTCAGGGC CATATGGACAGCGCGATCACCCTGTGGC
AATTCCTGCTGCAGCTGCTGCAAAAGCCGCAGAACAAACACATGATTTGCTGGACCAGCAACGACGGTCAATTTAAGCTGCTGCAAG
CGGAGGAAGTGGCGCGTCTGTGGGGCATCCGTAAGAACAAACCGAACATGAACTATGATAAGCTGAGCCGTGCGCTGCGTTACTACT
ACGTGAAGAACATCATTAAGAAAGTTAACGGTCAGAAGTTCGTGTATAAATTTG TTAGCTACCCGGAGATTCTGAACATG

MGSSHHHHHHSSGENLYFQ|GHMDSAITLWQFLLQLLQKPQNKHMICWTSNDGQFKLLQAEEVARLWGIRKNKPNMNYDKLSRALRYY
YVKNIIKKVNGQKFVYKFVSYPEILNM

Figure S19. pET28a(+) plasmid map (top), coding sequence (middle), and "SAP1" protein expression product (bottom) used in this study; | indicates the TEV protease cleavage site.

7.2.2.6 Bio-layer interferometry

7.2.2.6.1 Biotinylated ssDNA immobilization and dsDNA hybridization on streptavidin (SA) bionsensors and kinetic analysis with PU1 and SAP1proteins as analytes.

BLI measurements were carried out at 25 °C on an Octet R8 BLI Sartorius instrument, using streptavidin (SA) biosensors. To perform the ssDNA immobilization and hybridization, we took inspiration from the paper of Yang *et al.* were the authors reported SPR studies of DNA hybridization and protein-DNA binding.[61] The binding strongly depends on NaCl concentration and that [NaCl] was adjusted so that the binding and release kinetics can conveniently be monitored by BLI. All the steps, for setting-up these specific experiments, are exemplified in the figure S14. The buffer for the assembly of biotinylated ssDNA and dsDNA hybridization was PBS. After a baseline in buffer for 60 s (*A*), the biotinylated ssDNA **26** was

loaded on the biosensors (6) with a loading at 0.5 μM over 60 sto reach a 1-0.8 nm shift (*B*). The biosensors were then washed with PBS buffer for 60 sec (*C*), and the subsequent target DNA hybridization of the immobilized ssDNA was performed with a 0.5 μM target DNA for 120 sec to yield a signal plateau (*D*). The biosensors were washed again in PBS (*E*) before to record a second baseline for 120 s in the kinetic buffer: 25 mM Na_2HPO_4 (pH 7.5) containing 250 mM NaCl, 1 mM EDTA, 0.05% tween 20 (*F*). The association was next recorded for 120 sec with a range of PU1 concentration from 400 nM to 25 nM (*G*). This protein, being known to be prone to aggregation, we did not try to go higher in concentration.[62] For the dissociation, the sensors were moved back and dipped in the wells were the baseline was recorded (H). A global curve fitting with a 1:1 binding model was next performed using the reprocessing software embedded with the BLI Octet R8 instrument. The same approach was followed for the control dsDNA **27**. In the case of protein SAP1 as analyte, the same two phases loading process of the dsDNA was followed and the protein was assayed for its binding to dsDNA ligand sequences **26** and **27**, in the same kinetic buffer, in a protein concentration ranging from 2 to 0.125 mM. A global curve fitting was next performed to calculate the K_D values.

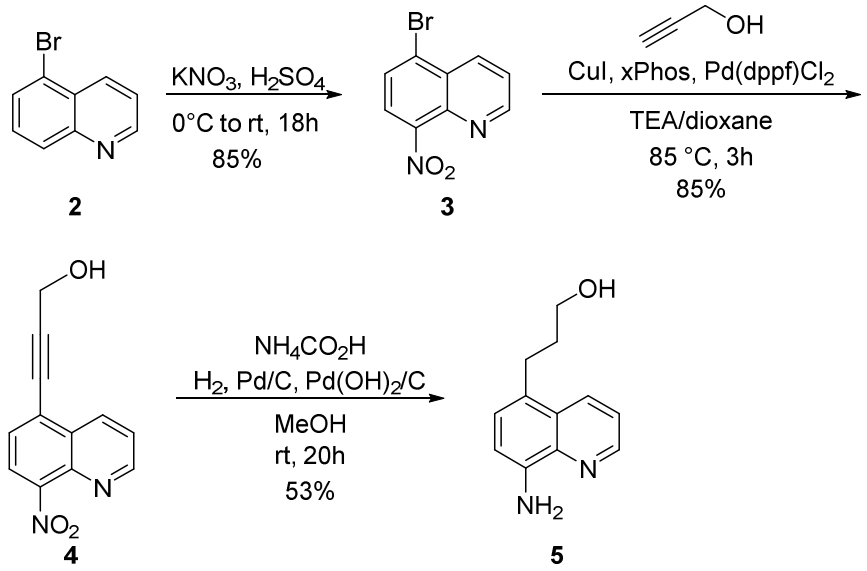
7.2.2.6.2 Loading of biotinylated hybrid hairpins 28 and 29 on SA-biosensors and kinetic assays

Biotinylated hairpins **28** and **29** were independently loaded on SA-biosensors and measurements for **29** are depicted in figure S15. After a baseline in 25 mM Na_2HPO_4 (pH 7.5) buffer (step *A*) containing 250 mM NaCl, 1 mM EDTA, 0.05% tween 20 for 60 sec, loading was performed at 2 μg/ml concentration of hybrid over 60 s (*B*). The sensors were then washed in the same buffer (*C*), a second baseline for 120 s was performed (*D*) and the association step was performed at different protein concentrations, from 500 nM to 25 nM for PUI and 2 μM to 0.125 μM for SAP1 using the same kinetic buffer (*E*). The dissociation step was recorded in the same well column were the second baseline was performed (*F*). The K_D values were calculated by global curve fitting with a 1:1 binding model using the reprocessing software embedded with the BLI instrument.

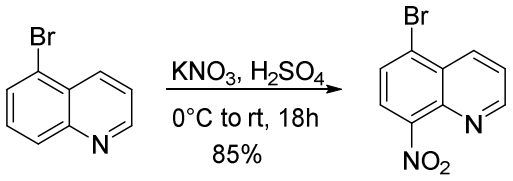
7.2.3 Synthetic procedures

7.2.3.1 Linker synthesis

7.2.3.1.1 Synthesis of 3'-terminal quinoline ring



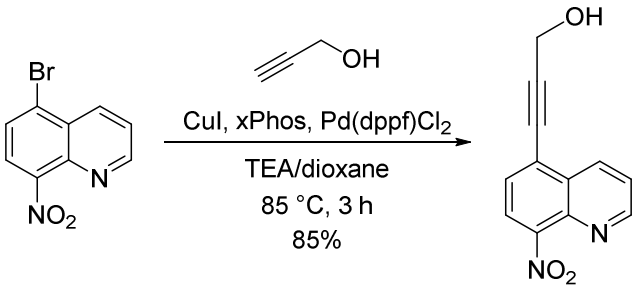
Compound **3**[63]



5-Bromoquinoline(**2**) (3.00 g, 14.4 mmol, 1.0 eq.) was dissolved in sulfuric acid (96%, 12.0 mL) and cooled to 0 °C. KNO_3 (2.33 g, 23.1 mmol, 1.6 eq.) was added portion wise over 15 min at 0°C and the reaction was stirred at rt for 22 h. The reaction mixture was poured on ice, filtered and washed with water. The product was lyophilized to give the title compound as a yellow solid (3.11 g, 12.3 mmol, 85%).

^{1}H NMR (500 MHz, DMSO-*d*$_6$) δ [ppm] = 9.12 (dd, *J* = 4.2, 1.5 Hz, 1H), 8.68 (dd, *J* = 8.6, 1.5 Hz, 1H), 8.26 (d, *J* = 8.0 Hz, 1H), 8.19 (d, *J* = 8.1 Hz, 1H), 7.90 (dd, *J* = 8.6, 4.2 Hz, 1H).

Compound **4**



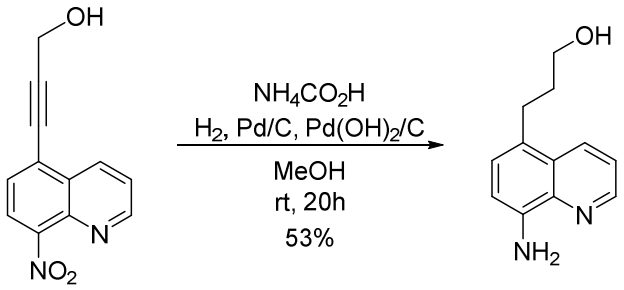
3 (1.74 g, 6.86 mmol, 1.0 eq.), XPhos (327 mg, 686 µmol, 0.1 eq.) and Copper(I) iodide (131 mg, 686 µmol, 0.1 eq.) were suspended in triethylamine (67.0 mL) and dioxane (67.0 mL) under N_2-atmosphere. The mixture was degassed for 10 min and propargyl alcohol (599 µl, 10.3 mmol, 1.5 eq.) and $Pd(dppf)Cl_2$ (502 mg, 686 µmol, 0.1 eq.) were added and the suspension further degassed for 10 min. The mixture was heated to 85 °C under reflux and N_2-atmosphere for 3 h.The solvent was removed *in vacuo*. The crude product was purified by column chromatography (EtOAc 50 → 70% EtOAc in cHex) to give the title compound as an orange solid (1.33 g, 5.83 mmol, 85%).

^{1}H NMR (500 MHz, DMSO-d_6) δ [ppm] = 9.11 (dd, *J* = 4.2, 1.7 Hz, 1H), 8.76 (dd, *J* = 8.5, 1.6 Hz, 1H), 8.28 (d, *J* = 7.8 Hz, 1H), 7.90 (d, *J* = 7.8 Hz, 1H), 7.87 (dd, *J* = 8.5, 4.2 Hz, 1H), 5.60 (t, *J* = 6.0 Hz, 1H), 4.50 (d, *J* = 6.0 Hz, 2H).

^{13}C NMR (126 MHz, DMSO-d_6) δ [ppm] = 153.3, 147.5, 138.3, 134.5, 129.5, 128.6, 124.3, 124.2, 123.05, 99.0, 79.2, 49.7.

HRMS: (ESI^+) m/z calc. for $C_{12}H_9N_2O_3$: 229.0608 $[M+H]^+$; found: 229.0606.

Compound **5**



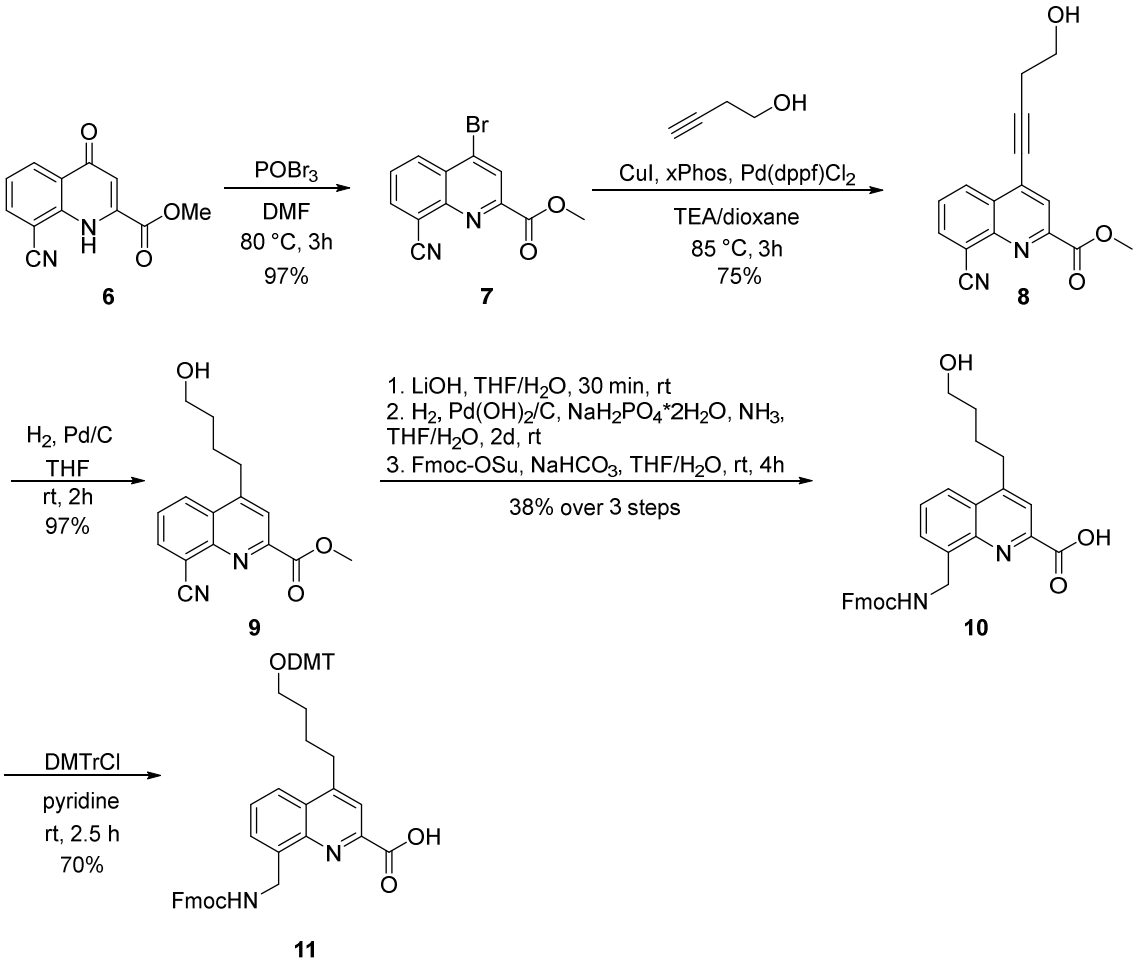
4 (767 mg, 3.36 mmol, 1 eq.) was dissolved in MeOH (150 mL) and degassed for 10 min by bubbling N_2 threw the solution. Pd/C (10% w/w, 150.0 mg) and NH_4HCO_2 (2.12 g, 33.6 mmol, 10 eq.) was added and the black suspension was further degassed for 10 min. The reaction mixture was vigorously stirred under H_2-atmosphere at rt for 20h. The black solid was filtered over celite and washed with MeOH (150.0 mL). The solvent was removed *in vacuo*. The crude product was purified by automated reversed-phase column chromatography (C18, 5 → 100% MeCN in H_2O) and further purified by column chromatography (20 → 40% acetone in DCM) to give the title compound (360 mg, 1.78 mmol, 53%) as a yellow oil.

^{1}H NMR (500 MHz, DMSO-d_6) δ [ppm] = 8.72 (dd, *J* = 4.1, 1.6 Hz, 1H), 8.34 (dd, *J* = 8.5, 1.7 Hz, 1H), 7.48 (dd, *J* = 8.5, 4.1 Hz, 1H), 7.13 (d, *J* = 7.7 Hz, 1H), 6.79 (d, *J* = 7.7 Hz, 1H), 5.72 (s, 2H), 4.48 (t, *J* = 5.1 Hz, 1H), 3.45 (td, *J* = 6.3, 5.1 Hz, 2H), 2.94 – 2.84 (m, 2H), 1.72 (ddt, *J* = 9.2, 7.7, 6.4 Hz, 2H).

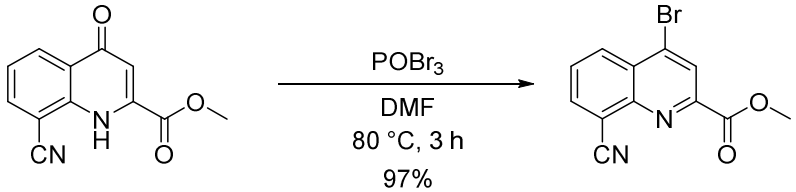
^{13}C NMR (126 MHz, DMSO-d_6) δ [ppm] = 146.4, 143.5, 137.9, 132.3, 127.1, 126.8, 124.3, 121.0, 108.3, 60.3, 34.3, 27.4.

HRMS: (ESI$^+$) m/z calc. for $C_{12}H_{15}N_2O$: 203.1179 $[M+H]^+$; found: 203.1177.

7.2.3.1.2 Synthesis of 5'-terminal quinoline ring



Compound **7**



6 (4.20 g, 16.8 mmol, 1.0 eq.) was dissolved in anhydrous DMF (50 mL) under N_2-atmosphere and $POBr_3$ (5.28 g, 18.4 mmol, 1.1 eq.) was added. The solution was vigorously stirred at 80 °C

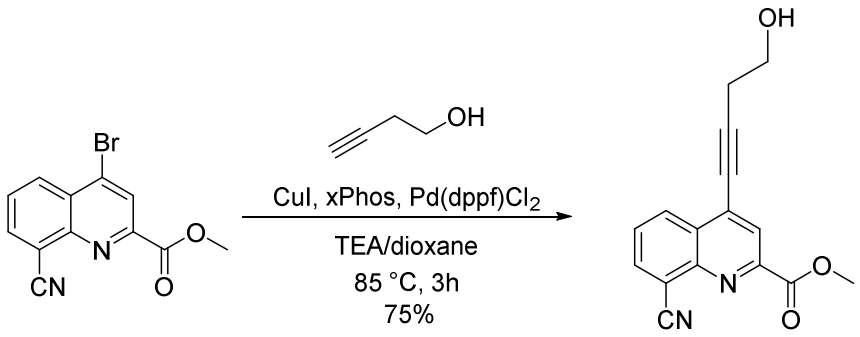
for 3 h under N_2-atmosphere to form a colorless precipitate. After cooling to rt, the reaction mixture was poured on ice and stirred for 30 min. The solid was filtered and washed well with H_2O until the washing solution gave a neutral pH and the residual solid dried in a vacuum oven at 60 °C over night (4.75 g, 16.3 mmol, 97%) as an off-white solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ [ppm] = 8.60 (s, 1H), 8.51 (dd, *J* = 8.5, 1.4 Hz, 1H), 8.27 (dd, *J* = 7.2, 1.4 Hz, 1H), 7.83 (dd, *J* = 8.6, 7.2 Hz, 1H), 4.10 (s, 3H).

^{13}C NMR (126 MHz, $CDCl_3$) δ [ppm] = 164.5, 149.6, 147.2, 137.43, 136.0, 131.85, 129.1, 126.8, 116.2, 115.0, 53.8.

HRMS: (ESI^+) m/z calcd. for $C_{12}H_7N_2O_2BrNa^+$ 312.9583 $[M+Na]^+$; found: 312.9582.

Compound **8**



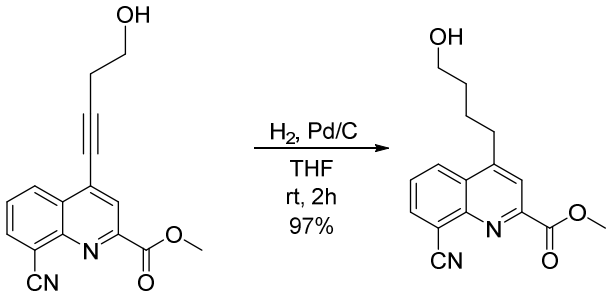
7 (4.75 g, 16.3 mmol, 1.0 eq.), XPhos (778 mg, 1.63 mmol, 0.1 eq.) and Copper(I) iodide (321 mg, 1.63 mmol, 0.1 eq.) were suspended in triethylamine (150 mL) and dioxane (150 mL) under N_2-atmosphere. The mixture was degassed for 10 min and but-3-yn-1-ol (1.85 mL, 24.5 mmol, 1.5 eq.) and $Pd(dppf)Cl_2$ (1.19 g, 1.63 mmol, 0.1 eq.) were added and the suspension further degassed for 10 min. The mixture was heated to 85 °C under reflux and N_2-atmosphere for 3 h to give a color gradient from yellow to orange to red to black.The solvent was removed *in vacuo*. The crude product was purified by column chromatography (80% EtOAc in cHex) and then precipitated from DCM (35 mL)/MeOH (35 mL) by slow evaporation of the DCM under reduced pressure. The title compound was isolated by filtration as an off-white solid (3.44 g, 12.2 mmol, 75%).

^{1}H NMR (500 MHz, $CDCl_3$) δ [ppm] = 8.55 (dt, *J* = 8.4, 1.4 Hz, 1H), 8.33 (s, 1H), 8.21 (dt, *J* = 7.2, 1.3 Hz, 1H), 7.75 (ddd, *J* = 8.4, 7.2, 1.3 Hz, 1H), 4.08 (s, 3H), 3.97 (q, *J* = 6.2 Hz, 2H), 2.90 (t, *J* = 6.3 Hz, 2H), 1.87 (t, *J* = 5.9 Hz, 1H).

^{13}C NMR (126 MHz, $CDCl_3$) δ [ppm] = 165.3, 149.6, 146.8, 136.9, 132.7, 131.1, 129.2, 128.21, 125.4, 116.6, 114.8, 100.4, 77.22, 60.9, 53.6, 24.3.

HRMS: (ESI^-) m/z calc. for $C_{16}H_{11}N_2O_3^-$ $[M-H]^-$: 279.0775; found: 279.0772.

Compound **9**



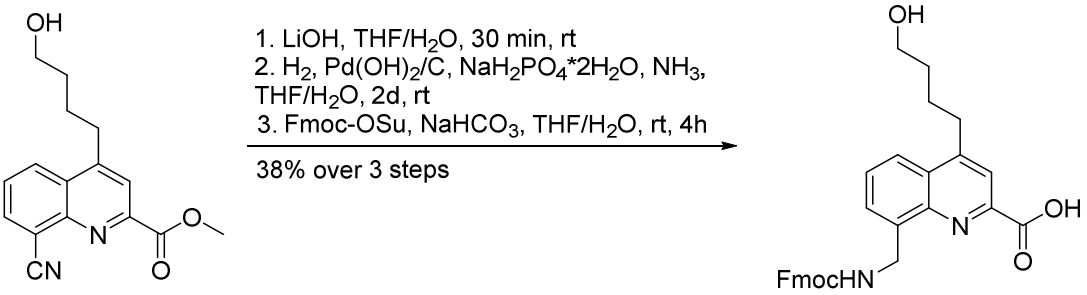
8 (3.44 g, 12.3 mmol, 1 eq.) was dissolved in THF (90 mL). The solvent was degassed by bubbling N_2 threw the solution for 15 min and put under N_2-atmosphere. To this, Pd/C (360 mg) was added and the N_2-atmopshere was replaced by H_2-atmosphere. The black suspension was stirred under H_2-atmosphere for 2 h at rt. The suspension was filtered through celite and the solvent was removed *in vacuo* then co-evaporated with toluene (1x) and Et_2O (1x) to give the title compound as a yellow solid (3.39 g, 11.9 mmol, 97%) as a yellow solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ [ppm] = 8.35 (dd, *J* = 8.5, 1.3 Hz, 1H), 8.26 – 8.10 (m, 2H), 7.73 (dd, *J* = 8.6, 7.2 Hz, 1H), 4.08 (s, 3H), 3.74 (td, *J* = 6.2, 4.8 Hz, 2H), 3.27 – 3.14 (m, 2H), 1.91 (tt, *J* = 9.4, 7.6 Hz, 2H), 1.81 – 1.67 (m, 2H), 1.32 (t, *J* = 5.0 Hz, 1H).

^{13}C NMR (126 MHz, $CDCl_3$) δ [ppm] = 165.9, 151.2, 149.6, 147.1, 136.3, 128.8, 128.7, 127.5, 122.3, 117.0, 115.2, 62.5, 53.5, 32.4, 32.1, 26.6.

HRMS: (ESI^-) m/z calc. for $C_{16}H_{11}N_2O_3^-$ $[M-H]^-$: 285.1088; found: 283.1086.

Compound **10**



9 (3.38 g, 11.9 mmol, 1 eq.) was dissolved in THF (200 mL). LiOH (370 mg, 15.5 mmol, 1.3 eq.) dissolved in water (100 mL) was added and the reaction mixture was stirred at rt for 30 min. The mixture was diluted with citric acid (5% w/v) and extracted with DCM (3x). The combined organic phases were dried over Na_2SO_4, filtered and the solvent was removed *in vacuo*, then coevaporated with toluene (1x) and Et_2O (1x). The resulting orange solid was used for the next step with no further purification. The orange solid was dissolved in THF (525 mL), NaH_2PO_4*$2H_2O$ (18.5 g, 119 mmol, 10 eq.) dissolved in water (240 mL) and NH_4OH solution (20% w/v, 6.80 mL, 35.7 mmol, 3.0 eq.) was added. The solution was degassed by bubbling N_2 threw the solution. To the mixture, $Pd(OH)_2$/C (20% loaded, 540 mg) was added and the

mixture was stirred under H_2-atmopshere at rt for 2d. To the mixture, $NaHCO_3$ (5.00 g, 59.5 mmol, 5 eq.) was added and the pH was checked to be neutral. To this Fmoc-Osu (3.30 g, 9.76 mmol, 0.82 eq., calculated based on HPLC-purity at 300 nm) dissolved in THF (50 mL) was added and the mixture stirred at rt for 4 h. The mixture was filtered threw celite, washed with THF (ca. 150 mL), concentrated to ca. 250 mL, diluted with H_2O to 500 mL and acidified to pH = 3 with citric acid (5% w/v). The mixture was extracted with DCM (3x), the combined organic phases dried over $MgSO_4$, filtered and the solvent was removed *in vacuo* and coevaporated with MeCN to give a dark green oil. The oil was dissolved in MeCN (15 mL) and precipitated under sonication. The compound was isolated by filtration and washed with MeCN and then further purified by column chromatography (silica, 1-3% MeOH/DCM + 0.1% AcOH) to give the title compound as a white solid (2.22 g, 4.47 mmol, 38%).

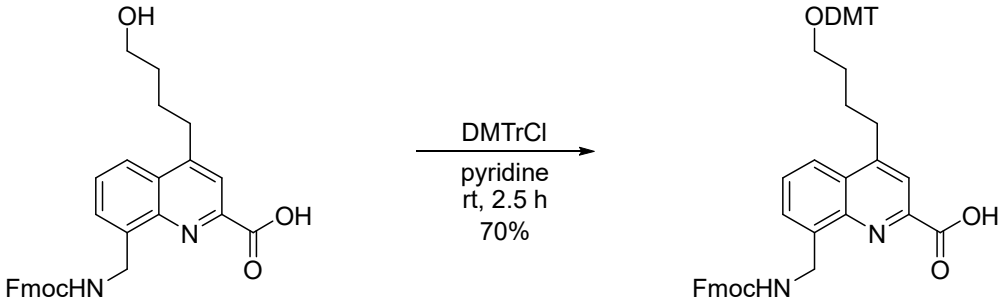
NMR does show 2 conformers ratio 85:15. Integrals are given with their respective integration. Overlapping NMRs are integrated as m and their integration is given.

^{1}H NMR (500 MHz, DMSO-*d*$_6$) δ [ppm] = 8.14 (dd, J = 8.5, 1.4 Hz, 1H), 8.01 (s, 1H), 7.91 (dd, J = 11.2, 6.8 Hz, 2.6H), 7.79 (d, J = 7.6 Hz, 0.3H), 7.73 – 7.67 (m, 2.6H), 7.64 (d, J = 7.8 Hz, 0.15H), 7.60 (dd, J = 7.0, 1.3 Hz, 0.85H), 7.53 (t, J = 6.5 Hz, 0.15H), 7.42 (t, J = 7.5 Hz, 1.8H), 7.35 (d, J = 7.6 Hz, 0.3H), 7.31 (td, J = 7.5, 1.2 Hz, 1.8H), 7.04 (t, J = 7.5 Hz, 0.3H), 4.87 (d, J = 6.2 Hz, 1.7H), 4.79 (d, J = 6.4 Hz, 0.3H), 4.44 (bs, 1H), 4.40 (d, J = 6.8 Hz, 1.7H), 4.30 (d, J = 6.6 Hz, 0.15H), 4.25 (t, J = 6.8 Hz, 0.85H), 4.13 (t, J = 6.8 Hz, 1.15H), 3.55 (m, J = 0.15H), 3.44 (t, J = 6.4 Hz, 2H), 3.17 (t, J = 7.7 Hz, 2H), 3.06 (q, J = 7.3 Hz, 0.15H), 1.80 – 1.69 (m, 2H), 1.57 – 1.49 (m, 2H), 1.25 – 1.19 (m, 1H).

^{13}C NMR (126 MHz, DMSO-d$_6$) δ [ppm] = 166.4, 156.6, 150.8, 147.02, 144.6, 143.9, 140.8, 138.7, 128.1, 128.0, 127.6, 127.5, 127.1, 125.2, 122.9, 120.2, 120.0, 65.3, 60.4, 46.9, 40.7, 32.2, 31.4, 26.5.

HRMS: (ESI$^+$) m/z calc. for $C_{30}H_{29}N_2O_5$: 497.2071 $[M+H]^+$; found: 497.2073.

Compound **11**



10 (295 mg, 594 µmol, 1.0 eq.) and DMTrCl (262 mg, 772 µmol, 1.3 eq.) were dissolved in anhydrous pyridine (1.8 mL) under N_2-atmosphere and stirred at rt for 3h. The solvent was removed *in vacuo*. The product was dissolved in DCM (50 mL) and the organic phase was washed with water (3x50 mL) and the combined aqueous phases were back extracted with

DCM (30 mL). The combined organic phases were dried over $MgSO_4$, filtered and the solvent was removed in vacuo. The crude product was purified by automated reversed-phase column chromatography (C18, 50 → 100% MeCN in H_2O) to give the title compound (332 mg, 416 µmol, 70%) as a white solid.

NMR does show 2 conformers ratio 85:15. Integrals are given with their respective integration. Overlapping NMRs are integrated as m and their integration is given.

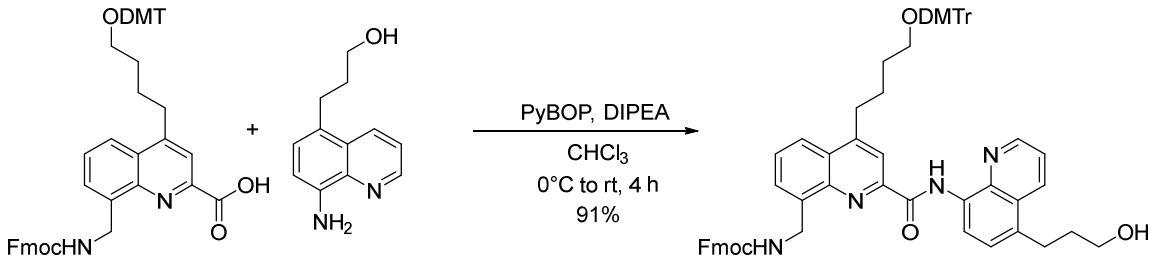
^{1}H NMR (500 MHz, DMSO-d_6) δ [ppm] = 13.10 (s, 1H), 8.08 (dd, *J* = 8.4, 1.6 Hz, 1H), 7.99 (s, 1H), 7.89 (d, *J* = 7.2 Hz, 2.5H), 7.75 (d, *J* = 7.7 Hz, 0.3H), 7.69 (d, *J* = 7.5 Hz, 1.7H), 7.67 – 7.57 (m, 1.7H), 7.59 – 7.50 (m, 0.3H), 7.41 (t, *J* = 7.5 Hz, 1.7H), 7.39 – 7.32 (m, 2H), 7.34 – 7.26 (m, 4H), 7.25 – 7.21 (m, 5H), 7.21 – 7.19 (m, 0.5H), 7.01 (t, *J* = 7.5 Hz, 0.3H), 6.91 – 6.83 (m, 4H), 4.87 (d, *J* = 6.2 Hz, 1.7H), 4.78 (d, *J* = 6.4 Hz, 0.3H), 4.40 (d, *J* = 6.8 Hz, 1.7H), 4.30 (d, *J* = 6.3 Hz, 0.3H), 4.25 (t, *J* = 6.7 Hz, 0.85H), 4.11 (t, *J* = 6.3 Hz, 0.15H), 3.71 (s, 6H), 3.10 (t, *J* = 7.7 Hz, 2H), 3.01 (t, *J* = 6.1 Hz, 2H), 1.78 (q, *J* = 8.0 Hz, 2H), 1.68 (q, *J* = 6.8 Hz, 2H).

^{13}C NMR (126 MHz, DMSO-d_6) δ [ppm] = 166.19, 157.97, 156.54, 150.64, 146.69, 145.19, 144.56, 143.87, 143.85, 140.78, 138.70, 135.96, 129.55, 128.04, 127.97, 127.79, 127.60, 127.58, 127.01, 126.56, 125.12, 122.79, 120.12, 119.90, 113.13, 85.21, 65.26, 62.18, 54.99, 46.85, 40.66, 31.23, 29.03, 26.68.

HRMS: (ESI^+) m/z calc. for $C_{51}H_{20}N_2O_7$: 799.3378 $[M+H]^+$; found: 799.3388.

7.2.3.1.3 Linker assembly & activation

Compound **12**



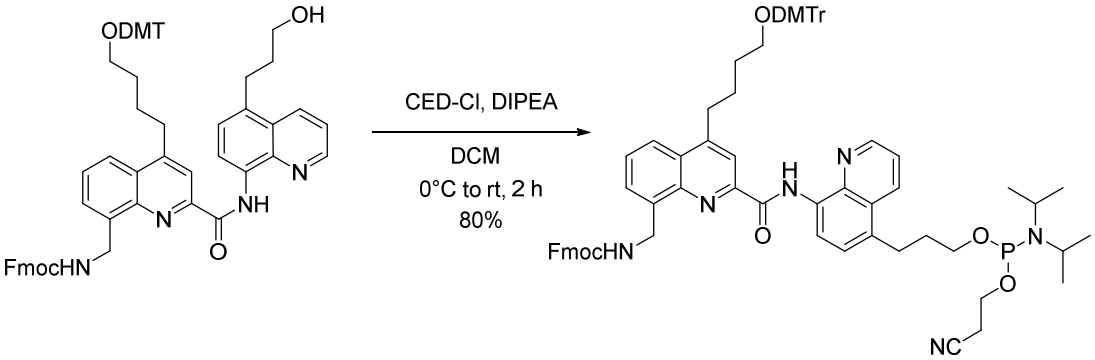
11 (315 mg, 394 µmol, 1.0 eq.) and PyBOP (410 mg, 789 µmol, 2.0 eq.) were dissolved in anhydrous chloroform (8.00 mL) under N_2-atmosphere and the solution was cooled to 0°C. **5** (87.7 mg, 0.43 mmol, 1.0 eq.) was dissolved in anhydrous chloroform and added to the reaction mixture. DIPEA was added and reaction mixture was stirred at rt for 4h. The solvent was removed *in vacuo* and the crude product was purified by column chromatography (60% EtOAc in cHex + 0.1% TEA) to give the title compound (351 mg, 357 µmol, 91%) as a pale-yellow solid. NMR shows 2 conformers ratio 85:15. Only major species in ^{13}C NMR listed.

^{1}H NMR (500 MHz, DMSO-d$_6$) δ [ppm] = 12.68 (s, 0.85H), 12.66 (s, 0.15H), 8.94 (d, *J* = 3.2 Hz, 2H), 8.73 (d, *J* = 7.8 Hz, 1H), 8.60 (d, *J* = 8.4 Hz, 1H), 8.18 (s, 1H), 8.11 (d, *J* = 6.3 Hz, 1H), 7.89 (t, *J* = 5.9 Hz, 0.85H), 7.86 (d, *J* = 7.6 Hz, 1.70H), 7.72 – 7.61 (m, 4.85H), 7.52 (d, *J* = 7.9 Hz, 1H), 7.47 – 7.43 (m, 0.30H), 7.41 – 7.34 (m, 3.80H), 7.29 (t, *J* = 7.6 Hz, 2.30H), 7.26 – 7.18 (m, 7H), 7.13 (t, *J* = 7.6 Hz, 0.30H), 6.87 (d, *J* = 8.8 Hz, 4H), 5.15 (d, *J* = 5.7 Hz, 1.70H), 5.06 – 5.01 (m, 0.30H), 4.58 (t, *J* = 5.1 Hz, 1H), 4.42 (d, *J* = 6.8 Hz, 1.70H), 4.35 (d, *J* = 6.1 Hz, 0.30H), 4.22 (t, *J* = 6.8 Hz, 0.85H), 4.07 (s, 0.15H), 3.71 (s, 6H), 3.51 (q, *J* = 5.8 Hz, 2H), 3.16 (t, *J* = 7.7 Hz, 2H), 3.08 (t, *J* = 7.8 Hz, 2H), 3.03 (t, *J* = 6.2 Hz, 2H), 1.88 – 1.78 (m, 4H), 1.72 – 1.66 (m, 2H).

^{13}C NMR (126 MHz, DMSO-d$_6$) δ [ppm] = 161.63, 157.97, 156.51, 151.58, 148.73, 147.86, 145.19, 143.84, 143.81, 143.70, 140.76, 138.71, 137.93, 135.97, 133.46, 133.26, 132.02, 129.56, 128.91, 128.04, 127.90, 127.79, 127.61, 127.54, 127.28, 126.93, 126.57, 126.52, 126.50, 125.11, 122.92, 122.07, 121.38, 120.06, 120.02, 119.80, 117.53, 115.20, 113.13, 109.76, 85.22, 65.30, 62.21, 60.16, 54.98, 46.86, 40.78, 34.03, 31.43, 29.07, 27.74, 26.71.

HRMS: (ESI$^+$) m/z calc. for $C_{63}H_{59}N_4O_7$: 983.4378 $[M+H]^+$; found: 983.4388.

Compound **1**



11 (439 mg, 447 mmol, 1.0 eq.) was dissolved in dry degassed DCM (4.5 mL) under argon atmosphere. DIPEA (311 mL, 1.79 mmol, 4.0 eq.) and CED-Cl (249 mL, 1.12 mmol, 2.5 eq.) were added at 0 °C and the reaction mixture was stirred for 2h. Subsequently the mixture was diluted with DCM (15 mL) and quenched with sat. sodium bicarbonate solution (10 mL). The aqueous phase was extracted with DCM (3×10 mL), the organic layers were combined, washed with brine (5 mL) and dried over Mg_2SO_4. The solvent was removed *in vacuo* and the product was purified by flash column chromatography (silica, EtOAc/nHex= 1:2) to yield **1** (421 mg 356 mmol, 80%) as a yellow powder after lyophilization from benzene (3×). $\mathbf{R_f}$ = 0.65 (EtOAc/nHex = 1:2)

^{31}P NMR (202 MHz, Acetone-d_6): δ(ppm) = 152.34, 143.69.

HRMS: (ESI$^+$) m/z calc. for $C_{72}H_{76}N_6O_8P$: 1183.5462 $[M+H]^+$; found: 1183.5322.

7.2.3.2 Foldamer synthesis

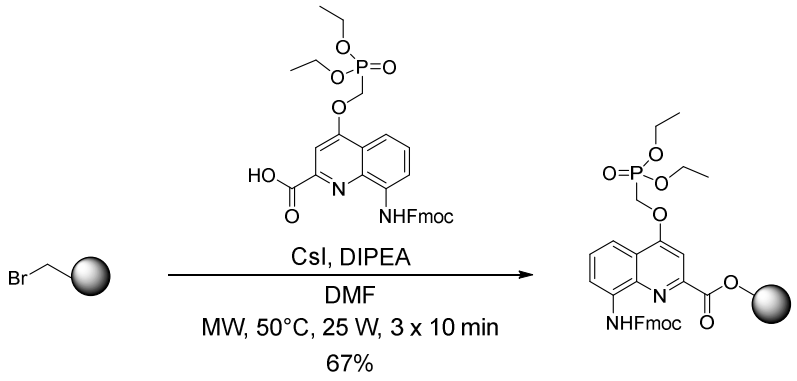
7.2.3.2.1 General solid-phase foldamer synthesis

Solid phase synthesis (SPS) was performed manually under MW-irradiation on a CEM Discover (Liberty Bio) microwave oven using a reaction vessel and an internal fiber optic probe for temperature control, or with a fully automated synthesizer followed by previously reported protocol.[2, 64]

- Bromination of low loading Wang resin

Low loading-(LL) Wang resin (Novabiochem, 100-200 mesh, 1.00 g, 0.37 mmol, 1.0 eq.) was swollen in DMF (6 mL) for 30 min. PPh_3 (970 mg, 3.70 mmol, 10.0 eq.) and CBr_4 (1.23 g, 3.70 mmol, 10 eq.) were quickly added and the suspension was stirred slowly for 20 h at RT. The resin was filtered, washed with DMF (3 × 3 mL) and DCM (3 × 3 mL), and dried by passing N_2 through the resin and stored at 4 °C until usage.

- Loading of the 1st monomer unit & resin loading estimation



LL-brominated Wang resin (Novabiochem, 100-200 mesh, 100 mg, 37.0 µmol, 1.0 eq.) was swollen in anhydrous DMF (3 mL) for 30 min under N_2. After washing with anhydrous DMF (3 mL), Fmoc-Q_4-COOH (42.5 mg, 74.0 µmol, 2.0 eq.) dissolved in anhydrous DMF (2 mL), CsI (19.2 mg, 74.0 µmol, 2.0 eq.) and freshly distilled DIPEA (9.56 mg, 12.9 µL, 74 µmol, 2.0 eq.) were quickly added and the reaction vessel was placed under microwave irradiation (25 W, ramp to 50°C over 5 min, hold at 50°C for 10 min) while bubbling N_2 through the solution. The resin was washed with anhydrous DMF (3 × 3 mL) and the process was repeated twice. The resin was washed with DMF (3 × 3 mL) and DCM (3 × 3 mL) and dried by passing N_2 through the resin. For resin loading estimation, a solution of 20% piperidine in DMF (v/v, 3 mL) was added to a known mass of the previously dried resin (1-2 mg) and agitated for 5 min. Meanwhile, the absorbance at 290 nm of the piperidine/DMF-solution was measured. After agitation, the absorbance of the solution that contained the resin was measured at 290 nm.

$$Resin\ loading\ \left[\frac{mmol}{g}\right] = \frac{[Abs_{final} - Abs_{initial}]}{[2.00 \times m_{resin}]}$$

- Fmoc deprotection

To the pre-swollen loaded Wang resin (75.0 mg, 0.26 mmol g^{-1}, 19.5 µmol), a 20% solution of piperidine in DMF (3 mL, v/v) was added and the resin was mixed by bubbling N_2-gas through the solution for 3 min. The resin was filtered and washed with DMF (2 x 3 mL) and the deprotection was repeated once for 7 min to give the respective amine NH_2-Q-Wang resin. The resin was filtered, washed with DMF (5 x 3 mL), and washed with anhydrous THF (3 x 3 mL) prior to coupling.

- *In situ* acid chloride activation, coupling, and capping

The NH_2-Wang resin (75.0 mg, 0.26 mmol g^{-1}, 19.5 µmol, 1.0 eq.) was suspended in anhydrous THF (0.9 mmol) and *2-4-6*-collidine (23.0 µL, 176 µmol, 9.0 eq.) was added. Concurrently, in a glass vial, **Monomer** (35 mg, 58.6 µmol, 3.0 eq.) and PPh_3 (41 mg, 156 µmol, 8 eq.) were mixed and dissolved in anhydrous $CHCl_3$ (0.9 mL). Subsequently, trichloroacetonitrile (18.0 µL, 176 µmol, 9 eq.) was added to the vial, which was quickly shaken and the mixture was added to the pre-swollen resin. After mixing, the reaction vessel was placed under microwave irradiation (25 W, ramp to 50°C over 5 min, hold at 50°C for 15 min). The resin was filtered off and washed with anhydrous THF (3 x 3 mL). The coupling step was repeated once. The resin was filtered off and washed with anhydrous THF (3 x 3 mL) and DCM (3 x 3 mL) prior to the capping step. The resin was suspended in a 50% solution of Ac_2O in DCM (v/v) and mixed by bubbling N_2-gas through the solution for 10 min. The resin was washed with DCM (2 x 3 mL) and DMF (2 x 3 mL).

- Resin cleavage

The resin-bound foldamer was placed in a syringe equipped with a filter and then suspended in TFA (3 mL). The resin was shaken for 2 h at RT. The resin was then filtered off and washed twice with TFA. TFA was removed *in vacuo* and the resulting solid was precipitated in cold Et_2O. The precipitate was centrifuged and the solvent was decanted to give a yellow solid. The decanted Et_2O was concentrated by rotary evaporation and the precipitation was repeated. The combined precipitates were dissolved in water/MeCN and then lyophilized to give the crude protected foldamer as a yellow solid and further purified as a diethyl-phosphonate protected compound.

7.2.3.2.2 Removal of the diethyl-phosphonate protecting groups

Purified protected foldamers were dissolved in anhydrous chloroform (1 mL per 10 mg compound) and cooled to 0 °C. TMSBr (0.2 mL per 10 mg compound) was diluted in anhydrous chloroform (0.8 mL per 10 mg compound). Diluted TMSBr-solution was added dropwise over 10 min to the reaction mixture. The reaction mixture was allowed to room

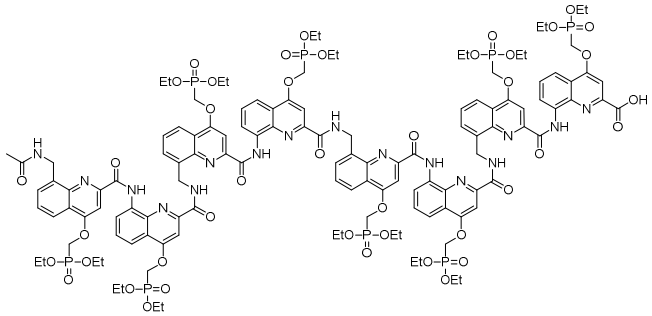
temperature and stirred under N_2-atmosphere for 1-3 d until an HPLC- and LCMS aliquot of the reaction mixture showed full cleavage of all phosphonate-diethyl esters. The reaction mixture was evaporated *in vacuo* (40°C water bath) to give a yellow oil, then coevaporated with DCM (2x) to give a yellow solid. The solid was suspended in water, basified to pH > 12 with triethylamine and stirred for 30 min. The suspension was filtered through nylon syringe filters to give a pale yellow solution that was freeze-dried to give the crude deprotected foldamers as yellow solids.

7.2.3.2.3 Cation exchange chromatography

Cation exchange chromatography was performed on Dowex© 50WX4 200-400 (H) resin. The resin was swollen in H_2O the orange solution was decanted. The resin was transferred into a column and washed with H_2O (all with gravity flow) followed by two column volumes (CV) of 2 M HCl solution, washed with H_2O until pH = 6-7 (ca. 5-10 CV), washed with two CV of 2 M NH_4OAc solution, washed again with 5 CV of H_2O. Purified triethylammonium salts of foldamers were dissolved in water and loaded on the column, the column was closed without flow for 2 h, the column was eluted with water (20 mL) and the compound was lyophilized to give the purified foldamer as NH_4^+-salt as yellow solid.

7.2.3.2.4 Synthesized foldamers

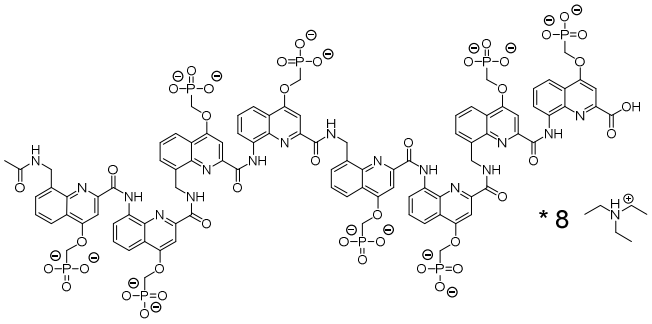
7.2.3.2.4.1 Compound **22a** protected acetylated octamer



Oligomer **22a** was synthesized on LL-Wang resin (23.2 µmol). The compound was isolated by semi-preparative RP-HPLC (acidic conditions, linear gradient 30-100% B in A) to give the title compound (35.0 mg, 12.5 µmol, 54%) as a yellow solid.

(ESI$^+$) m/z calc. for $C_{126}H_{150}N_{16}O_{42}P_8$: 1403.399 $[M+2H]^{2+}$; found: 1403.393.

7.2.3.2.4.2 Compound **22** unprotected acetylated octamer



Compound **22a** (30.0 mg, 13.7 µmol, 1 eq.) was deprotected according to section 4.2.2 of the SI. Stirring time: 2 d at rt. The crude foldamer was purified by semi-preparative RP-HPLC (TEAA buffers, 0-50% B in A) to give the title compound (30.2 mg, 12.5 µmol, 89%) as a octa-triethylammonium salt as a pale yellow solid. (ESI$^-$) m/z calc. for $C_{94}H_{82}N_{16}O_{42}P_8$: 1177.1342 $[M-2H]^{2-}$; found: 1177.1368.

Synthesis of **23** is described here.[2]

7.2.3.3 Conjugate synthesis and oligonucleotides

7.2.3.3.1 DNA hairpin biotinylation

Isolated desalted DNA strand **20** (1.83 mg, 276 nmol, determined by UV-spectroscopy) were dissolved in degassed H_2O/DMF (250 µL, 1:4, v/v). Biotin-PEG4-NHS (5 eq.) ester was dissolved in degassed DMF (50 µL) and added to the reaction mixture. The reaction was followed by consumption of the starting material of an aliquot of the reaction by RP-HPLC (basic conditions) after 1 day. If incomplete, further Biotin-PEG_4-NHS (25 eq.) in DMF (100 µL) was added. Upon completion, the reaction mixture was diluted with degassed H_2O (15 mL) and lyophilized to give the crude biotinylated DNA-strand that was purified by basic HPLC (TEAA buffer system) to give the respective biotinylated DNA-strand **28** that was used as a triethylammonium salt for further usage (1.31 mg, 185 nmol, 67%). For **21** the reaction was performed on (520 µg, 78.2 nmol) to yield **29** (182 µg, 25.6 nmol, 33%).

7.2.4 By EDC hydrochloride

Desalted DNA-strand (2.28 mg, 548 nmol, 1eq.) was dissolved in water (500 µL). A polytriethyl ammonium salt of respective foldamer (4.35 mg, 1.25 µmol, 2.3 eq.) was dissolved in water (100 µL). EDC hydrochloride (3.15 mg, 16.4 µmol, 30 eq.), N-hydroxysuccinimide (1.89 mg, 16.4 µmol, 30 eq.) and DIPEA (2.86 µL, 16.4 µmol, 30 eq.) were dissolved in water (50 µL). EDC hydrochloride, NHS and DIPEA were added to dissolved foldamer. The activated foldamer solution was added to the DNA-strand and the mixture was shaken at 700 rpm at

21 °C for 24 h. The mixture was directly injected into semi-preparative RP-HPLC for purification (basic conditions, gradient 0-30% A in B, NH_4OAc buffer system) to obtain the foldamer-DNA conjugate as an NH_4^+-salt as pale yellow solid (2.01 mg, 307 nmol, 56%).

7.2.4.1.1 By DMTMM hydrochloride

Desalted DNA-strand (493 µg, 118 nmol, 1eq.) and polytriethyl ammoniumsalt foldamer (1.88 mg, 544 nmol, 4.6 eq.) were dissolved in water (90 µL) and diluted with MES-buffer (100 µL, 300 mM, pH = 6) to a volume of 200 µL. DMTMM hydrochloride (983 µg, 3.55 µmol, 30 eq.) was dissolved in water (10 µL) and added and the mixture was shaken at 700 rpm at 20 °C for 24 h. The mixture was directly injected into semi-preparative HPLC for purification (gradient 0-20% B in A, NH_4OAc buffer system) to obtain the foldamer-DNA conjugate as an NH_4^+-salt as pale yellow solid (338 µg, 49.8 nmol, 42%).

7.2.5 Spectra & Chromatogramms

7.2.5.1 NMR spectra

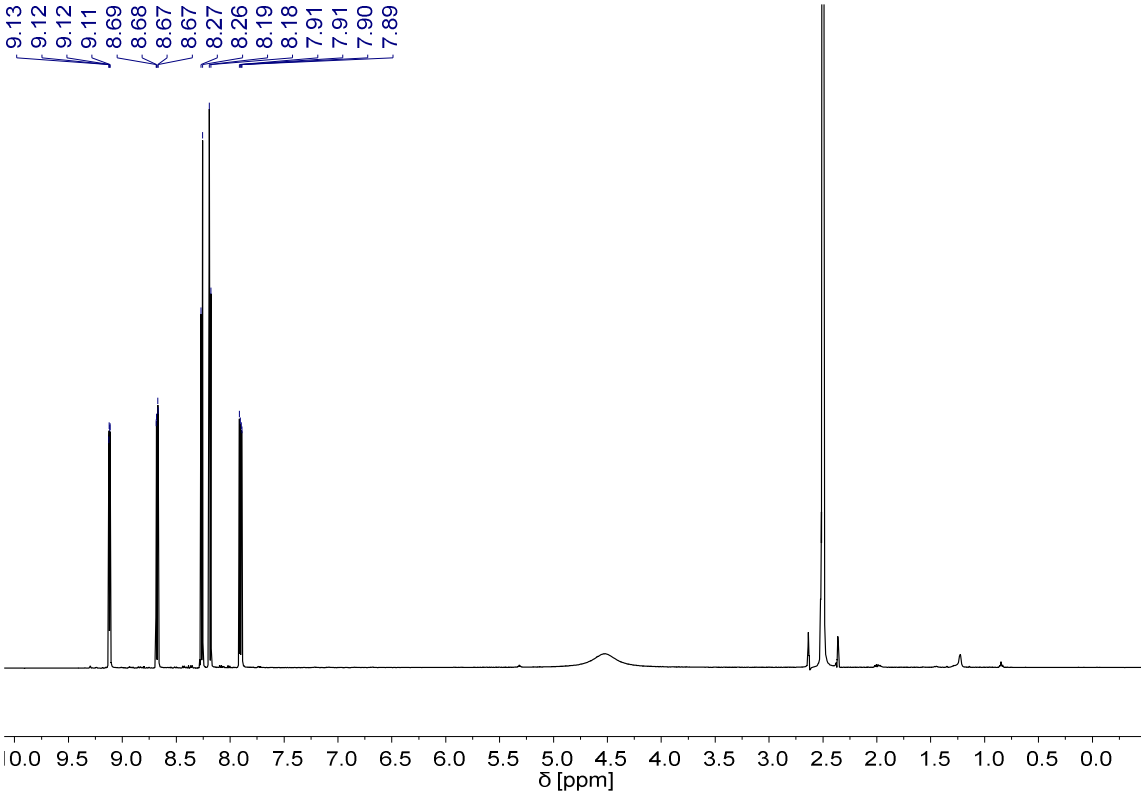


Figure S20 ^{1}H NMR spectrum of compound **3**.

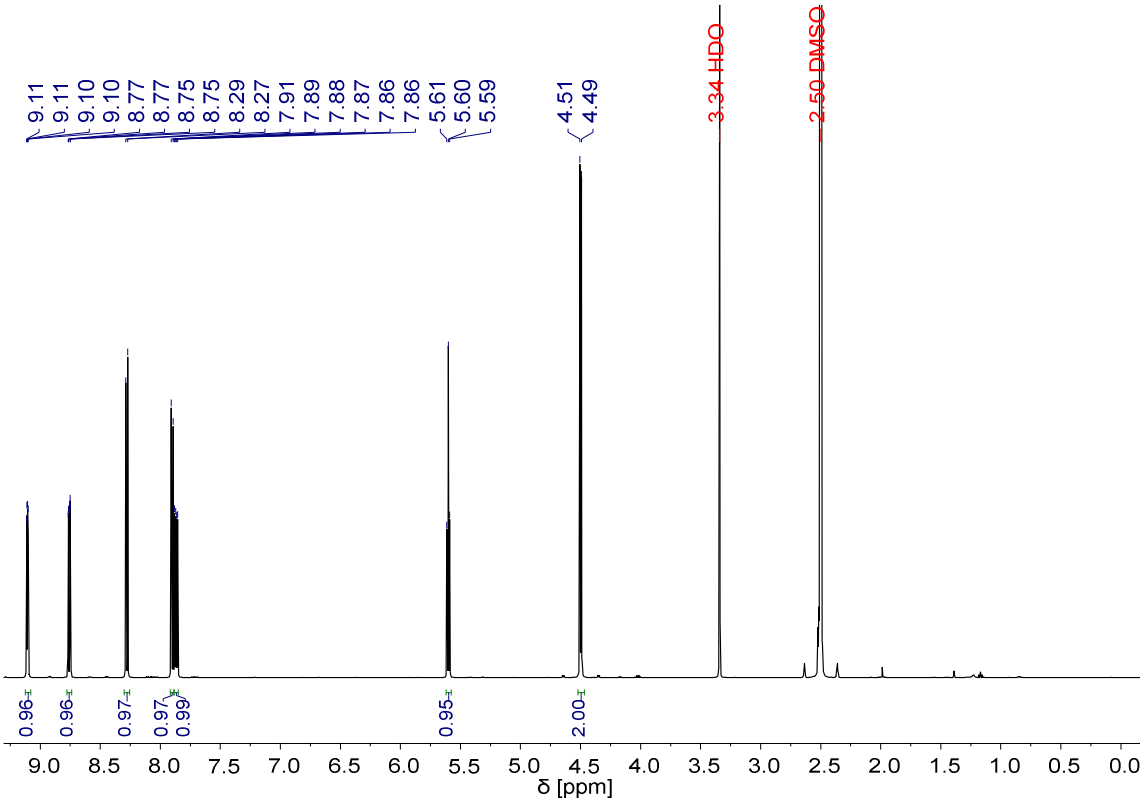


Figure S21 [1]H NMR spectrum of compound **4**.

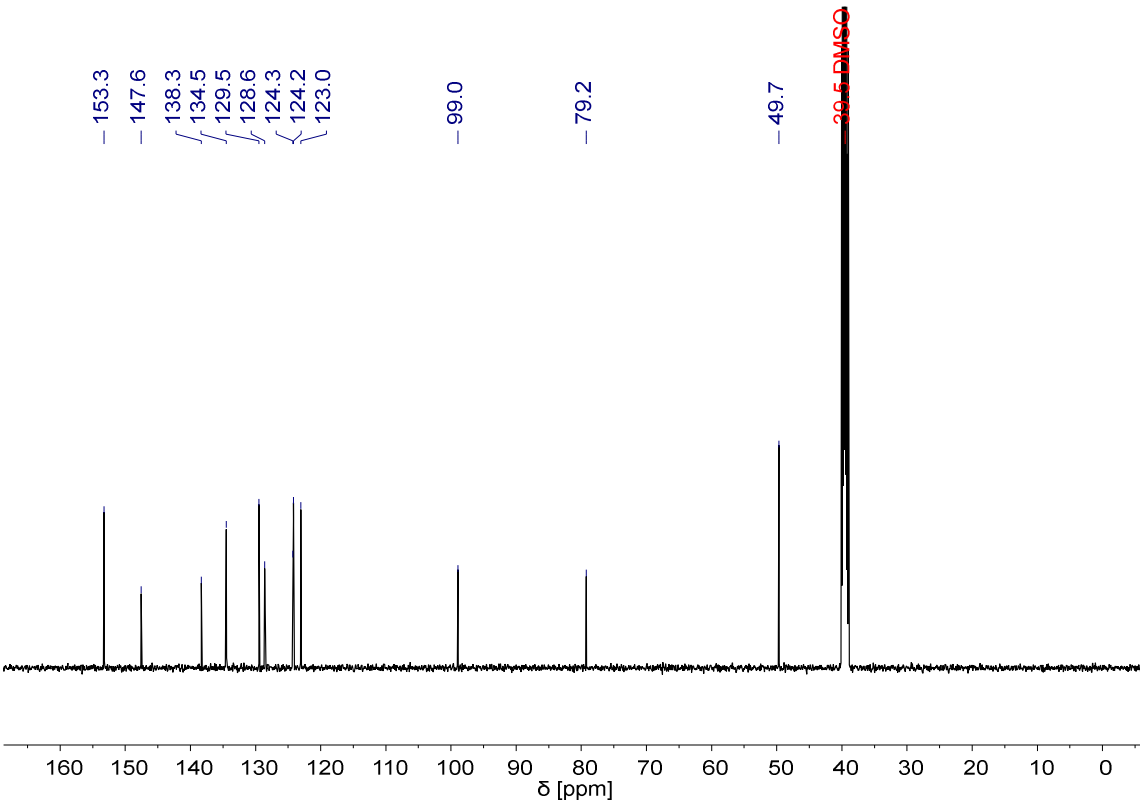


Figure S22 [13]C NMR spectrum of compound **4.**

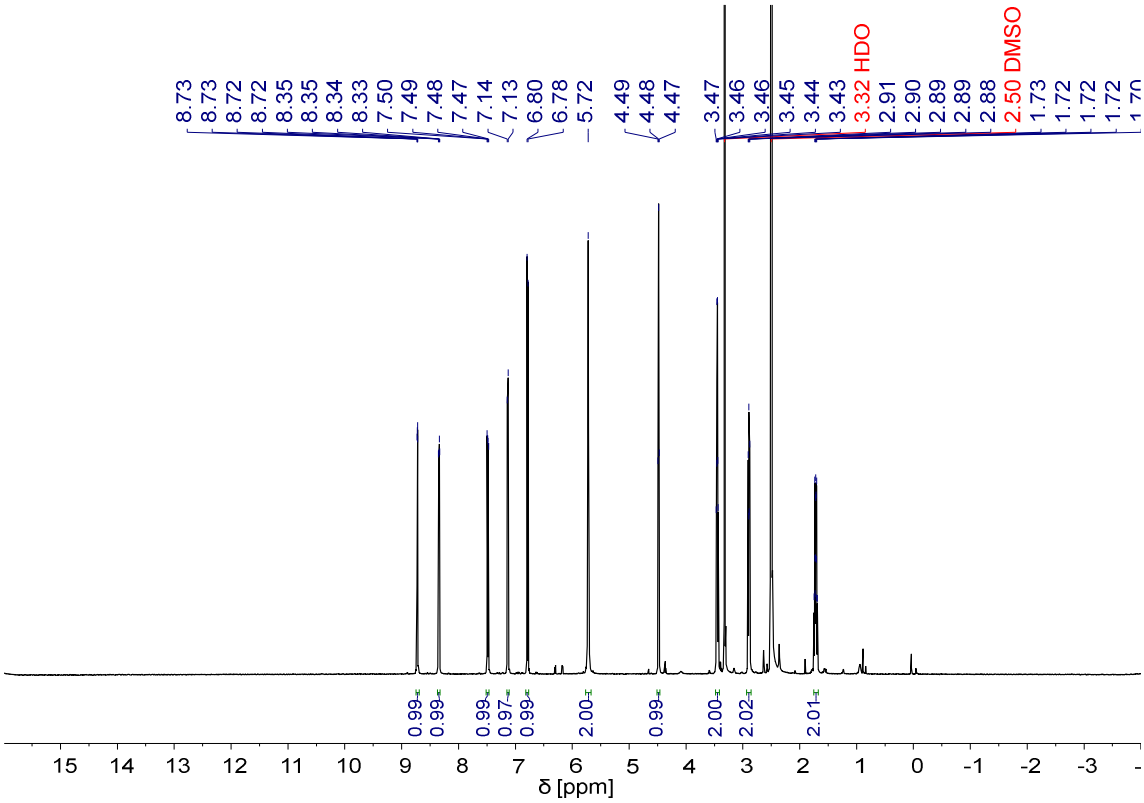


Figure S23 [1]H NMR spectrum of compound **5**.

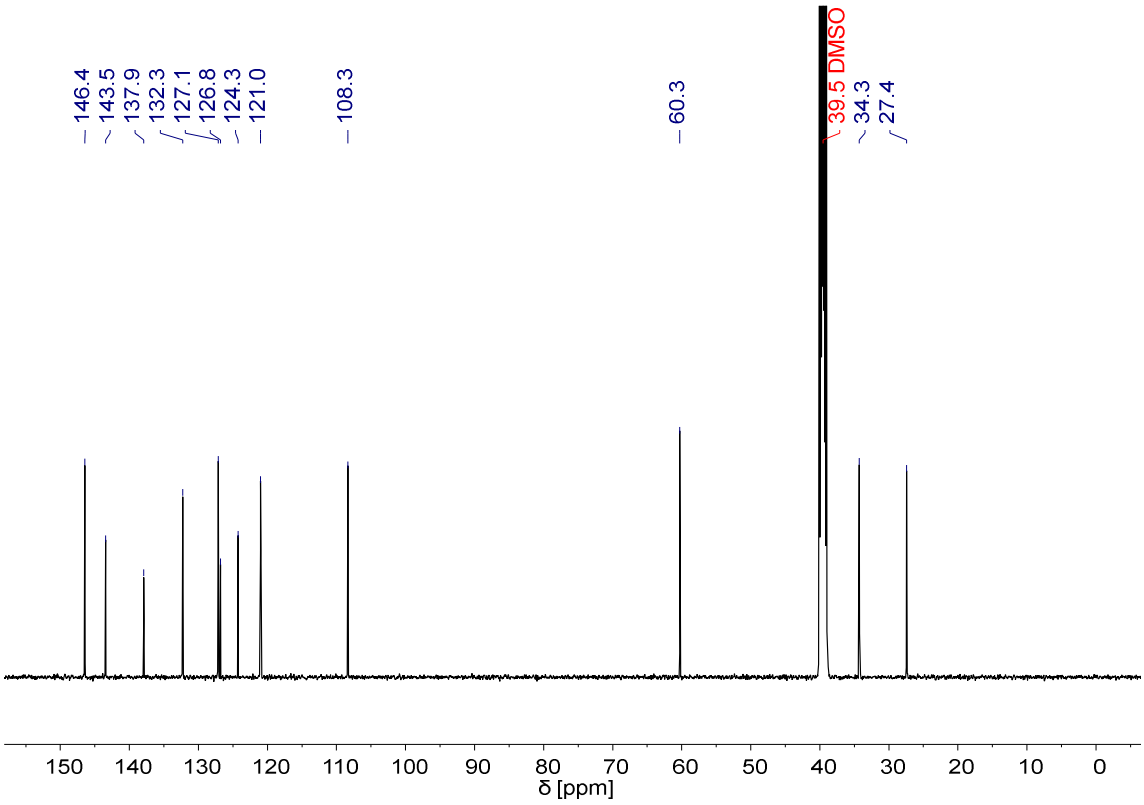


Figure S24 [13]C NMR spectrum of compound **5**.

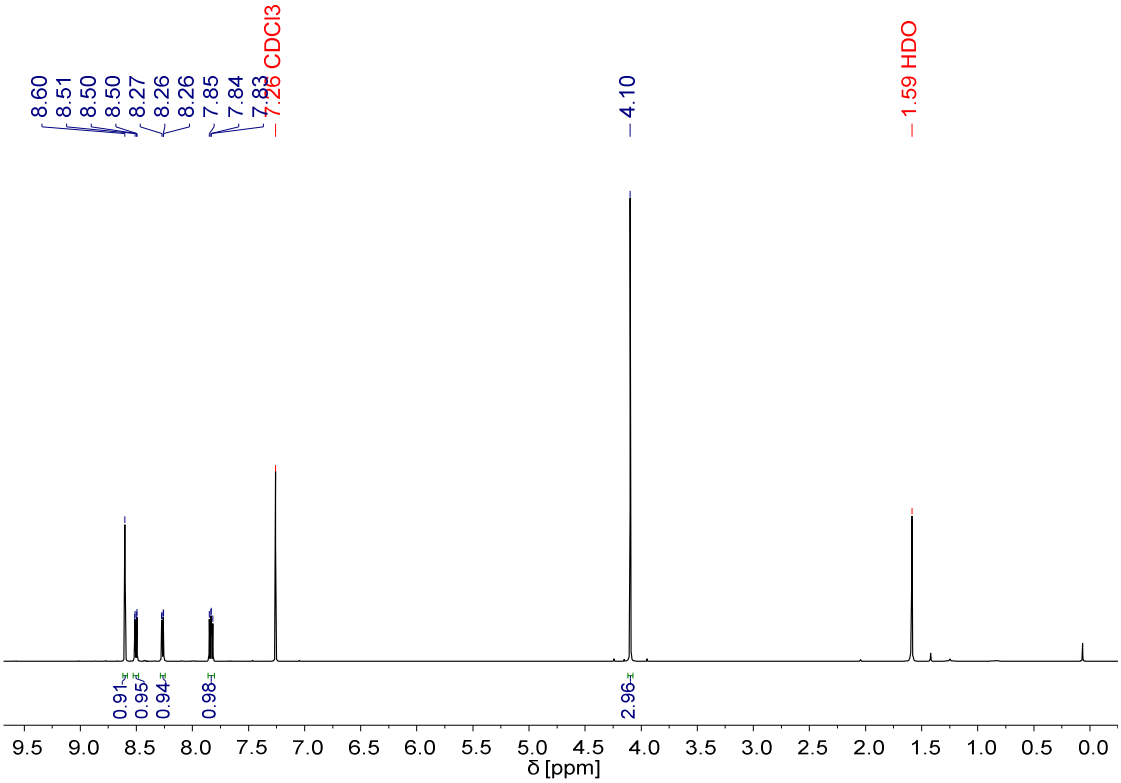


Figure S25 ^{1}H NMR spectrum of compound **7**.

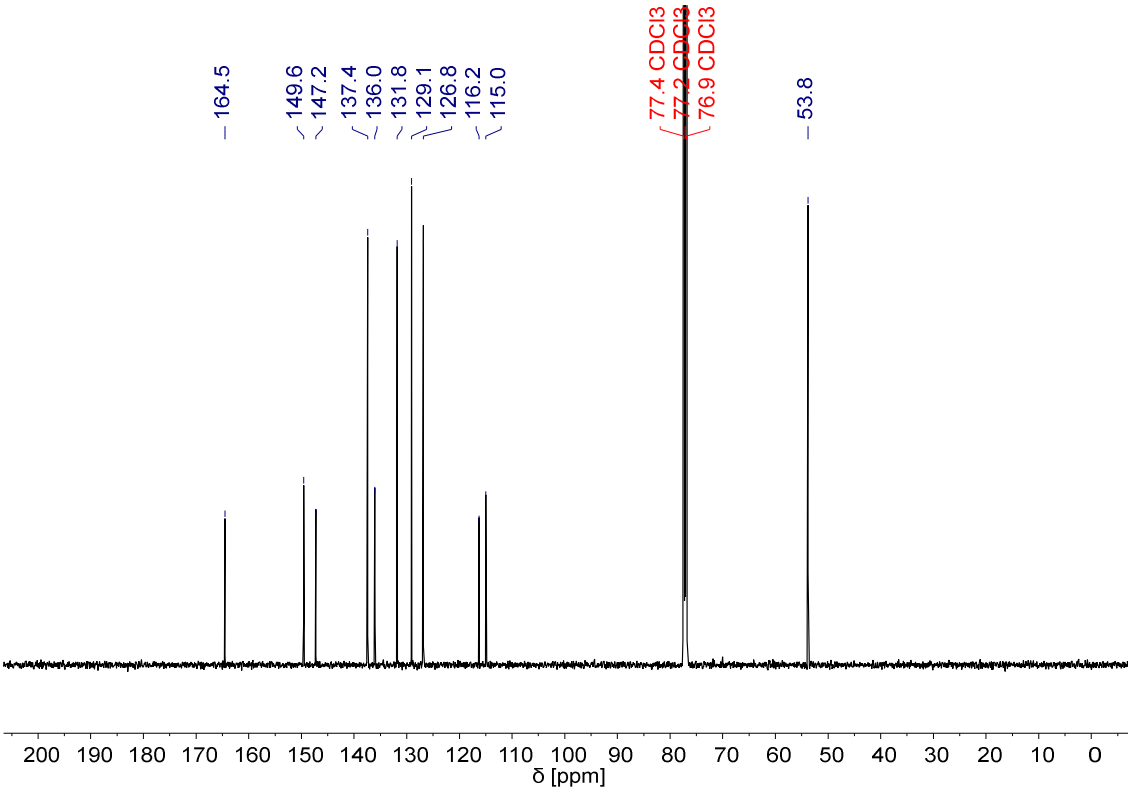


Figure S26 ^{13}C NMR spectrum of compound **7**.

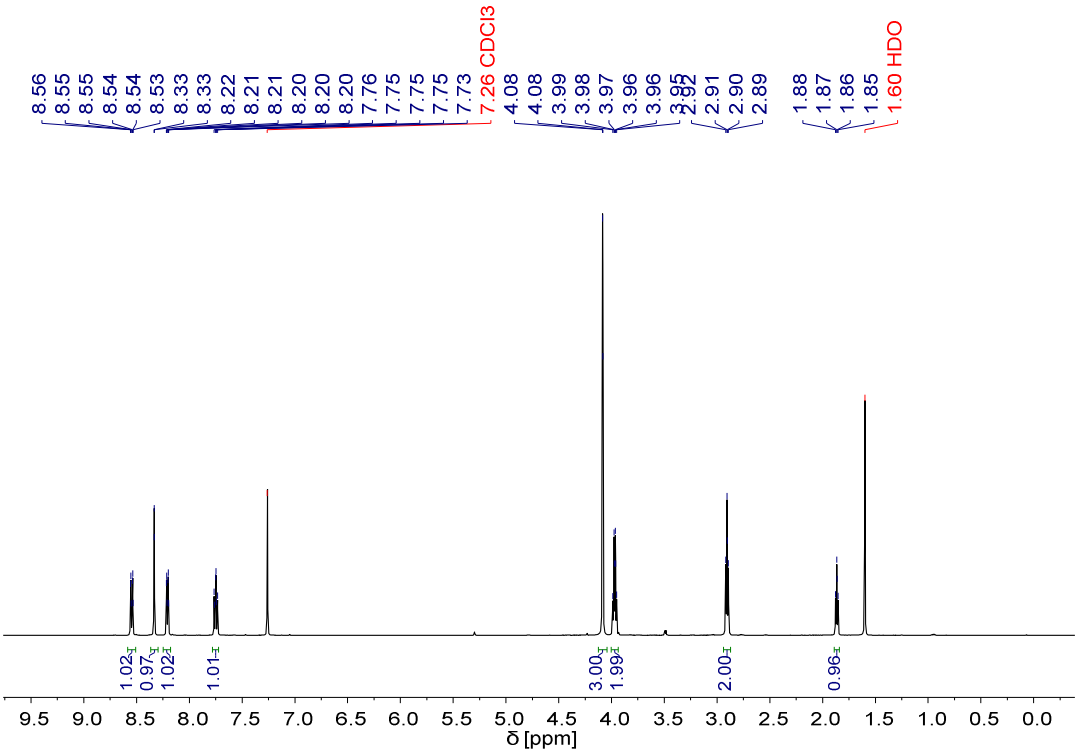


Figure S27 ^{1}H NMR spectrum of compound **8**.

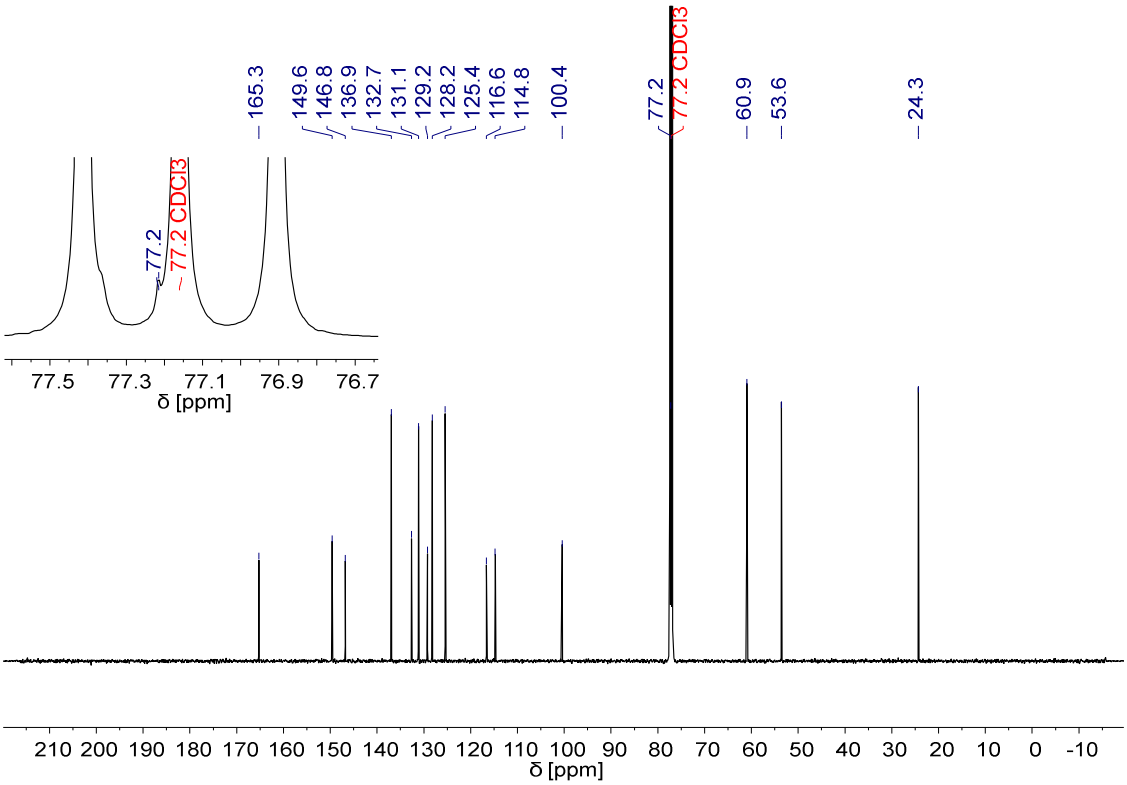


Figure S28 ^{13}C NMR spectrum of compound **8**.

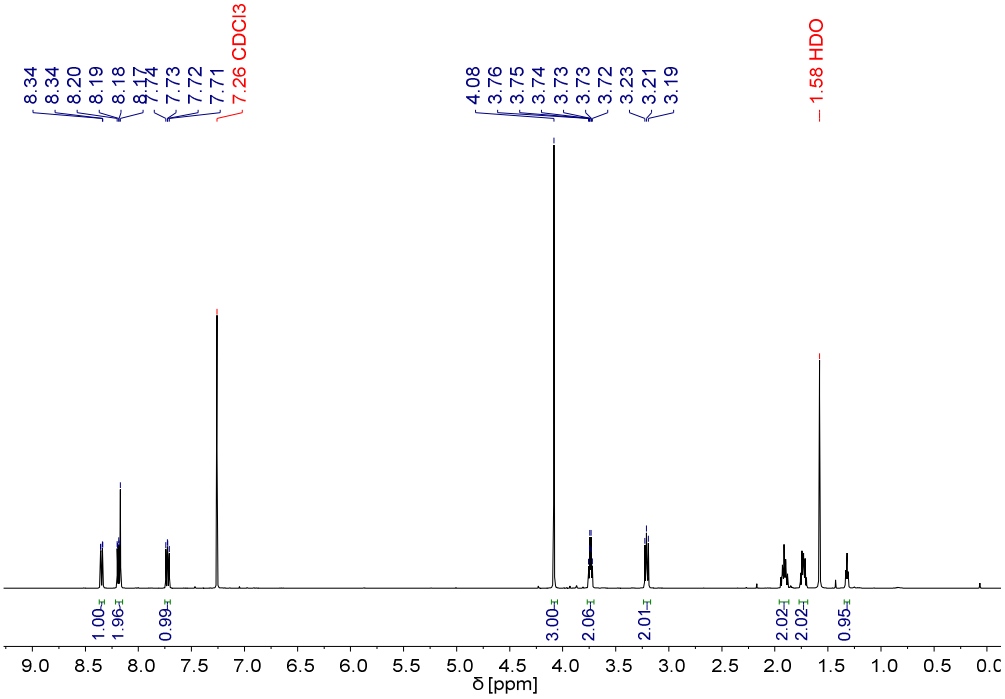


Figure S29 ^{1}H NMR spectrum of compound **9**.

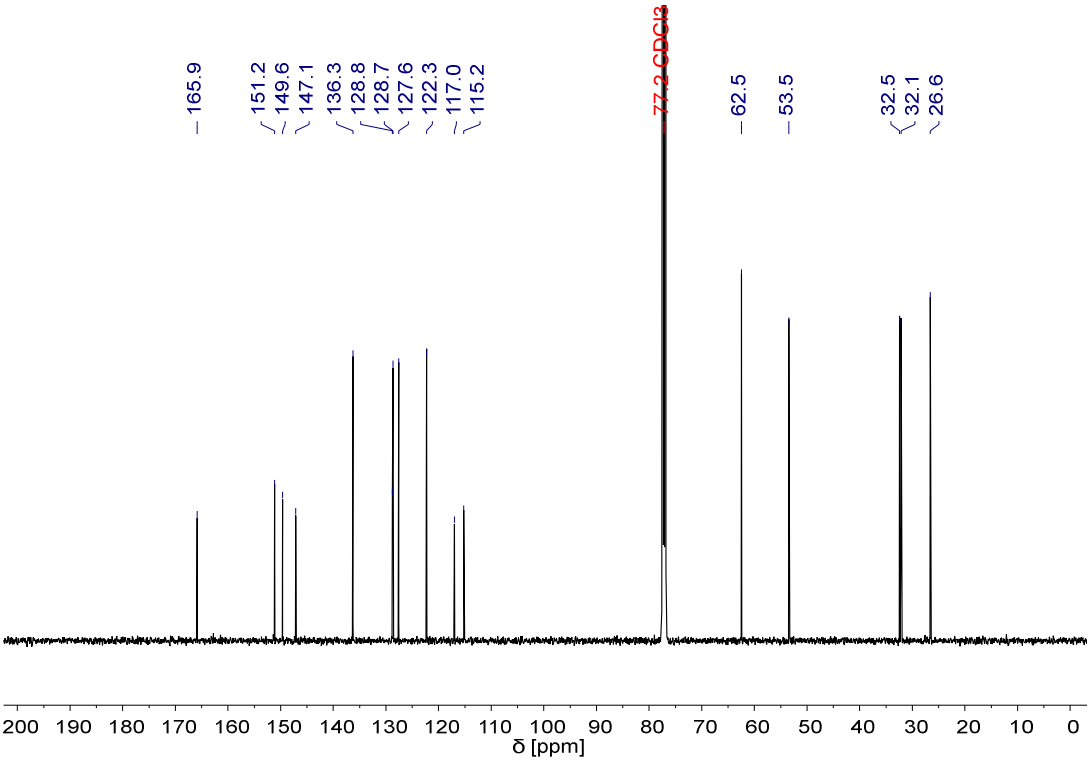


Figure S30 ^{13}C NMR spectrum of compound **9**.

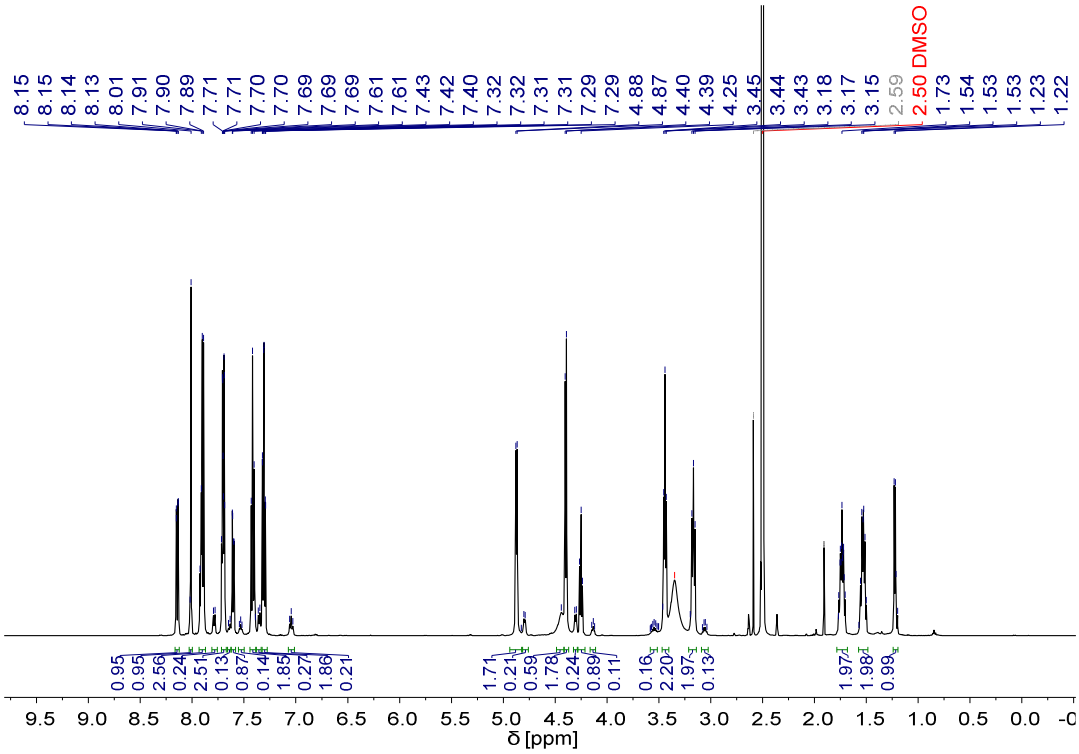


Figure S31 [1]H NMR spectrum of compound **10**.

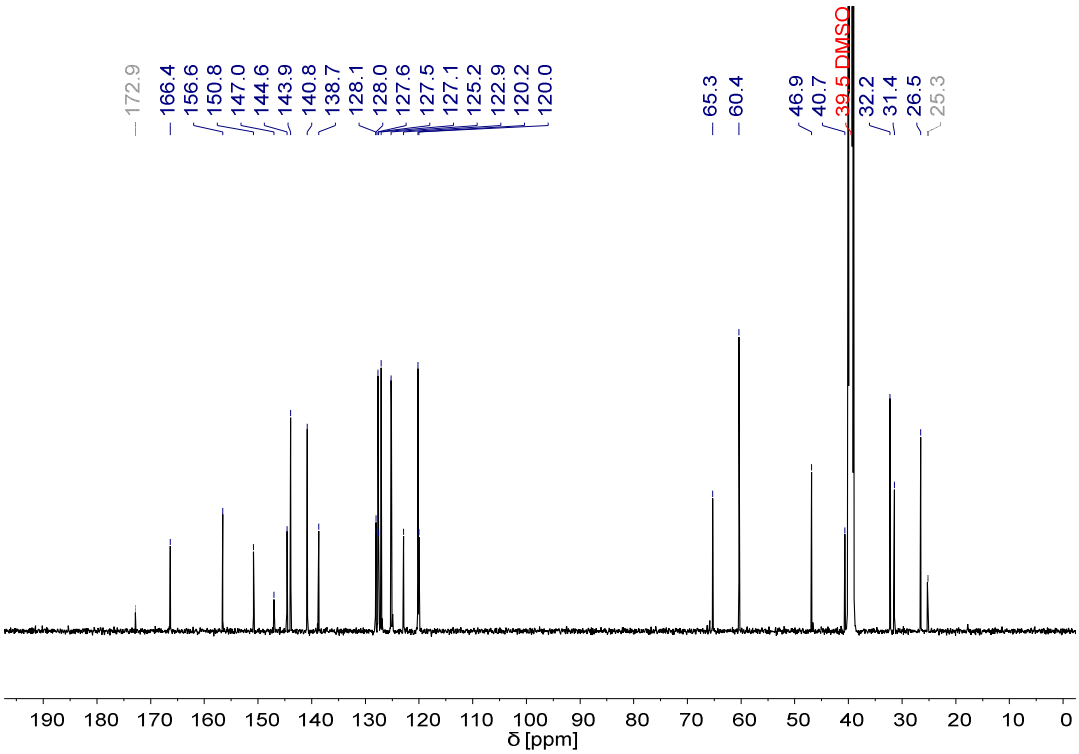


Figure S32 [13]C NMR spectrum of compound **10**.

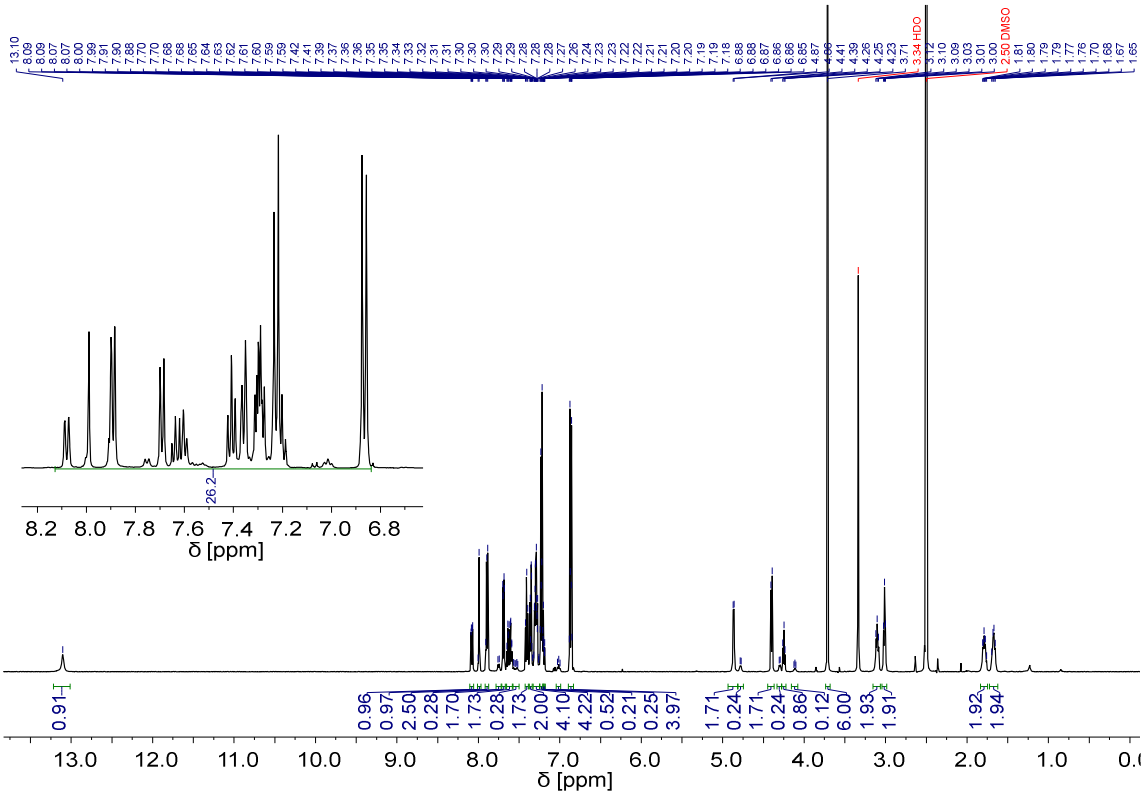


Figure S33 ^{1}H NMR spectrum of compound **11**.

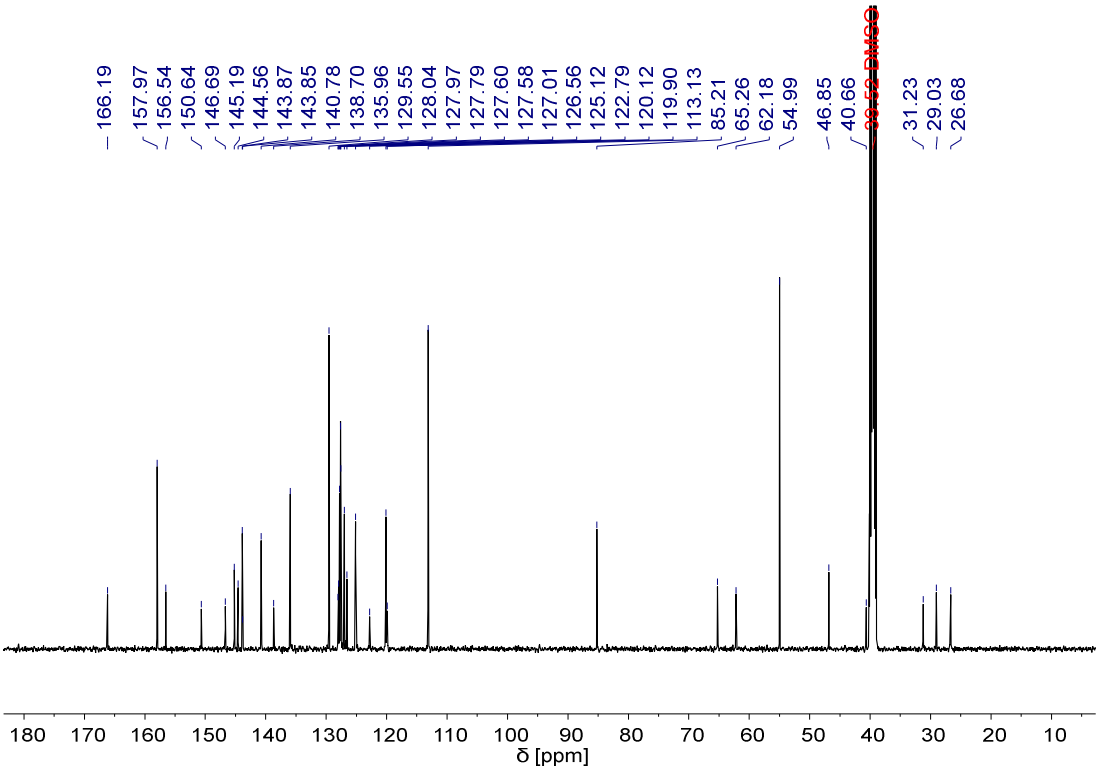


Figure S34 ^{13}C NMR spectrum of compound **11.**

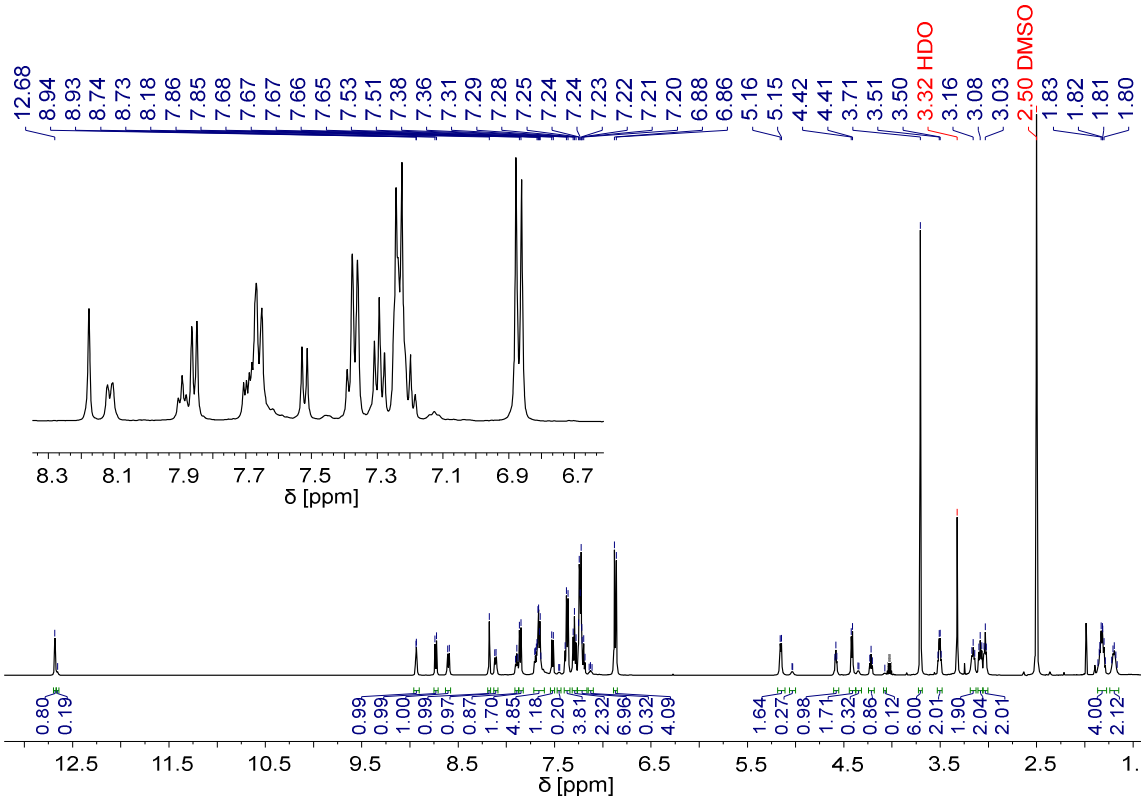


Figure S35 ^{1}H NMR spectrum of compound **12**.

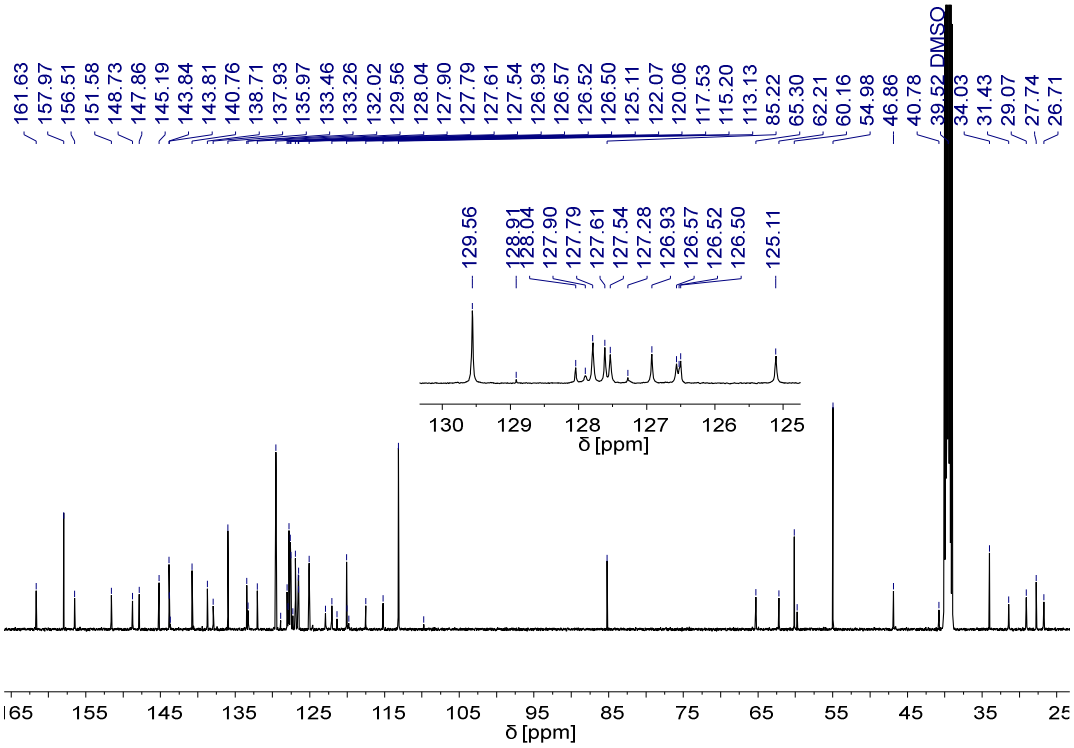


Figure S36 ^{13}C NMR spectrum of compound **12**.

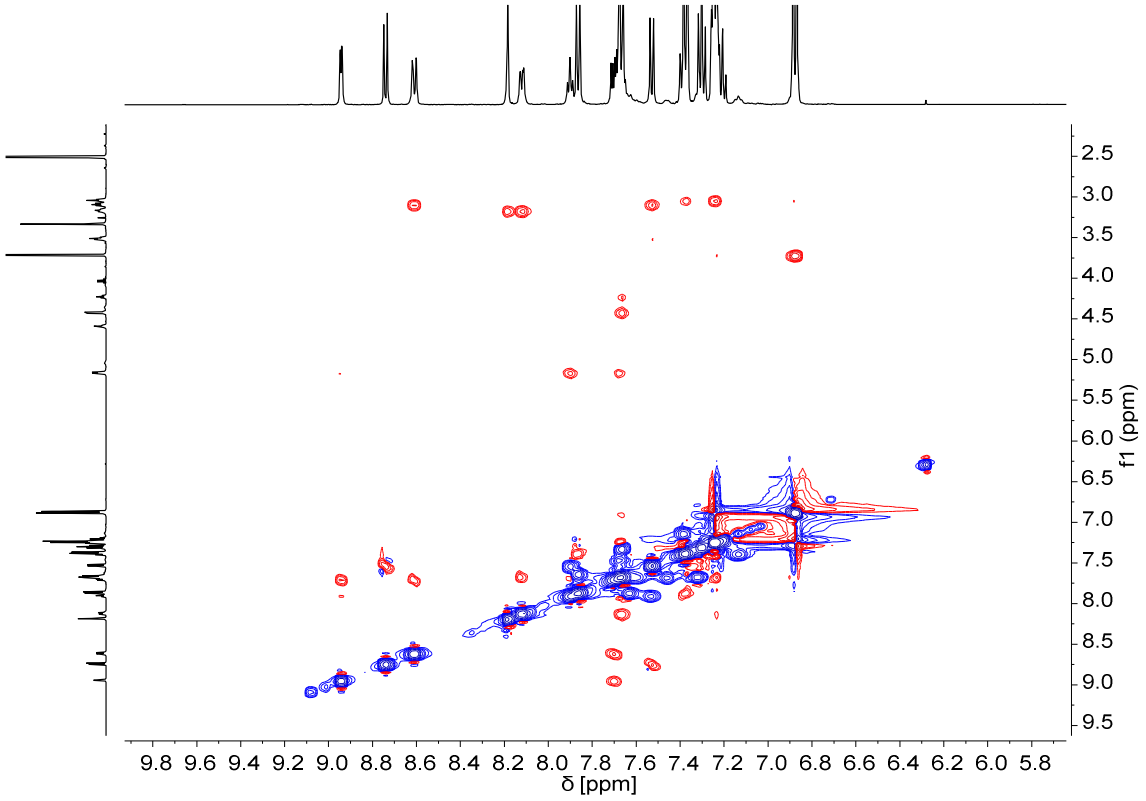


Figure S37 ROESY-EXSY NMR spectrum of compound **12.**

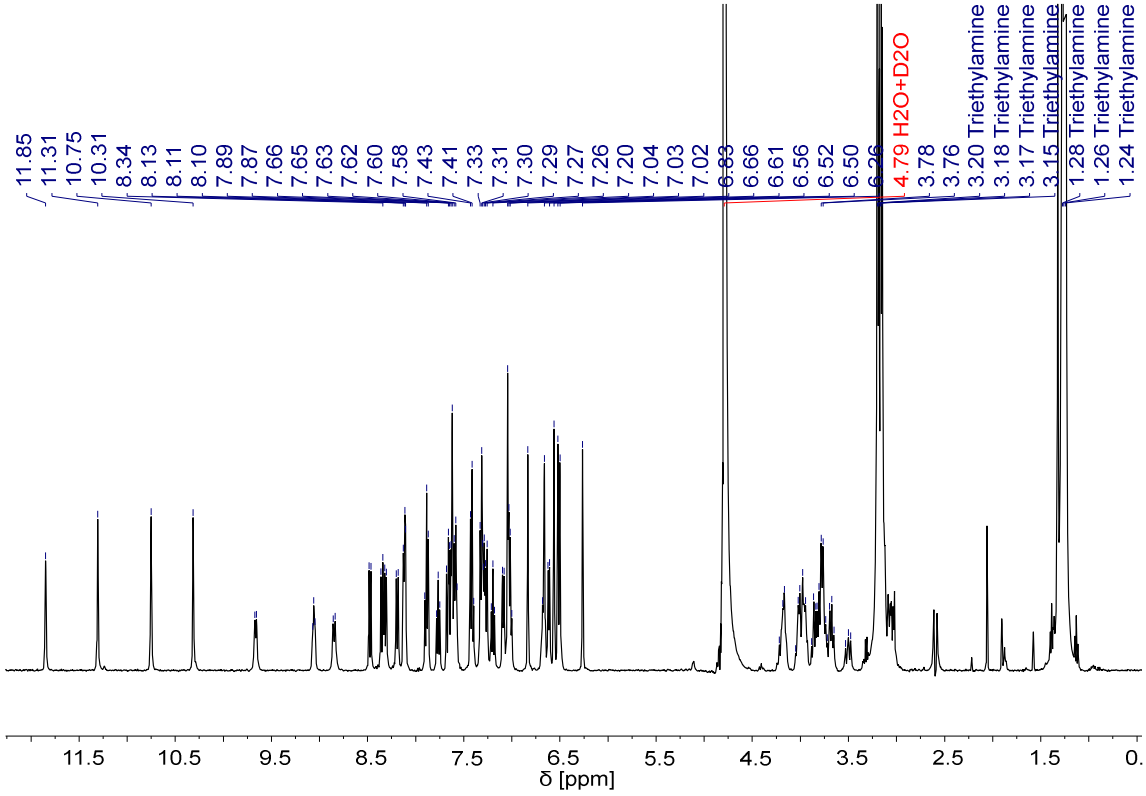


Figure S38 ^{1}H NMR spectrum of **22** in H_2O/ D_2O (v/v; 9:1; 50 mM NH_4HCO_3), water suppression.

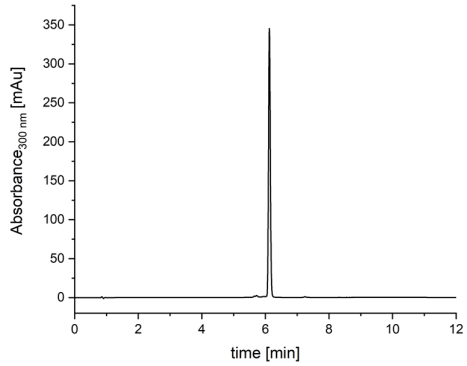
7.2.5.2 HPLC chromatograms

7.2.5.2.1 Foldamer chromatograms

Compound **22a**

Gradient: 30-100 % B in A, 0.1% TFA buffer buffer

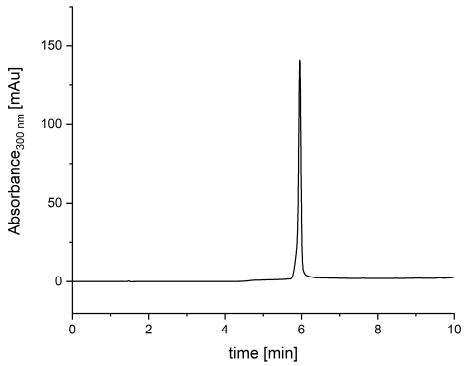
Column: Nucleodur C18



Compound **22**

Gradient: 0-100 % B in A, TEAA

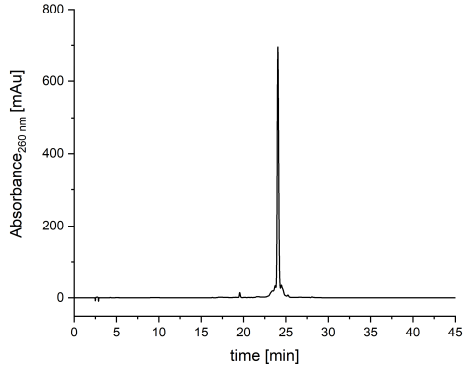
Column: Nucleodur C18



7.2.5.2.2 Oligonucleotide chromatograms

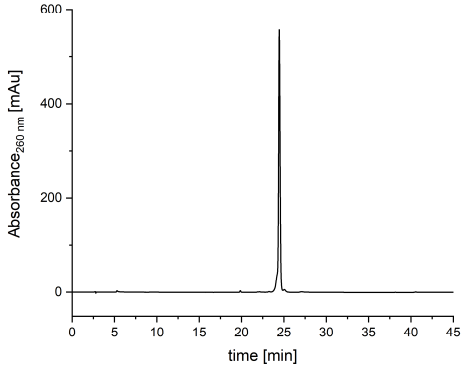
Compound **13**

D-(5'-GTT TTG **X**CA AAA C-3')



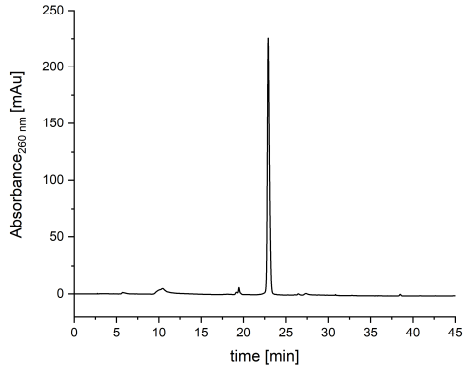
Compound **14**

L-(5'-GTT TTG **X**CA AAA C-3')



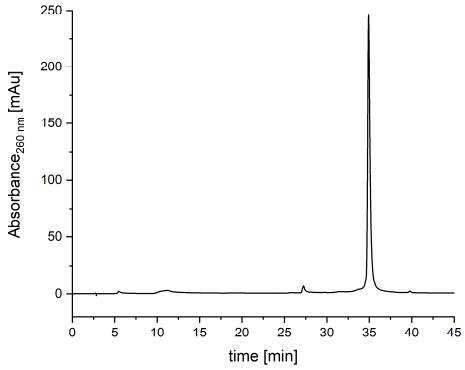
Compound **15**

5'-ACA GGA T**X**A TCC TGT-3'



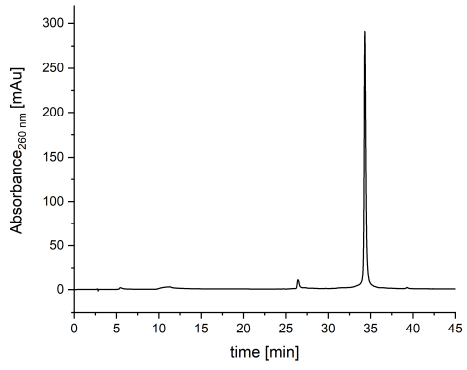
Compound **16**

5'-pC**X**G TCC TA-3'



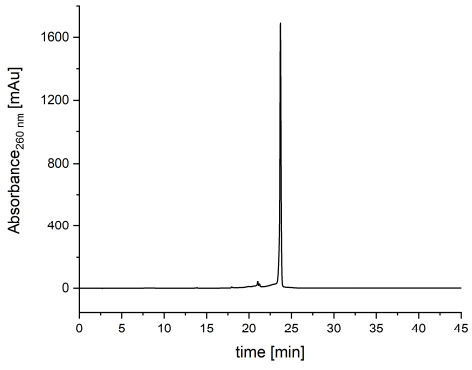
Compound **17**

5'-pAGG ATG **X**C-3'



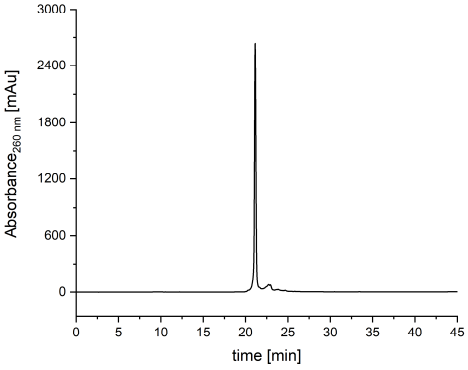
Compound **18**

5'-CTG **X**CA GGA TG**X** CAT Cp-3'



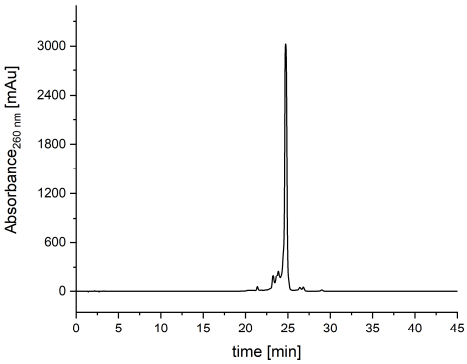
Compound **19**

5'-CTG T**X**A CAG GAT GT**X** ACA TCp-3'



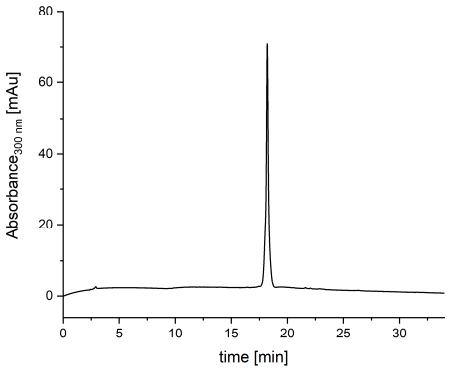
Compound **20**

5'-CAC ATC CTG T**X(H)**A CAG GAT GTG-3'



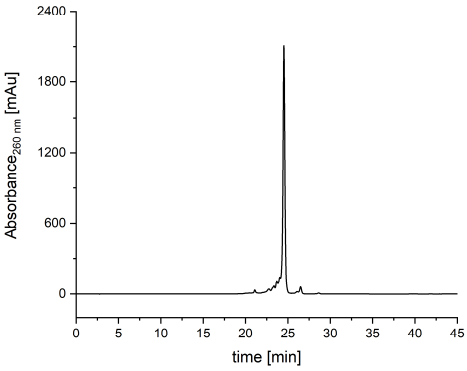
Compound **28**

5'-CAC ATC CTG T**X(Bio)**A CAG GAT GTG-3'



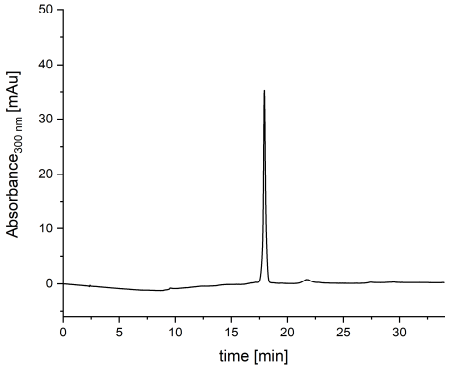
Compound **21**

5'-GAC AGG ATG T**X(H)**A CAT CCT GTC-3



Compound **29**

5'-GAC AGG ATG T**X(Bio)**A CAT CCT GTC-3

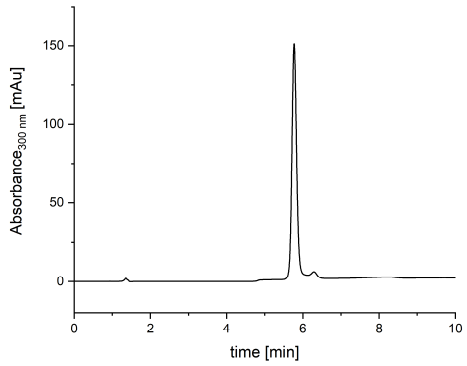


7.2.5.2.3 Conjugate chromatograms

Compound **25**

Gradient: 0-50 B in A; NH_4OAc buffer system

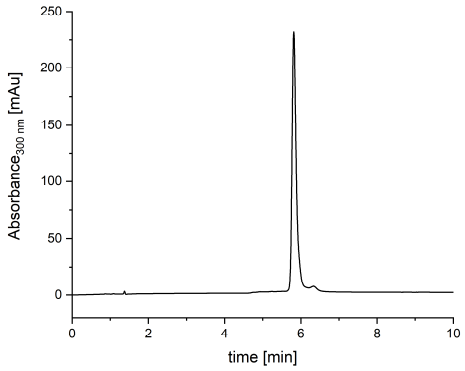
Column: Nucleodur C18



Compound **26**

Gradient: 0-50 B in A; NH_4OAc buffer system

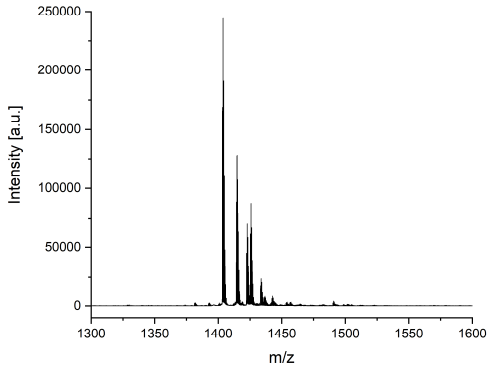
Column: Nucleodur C18



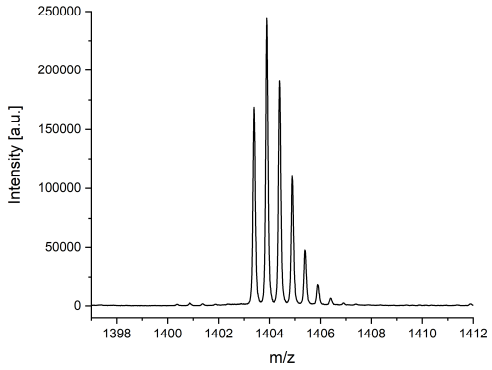
7.2.5.3 Mass spectra

7.2.5.3.1 Foldamer mass spectra

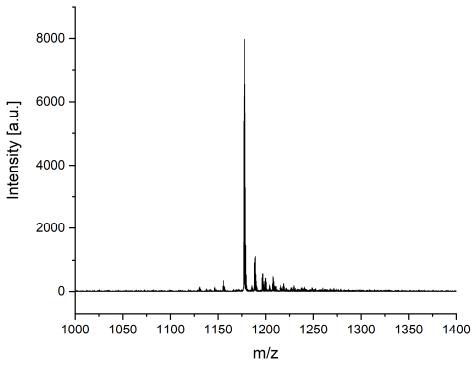
Compound **22a**



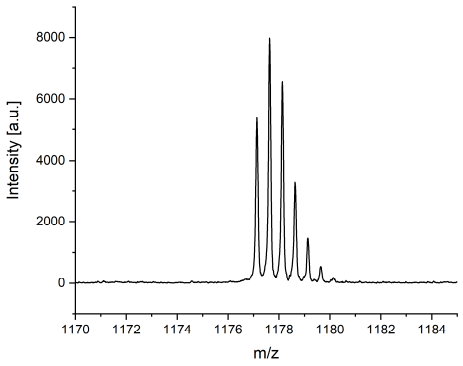
Compound **22a** Zoom



Compound **22**



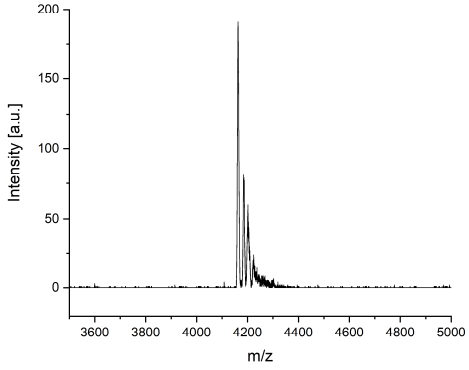
Compound **22** Zoom



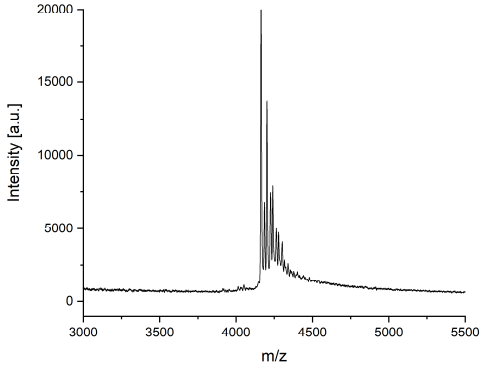
7.2.5.3.2 Oligonucleotide mass spectra

For conditions, see the general section for oligonucleotide synthesis

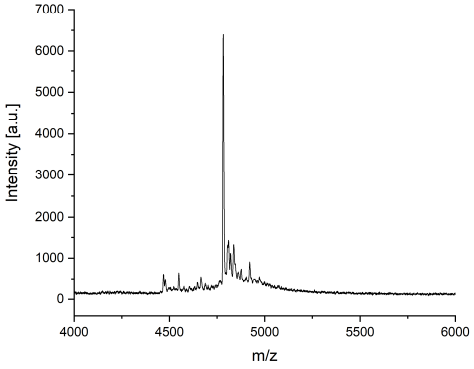
Compound **13**



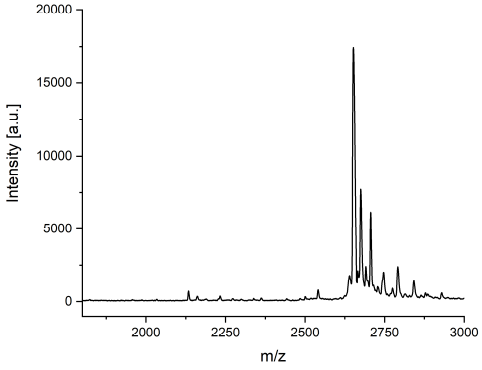
Compound **14**



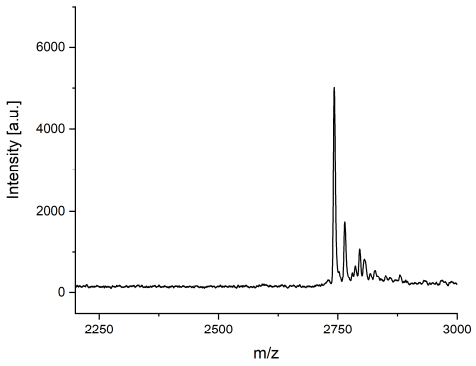
Compound **15**



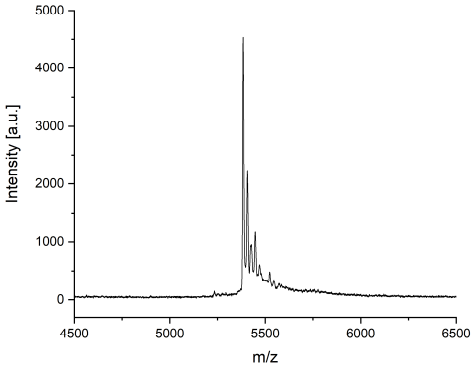
Compound **16**



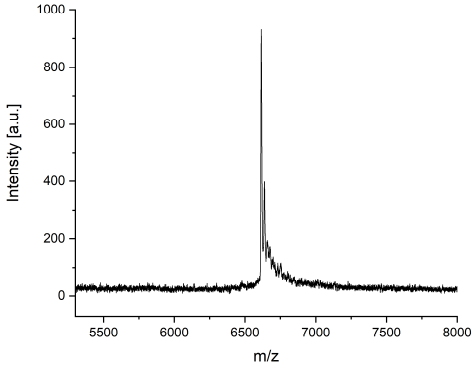
Compound **17**



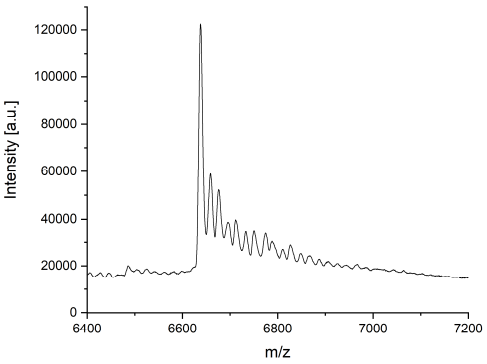
Compound **18**



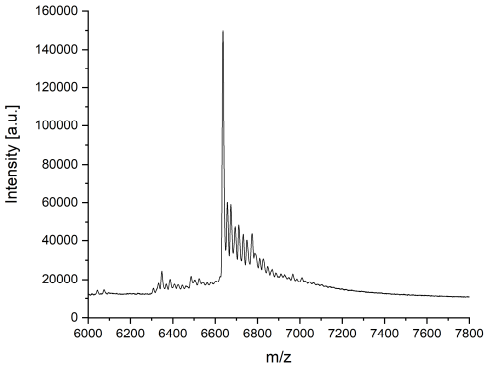
Compound **19**



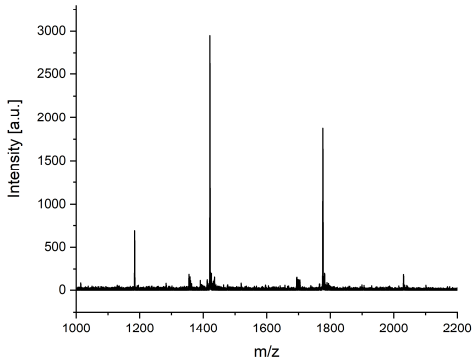
Compound **20**



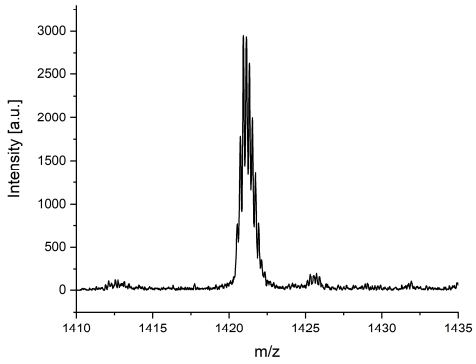
Compound **21**



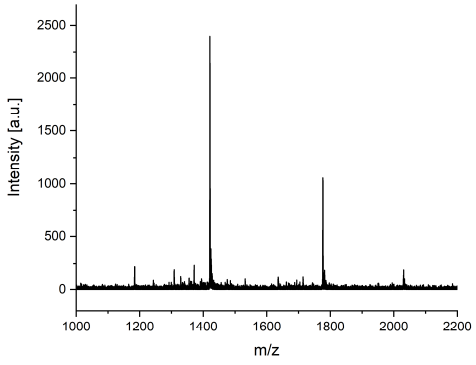
Compound **28**



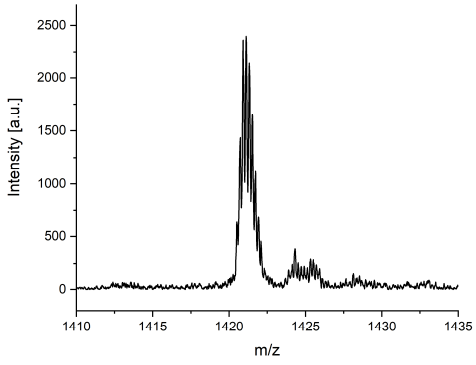
Compound **28** Zoom



Compound **29**

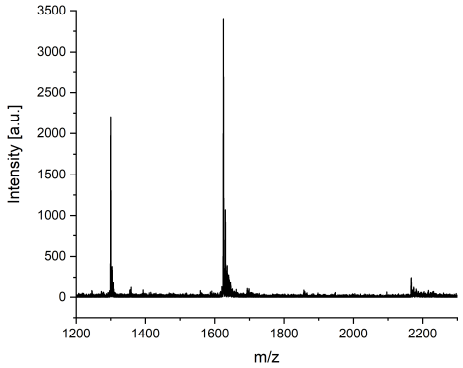


Compound **29** Zoom

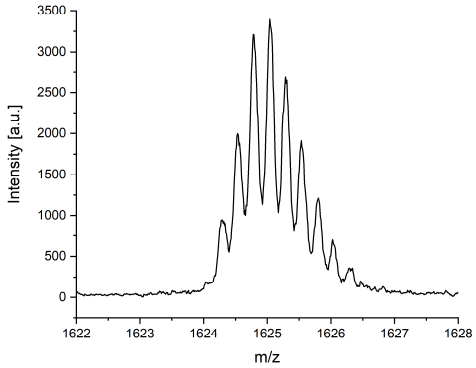


7.2.5.3.3 Conjugate mass spectra

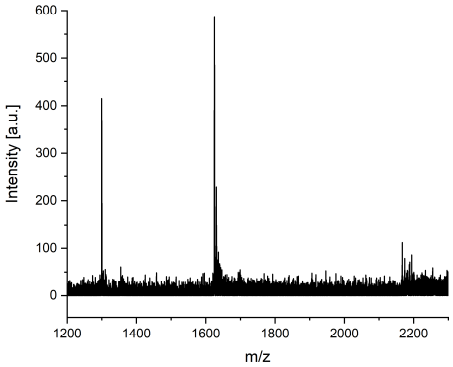
Compound **24**



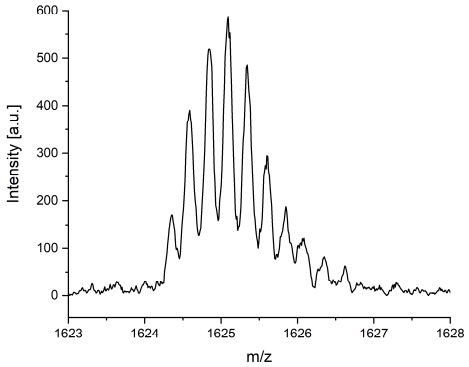
Compound **24** Zoom



Compound **25**



Compound **25** Zoom



7.2.6 References

1 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.

2 V. Corvaglia, J. Wu, D. Deepak, M. Loos, I. Huc, *Chem. - Eur. J.* **2024**, *30*, e202303650.

23 M. Egli, V. Tereshko, G. N. Mushudov, R. Sanishvili, X. Liu, F. D. Lewis, *J. Am. Chem. Soc.* **2003**, *125*, 10842-10849.

25 P. K. Mandal, G. W. Collie, B. Kauffmann, I. Huc, *Angew. Chem., Int. Ed.* **2014**, *53*, 14424-14427.

36 J. R. Terrell, S. J. Taylor, A. L. Schneider, Y. Lu, T. N. Vernon, S. Xhani, R. H. Gumpper, M. Luo, W. D. Wilson, U. Steidl, G. M. K. Poon, *Cell Rep.* **2023**, *42*, 112671.

37 Y. Mo, B. Vaessen, K. Johnston, R. Marmorstein, *Mol. Cell* **1998**, *2*, 201-212.

40 R. E. Franklin, R. G. Gosling, *Nature (London, U. K.)* **1953**, *172*, 156.

41 G. R. Fulmer, A. J. M. Miller, N. H. Sherden, H. E. Gottlieb, A. Nudelman, B. M. Stoltz, J. E. Bercaw, K. I. Goldberg, *Organometallics* **2010**, *29*, 2176-2179.

42 M. Cianci, G. Bourenkov, G. Pompidor, I. Karpics, J. Kallio, I. Bento, M. Roessle, F. Cipriani, S. Fiedler, T. R. Schneider, *J. Synchrotron Radiat.* **2017**, *24*, 323-332.

43 P. Evans, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2006**, *D62*, 72-82.

44 P. R. Evans, G. N. Murshudov, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *69*, 1204-1214.

45 W. Kabsch, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 125-132.

46 C. Vonrhein, C. Flensburg, P. Keller, A. Sharff, O. Smart, W. Paciorek, T. Womack, G. Bricogne, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 293-302.

47 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.

48 M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 235-242.

49 P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 486-501.

50 M. D. Winn, G. N. Murshudov, M. Z. Papiz, *Methods Enzymol.* **2003**, *374*, 300-321.

51 G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 355-367.

52 A. A. Vagin, R. A. Steiner, A. A. Lebedev, L. Potterton, S. McNicholas, F. Long, G. N. Murshudov, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *D60*, 2184-2195.

53 A. T. Bruenger, *Nature (London)* **1992**, *355*, 472.

54 A. W. Schuettelkopf, D. M. F. van Aalten, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *D60*, 1355-1363.

55 A. Wilson, *Acta Crystallogr.*, **1950**, *3*, 397-398.

56 H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res.* **2000**, *28*, 235-242.

57 W.L. Delano, 2002 The PyMOL Molecular Graphics System, Delano Scientific, San Carlos **2002**.

58 E. Krissinel, K. Henrick, *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *D60*, 2256-2268.

59 X. -J. Lu, W. K. Olson, *Nat. Protoc.* **2008**, *3*, 1213-1227.

60 Schrödinger, LLC. *Maestro*; Schrödinger, LLC: New York, NY, **2021**.

61 N. Yang, X. Su, V. Tjong, W. Knoll, *Biosens. Bioelectron.* **2007**, *22*, 2700-2706.

62 S. Wang, G. M. K. Poon, W. D. Wilson, *Methods Mol. Biol. (N. Y., NY, U. S.)* **2015**, *1334*, 313-332.

63 C. Perez, A. M. Barkley-Levenson, B. L. Dick, P. F. Glatt, Y. Martinez, D. Siegel, J. D. Momper, A. A. Palmer, S. M. Cohen, *J. Med. Chem.* **2019**, *62*, 1609-1625.

64 V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat, I. Huc, *Chem. - Eur. J.* **2023**, *29*, e202300898.

8 Towards TFA-labile phosphonic acid building blocks (unpublished)

The current synthesis of DMFOs is based on a SPFS strategy in which TFA cleavage does release the DMFO from the resin but does not cleave any of the ethyl ester side chain protecting groups (PGs). This has its own advantages, e.g. it was exploited in a pre-purification strategy as described in chapter 6 to produce particularly long DMFOs. The following global deprotection is TMSBr-mediated, which is known to be an additive in some peptide-cleavage cocktails initially designed as a less toxic alternative to replace HF.[1, 2] Complete TMSBr-mediated deprotection of diethyl phosphonate esters in DMFO sequences requires long reaction times and elevated temperatures if sequences are long. In practice, ambient moisture causes hydrolysis of unreacted bromotrimethylsilane. The resulting hydrobromic acid may lead to many side reactions like the cleavage of side chain ethers or even main chain amides when they are more sterically accessible. Even though diethyl phosphonate esters are perfectly suitable PGs for DMFO sequences, when phosphonic acid monomers are to be incorporated together with the wider scope of available side chains, main chain variations or terminal functionalities, TMSBr-mediated conditions for their removal are no longer a viable synthetic strategy and monomers bearing alternative phosphonic acid PGs become necessary. Previously, the synthesis of bis-*tert*-butyl protected phosphonic acid derivatives of M and Q^4 as well as SPFS protocols for oligomer synthesis have been reported.[3] However, their routine usage has been challenging due to their high susceptibility to acidic hydrolysis, including e.g. silica gel purification, as well as the lack of a synthetic strategy to re-esterify the phosphonic acid in presence of a quinoline-2-carboxylic acid, once they got cleaved.[4] Benzyl phosphonate esters have been used previously as a PG for phosphonic acids, including TFA-mediated deprotection and presents a suitable middle ground between bis-*tert*-butyl esters as well as diethyl esters (Fig. 24).[5] Monomers for Q^4 and M residues that may be readily combined with a wide range of other side chains will enable the synthesis of oligomers like **8.4** that encode sequence information, which is not available in the current generation of DMFOs.

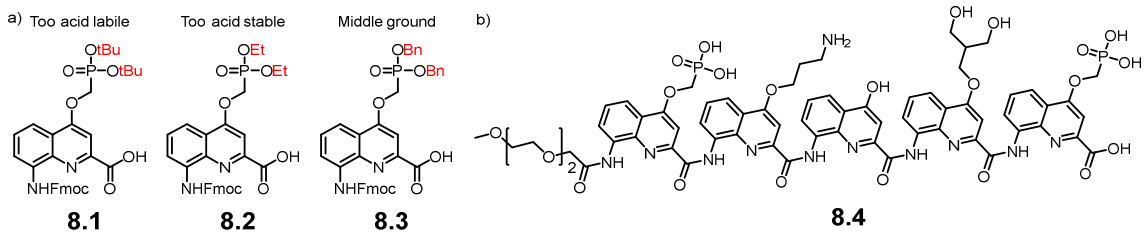
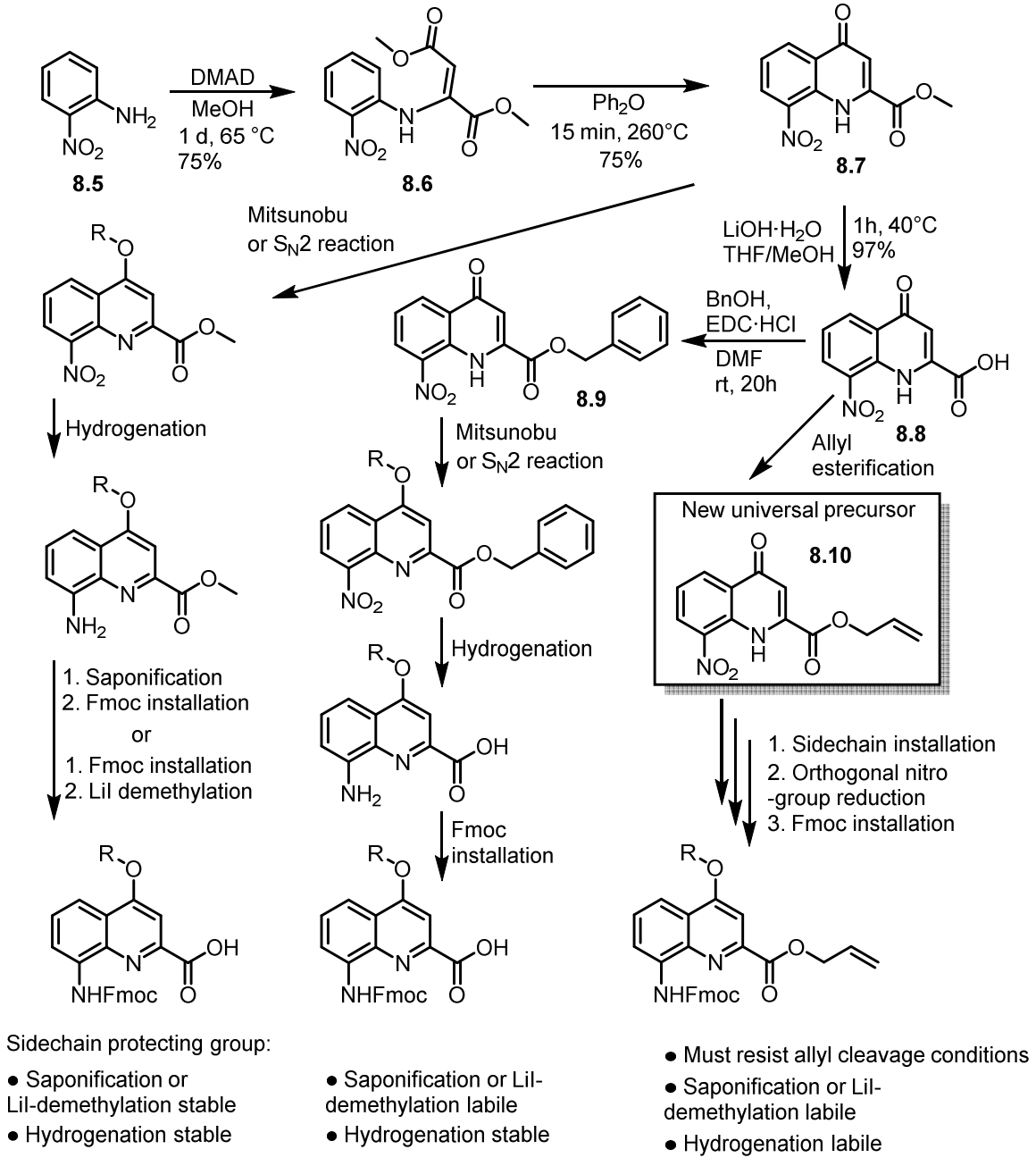


Figure 24 a) Different side chain PGs depicted as derivatives of Q^4. b) Example sequence of a pentamer **8.4** with mixed side chains as a proof-of-concept molecule.

8.1 Dibenzyl protected derivatives of Q4

The following synthesis was developed together with Julian Heeg during his organic chemistry internship (LMUs chemistry master program) supervised by me.

Most 4-ether substituted monomers in the Huc group follow a rather straight-forward synthetic strategy, that involves the construction of the quinoline ring (as a quinolinone tautomer) from 2-nitroaniline **8.5** by a Michael-type addition of the aromatic amine to DMAD to yield fumarate **8.6** , that is followed by an intramolecular cyclization quinolinone



Scheme 1 Summary of the different synthetic routes towards Fmoc-quinoline monomers for SPFS with their side chains installed in position 4 of the quinoline ring by an ether linkage. Dimethyl acetylene dicarboxylate (DMAD), 1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC).

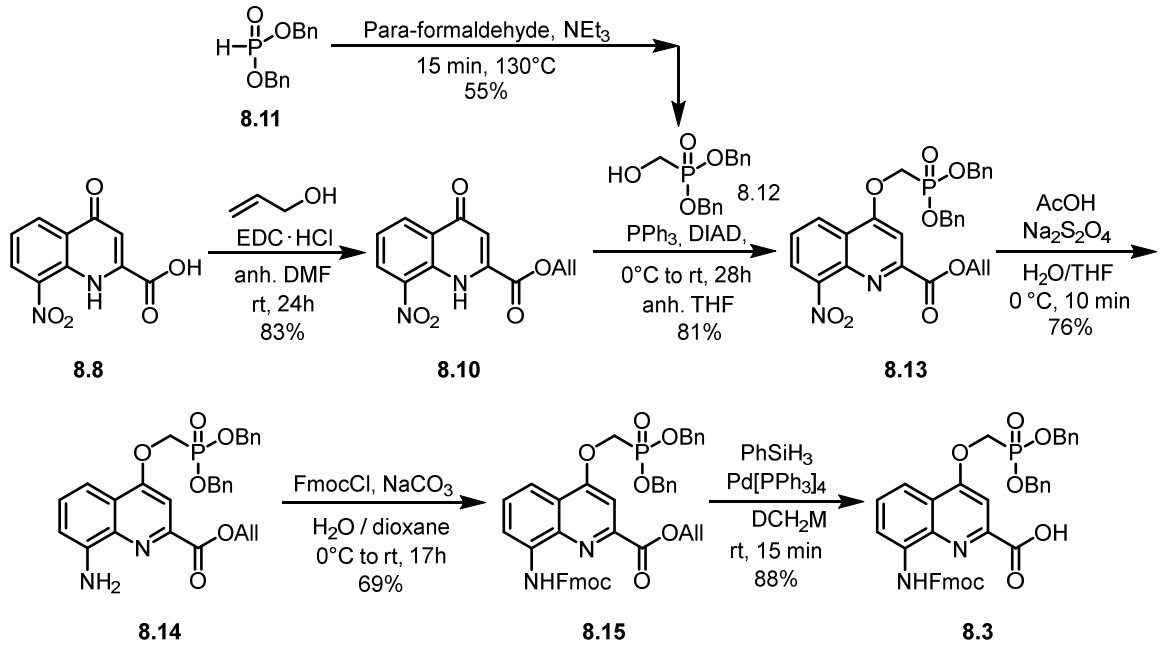
tautomer **8.7** (when dissolved in $CHCl_3$) of 8-nitro kynurenic acid methyl ester (8-nitro-4-oxo-1,4-dihydroquinoline-2-carboxylic acid methyl ester).[6]

From **8.7**, many transformations towards different building blocks can be achieved, such as producing the 4-halide derivates (chloro- and bromo-) that give access to methylene linked side chains by e.g. Sonogashira or Suzuki couplings (not depicted).[7] Quinolinone **8.7** can be directly functionalized by either S_N2 with halides as leaving group or under Mitsunobu conditions with alcohols. Similarly to peptide SPS, side chains require distinct PGs that i) must be orthogonal to SPS conditions ii) are cleavable by TFA and iii) are orthogonal to synthetic conditions required for their synthesis. Many side chains and their protecting groups commonly used in peptide SPS are also applicable to oligoquinoline synthesis, though the synthetic routes to access them differ. For quinoline monomers, the C-terminal PG must be removed at a late stage during the synthesis (after side chain installation) as a free carboxylic acid reacts under conditions utilized for side chain insertion. The choice of C-terminal ester is determined by the chemical stability of the side chain protection. Three pathways from **8.7** to SPS-ready monomers are depicted in Scheme **1**. Since cyclization of **8.6** directly produces the methyl ester **8.7**, it is the most efficient way for further functionalization but later requires either saponification or LiI-mediated demethylation to be cleaved, incompatible with e.g. many base labile PGs such as *tert*-butyl esters. In this case **8.7** may be saponified to give **8.8**, and re-esterified to yield the benzyl ester **8.9**. The side chain then is installed similarly, and the benzyl ester is removed by catalytic hydrogenation together with the reduction of the nitro-group to produce an amino-acid. This amino acid must be Fmoc protected to yield the SPFS ready monomer. However, both strategies are incompatible, when the side chain protection is labile towards demethylation, saponification and hydrogenation, as is the case for the desired dibenzyl protection. For this reason, an alternative ester was necessary, that could be cleaved orthogonally to Fmoc and the side chain PG and must further resist the nitro to amine reduction. Of the available esters, allyl groups were deemed to fulfill these criteria. Hence, **8.7**. was esterified in DMF with EDC·HCl as a dehydrating reagent to give quinolinone allyl ester **8.10**, that may serve as an intermediate to branch into monomers, which side chains do not resist hydrogenation reactions.

To install the dibenzyl phosphonates under Mitsunobu condition, respective alcohol **8.12** was necessary. Its synthesis has been performed according to literature procedures. [8, 9] Purification was conducted on an automated chromatography system with a RP stationary phase, due to strong tailing of the compound on silica. Since the yields are lower than reported (55% vs 71%), silica gel purification might have to be revisited.

The side chain was installed under Mitsunobu conditions and purified by column chromatography to obtain intermediate **8.13** in good yield. Next, the nitro group was intended to be reduced to the respective amine. In addition to the dibenzyl phosphonate ester, the double bond of the allyl ester is also subject to be saturated by hydrogenation, therefore

alternative conditions are required to enable this reaction. This can be achieved in good yields through radical based reagents, such as $Na_2S_2O_4$ as well as B_2OH_4 (not shown) to give the respective amine **8.14**. Finally, Fmoc installation via FmocCl was achieved in modest yields to give **8.15**, followed by the cleavage of the C-terminal allyl ester using Pd^0 as a catalyst with phenyl silane to scavenge allyl cations to afford the final Fmoc-monomer **8.3** with the desired protection scheme.

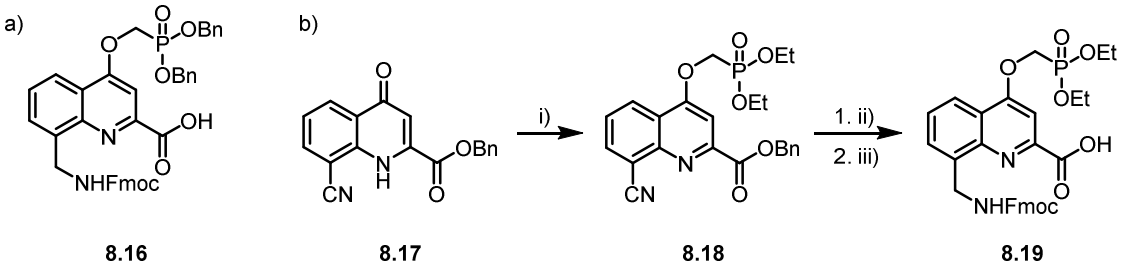


Scheme 2 Synthesis of Fmoc and dibenzyl phosphonate protected derivative of Q^4 **8.3**. 1-Ethyl-3-(3-dimethyl-aminopropyl)carbodiimide (EDC), Fluorenyl methyl oxycarbonyl (Fmoc), diisopropyl azodicarboxylate (DIAD), allyl (All).

8.2 Dibenzyl protected derivatives of M

The following synthesis were developed together with Leon Stanislawski (LMUs chemistry master program) during his organic chemistry internship supervised by me as well as Marla Dorfer during her bachelor thesis (LMU Bachelor Program Pharmaceutical sciences) supervised by Nikolas Schneider and me.

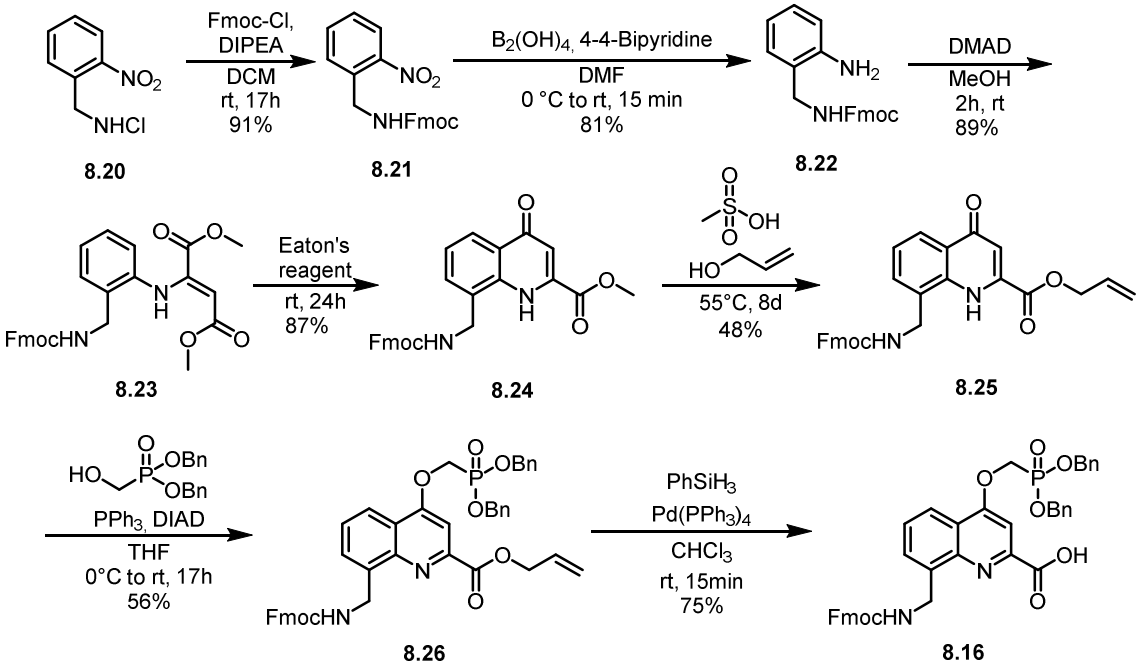
To synthesize DMFOs with sequence diversity, a dibenzyl derivative of M (**8.16**) is also necessary. On paper, this may be achieved with a similar synthetic strategy than for **8.3** but could not be realized in practice. In the reported synthesis of the diethyl derivative of M (**8.19,** Scheme **3**), the starting material **8.17** was synthesized similarly than **8.7** from 2-cyanoanilin via addition of DMAD to give the respective fumarate followed by an intramolecular cyclization to the quinolinone.[10] A transesterification with benzyl alcohol generates the benzyl ester **8.17** on which the side chain was installed by Mitsunobu reaction to yield the intermediate **8.18**.[11] The nitrile group is reduced together with the benzyl ester by heterogenous catalytic



Scheme 3 a) Fmoc and dibenzyl protected derivate of M **8.16**. b) Reported synthesis of diethyl protected derivative of M **8.19**.[12] i) Diethyl hydroxymethyl phosphonate, PPh_3, DIAD, THF, 0 °C to 40 °C overnight, 76%. ii) H_2, Pd(OH)2/C, $NaH_2PO_4 \cdot H_2O$, NH_4OH, H_2O/THF, rt, 30h, quant. iii) FmocCl, $NaHCO_3$, H_2O/dioxane, 0 °C to rt, 21 h, 59% over 2 steps.

hydrogenation utilizing $Pd(OH)_2$, which is more reactive than Pd^0 on carbon a good alternative to Raney-Ni, that was originally used for this transformation.[10]

Similarly, in the synthesis of Q^4-derivative **8.3**, the hydrogenation step is incompatible with dibenzyl phosphonate ester protection of the phosphonic acid side chain. This limitation was circumvented by employing a hydrogen-free reduction of the nitro group by radical mechanism **8.14**. However, nitriles require much harsher conditions to reduce than aromatic nitro groups due to the stability of the carbon-nitrogen triple bond. This was not achieved orthogonally to dibenzyl phosphonate esters and therefore was not further pursued.[13, 14] Instead, the synthetic route was redesigned (Scheme 4), starting from the corresponding hydrochloride salt **8.20** that already possessed the benzylic amine functionality, a synthetic strategy that was also performed in polydeutered derivates of Q.[15] This shifts the Fmoc protection from the end of the synthesis to the start and gave the respective Fmoc protected precursor **8.21** in excellent yield. The amine **8.22** is generated by hypodiboric acid reduction of the nitro group in *ortho*-position

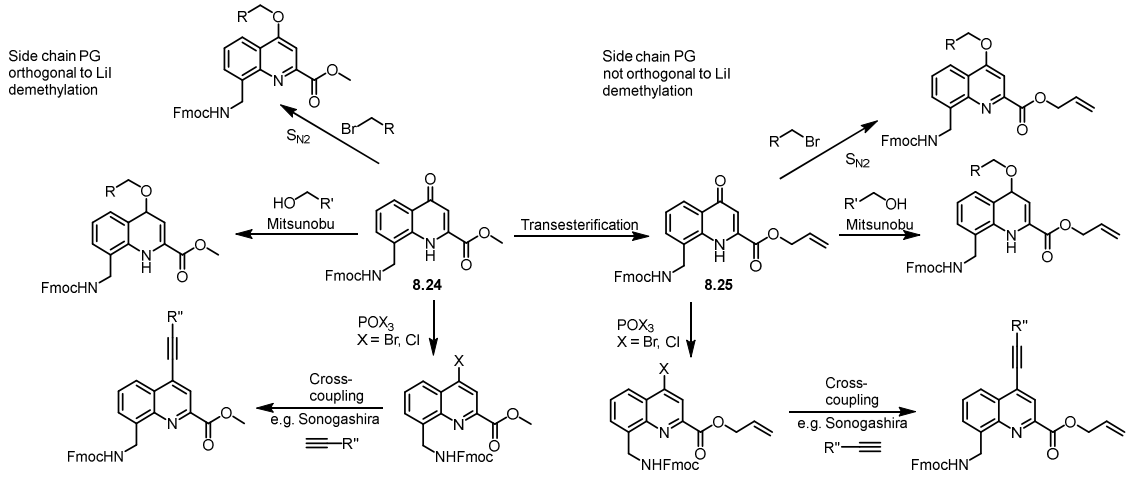


Scheme 4 Synthesis of Fmoc and dibenzyl phosphonate protected derivative of M **8.16**.

and then can undergo a Michael-type addition on DMAD to yield fumarate **8.23**, that readily precipitates from the reaction mixture upon completion of the reaction.

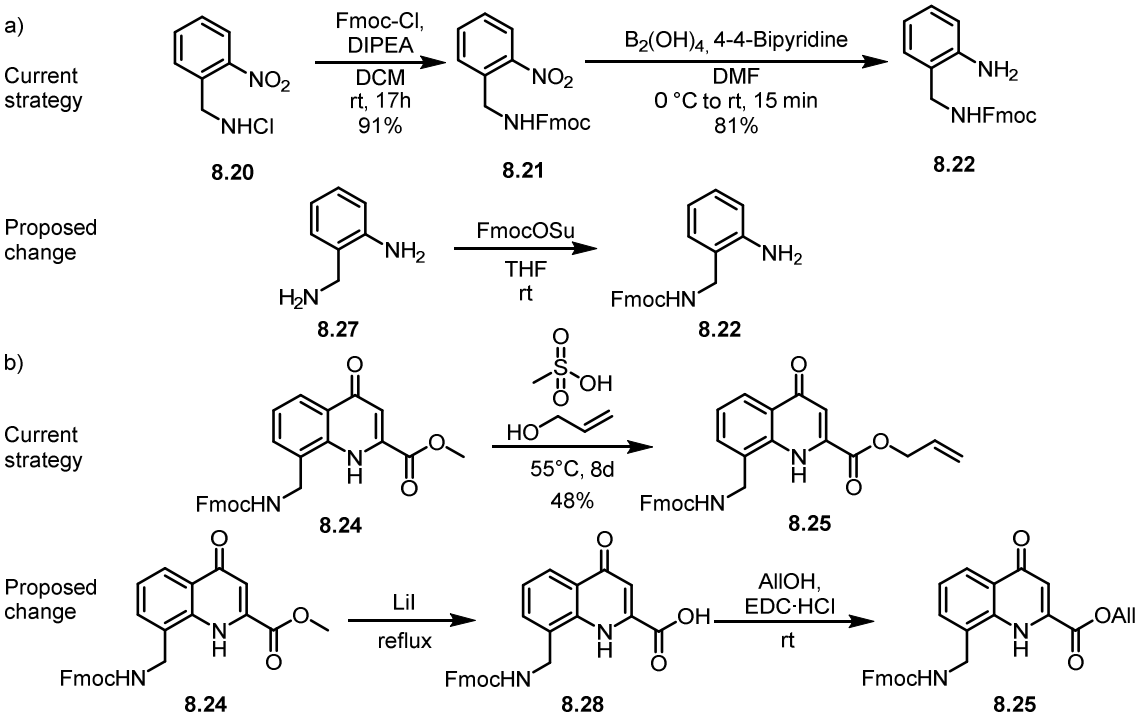
The absence of electron withdrawing groups on the quinoline ring allows for a cyclization in Eaton's reagent at room temperature that yields the Fmoc-protected aminomethyl quinolinone **8.24**. Due to the lability of Fmoc against basic conditions, saponification of the methyl ester or a base-mediated transesterification was disregarded. The transesterification of **8.24** to **8.25** can be achieved by methanesulfonic acid mediated transesterification, however the long reaction time in the presence of a strong acid is likely to cause either decomposition or polymerization. This results in a moderate yield of only 48%, despite the fact that no side products could be identified. Similarly to other building blocks, the side chain was attached with respective dibenzyl phosphonate alcohol **8.12** under Mitsunobu conditions to form **8.26**. Even though TLC and RP-HPLC analysis complete conversion of the starting material with no significant side products. An isolated yield of 56% indicates that further optimization is necessary. The isolation of the compound was conducted by precipitation from isopropanol, and it is likely that some product was lost at this stage. The following final Pd^0 catalyzed deallylation step yielded the monomer **8.16** in a good yield.

A future endeavor is to synthesize sequences containing both phosphonic acids and alternative side chains to encode sequence information while reducing overall negative charge, thereby minimizing non-specific binding to polycationic DBPs. For Q^4 derivatives, many different side chains have been reported of which all are derived from **8.7**.[16]. For derivatives of M, this depth of side chains is underexplored. Both **8.24** and **8.25** represent very advanced intermediates from which branching towards many different side chains is possible (Scheme **5**) in two synthetic steps to yield an Fmoc acid monomer. If the side chain (including its PG) does tolerate LiI mediated demethylation, **8.24** can directly be used for side chains like methoxy, isobutoxy or ethylene glycols. If the side chain does not resist this condition, a synthetic route via **8.25** represents an alternative as described above for dibenzyl phosphonate esters monomer **8.16**.



Scheme 5 Schematic depiction of the versatility of compounds **8.21** and **8.22** for a diverse set of side chains.

This strategy to produce derivatives of M represents alternative for side chains that would resist the catalytic hydrogenation of the nitrile of **8.17**. This hydrogenation, even though robust, does require high dilutions that may hamper attempts to scale up this synthetic step. For this route to be a viable alternative, some low-yielding steps of the current route need to be improved (Scheme 6). Starting the synthesis from **8.20** requires a later reduction of the nitro group to the aromatic amine of **8.22**. With the concurrent protocol, the scale-up of this reaction may cause significant amount of heat that limits the reaction scale. This can be circumvented by directly using **8.27** in combination with FmocOSu, a less reactive electrophile than FmocCl, to selectively react with the benzylic amine as opposed to the aromatic amine directly yielding **8.22** in a highly scalable reaction. One further issue represents the transesterification of **8.24** to **8.25**. Saponification of **8.24** to carboxylic acid **8.28** with strong bases such as NaOH cannot be achieved reliably in the presence of Fmoc but may be achieved with LiI. If successful, esterification similarly to the allyl esterification towards **8.10** may significantly improve the overall yields towards **8.25**, even though this expands the synthetic route by an additional step.



Scheme 6 a) Proposed changes toward precursor **8.22**. *N*-Hydroxysucciniimid (OSu). b) Proposed changes towards precursor **8.25**. Allyl (All); 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC).

In summary, to combine the chemistry of DMFOs with other quinoline monomers, TFA labile side chain PGs must be incorporated of which dibenzyl phosphonate esters were deemed the most viable. These dibenzyl phosphonate derivatives of Q[4] and M **8.3** and **8.16** required the development of new synthetic routes. In the process, three intermediates **8.10**, **8.24** and **8.25** were produced, which give access to a greater scope of potential side chains for DMFOs.

8.3 Synthesis of foldamers with diverse side chains

The following synthesis was developed together with Julian Heeg during his organic chemistry internship (LMUs chemistry master program) supervised by me.

For an initial proof of concept synthesis of sequences that contain novel dibenzyl phosphonate ester monomers, only **8.3** was tested to synthesize model pentamer **8.4** (further necessary monomers are described here[17]) because it was more readily available than **8.16**. Additionally, the target molecule **8.4** was required for different projects that are distinct from this thesis. Nevertheless, **8.4** contains two phosphonic acid monomers, both at two different positions in the sequence. This synthesis validates both the resin loading protocol using diisopropyl carbodiimide and OxymaPure® (Section 6.3.2.1.1) as well as a coupling protocol that *in situ* generates acyl chlorides.[11] As it is present at the beginning of the synthesis, any PG erosion is expected to result in complex crude mixtures, which was not observed. Lastly, the sequence is finished by another phosphonic acid residue to confirm that coupling of the carboxylic acid of **8.3** with *in situ* activation is also tolerated. The bis-silyl side chain PG of the bis-alcohol containing monomer of the sequence **8.29** requires a tetrabutyl ammonium fluoride (TBAF) deprotection step (Fig. 25a), also compatible with dibenzyl phosphonates. Final resin cleavage was conducted with TFA without additives, the crude material was purified by HPLC and identified by LC-MS (Fig. 25 b, c). It will be necessary to optimize the resin cleavage cocktail, especially to scavenge benzylium cations when many dibenzyl phosphonates will be incorporated into a sequence or in the presence of more reactive nucleophiles such as thiols to prevent thioether formation. This synthesis demonstrates that the dibenzyl protection of phosphonic acids does enable the integration of phosphonic acids with other building blocks in an all-TFA-based global deprotection as is common in peptide chemistry.

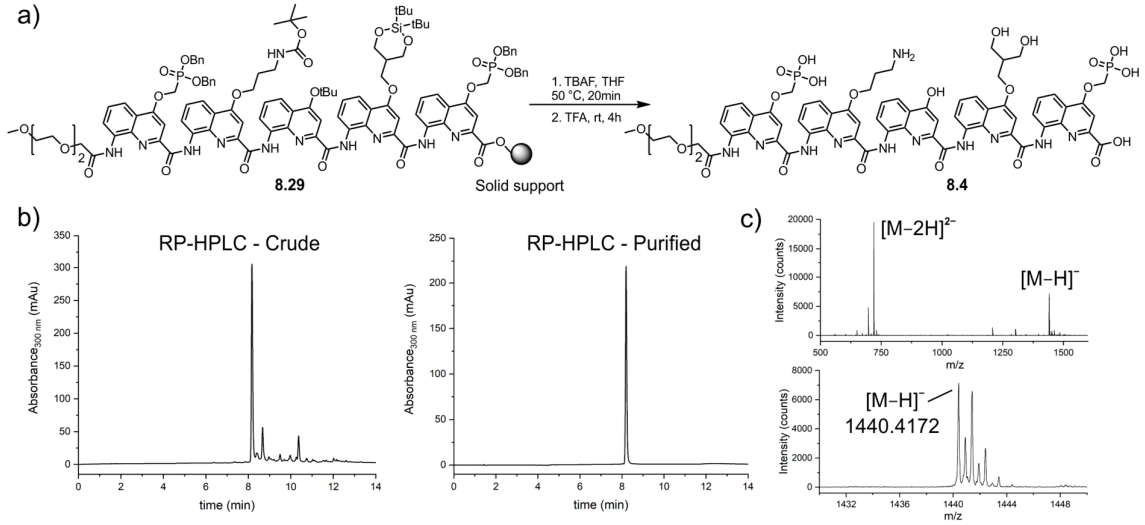


Figure 25 a) Proof of concept synthesis of AOFs **8.29** utilizing **8.3**. Resin-bound pentamer **8.29** was treated to give the final desired pentamer **8.4**. b) RP-HPLC-chromatograms of crude and purified product. c) ESI-TOF mass spectrum of the isolated target pentamer **8.4**.

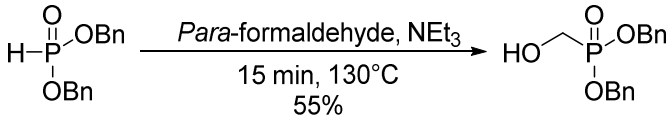
8.4 Experimental

8.4.1 Materials and methods

All materials and methods used, including suppliers, treatment of commercial compounds, analytical methods, and purification methods are listed within section 6.2 of this thesis.

8.4.2 Synthetic protocols

Synthesis of compound **8.12**



Para-formaldehyde (2.52 g, 83.9 mmol, 1.1 eq.) was suspended in compound **8.11** (20.0 g, 76.3 mmol, 1.0 eq.) in a sealed pressure flask. Triethylamine was added and the mixture was stirred at rt for 10 min, then the mixture was heated up to 130 °C under continuous stirring, until the solution turned colorless and then for 10 more minutes (circa 15 min total). After that, the solution was cooled down to rt and purified via reverse-phase column chromatography (40-100% acetonitrile in water). Acetonitrile was removed under reduced pressure. The received mixture was three times extracted with ethyl acetate and the phases separated. The combined organic phases were dried over $MgSO_4$, filtered, and the solvent removed under reduced pressure. Compound **8.12** (12.33 g, 42.2 mmol, 55%) was received as a colorless oil.

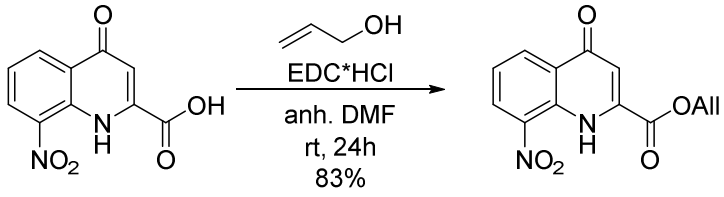
^{1}H NMR (500 MHz, $CDCl_3$): δ (ppm) = 7.40 – 7.30 (m, 10H), 5.12 (dd, J = 11.7, 8.9 Hz, 2H), 5.05 (dd, J = 11.7, 8.4 Hz, 2H), 3.87 (d, J = 6.3 Hz, 2H).

^{13}C NMR (126 MHz, $CDCl_3$): δ (ppm) = 136.16 (d, *J* = 5.5 Hz), 128.80, 128.75, 128.21, 68.17 (d, *J* = 6.8 Hz), 57.71 (d, *J* = 159.8 Hz).

^{31}P NMR (202 MHz, $CDCl_3$): δ (ppm) = 24.40.

HRMS: (ESI$^+$) m/z calc. for $C_{15}H_{17}O_4P$: 315.0762 $(M+Na)^+$; found: 315.0757.

Synthesis of compound **8.10**



In a dry flask, compound **8.8** (5.27 g, 22.3 mmol, 1.0 eq.) and EDC-HCl (6.42 g, 33.5 mmol, 1.5 eq.) dissolved in anhydrous DMF (72 mL) under N_2-atmosphere. Allyl alcohol (1.43 g, 1.68 mL, 24.5 mmol, 1.1 eq.) was added at 0 °C over 5 min, warmed to rt and stirred for 20 h. The mixture was poured on ice, the precipitate was vacuum filtered and washed with cold

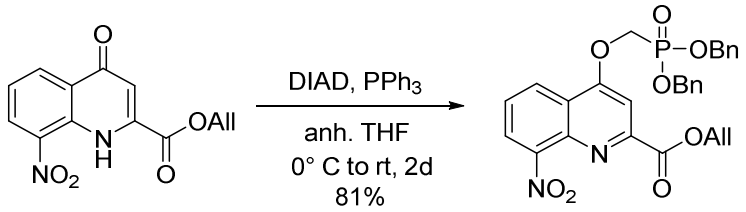
water. The received solid was dried in a vacuum oven at 60 °C over night to give the title compound (5.09 g, 18.6 mmol, 83%) as a yellow solid.

^{1}H NMR (500 MHz, $CDCl_3$) δ (ppm) = 11.80 (s, 1H), 8.78 – 8.69 (m, 2H), 7.49 (t, J = 8.0 Hz, 1H), 7.10 (d, J = 1.8 Hz, 1H), 6.04 (ddt, J = 16.5, 11.0, 5.9 Hz, 1H), 5.50 (dd, J = 17.3, 1.7 Hz, 1H), 5.41 (d, J = 10.1 Hz, 1H), 4.96 (d, J = 6.1 Hz, 2H).

^{13}C NMR (126 MHz, $CDCl_3$) δ (ppm) = 178.09, 161.63, 137.58, 136.09, 135.36, 134.41, 131.29, 130.49, 128.42, 122.92, 120.46, 113.69, 68.02.

HRMS: (ESI^-) m/z calc. for $C_{13}H_9N_2O_5^-$ $(M-H)^-$: 273.0517; found: 273.0515.

Synthesis of compound **8.13**



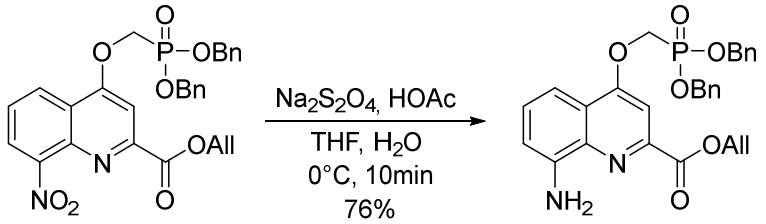
Freshly dried compound **8.10** (1.66 g, 6.05 mmol, 1.0 eq.), $P(Ph)_3$ (2.06 g, 7.87 mmol, 1.3 eq.), and compound **8.12** (1.95 g, 6.66 mmol, 1.1 eq.) were suspended in anhydrous THF under N_2-atmosphere and cooled to 0 °C. Diisopropyl azodicarboxylate (1.54 mL, 7.87 mmol, 1.3 eq.) was added dropwise over a period of 20 min at 0 °C. The mixture was stirred at 0 °C for an additional 30 min and then at rt overnight. After that, an additional 0.1 eq. of dibenzyl hydroxymethyl phosphonate, and 0.2 eq. of each DIAD and $P(Ph)_3$ was added, and the mixture was stirred at rt for 24h. Then, the solvent was removed under reduced pressure, and the residue was purified by column chromatography (silica, 50% EtOAc in cHex). To give the title compound (2.67 g, 4.88 mmol, 81%) as a yellow solid.

^{1}H NMR (500 MHz, $CDCl_3$): δ (ppm) = 8.25 (dd, *J* = 8.5, 1.4 Hz, 1H), 8.10 (dd, *J* = 7.5, 1.4 Hz, 1H), 7.61 (dd, *J* = 8.5, 7.5 Hz, 1H), 7.54 (s, 1H), 7.39 – 7.33 (m, 4H), 7.33 – 7.27 (m, 6H), 6.14 – 6.00 (m, 1H), 5.52 (dq, *J* = 17.1, 1.5 Hz, 1H), 5.35 (dq, *J* = 10.4, 1.3 Hz, 1H), 5.25 – 5.08 (m, 4H), 4.92 (dt, *J* = 5.6, 1.4 Hz, 2H), 4.42 (d, *J* = 10.2 Hz, 2H).

^{13}C NMR (125 MHz, $CDCl_3$): δ (ppm) = 164.49, 162.12 (d, *J* = 13.4 Hz), 151.26, 148.46, 140.12, 135.54 (d, *J* = 5.0 Hz), 131.46, 129.05, 128.87, 128.37, 126.52, 126.33, 125.35, 122.82, 119.20, 102.20, 68.78 (d, *J* = 6.4 Hz), 67.04, 62.93 (d, *J* = 171.4 Hz).

HRMS: (ESI^+) m/z calc. for $C_{28}H_{25}N_2O_8P$: 571.1246 $(M+Na)^+$; found: 571.1245.

Synthesis of compound **8.14**



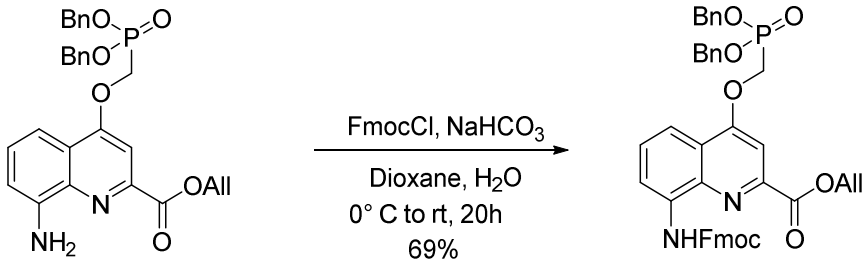
Compound **8.13** (2.4 g, 4.4 mmol, 1.0 eq.) was dissolved in a THF/H_2O (30 mL/30 mL) mixture, cooled to 0 °C and acetic acid (2.6 g, 2.5 mL, 44 mmol, 10 eq.) was added. Then sodium dithionite (2.3 g, 13 mmol, 3.0 eq.) was added and the mixture was stirred at 0 °C for 10 min. The mixture was diluted with ethyl acetate (20 mL) and the organic phase was washed with citric acid (5%, 2 x 20 mL), $NaHCO_3$ (5%, 2 x 20 mL) and brine (20 mL). The organic phase was dried over $MgSO_4$ filtered and the solvent was removed under reduced pressure. Compound **8.14** (1.7 g, 3.3 mmol, 76%) was received as a yellow oil.

^{1}H NMR (500 MHz, $CDCl_3$): δ (ppm) = 7.43 – 7.33 (m, 8H), 7.32 – 7.27 (m, 6H), 6.95 (dd, *J* = 7.4, 1.4 Hz, 1H), 6.10 (ddt, *J* = 17.2, 10.4, 5.8 Hz, 1H), 5.47 (dq, *J* = 17.2, 1.5 Hz, 1H), 5.34 (dq, *J* = 10.4, 1.3 Hz, 1H), 5.18 (qd, *J* = 11.7, 9.1 Hz, 5H), 4.91 (dt, *J* = 5.8, 1.4 Hz, 2H), 4.41 (d, *J* = 10.3 Hz, 2H).

^{13}C NMR (126 MHz, $CDCl_3$): δ (ppm) = δ 165.17, 161.98 (d, *J* = 13.9 Hz), 145.64, 144.96, 138.50, 135.72 (d, *J* = 5.4 Hz), 132.05, 129.28, 128.85, 128.80, 128.26, 122.61, 119.04, 111.24, 109.65, 100.80, 68.66 (d, *J* = 6.4 Hz), 66.59, 62.55 (d, *J* = 171.2 Hz).

HRMS: (ESI$^+$) m/z calc. for $C_{28}H_{27}N_2O_8P$: 519.1685 $(M+H)^+$; found: 519.1680.

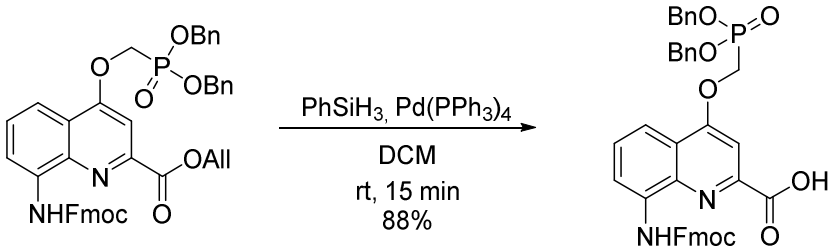
Synthesis of compound **8.15**



Compound **8.14** (3.82 g, 7.37 mmol, 1.0 eq.) and $NaHCO_3$ (13.0 g, 155 mmol, 21.0 eq.) was dissolved in a mixture of H_2O (125 mL) and 1,4-dioxane (75mL) and cooled to 0 °C. Fmoc-Cl (1.91 g, 7.37 mmol, 1.0 eq.) dissolved in 1,4-dioxane (114 mL) was added dropwise over 1h at 0 °C. The mixture was stirred at 0 °C, allowed to warm to rt and then stirred at rt for 20h. 1,4-dioxane was removed under reduced pressure and diluted with H_2O (30 mL) and ethyl acetate (150 mL). The phases were allowed to separate, and the aqueous phase was extracted with ethyl acetate (3x 120 mL). The combined organic phases were dried over $MgSO_4$, filtered, and the solvent removed under reduced pressure. The crude product was dissolved in acetonitrile

(10 mL) and water (2 mL) and purified using reversed-phase chromatography (C18, 50-100% acetonitrile in water) to receive the title compound (3.77 g, 5.09 mmol, 69%) as a yellow solid.
^{1}H NMR (500 MHz, $CDCl_3$): δ (ppm) = 9.32 (s, 1H), 8.50 (bs, 1H), 7.80 (d, *J* = 7.5 Hz, 1H), 7.76 – 7.69 (m, 3H), 7.55 (t, *J* = 8.2 Hz, 1H), 7.48 (s, 1H), 7.43 (t, *J* = 7.5 Hz, 1H), 7.39 – 7.32 (m, 6H), 7.32 – 7.28 (m, 6H), 6.14 (ddt, *J* = 17.2, 10.4, 5.7 Hz, 1H), 5.55 (dq, *J* = 17.2, 1.5 Hz, 1H), 5.38 (dq, *J* = 10.3, 1.2 Hz, 1H), 5.25 – 5.12 (m, 4H), 4.96 (dt, *J* = 5.7, 1.4 Hz, 2H), 4.55 (d, *J* = 7.0 Hz, 2H), 4.43 (d, *J* = 10.2 Hz, 2H), 4.38 (t, *J* = 7.3 Hz, 1H).
^{13}C NMR (126 MHz, $CDCl_3$): δ (ppm) = 164.82, 162.40 (d, *J* = 13.4 Hz), 153.50, 146.80, 143.97, 141.49, 138.56, 135.66 (d, *J* = 5.3 Hz), 135.43, 131.83, 129.02, 128.95, 128.87, 128.85, 128.32, 127.96, 127.31, 125.37, 121.86, 120.20, 119.12, 116.20, 114.83, 101.31, 68.73 (d, *J* = 6.4 Hz), 67.59, 66.75, 62.76 (d, *J* = 171.2 Hz), 47.23.
HRMS: (ESI$^+$) m/z calc. for $C_{43}H_{37}N_2O_8P$: 741.2366 $(M+H)^+$; found: 741.2351.

Synthesis of compound **8.3**

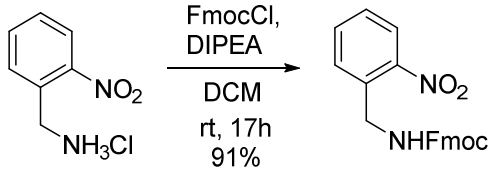


In a light protected flask, compound **8.15** (1.05 g, 1.42 mmol, 1.00 eq.) was dissolved in dichloromethane (33.0 mL). Then $PhSiH_3$ (199 mg, 227 µL, 1.84 mmol, 1.30 eq.) and $Pd(PPh_3)_4$ (81.9 mg, 70.9 µmol, 0.05 eq.) were added and the mixture was stirred 15 min at rt. The mixture was poured on citric acid (5%, 50 mL) and aerated well. The phases were separated, and the organic phase was dried over $MgSO_4$, filtered, and the solvent removed under reduced pressure. Toluene (30 mL) was added, and the mixture was sonicated for 30min and left in the fridge (4°C) for 1h. The precipitant was vacuum filtered, washed with cold toluene, dissolved in DCM/MeOH/acetone, and purified by column chromatography (silica, DCM/MeOH/AcOH, 97:2.9:0.1). Fractions containing the product were concentrated under reduced pressure to a volume of 100 mL and the organic phase was washed with H_2O (2x). The phases were separated, the organic phase was concentrated *in vacuo*, then freeze-dried to give compound **1** (875 mg, 1.25 mmol, 88%) as a yellow solid.
^{1}H NMR (500 MHz, DMSO-d_6): δ (ppm) = 13.54 (s, 1H), 10.44 (s, 1H), 8.35 (bs, 1H), 7.93 (d, *J* = 7.6 Hz, 1H), 7.82 – 7.76 (m, 3H), 7.67 – 7.60 (m, 1H), 7.44 (t, *J* = 7.4 Hz, 2H), 7.42 – 7.36 (m, 5H), 7.37 (d, *J* = 1.2 Hz, 1H), 7.37 – 7.29 (m, 7H), 5.26 – 5.14 (m, 4H), 5.02 (d, *J* = 9.9 Hz, 2H), 4.62 (d, *J* = 6.9 Hz, 2H), 4.45 (t, *J* = 6.8 Hz, 1H).
^{13}C NMR (126 MHz, DMSO-d_6): δ (ppm) = 165.26, 162.28 (d, *J* = 13.8 Hz), 153.50, 146.64, 143.70, 140.81, 137.48, 136.19 (d, *J* = 6.1 Hz), 135.69, 128.80, 128.75, 128.71, 128.42, 128.30,

127.80, 127.22, 125.19, 125.15, 121.55, 120.25, 116.64, 114.35, 101.25, 67.49 (d, *J* = 5.9 Hz), 66.41, 61.93 (d, *J* = 164.7 Hz), 46.58.

HRMS: (ESI$^-$) m/z calc. for $C_{40}H_{33}N_2O_8P$: 699.1902 $(M-H)^-$; found: 699.1902.

Synthesis of compound **8.21**



(2-Nitrophenyl)methanamine hydrochloride (25.0 g, 133 mmol, 1 eq.) was dissolved in a mixture of DCM (400 mL) and DIPEA (46.4 mL, 266 mmol, 2 eq.) and cooled to 0 °C. Fmoc-Cl (35.0 g, 135 mmol, 1.02 eq.) dissolved in DCM (250 mL) was added dropwise at 0 °C. After addition, the mixture was allowed to warm to rt and stirred at rt for 17 h. The organic phase was washed with citric acid (5% w/v, 2x) and water (1x). The organic phase was dried over $MgSO_4$, filtered, and concentrated *in vacuo*. Upon starting precipitation, diisopropyl ether (300 mL) was added and the mixture was sonicated for 10 min. The colorless precipitate was filtered, washed with cold diisopropyl ether and was dried on air over night to give the title compound as an off-white solid (45.2 g, 121 mmol, 91%).

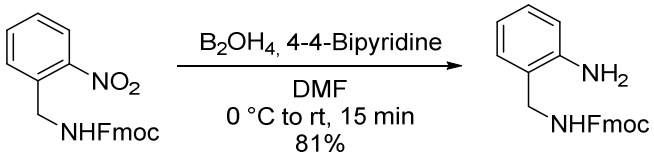
NMR shows 2 conformers. Major 85%, minor 15%. Resolved signals are given with their respective integration. For unresolved signals, total integration is given. ^{13}C shifts of the minor species are not listed.

^{1}H NMR (500 MHz, $CDCl_3$): δ (ppm) = 8.11 – 8.02 (m, 1H), 7.76 (d, *J* = 7.6 Hz, 2H), 7.65 – 7.54 (m, 3.40H), 7.50 – 7.44 (m, 1.15H), 7.40 (t, *J* = 7.5 Hz, 2.30H), 7.30 (t, *J* = 7.4 Hz, 2H), 6.89 (t, *J* = 7.4 Hz, 0.15H), 5.61 (t, *J* = 6.7 Hz, 0.85H), 5.24 (t, *J* = 6.7 Hz, 0.15H), 4.64 (d, *J* = 6.6 Hz, 2H), 4.42 (d, *J* = 6.9 Hz, 1.7H), 4.31 (d, *J* = 6.0 Hz, 0.3H), 4.20 (t, *J* = 6.8 Hz, 1H).

^{13}C NMR (126 MHz, $CDCl_3$): δ (ppm) = 156.57, 148.33, 143.93, 141.46, 134.26, 134.10, 132.07, 128.93, 127.85, 127.19, 125.35, 125.14, 120.14kkf, 67.01, 47.37, 43.08.

HRMS: (ESI$^+$) m/z calc. for $C_{22}H_{18}N_2O_4Na$ $(M+Na)^+$: 397.1164; found: 397.1166.

Synthesis of compound **8.22** (CAUTION: BE CAREFUL UPON SCALEUP)



Compound **8.21** (10.0 g, 26.7 mmol, 1 eq.) and 4-4-bipyridine (20.9 mg, 0.133 mmol, 0.005 eq.) were dissolved in DMF (200 mL) to give a yellow solution and cooled to 0 °C. To this, hypodiboric acid (7.23 g, 80.7 mmol, 3.02 eq.) was added portion wise as a solid within 3 min. The mixture was allowed to warm to room temperature. Upon start of the reaction, a lot

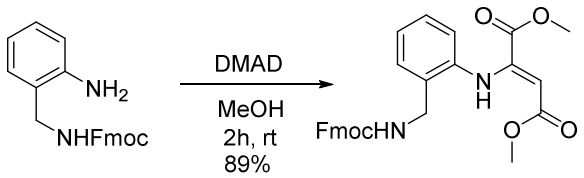
of heat will be produced, and the solution will turn dark blue for ca. 1 min and will form a yellow solution afterwards. The solution was stirred for 15 min at rt and then poured on water (600 mL). The aqueous phase was extracted with EtOAc (500 mL). The organic phase was washed with LiCl (5%, 5x), dried over $MgSO_4$ filtered and the solvent removed *in vacuo*. The resulting solid was sonicated in diisopropyl ether (150 mL), the resulting colorless precipitate was filtered and washed with cold diisopropyl ether and the colorless solid was dried on air to give the title compound (7.43 g, 21.6 mmol, 81%).

^{1}H NMR (500 MHz, $CDCl_3$): δ (ppm) = 7.76 (d, *J* = 7.6 Hz, 2H), 7.58 (d, *J* = 7.5 Hz, 2H), 7.40 (t, *J* = 7.5 Hz, 2H), 7.30 (td, *J* = 7.5, 1.2 Hz, 2H), 7.12 (td, *J* = 7.6, 1.6 Hz, 1H), 7.04 (d, *J* = 7.3 Hz, 1H), 6.78 – 6.65 (m, 2H), 5.04 (bs, 1H), 4.44 (d, *J* = 6.9 Hz, 2H), 4.32 (d, *J* = 6.2 Hz, 2H), 4.21 (t, *J* = 7.0 Hz, 1H), 4.14 (bs, 2H).

^{13}C NMR (126 MHz, $CDCl_3$): δ (ppm) = 157.10, 145.55, 143.95, 141.46, 130.45, 129.41, 127.84, 127.19, 125.16, 122.16, 120.13, 118.16, 116.04, 67.10, 47.35, 42.63.

HRMS: (ESI$^+$) m/z calc. for $C_{22}H_{20}N_2O_2Na$ $(M+Na)^+$: 367.1422; found: 367.1424.

Synthesis of compound **8.23**



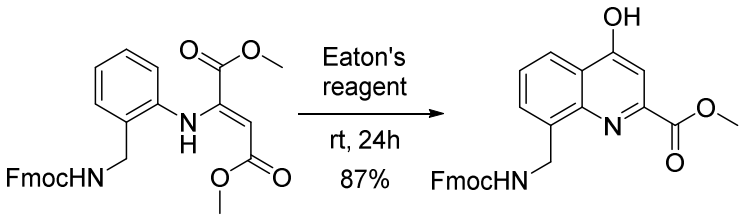
Compound **8.22** (7.43 g, 21.5 mmol, 1 eq.) was suspended in MeOH (75 mL). To this, dimethyl acetylene dicarboxylate (3.58 mL, 23,7 mmol, 1.1 eq.) was added and the suspension was stirred for 2 h at rt. The formed product was filtered off and washed with cold MeOH (2x 50 mL) to give the title compound (7.63 g, 15.7 mmol, 73%) as a colorless solid. The filtrate was kept in the freezer for 7 d, to form more precipitate. Isolation by filtration yielded another batch of the title compound t (1.71 g, 3.51 mmol, 16%) as a pale-yellow solid.

^{1}H NMR (500 MHz, $CDCl_3$): δ (ppm) = 9.57 (s, 1H), 7.75 (d, *J* = 7.5 Hz, 2H), 7.60 (d, *J* = 7.4 Hz, 2H), 7.39 (t, *J* = 7.5 Hz, 2H), 7.34 (d, *J* = 7.5 Hz, 1H), 7.28 (t, *J* = 7.4 Hz, 1H), 7.22 (td, *J* = 7.6, 1.7 Hz, 1H), 7.14 (t, *J* = 7.5 Hz, 1H), 5.50 (s, 1H), 5.40 (s, 1H), 4.47 (d, *J* = 5.7 Hz, 2H), 4.42 (d, *J* = 7.2 Hz, 2H), 4.25 (t, *J* = 7.1 Hz, 1H), 3.69 (s, 3H), 3.63 (s, 3H).

^{13}C NMR (126 MHz, $CDCl_3$): δ (ppm) = 170.02, 164.83, 156.55, 148.91, 144.15, 141.43, 139.51, 131.55, 130.24, 130.22, 128.71, 127.77, 127.19, 127.15, 125.84, 125.32, 123.31, 120.06, 94.77, 67.03, 52.93, 51.43, 47.40, 42.37.

HRMS: (ESI$^+$) m/z calc. for $C_{28}H_{26}N_2O_6Na$ $(M+Na)^+$: 509.1689; found: 509.1700.

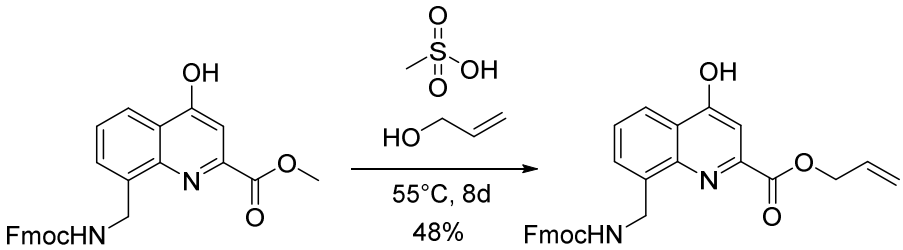
Synthesis of compound **8.24**



Compound **8.23** (5.15 g, 10.6 mmol, 1 eq.) was dissolved portion wise over 20 min in Eaton's reagent (100 mL) and the solution was stirred at rt for 24h. The solution was slowly poured in $NaHCO_3$ (5% w/v, 500 mL) and basified to a pH = 7 by adding further solid $NaHCO_3$. The aqueous solution was extracted with $CHCl_3$ (3 x 500 mL), the combined organic phases were dried over $MgSO_4$, filtered, and the solvent was removed *in vacuo*. The resulting oil was co-evaporated with DCM. The resulting solid was sonicated in diisopropyl ether (200 mL), filtered and washed with diisopropyl ether (2x) and diethyl ether (1x) to give the title compound as colorless solid (4.19 g, 9.22 mmol, 87%).

^{1}H NMR (500 MHz, DMSO-d_6): δ (ppm) = 10.70 (s, 1H), 8.34 (dd, *J* = 8.3, 1.6 Hz, 1H), 7.73 (dt, *J* = 7.6, 0.9 Hz, 2H), 7.54 – 7.46 (m, 3H), 7.28 (dd, *J* = 8.2, 7.1 Hz, 1H), 7.21 (td, *J* = 7.5, 1.1 Hz, 2H), 7.00 (d, *J* = 1.7 Hz, 1H), 5.47 (s, 1H), 4.64 (d, *J* = 6.7 Hz, 2H), 4.55 (d, *J* = 6.8 Hz, 2H), 4.21 (t, *J* = 6.8 Hz, 1H), 3.95 (s, 3H).
^{13}C NMR (126 MHz, DMSO-d_6): δ (ppm) = ^{13}C NMR (126 MHz, $CDCl_3$) δ 179.67, 162.95, 143.65, 141.44, 138.23, 137.63, 134.45, 127.96, 127.46, 127.18, 127.15, 126.99, 124.98, 123.89, 120.15, 112.11, 77.36, 67.62, 53.66, 47.26, 42.08.
HRMS: (ESI$^-$) m/z calc. for $C_{27}H_{21}N_2O_5$ $(M-H)^-$: 453.1456; found: 453.1453.

Synthesis of compound **8.25**



Compound **8.24** (1.78 g, 3.92 mmol, 1 eq.) was suspended in allyl alcohol (50.6 mL, 744 mmol, 190 eq.). Methanesulfonic acid (5.09 mL, 78.3 mmol, 20 eq.) was added dropwise over 5 min. The solution was stirred in an open flask at 55 °C for 8 d. The dark solution was neutralized with $NaHCO_3$ (5% w/v) to pH = 6-8. The aqueous solution was extracted with $CHCl_3$ (4x) . The combined organic phases were dried over $MgSO_4$, filtered, and the organic phase was removed *in vacuo*. The dark oil was co-evaporated with MeCN (1x), and then the residue was redissolved in MeCN (ca. 20 mL). The product was sonicated in MeCN and left at 4 °C over night to allow

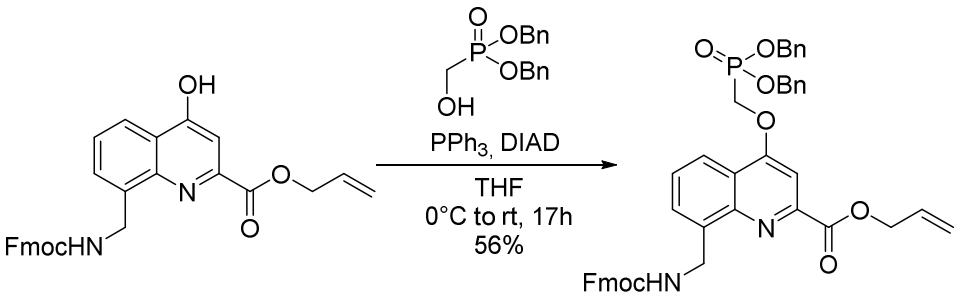
further precipitation. The product was filtered off and washed with cold MeCN and dried in the filter to yield the title compound as a grey solid (839 mg, 1.75 mmol , 48%).

^{1}H NMR (500 MHz, DMSO-d_6): δ (ppm) = 10.74 (s, 1H), 8.28 (d, *J* = 7.8 Hz, 1H), 7.72 (d, *J* = 7.6 Hz, 2H), 7.52 (d, *J* = 7.5 Hz, 2H), 7.47 (d, *J* = 6.8 Hz, 1H), 7.35 (t, *J* = 7.5 Hz, 2H), 7.25 – 7.15 (m, 3H), 7.01 (d, *J* = 1.3 Hz, 1H), 5.97 (ddt, *J* = 17.2, 10.4, 5.9 Hz, 1H), 5.82 (s, 1H), 5.42 – 5.34 (m, 1H), 5.26 (dt, *J* = 10.4, 1.2 Hz, 1H), 4.82 (dt, *J* = 5.8, 1.4 Hz, 2H), 4.52 (d, *J* = 7.0 Hz, 2H), 4.21 (t, *J* = 7.0 Hz, 1H).

^{13}C NMR (126 MHz, DMSO-d_6): δ (ppm) = δ 179.64, 162.11, 157.76, 143.67, 141.40, 138.20, 137.66, 134.49, 130.99, 127.91, 127.34, 127.12, 126.77, 125.03, 123.80, 120.11, 119.76, 112.09, , 77.36, 67.65, 67.42, 47.22, 42.07.

HRMS: (ESI$^+$) m/z calc. for $C_{27}H_{21}N_2O_5$ $(M+H)^+$: 481.1758; found: 481.1753.

Synthesis of compound **8.26**



Compound **8.25** (3.91 g, 8.14 mmol, 1 eq.), PPh_3 (2.77 g, 10.6 mmol, 1.3 eq.) and compound **8.12** (2.62 g, 8.95 mmol, 1.1 eq.) was dissolved in anhydrous THF and cooled to 0 °C. Diisopropyl azodicarboxylate (2.08 mL, 10.6 mmol, 1.3 eq.) was added dropwise over 30 min at 0 °C, the mixture was stirred for 1 h at 0 °C and then at rt for 17 h. Further PPh_3 (320 mg, 1.22 mmol, 0.15 eq.) and diisopropyl azodicarboxylate (240 µL, 1.22 mmol, 1.3 eq.) was added and the mixture was stirred for further 6 h. THF was removed *in vacuo*, and co-evaporated twice with DCM to receive a dark oil. To this isopropanol (50 mL) was added to cause precipitation. The mixture was sonicated for 3 min and left at −24 °C over night. The precipitation was filtered off, washed with cold isopropanol at then further purified by column chromatography (silica, loading solvent: DCM, 40% EtOAc in cHex) to yield the title compound as a colorless solid (3.41 g, 4.52 mmol, 56%).

NMR shows 2 conformers. Major 85%, minor 15%. Resolved signals are given with their respective integration. For unresolved signals, total integration is given. ^{13}C shifts of the minor species are not listed.

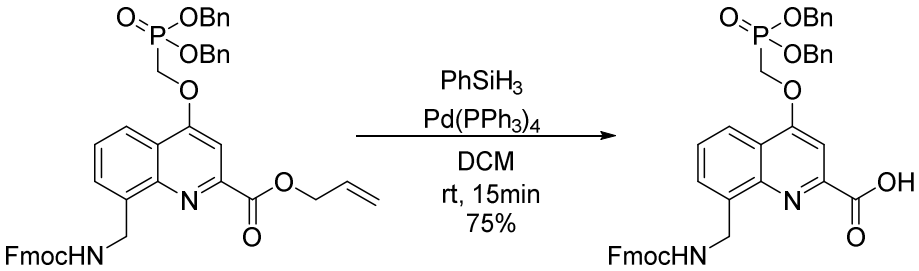
^{1}H NMR (500 MHz, $CDCl_3$): δ (ppm) = 8.10 – 8.01 (m, 1H), 7.82 – 7.69 (m, 2.8H), 7.57 – 7.38 (m, 4.3H), 7.38 – 7.32 (m, 5.7H), 7.32 – 7.26 (m, 6.2H), 7.23 (t, *J* = 7.4 Hz, 2H), 6.56 (t, *J* = 6.4 Hz, 0.85H), 6.11 (ddt, *J* = 16.3, 10.8, 5.6 Hz, 1H), 5.93 – 5.87 (m, 0.15H), 5.55 – 5.47 (m, 1H), 5.34 (dt, *J* = 10.5, 1.3 Hz, 1H), 5.25 – 5.11 (m, 4H), 4.98 – 4.89 (m, 3.7H), 4.70 (d, *J* = 6.8

Hz, 0.3H), 4.52 (d, *J* = 5.9 Hz, 0.3H), 4.44 (d, *J* = 10.2 Hz, 2H), 4.34 (d, *J* = 7.3 Hz, 1.7H), 4.25 (t, *J* = 7.0 Hz 0.15H), 4.19 (t, *J* = 7.3 Hz, 0.85H).

^{13}C NMR (126 MHz, $CDCl_3$): δ (ppm) = 164.99, 162.53 (d, *J* = 13.5 Hz), 156.78, 147.90, 147.21, 144.16, 141.37, 137.30, 135.66 (d, *J* = 5.2 Hz), 131.82, 130.43, 128.94, 128.87, 128.84, 128.80, 128.30, 128.27, 127.97, 127.71, 127.07, 125.27, 122.22, 121.58, 120.01, 119.00, 100.78, 77.36, 68.70 (d, *J* = 6.4 Hz), 66.80, 66.69, 62.67 (d, *J* = 171.4 Hz), 47.40, 43.40.

HRMS: (ESI$^+$) m/z calc. for $C_{27}H_{21}N_2O_5$ $(M+H)^+$: 755.2517; found: 755.2501.

Synthesis of compound **8.16**



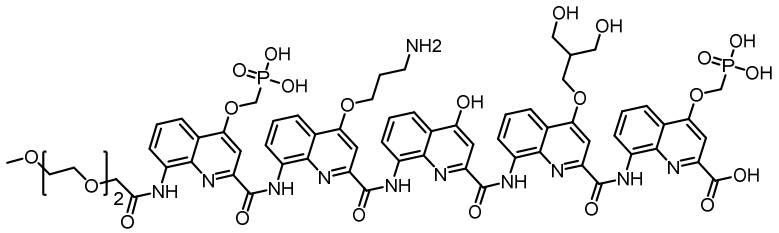
In a light-protected flask, compound **8.26** (1.8 g, 2.38 mmol, 1.00 eq.) was dissolved in degassed dichloromethane (50.0 mL). Then $PhSiH_3$ (882 µL, 7.15 mmol, 3.0 eq.) and $Pd(PPh_3)_4$ (82.7 mg, 71.5 µmol, 0.03 eq.) were added and the mixture was stirred 15 min at rt under N_2 atmosphere. The mixture was poured on citric acid (5%, 50 mL) and aerated well. The phases were separated, and the aqueous phase was extracted with DCM (2x50 mL). The combined organic phases were dried over $MgSO_4$, filtered, and the solvent removed under reduced pressure. Acetonitrile (50 mL) was added, and the mixture was sonicated for 1 min and left in the fridge (4°C) to precipitate for 5h. The precipitate was isolated by filtration and washed with cold MeCN (−24 °C) to receive the title compound as a dark-grey solid (1.28 g, 1.29 mmol, 75%). NMR shows 2 conformers. Major 85%, minor 15%. Resolved signals are given with their respective integration. For unresolved signals, total integration is given. ^{13}C shifts of the minor species are not listed.

^{1}H NMR (500 MHz, DMSO-d_6): δ (ppm) = δ 13.13 (s, 1H), 7.97 – 7.87 (m, 3.3H), 7.78 (d, *J* = 7.6 Hz, 0.3H), 7.74 (s, 1H), 7.70 (d, *J* = 7.5 Hz, 1.7H), 7.66 – 7.60 (m, 1.7H), 7.46 – 7.36 (m, 6H), 7.36 – 7.27 (m, 7.7H), 7.08 (t, *J* = 7.5 Hz, 0.3H), 5.26 – 5.14 (m, 4H), 5.00 (d, *J* = 9.8 Hz, 2H), 4.84 (d, *J* = 6.2 Hz, 1.7H), 4.74 (d, *J* = 6.4 Hz, 0.3H), 4.40 (d, *J* = 6.8 Hz, 1.7H), 4.32 (d, *J* = 6.4 Hz, 0.3H), 4.26 (d, *J* = 6.7 Hz, 0.85H), 4.13 (t, *J* = 6.4 Hz, 0.15H).

^{13}C NMR (126 MHz, DMSO-d_6): δ (ppm) = 166.00, 162.02 (d, *J* = 13.7 Hz), 156.56, 148.55, 145.25, 143.89, 140.80, 137.92, 136.23 (d, *J* = 6.0 Hz), 128.45, 128.33, 127.82, 127.70, 127.63, 127.07, 125.15, 121.11, 120.16, 120.15, 101.19, 67.50 (d, *J* = 5.9 Hz), 65.30, 61.76 (d, *J* = 164.8 Hz), 46.86, 40.57.

HRMS: (ESI$^-$) m/z calc. for $C_{27}H_{21}N_2O_5$ $(M-H)^-$: 713.2058; found: 713.2034.

Synthesis of compound **8.4**



Compound **8.4** was synthesized on LL-Wang resin (scale 30µmol), according to published procedures using common stock monomers, that were synthesized according to published procedures.[11, 16, 18, 19] The first monomer was loaded and the loading was determined according to section 7.3.2.3 of this thesis. The *N*-terminal peg capping group was installed using the same conditions as monomer coupling from the commercial acid. For silyl ester cleave, the resin was washed with DCM (5x) and then anhydrous THF (3x). Afterward, TBAF (4.2 mg, 160 µmol, 8 eq.) in anhyd. THF (4 mL) was added and heated by microwave irradiation 50 °C for 20 min. For resin cleavage & remaining side chain protection group cleavage, the resin was shaken in TFA for 4 h. Then the pentamer was precipitated with diethyl ether, dissolved in 200 mM NH_4OAc-buffer (pH = 8.5) and lyophilized to give a yellow solid. The crude product was purified via a basic semi-preparative HPLC with a gradient of 0-25 A in B, A = NH4OAc (12.5 mM, pH = 8.5), B = MeCN, and then lyophilized to give compound **9.4** (13 mg, 8.8 µmol, 29%) as a yellow solid.

8.4.3 NMR spectra

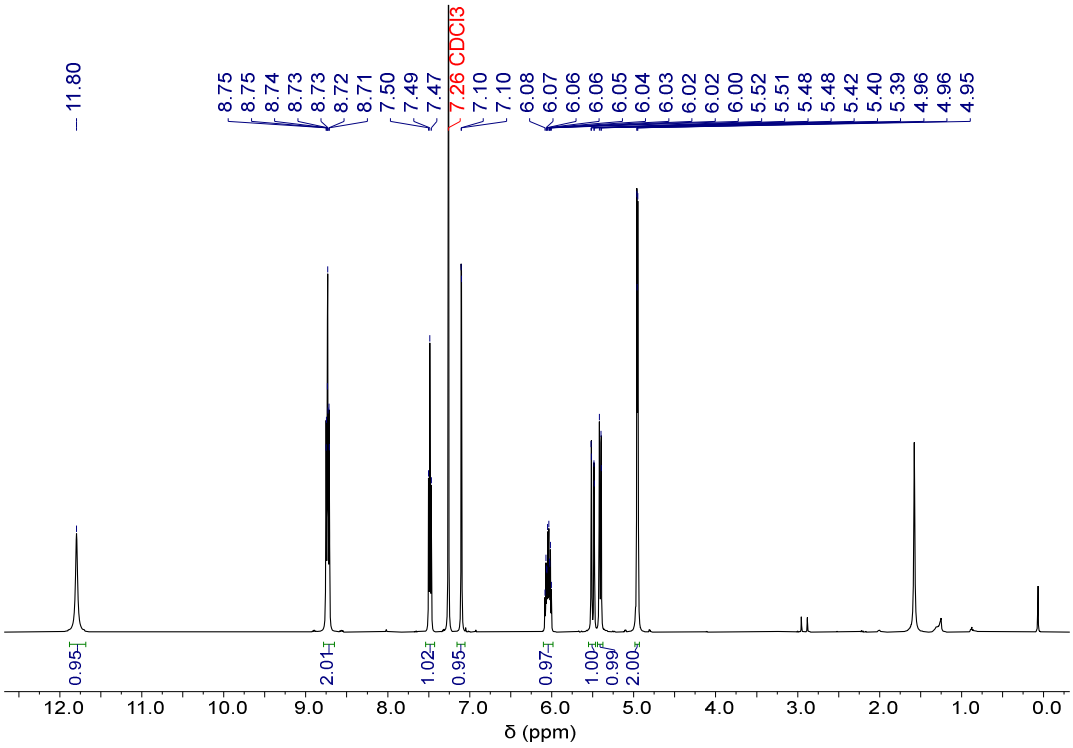


Figure S1 ^{1}H NMR spectrum of compound **8.10**.

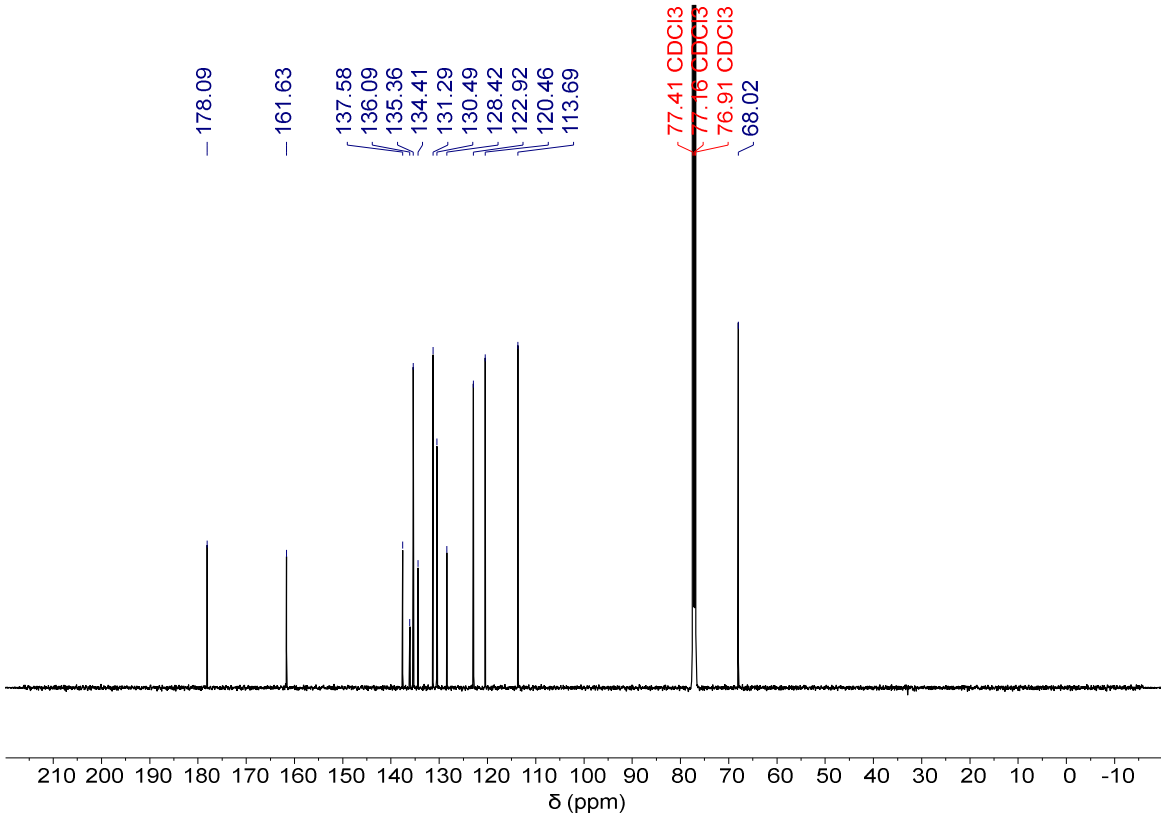


Figure S2 ^{13}C NMR spectrum of compound **8.10.**

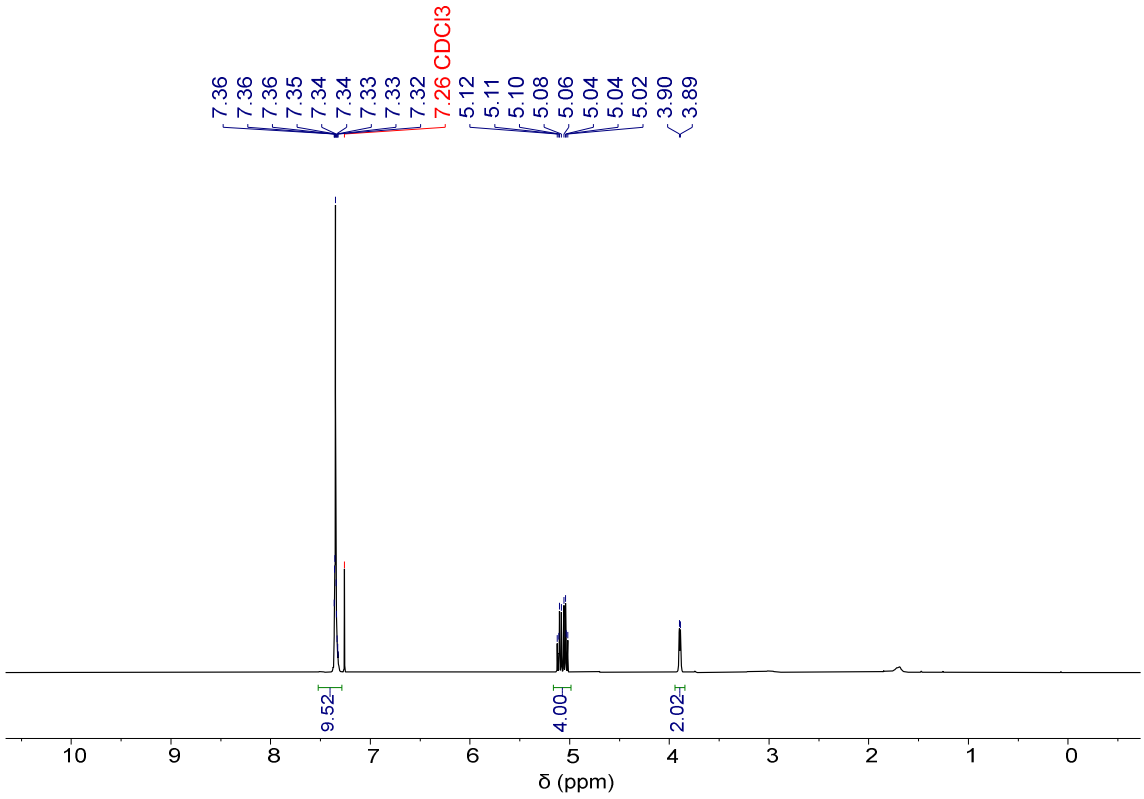


Figure S3 ^{1}H NMR spectrum of compound **8.12**.

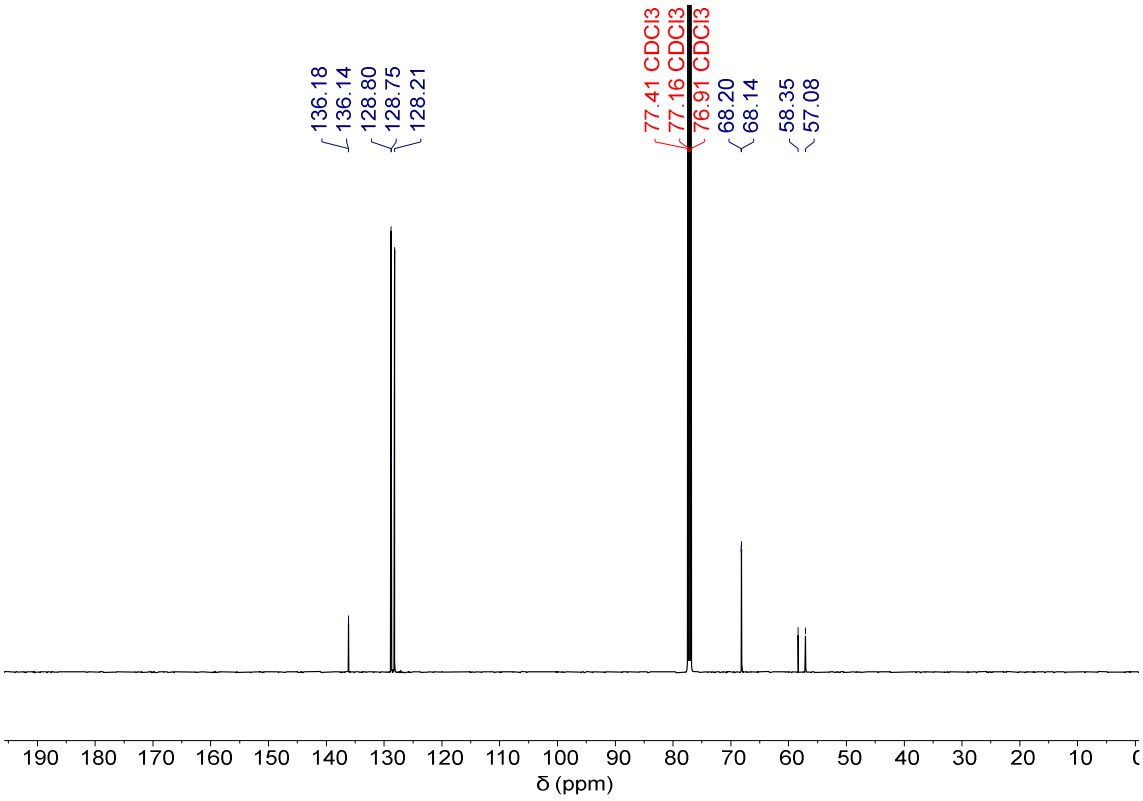


Figure S4 ^{13}C NMR spectrum of compound **8.12**.

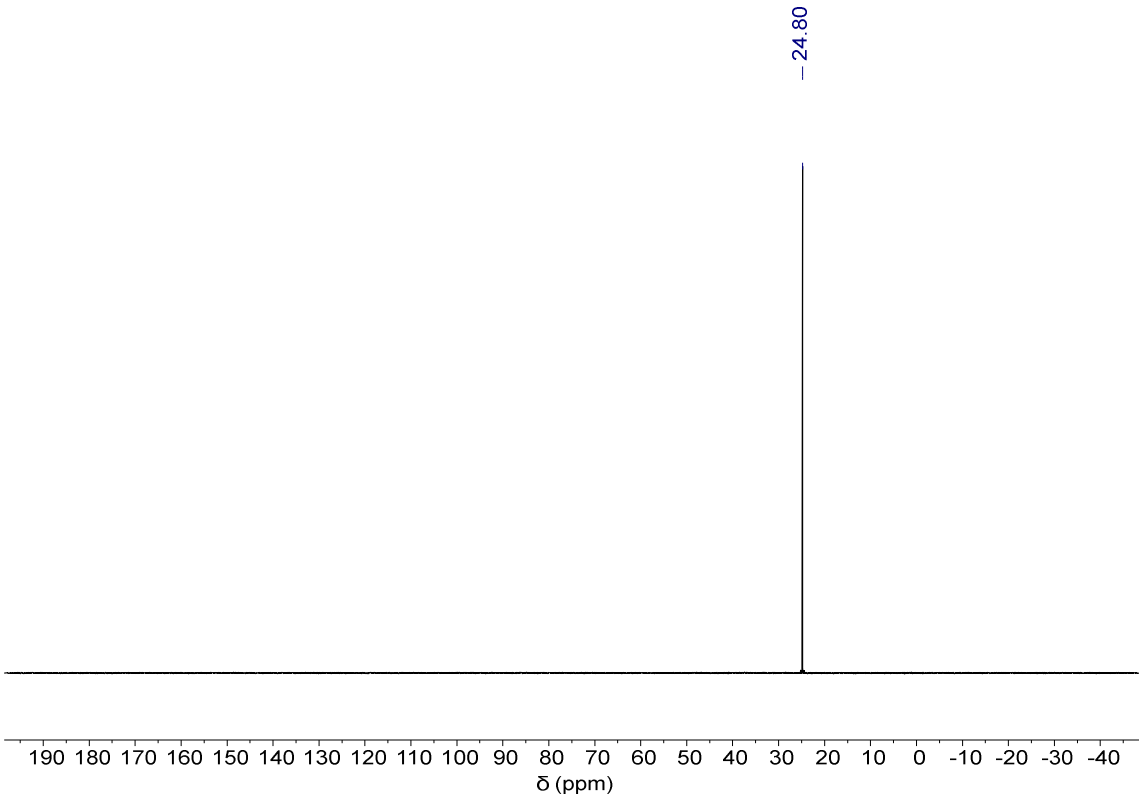


Figure S5 ^{31}P NMR spectrum of compound **8.12**.

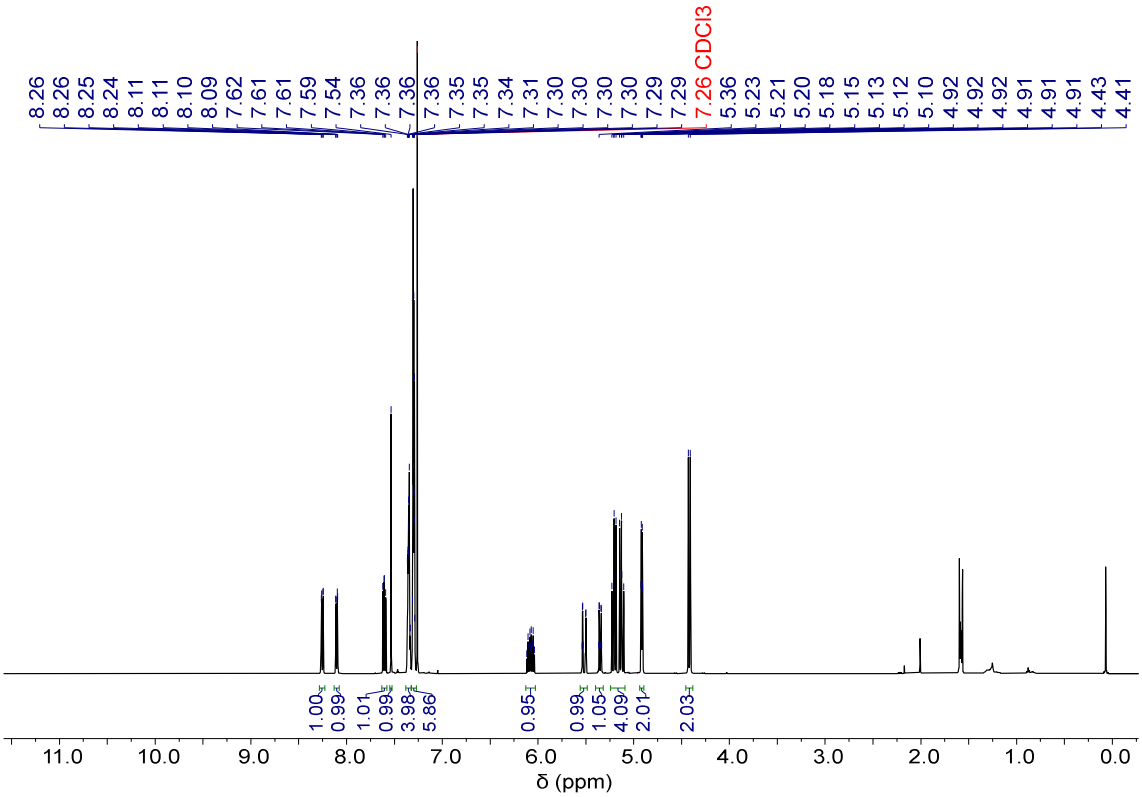


Figure S6 [1]H NMR spectrum of compound **8.13**.

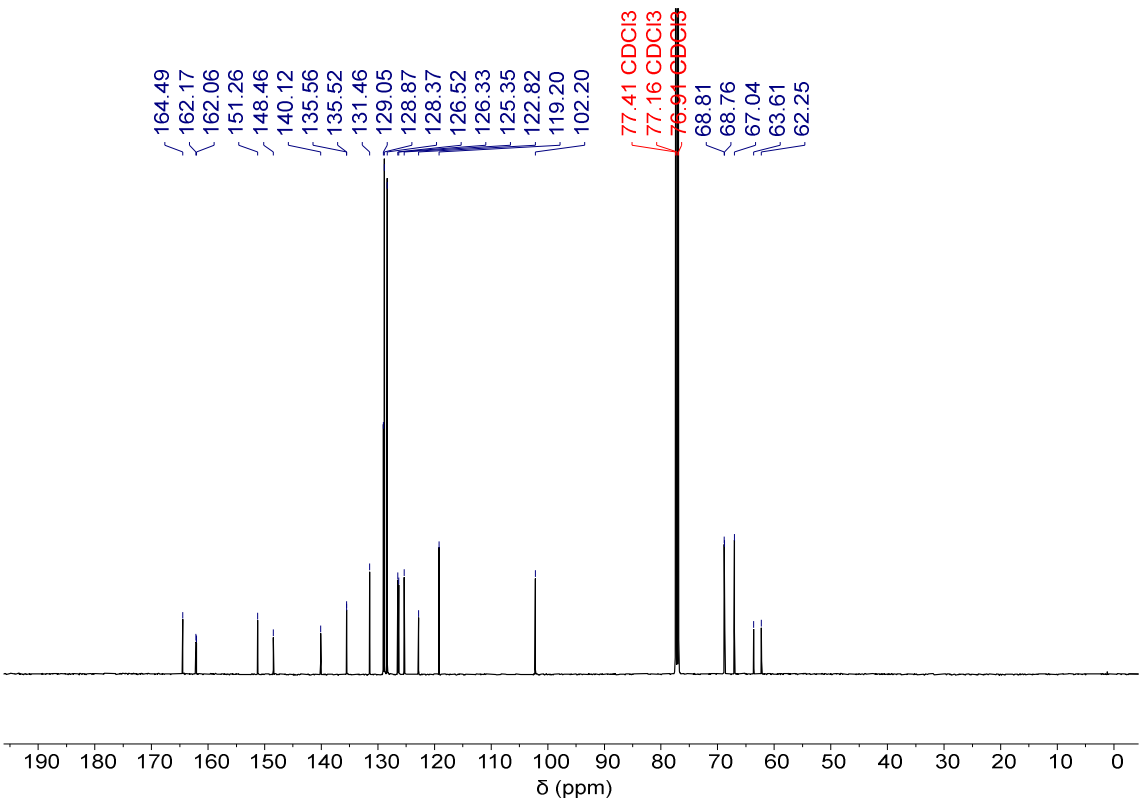


Figure S7 [13]C NMR spectrum of compound **8.13**.

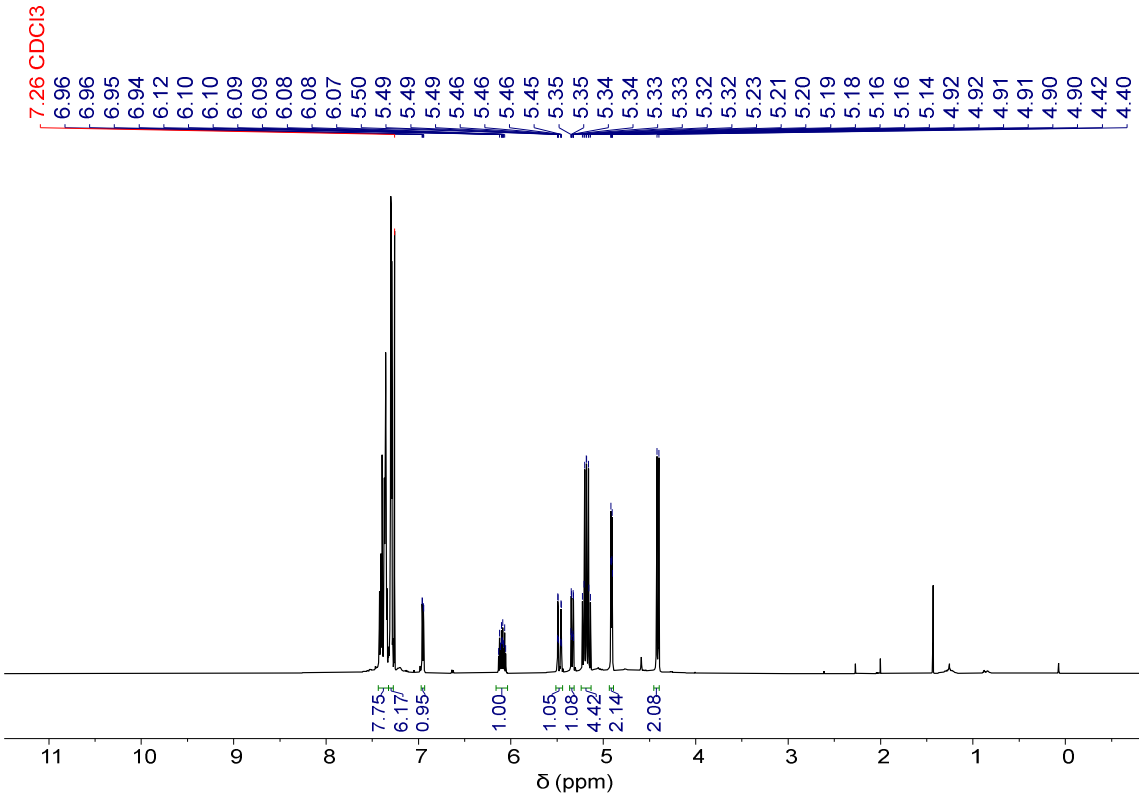


Figure S8 [1]H NMR spectrum of compound **8.14.**

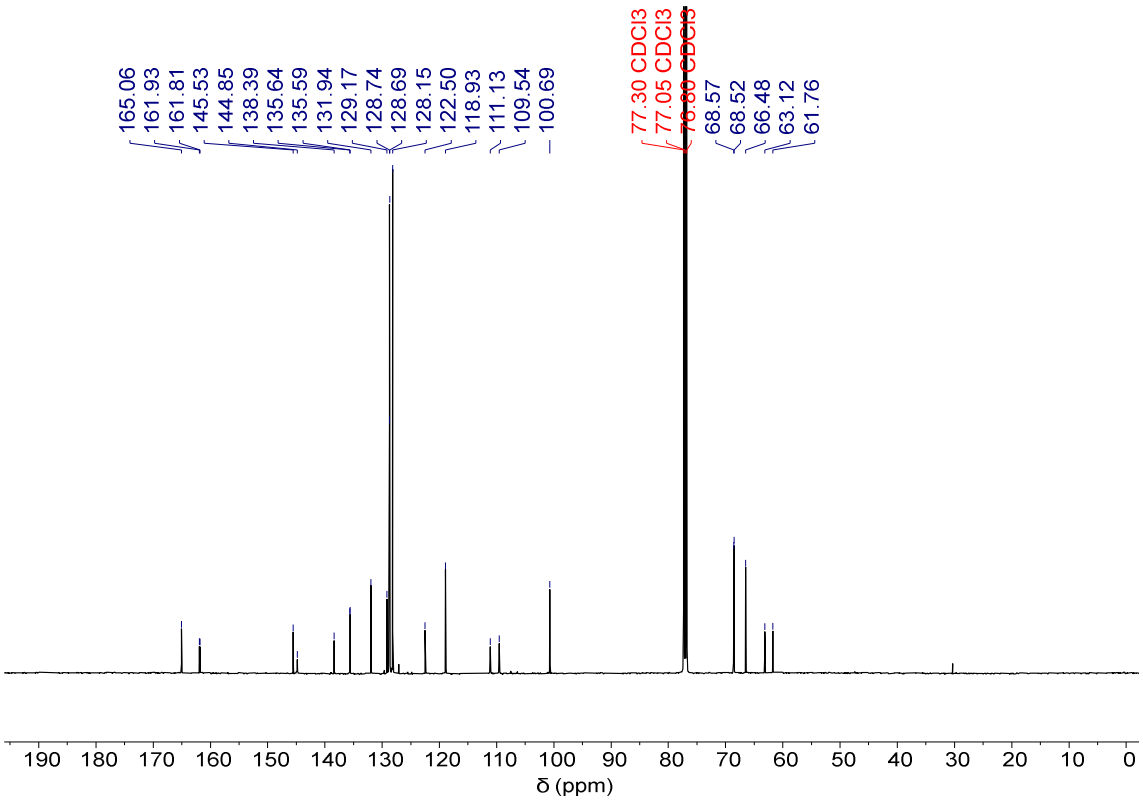


Figure S9 [13]C NMR spectrum of compound **8.14**.

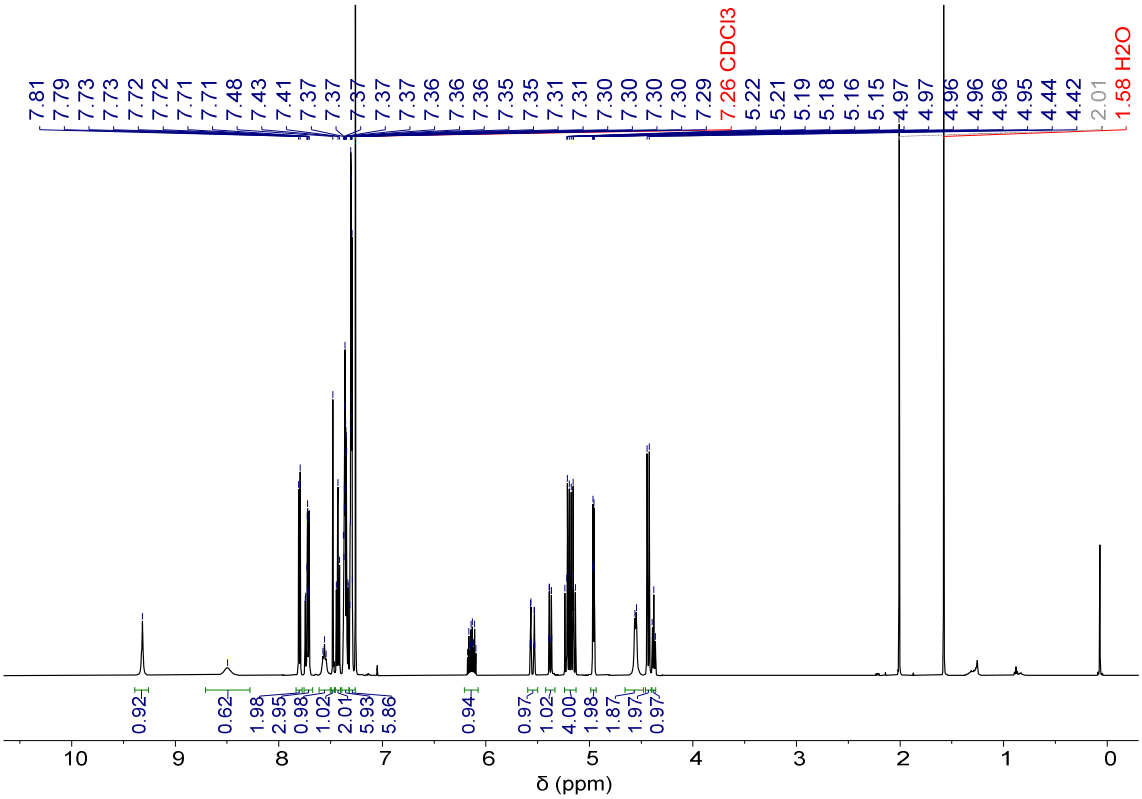


Figure S10 [1]H NMR spectrum of compound **8.15**.

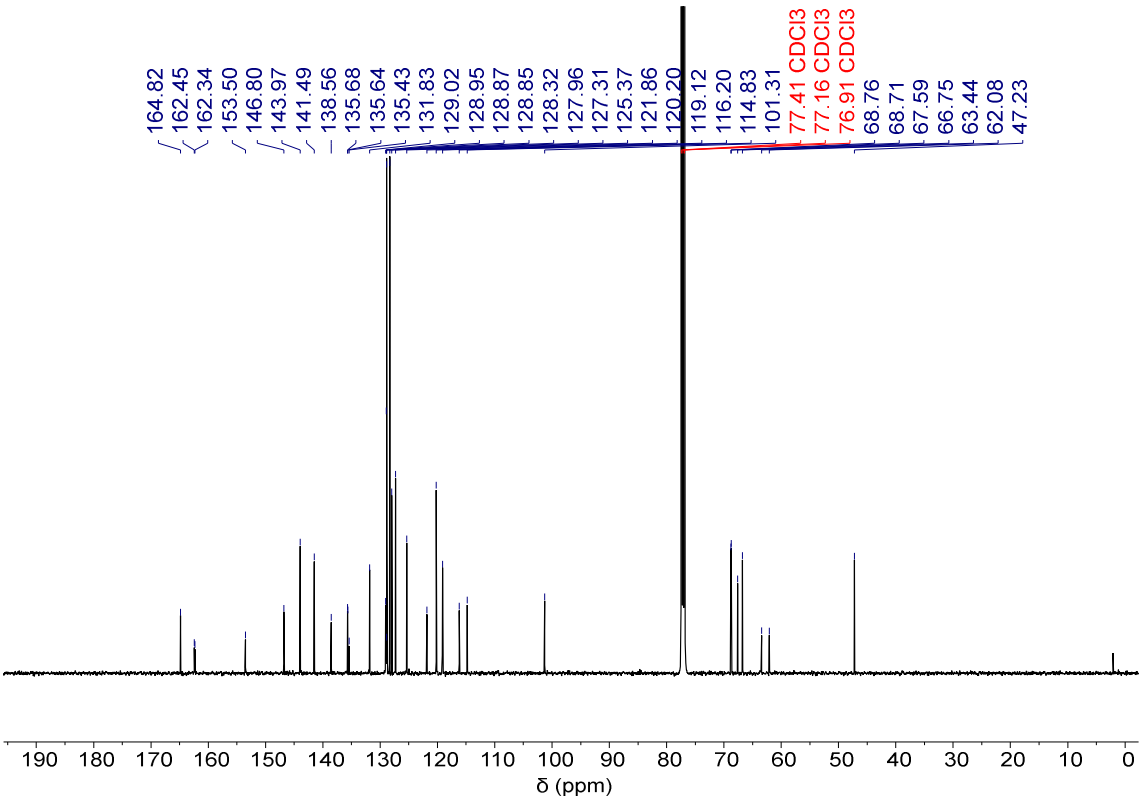


Figure S26 [13]C NMR spectrum of compound **8.15**.

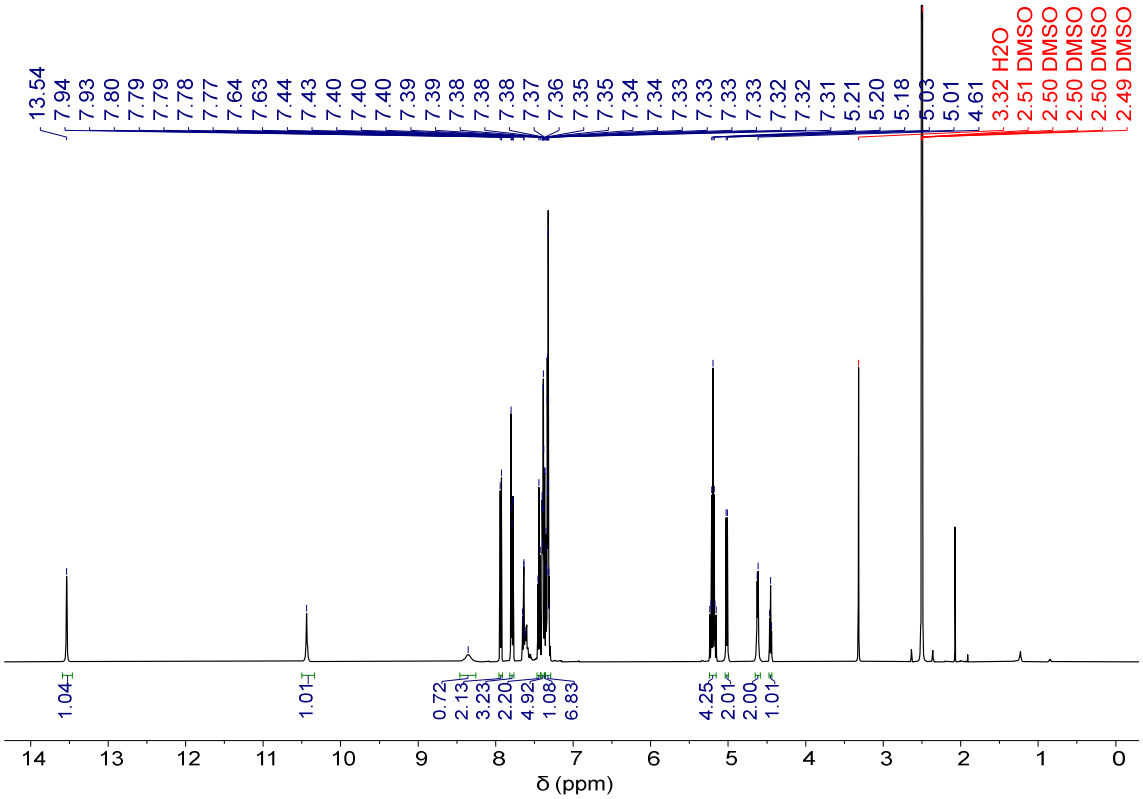


Figure S12 [1]H NMR spectrum of compound **8.3**.

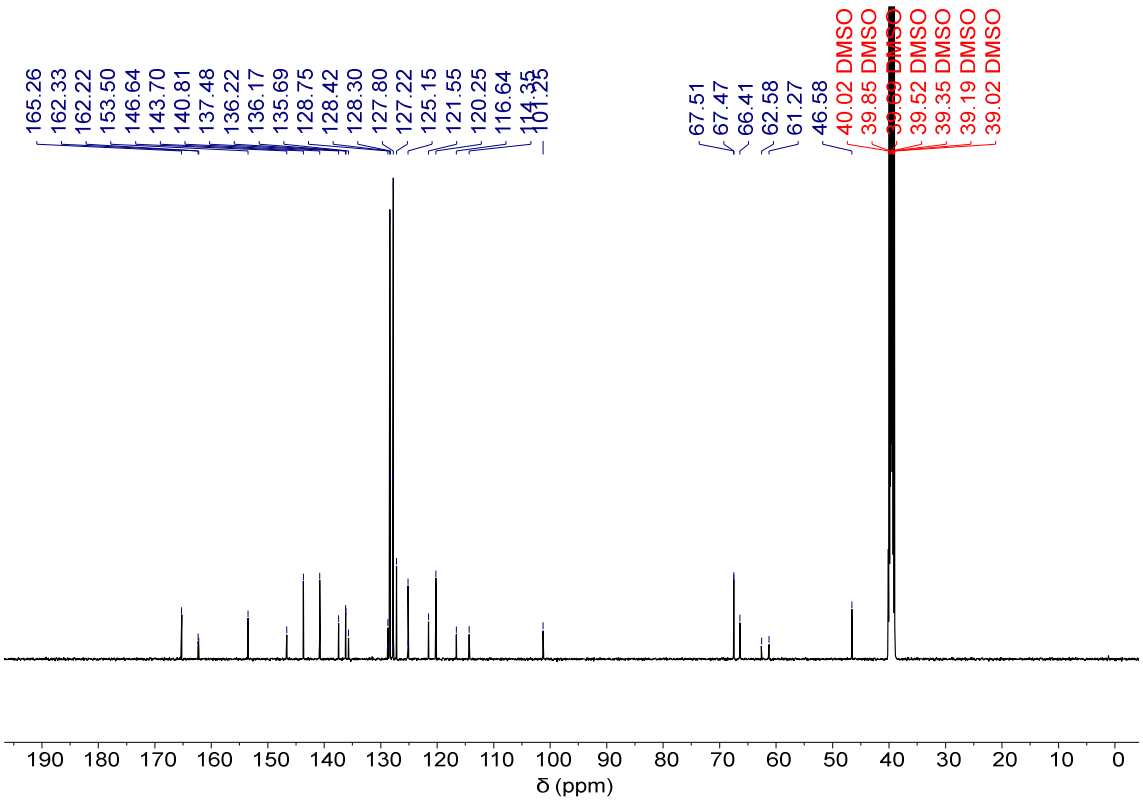


Figure S13 [13]C NMR spectrum of compound **8.3.**

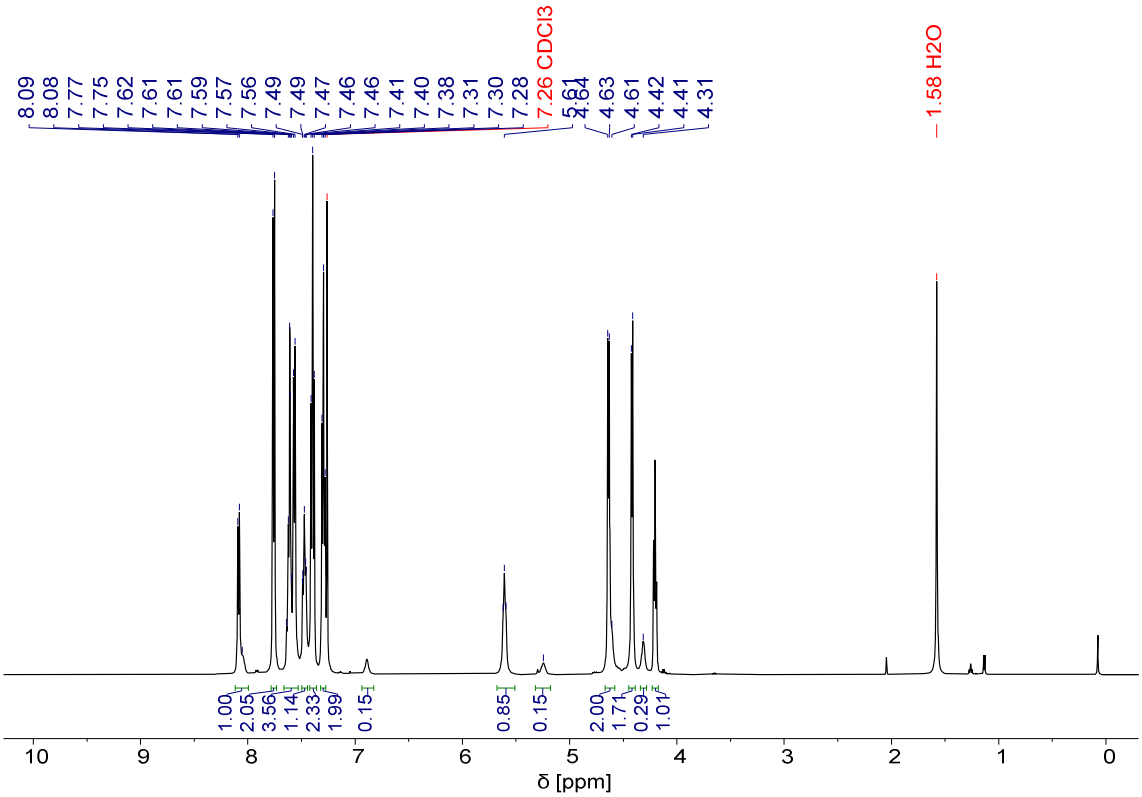


Figure S14 [1]H NMR spectrum of compound **8.21**.

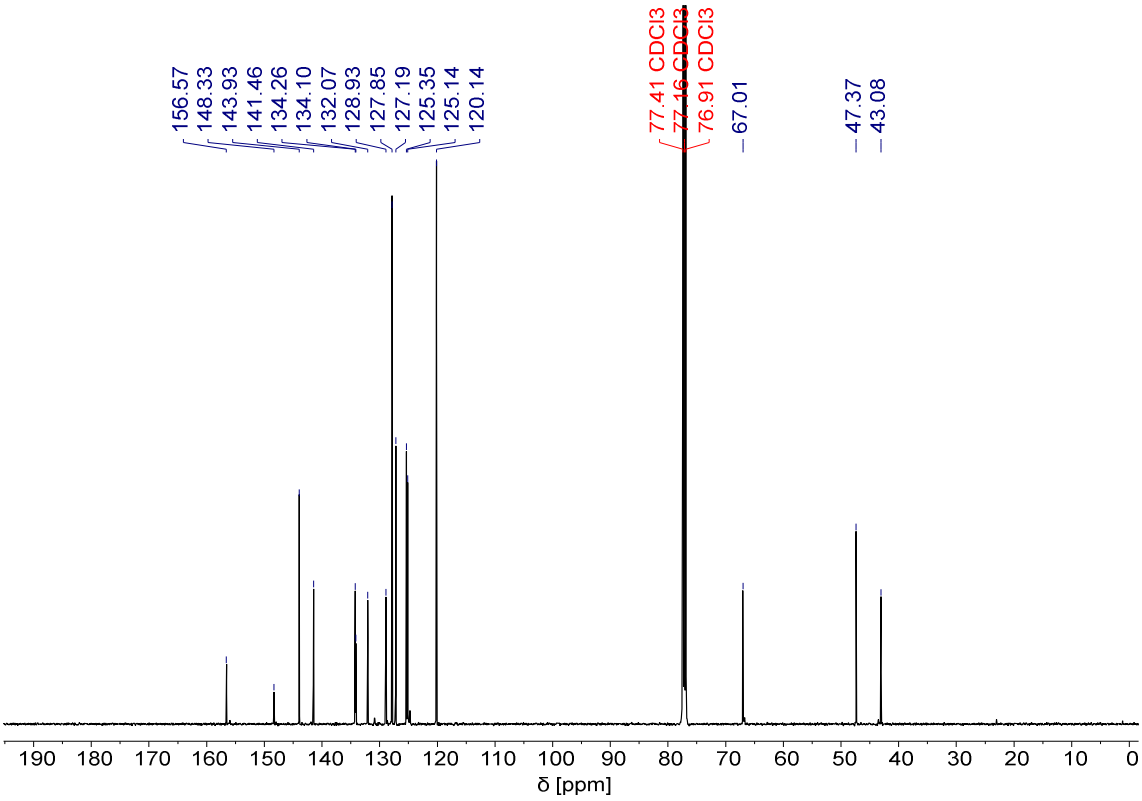


Figure S15 [13]C NMR spectrum of compound **8.21**.

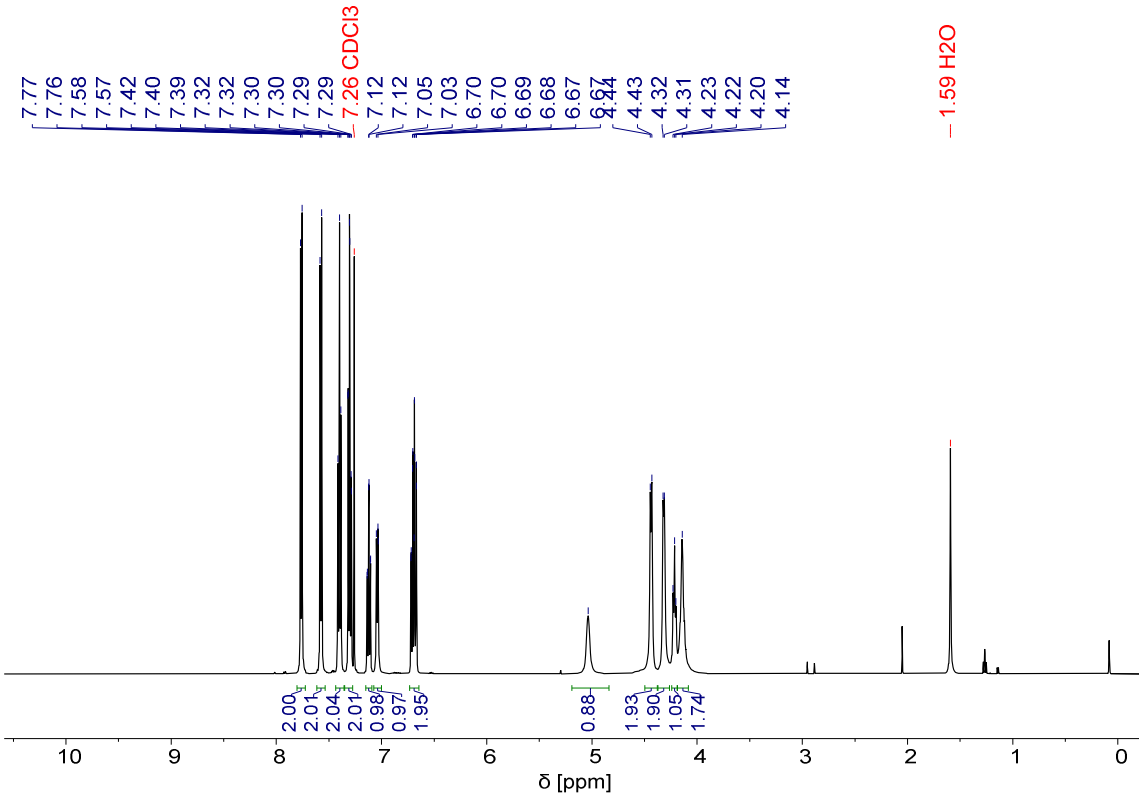


Figure S16 [1]H NMR spectrum of compound **8.22**.

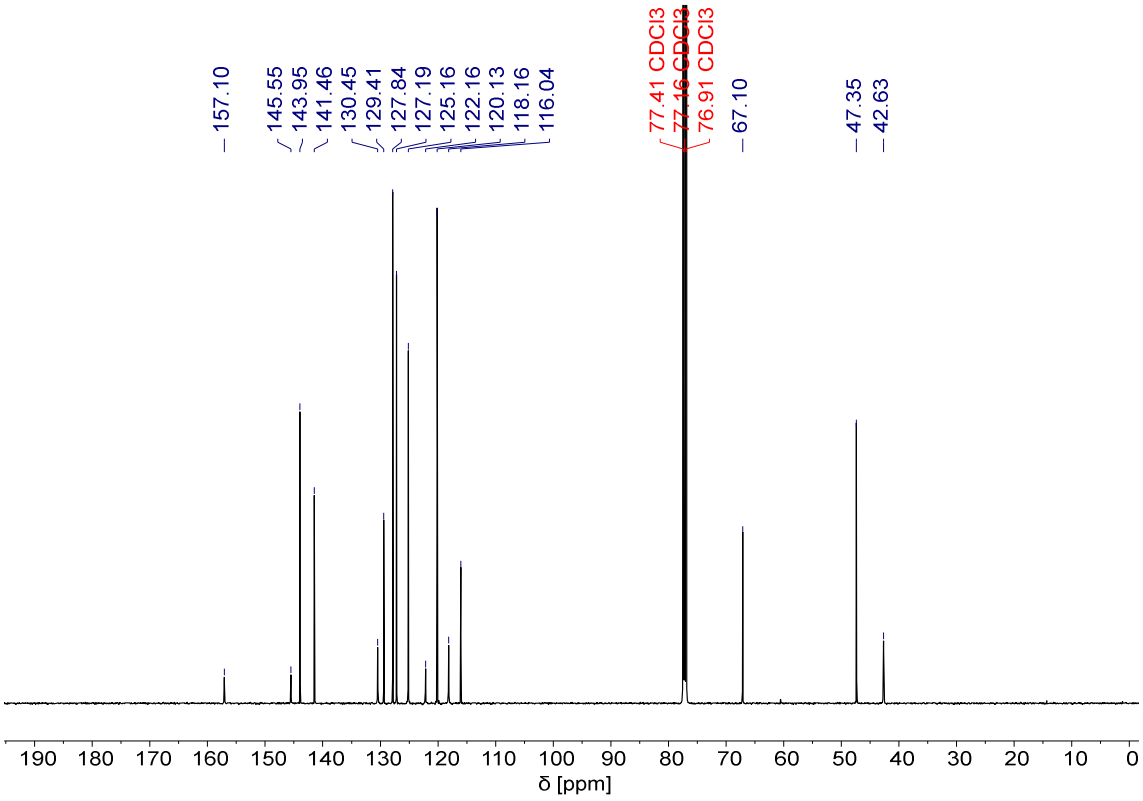


Figure S17 [13]C NMR spectrum of compound **8.22**.

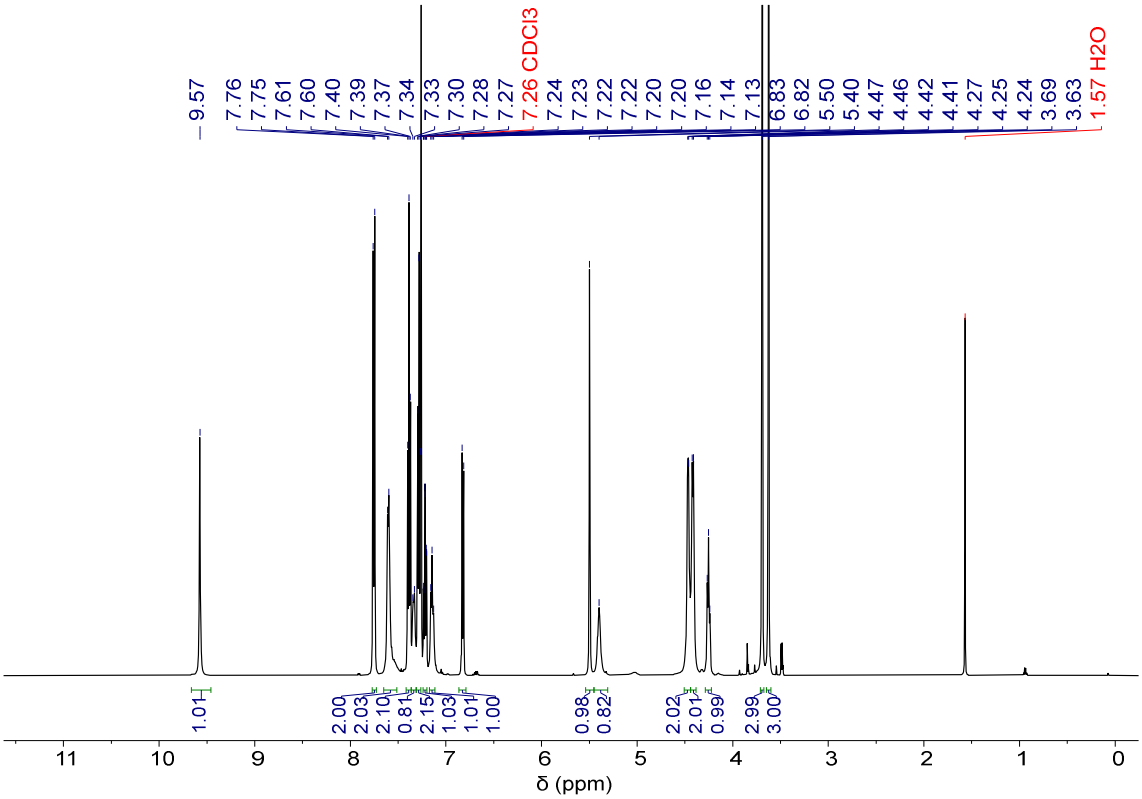


Figure S18 [1]H NMR spectrum of compound **8.23**.

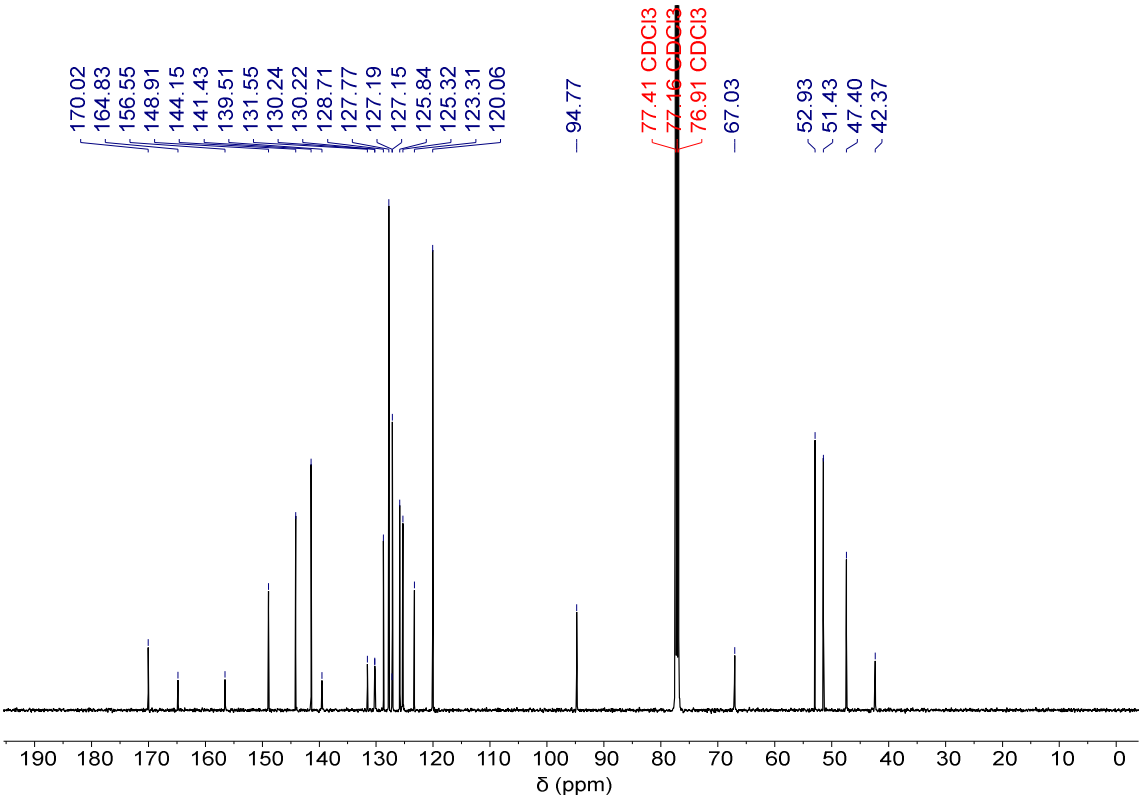


Figure S19 [13]C NMR spectrum of compound **8.23**.

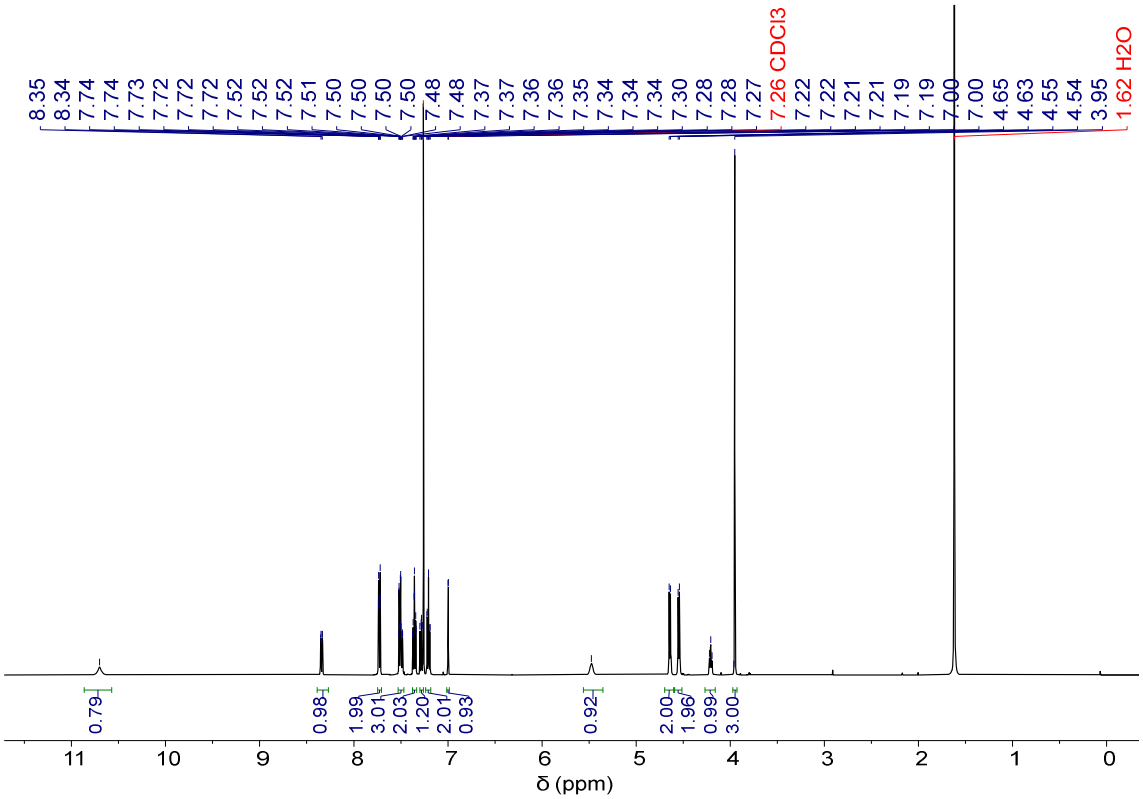


Figure S20 [1]H NMR spectrum of compound **8.24**.

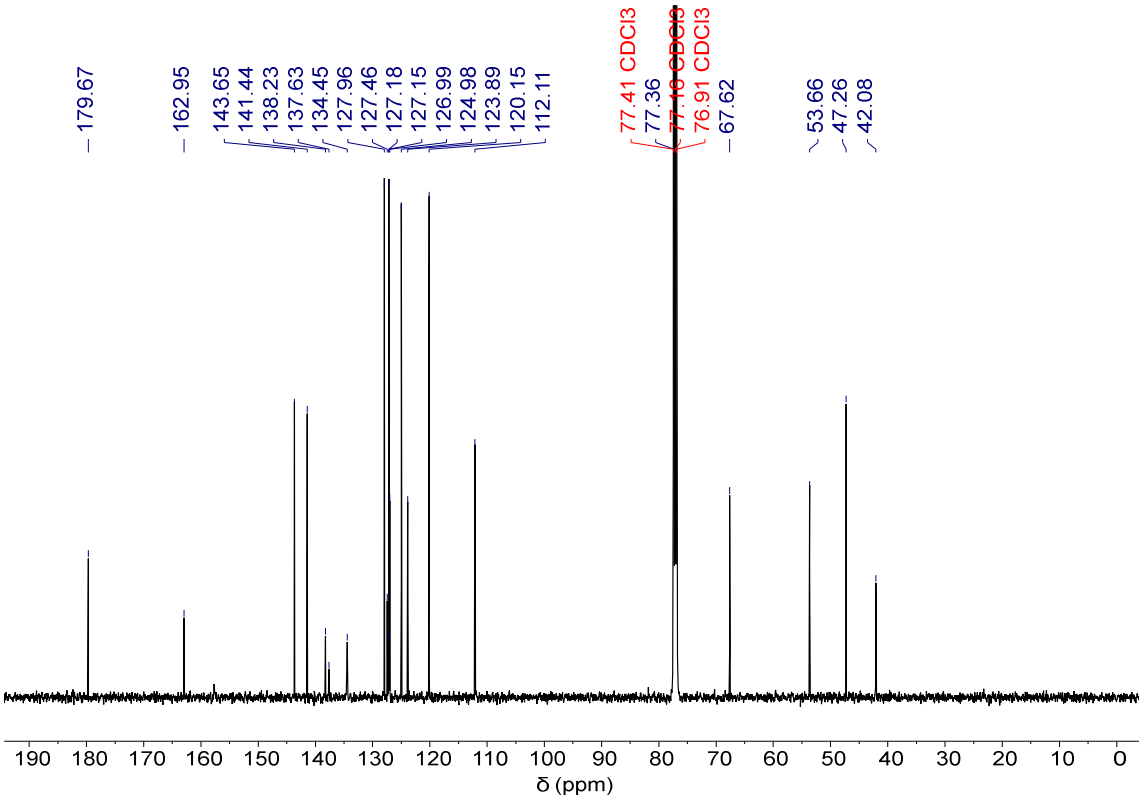


Figure S21 [13]C NMR spectrum of compound **8.24**.

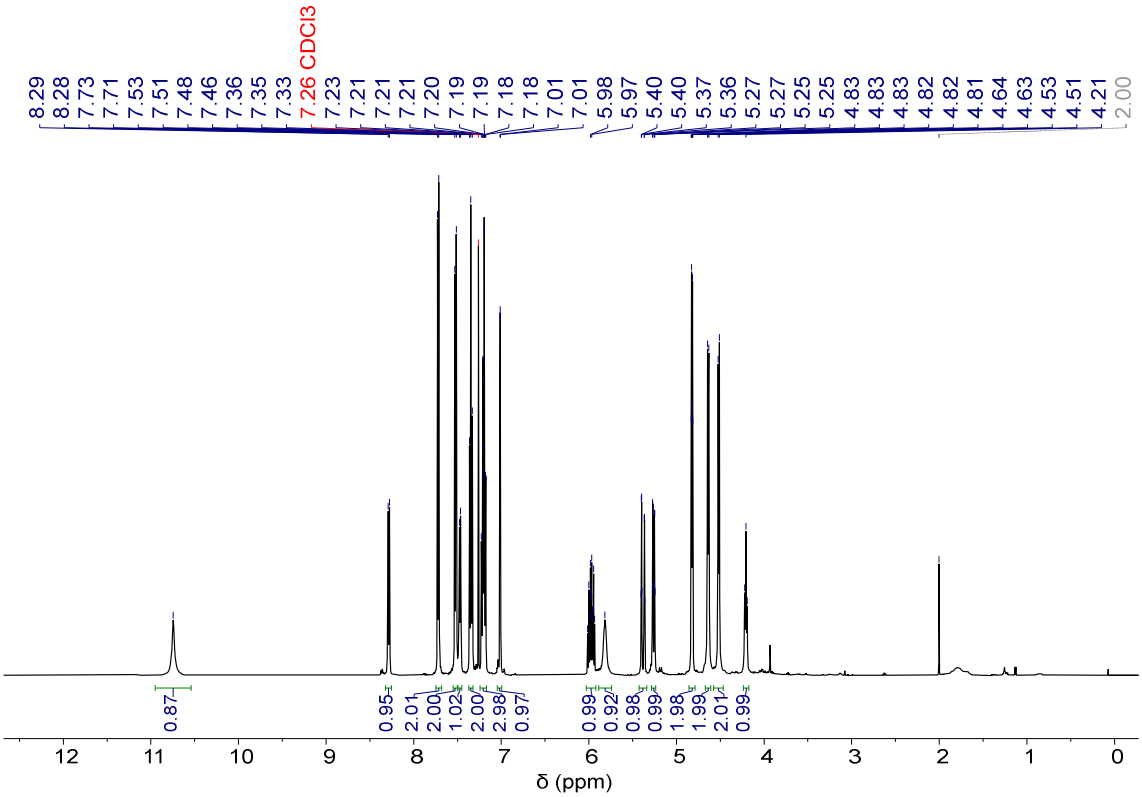


Figure S22 [1]H NMR spectrum of compound **8.25**.

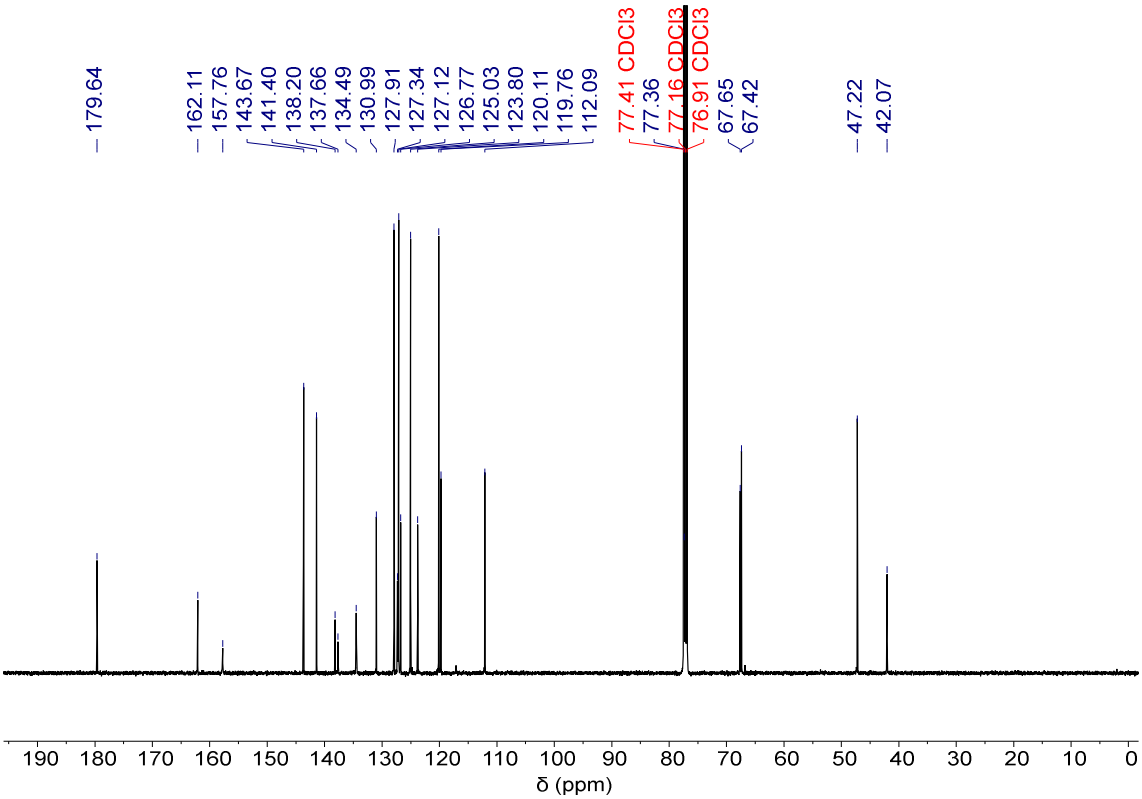


Figure S23 [13]C NMR spectrum of compound **8.25**.

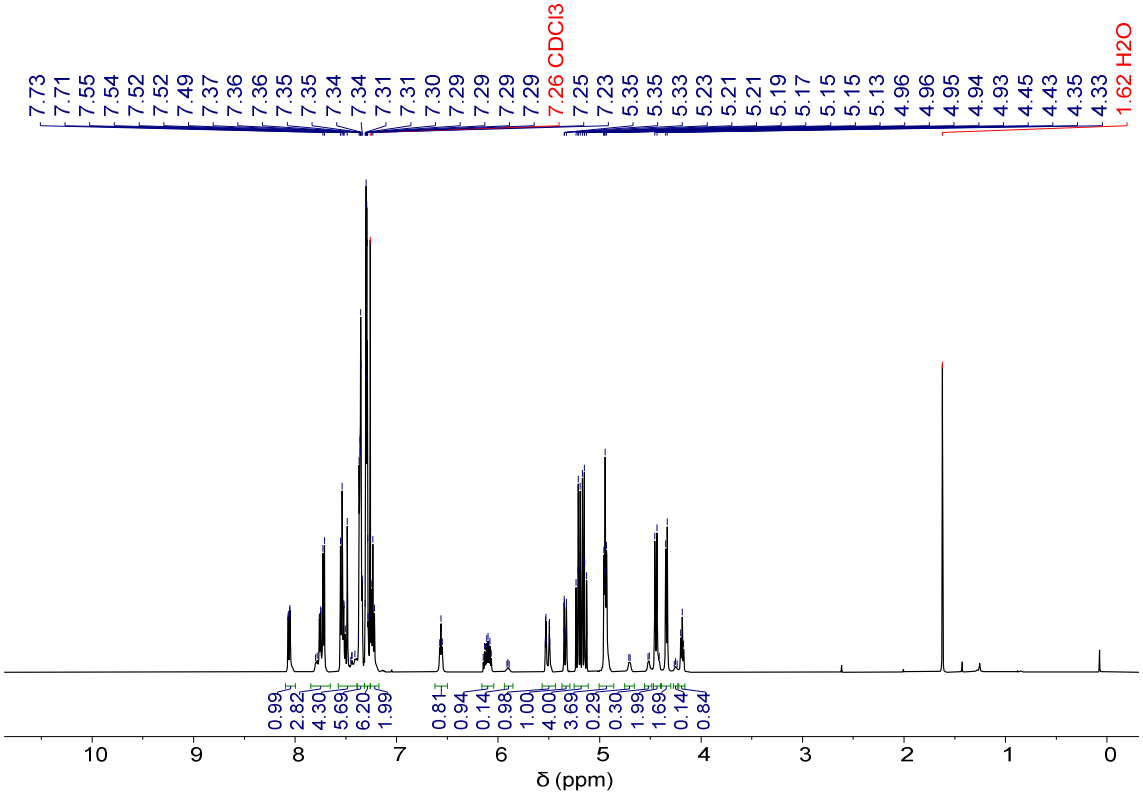


Figure S24 [1]H NMR spectrum of compound **8.26.**

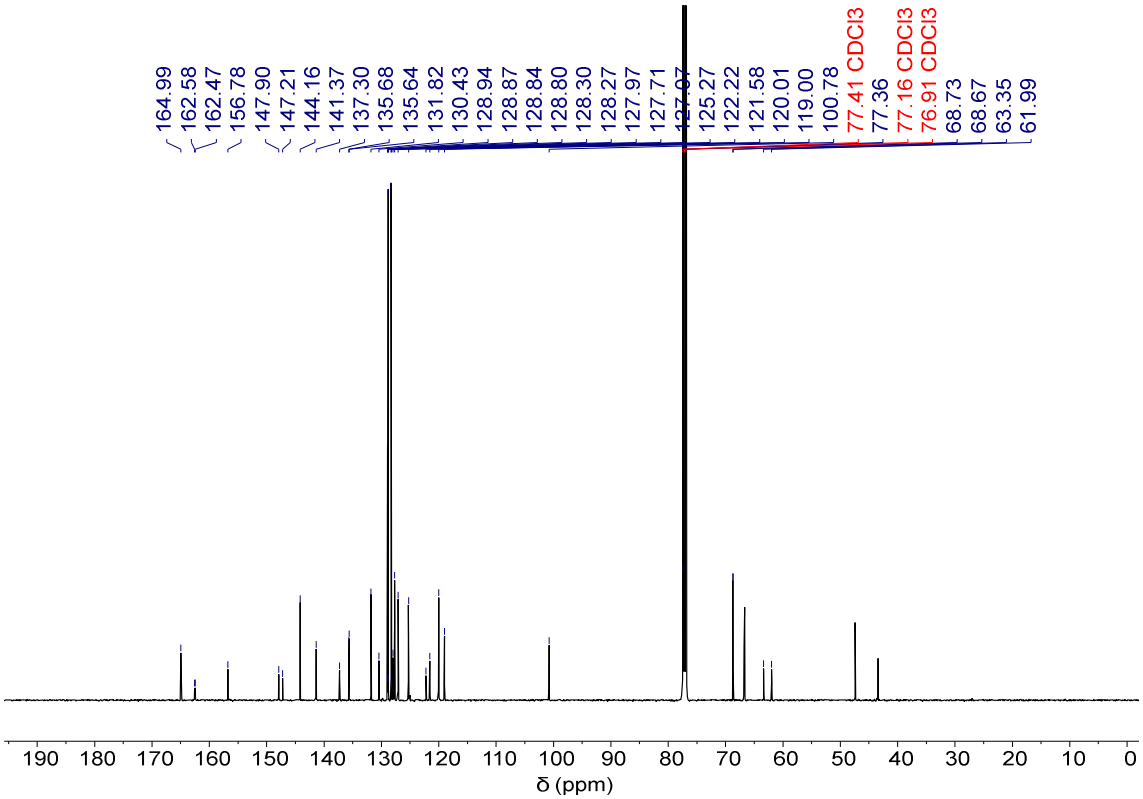


Figure S25 [13]C NMR spectrum of compound **8.26**.

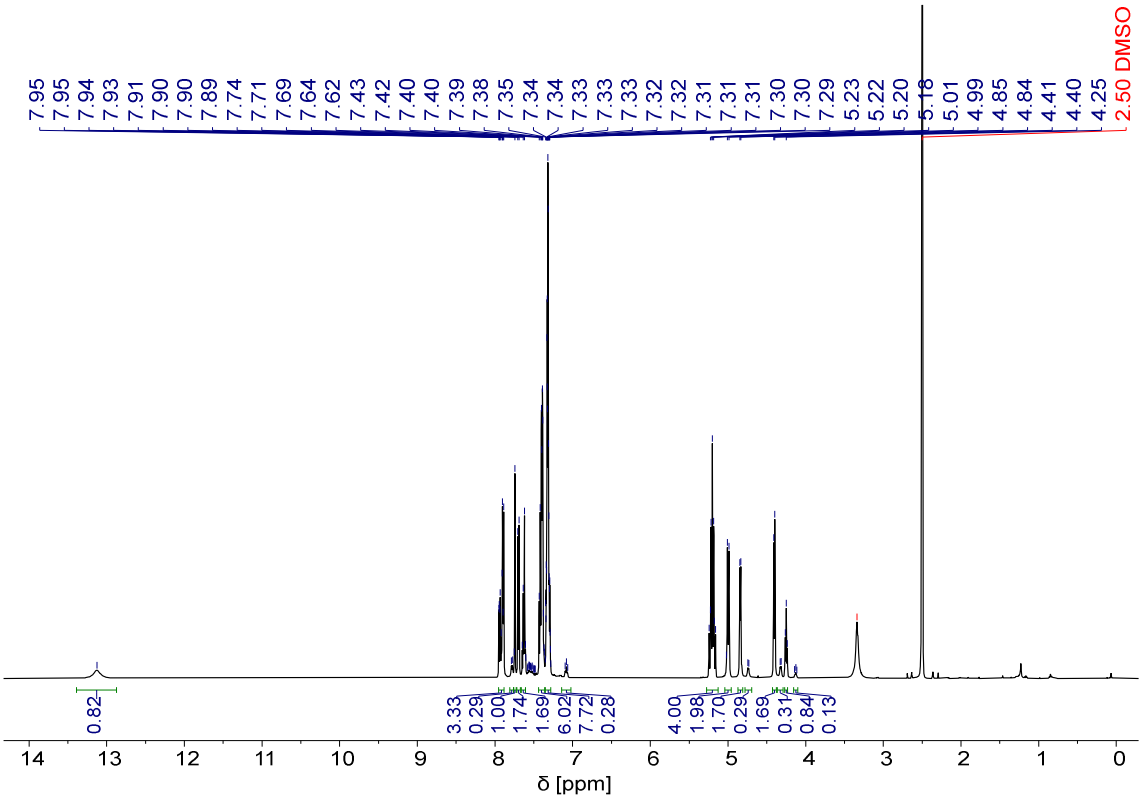


Figure S26 [1]H NMR spectrum of compound **8.16**.

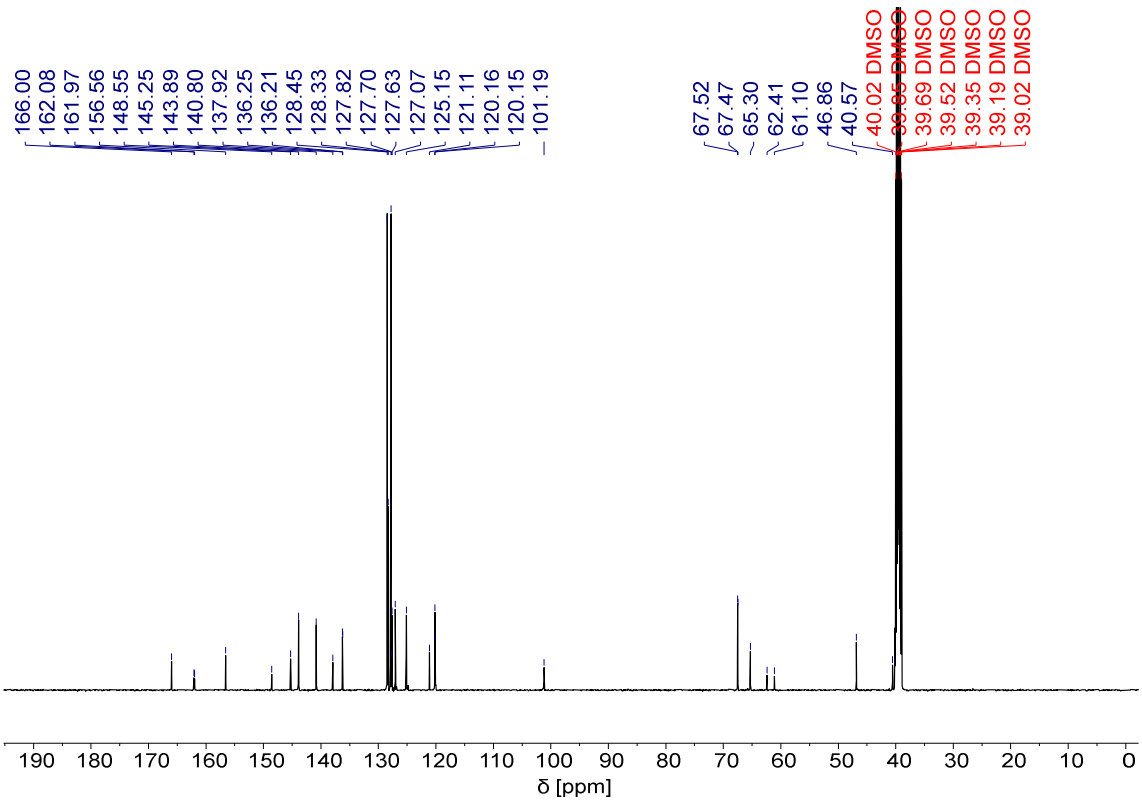


Figure S27 [13]C NMR spectrum of compound **8.16.**

8.5 References

1. N. Fujii, A. Otaka, N. Sugiyama, M. Hatano and H. Yajima, *Chem. Pharm. Bull.*, 1987, **35**, 3880.
2. J. T. Sparrow and O. D. Monera, *Pept. Res.*, 1996, **9**, 218-222.
3. V. Corvaglia, D. Carbajo, P. Prabhakaran, K. Ziach, P. K. Mandal, V. Dos Santos, C. Legeay, R. Vogel, V. Parissi, P. Pourquier*, et al.*, *Nucleic Acids Res.*, 2019, **47**, 5511-5521.
4. in *Greene's Protective Groups in Organic Synthesis*, 2006, DOI: https://doi.org/10.1002/9780470053485.ch9, pp. 934-985.
5. J. Picha, M. Budesinsky, I. Hanclova, M. Sanda, P. Fiedler, V. Vanek and J. Jiracek, *Tetrahedron*, 2009, **65**, 6090-6103.
6. T. Qi, T. Deschrijver and I. Huc, *Nat. Protoc.*, 2013, **8**, 693.
7. M. Zwillinger, P. Soregi, F. Sanchez, C. Douat, M. Csekei, I. Huc and A. Kotschy, *J. Org. Chem.*, 2025, **90**, 3043-3052.
8. WO2008098056, 2008.
9. WO2019209850, 2019.
10. K. Ziach, C. Chollet, V. Parissi, P. Prabhakaran, M. Marchivie, V. Corvaglia, P. P. Bose, K. Laxmi-Reddy, F. Godde, J.-M. Schmitter*, et al.*, *Nat. Chem.*, 2018, **10**, 511-518.
11. V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat and I. Huc, *Chem. - Eur. J.*, 2023, **29**, e202300898.
12. V. Corvaglia, J. Wu, D. Deepak, M. Loos and I. Huc, *Chem. - Eur. J.*, 2024, **30**, e202303650.
13. Q. Lu, J. Liu and L. Ma, *J. Catal.*, 2021, **404**, 475-492.
14. F. A. Carey and R. J. Sundberg, *Advanced Organic Chemistry: Part B: Reaction and Synthesis*, Springer US, 2007.
15. M. Zwillinger, L. Fischer, G. Salyi, S. Szabo, M. Csekei, I. Huc and A. Kotschy, *J. Am. Chem. Soc.*, 2022, **144**, 19078-19088.
16. X. Hu, S. J. Dawson, P. K. Mandal, X. de Hatten, B. Baptiste and I. Huc, *Chem. Sci.*, 2017, **8**, 3741-3749.
17. S. Kwon, V. Morozov, L. Wang, P. K. Mandal, S. Chaignepain, C. Douat and I. Huc, *Org. Biomol. Chem.*, 2024, **22**, 9342-9347.
18. B. Baptiste, C. Douat-Casassus, K. Laxmi-Reddy, F. Godde and I. Huc, *J. Org. Chem.*, 2010, **75**, 7175-7185.
19. D. Mazzier, S. De, B. Wicher, V. Maurizot and I. Huc, *Angew. Chem., Int. Ed.*, 2020, **59**, 1606-1610.

9 Summary and perspectives

The work conducted during the period of this thesis and my master's thesis previously have broadened many aspects of our understanding of DMFOs, how they may be synthesized efficiently, their structural and dynamic behaviors, and how they interact with diverse protein targets.

Synthetic advancements give access to DMFOs of unprecedented length

Potentially the biggest achievement directly derived from this work is the shift from an all-solution-based synthetic approach towards SPFS by emulating peptide chemistry and applying them to DMFOs.

The potential to rapidly change the molecular design was directly utilized in a study to introduce chirality into foldamer sequences and probe the efficiency of chiral induction. Additionally, these fragments could be employed in the synthesis of C_2-symmetrical foldamers and pseudo-sticky ends that allowed for the crystallization of the first DMFO bearing phosphonic acids from aqueous solution.[1] This successful paradigm shift has then inspired us to synthesize all other Fmoc-protected Q and M building blocks as diethyl phosphonate esters (or *t*Bu esters if the side chain consists of a carboxylate moiety). Building block synthesis was expanded by dimeric building blocks, which serve as macromonomers. They combine a reactive benzylic amine to simplify AOF synthesis further. This has revolutionized the synthesis of medium length DMFOs (16-32 units) due to its overall efficiency and gave access to DMFOs of unprecedented length.

Probing and visualization of long DMFOs with multi-domain DBP by cryo-EM

These much larger objects have allowed us to probe for foldamer interaction with protein complexes that interact with significantly longer stretches of DNA. We obtained a cryo-EM structure with the A-module of INO80, a multi-protein complex that serves as a DNA sensor to position the ATPase of INO80 correctly for DNA translocation.[2] At the same time, DMFOs of similar length are investigated in studies with another chromatin remodeler ATP-utilizing chromatin assembly and remodeling factor (ACF) as well as transcriptional repressor heterochromatin protein 1 (HP1).[3, 4]

DMFOs to reconstitute a nucleosome core particle without DNA

In future endeavors, it might be possible to assemble them to a nucleosome core particle (NCP), albeit where the entire DNA wrapped around a histone octamer is replaced by a synthetic DMFO construct. When dissecting the crystal structure of the NCP, one finds that DNA consists of 147 bps wrapped around the histone core in a left-handed 1.65 superhelical turns. Along the structure, there are primarily 14 contact points between the histone core's basic residues and DNA's minor groove where charge density is highest.[5] The entire construct is pseudo-C_2 symmetric about its dyad axis where its superhelical location (SHL) is zero (Fig. 26a). DNA bending and local curvature differ on each side reflected by pseudo-symmetry. At

four of these contact points, two arginine residues of H3 (SHL = −1.5 and −6.5) and H2B (SHL = approximately −3 and −5) respectively penetrate the DNA minor groove causing significantly greater bending at these four positions, something that must be accounted for when trying to reconstitute such an object with a DMFO.[5-7]

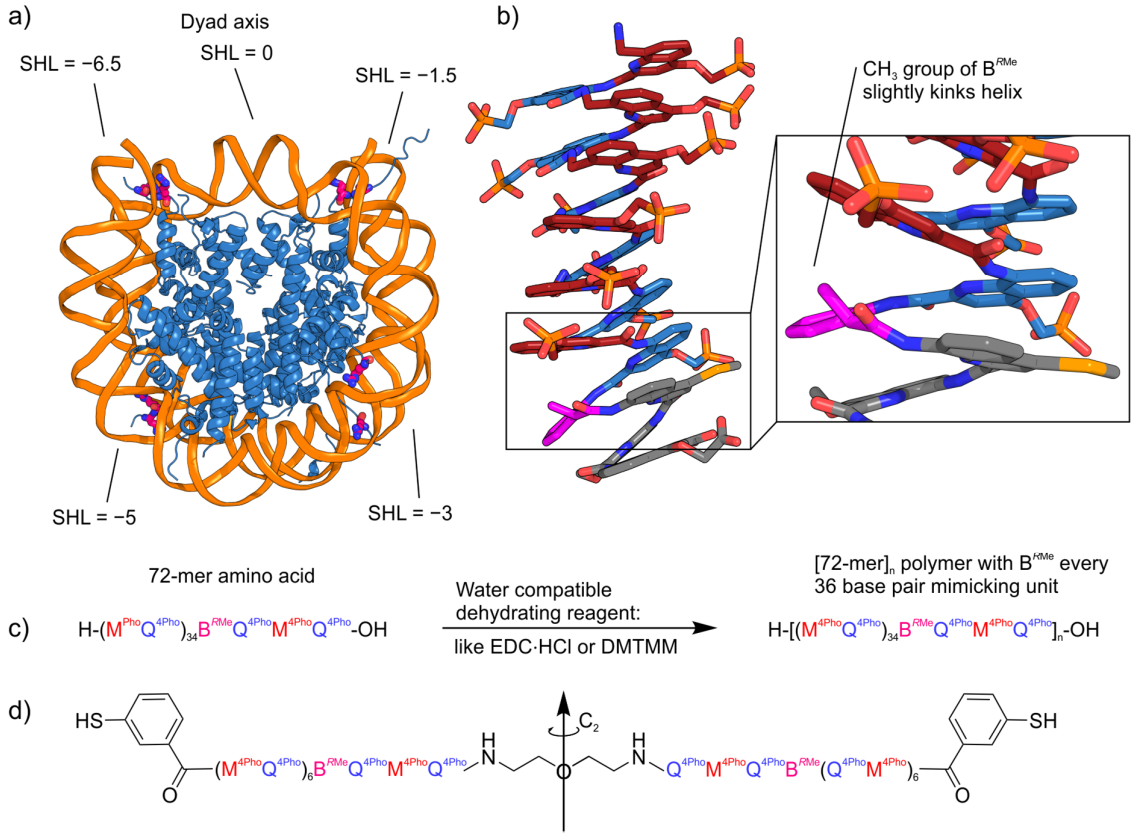


Figure 26 a) Crystal structure of the NCP with penetrating arginine residues shown with carbon atoms highlighted as pink sticks and their SHL marked in respect to the dyad axis.[5] b) Crystal structure of a polyphosphonic acid DMFO crystallized from water that contains a B^{RMe} unit (pink carbon atoms).[1] The insert zooms onto the bend location. c) Schematic polymerization of a $(MQ)_n$-DMFO. EDC = 1-Ethyl-3-(3-dimethylaminopropyl)carbodi-imide; DMTMM = (4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride). d) C_2-symmetric DMFO linked from two fragments by their respective C-termini and containing a thiol at each N-terminus.

Therefore, at these respective positions, kinks might have to be programmed into the AOF architecture. One potential building block for this purpose might already exist in the B^{RMe} unit. In $(Q)_n$ sequences that contain this chiral unit, effects on helix integrity and curvature, at least in the solid state, were limited.[8, 9] In the XRD structure of a polyphosphonic acid $(MQ)_n$ sequence, the AOF is bent at the position where B^{RMe} is incorporated. Due to the steric bulk of the methyl group pointing out of the aromatic plane, the M unit at the *i+2* position does not π-π stack with the benzene ring (Fig. 26b). The described chiral 72-unit DMFOs (Chapter 6) could be polymerized (Fig. 26c) from their respective amino acid precursors, which would yield polymers with a programmed kink every 36 pseudo-bps.

Dynamic analysis gives necessary insights of DMFO helical parameters

Design ideas like this make it essential gain the knowledge needed to both predict and quantify the effects that novel building blocks may exert on the helical integrity of an AOF main chain. This sequence-structure paradigm was initially circumvented by synthesizing $(Q)_n$ AOFs with highly rigid and thus highly conserved, main chain folds. It is likely to play a crucial role when reintroducing the necessary flexibility into these systems to better align with properties found in bio-oligomers. For these reasons, we have investigated the dynamics and the influence of external parameters like salt concentration, pH and temperature on $(MQ)_n$ sequences with phosphonic acid side chains (including some variants). We did aim for both *in silico* methods like molecular dynamics (MD) with a general Amber force field with reparametrized torsional parameters as well as experimental studies to confirm the observations from MD simulations.[10] Compared to DNA, DMFOs show higher bending persistence, but simultaneously are prone to minor unstacking and major transient kinking events that are more pronounced when phosphonic acids are doubly deprotonated (pH > 12).

Experimentally, DMFOs show outstanding thermal stability. As a single-stranded main chain, they cannot dissociate into two single strands like DNA, but their melting behavior can be conceptually compared with α-helix denaturation of proteins. NMR, CD and UV-Vis measurements confirmed that DMFOs remain, on average, fully helical until at least 85 °C and stay folded until at least pH 12.5.

The handedness bias towards M-helix, that is exerted by the chiral B^{RMe} unit could be converted to a *P*/*M* diastereomeric mixture upon certain content of organic solvent in a DMF/H_2O mixture. The diastereomeric enrichment toward the *M*-helix could be monitored to quantify the dynamics of helix-handedness interconversion under different conditions. In general, *P*- to *M*-helix interconversion is faster at higher temperatures and at higher pH but remains constant upon addition of monovalent cations. Interestingly, when a phosphonic acid was replaced with an uncharged side chain, the helix interconversion rate increases drastically. This suggests that side-chain interactions of two phosphonic acids, albeit both negatively charged, stabilize the helix contrary to our predictions. This can also be observed in MD simulations via transient hydrogen-bonding. This assay is a valuable tool to assess changes in helix dynamics in future variations (side chain or main chain) of DMFO sequences. This is especially important in the context of programming more deformable regions for bigger architectures, e.g., a B^{RMe} unit could be flanked by residues bearing methoxy side chains to further destabilize the helix at this position.

Other synthetic strategies to increase DMFO length

A more extreme strategy for introducing kinks involves incorporating flexible stretches, such as two C-terminally linked DMFOs, which are equipped with two N-terminal thiols. Oxygen mediated polymerization directly yields a head-to-head to tail-to-tail polymer with two C_2-axes (second C_2-axis not shown in Fig. 26d is through the disulfide bond). This thiol-based

polymerization strategy is oxygen-mediated and therefore produces no byproducts that require removal. In principle, reaction can even be performed in the presence of a histone octamer. However, the hinges introduced by both the disulfide bond and the flexible linker disrupt the continuity DMFO grooves. Both strategies have the potential to construct molecules large enough to assemble around a histone core and may be pursued in parallel.

Towards sequence selective protein binders with DNA-DMFO chimeras

Aside from constructing even larger DMFO-DBP complexes, another endeavor is interfering with sequence-selective PNIs such as DNA-TF interactions, especially those of pharmaceutical interest. One way to achieve this is with chimeric molecules that combine features of DNA and DMFOs into a single synthetic construct.

So far, we have developed the synthetic strategy towards these chimeras, including the conceptualization of a MQ-dimer not bearing phosphonic acids but containing all necessary moieties to function as a DNA hairpin turn, a ligation point for a pre-synthesized foldamer as well as a unit that aligns the grooves of both oligomers in register. We have further expressed two TFs from the ETS family that recognize a GGA core-consensus sequence and tested the binding of DNA hairpins against them. Satisfactorily, binding of the hairpins is conserved when compared to canonical ds-DNA without covalent tethering of the two strands.

With this as a starting point, future studies will revolve around replacing significant parts, up until the core consensus sequence (Fig. 27b) and analyze their binding behavior both in terms of affinity and kinetics. With this binding assay, these chimeras can be evolved iteratively and optimized simultaneously for selectivity and affinity to obtain a novel class of DNA decoys. In the next step, sequence features can be incorporated into the foldamer itself by either backbone or side-chain modifications. For the latter, TFA-labile side chain PGs for Q and M will be required.

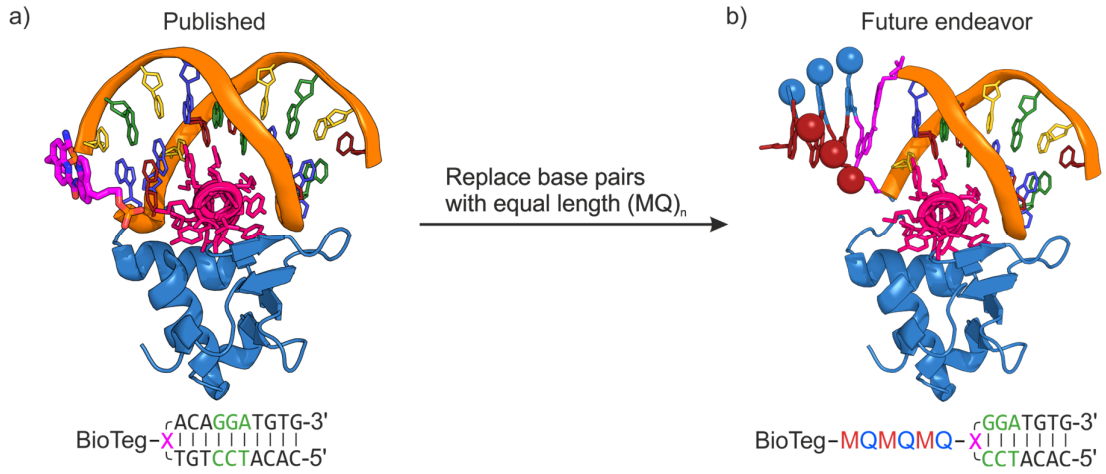


Figure 27 a) Crystal structure of SAP1 with the terminal CG bp removed and the MQ-dimer hairpin turn modeled in its place.[11] b) Same as a) but four terminal bps removed and replaced with a DMFO fragment of comparable length.

Exploring the mimicry of other DNA folds – G-Quadruplexes

Whereas A-, B- and Z-DNA resemble each other, the tetrads of G-quadruplexes are more structurally distinct than other nucleic acid folds, and nature has found different ways to utilize this, such as in the regulation of gene expression, control of DNA replication and telomere maintenance. [12-14] Crystal structures of G-quadruplexes in complex with proteins are rare. The available structural insight of G-quadruplexes with proteins suggests that most commonly, the interaction occurs between the hydrophobic cross-section of a G-tetrad and the aromatic residues of protein α-helices.[15-18] Fig. 28a) displays the interactions of a G-quadruplex including a single-stranded DNA overhang that is part of the c-MYC promoter region in complex with the DEAH-box helicase 36 (DHX36).[15] DHX36 is an enzyme that unwinds parallel G-quadruplexes, e.g., in the c-MYC promoter region and its inhibition was shown to reduce expression of c-MYC, a key oncogenic target in pharmaceutical research.[19, 20]

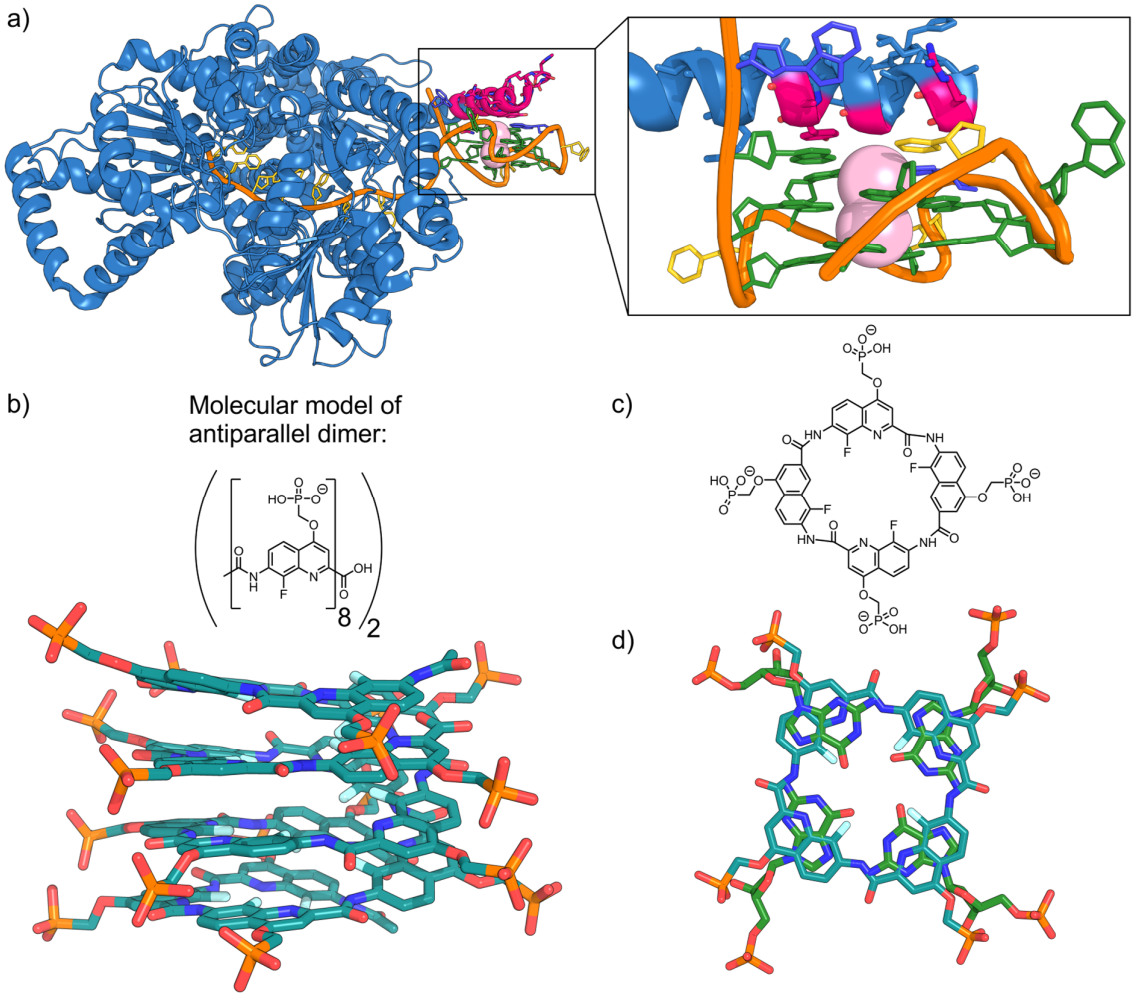


Figure 28 a) Crystal structure of the DHX36 helicase and a c-MYC quadruplex (incl. 3' single-stranded DNA) with the interacting α-helix highlighted in pink. G-quadruplex central ions are colored pale pink. Other motifs are colored based on the established color code.[15] b) Molecular model based on the crystal structure of a antiparallel dimer of a 8-fluoroquinoline AOF.[21] c) 8-Fluoroquinoline cyclic tetramer. d) Overlay of a G4-tetrad and a molecular model of the cyclic tetramer displayed in c). The tetrad was extracted from the following structure.[17]

DHX36 is essential for normal cells, as shown by embryonic lethality in knockout mice, hence it is a niche target for therapeutics. This is because the helicase ATP pocket is typically highly conserved and small molecule inhibitors are very likely to cause off-target effects.[20, 22] Thus, therapeutic approaches focus more on indirect way to interfere with these processes, e.g. drug candidate CX-5461 (pidnarulex) stabilizes G-quadruplexes to prevent G4 unwinding or can poison Topoisomerase II (Top2) at transcribed regions that contain G4 structures.[23]

An alternative approach that could achieve similar outcomes is not to stabilize endogenous quadruplexes but to replicate their structure by using artificial backbones. G4 mimics aim to bind protein regions that interact with quadruplexes. This might achieve similar functional consequences on the transcription machinery originating from a different MoA. An octamer of 8-fluoroquinoline was shown to dimerize and adopt an antiparallel arrangement in organic solvents and in the solid state.[21] Based on this crystal structure, a molecular model of the octamer functionalized with phosphonic acid side chains (Fig. 28b) was constructed. It overall matches well with the dimensions and *exo*-helix sense of a parallel G-quadruplex. Synthesis, structural elucidation, and biophysical analysis of compounds of this kind can be conducted against the N-terminal isolated G-quadruplex-recognizing peptide fragment of DHX36 called RHAU29 that has been co-crystallized with a canonical parallel G-quadruplex.[17] A cyclic tetramer of 8-fluoroquinoline that bears phosphonic acid side chains (Fig. 28c) could function as a mimic of an isolated G4-tetrad. The resemblance of their cross-sections is displayed in the overlay of a G4-tetrad and the cyclic tetramer (Fig. 28d). The four quinolines of this macrocycle that together compose the hydrophobic surface of the G4 mimic have the potential to be recognized by typical G-quadruplex sensing domains of enzymes.

Concluding remarks

Mimics of DNA including DMFOs based on AOFs, whether they mimic B-DNA or other DNA isoforms are an underexplored compound class when compared to their natural analogues. They have certainly shown potential for therapeutic application or may serve as tools to investigate DNA-protein interplay in PNIs. The established synthetic advances, the analysis of their dynamic behavior and the first steps toward sequence-selective binders represent key milestones in their development. Especially their properties to bind to DNA binding proteins, often much better than DNA itself represents a core unique property for developing DNA decoys based on an artificial backbone. Exploring them further, starting with the production of new monomers, up until their investigation in biological systems might elevate them to drug candidates of the next generation.

References

1. V. Corvaglia, J. Wu, D. Deepak, M. Loos and I. Huc, *Chem. - Eur. J.*, 2024, **30**, e202303650.
2. S. Brahma, M. Ngubo, S. Paul, M. Udugama and B. Bartholomew, *Nat. Commun.*, 2018, **9**, 1-10.
3. W. Zeng, A. R. Ball, Jr. and K. Yokomori, *Epigenetics*, 2010, **5**, 287-292.
4. O. Z. Aydin, W. Vermeulen and H. Lans, *Cell Cycle*, 2014, **13**, 3016-3025.
5. K. Luger, A. W. Mader, R. K. Richmond, D. F. Sargent and T. J. Richmond, *Nature (London)*, 1997, **389**, 251-260.
6. C. A. Davey, D. F. Sargent, K. Luger, A. W. Maeder and T. J. Richmond, *J. Mol. Biol.*, 2002, **319**, 1097-1113.
7. R. K. McGinty and S. Tan, *Chem. Rev. (Washington, DC, U. S.)*, 2015, **115**, 2255-2273.
8. D. Bindl, P. K. Mandal, L. Allmendinger and I. Huc, *Angew. Chem., Int. Ed.*, 2022, **61**, e202116509.
9. D. Bindl, E. Heinemann, P. K. Mandal and I. Huc, *Chem. Commun. (Cambridge, U. K.)*, 2021, **57**, 5662-5665.
10. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.
11. Y. Mo, B. Vaessen, K. Johnston and R. Marmorstein, *Mol. Cell*, 1998, **2**, 201-212.
12. D. Rhodes and H. J. Lipps, *Nucleic Acids Res.*, 2015, **43**, 8627-8637.
13. Y. Xu and M. Komiyama, *Molecules*, 2024, **29**, 174.
14. W. Yoo, Y. W. Song, V. Bansal and K. K. Kim, *Nucleic Acids Res.*, 2025, **53**, gkaf827.
15. M. C. Chen, R. Tippana, N. A. Demeshkina, P. Murat, S. Balasubramanian, S. Myong and A. R. Ferre-D'Amare, *Nature (London, U. K.)*, 2018, **558**, 465-469.
16. L. Chen, J. Dickerhoff, K.-w. Zheng, S. Erramilli, H. Feng, G. Wu, B. Onel, Y. Chen, K.-B. Wang, M. Carver*, et al.*, *Science (Washington, DC, U. S.)*, 2025, **388**, eadr1752.
17. B. Heddi, V. V. Cheong, E. Schmitt, Y. Mechulam and A. T. Phan, *J. Struct. Biol.*, 2020, **209**, 107399.
18. K. H. Ngo, C. W. Liew, B. Heddi and A. T. Phan, *J. Am. Chem. Soc.*, 2024, **146**, 13709-13713.
19. S. K. Madden, A. D. de Araujo, M. Gerhardt, D. P. Fairlie and J. M. Mason, *Mol Cancer*, 2021, **20**, 3.
20. C. Yang, J. Yao, H. Yi, X. Huang, W. Zhao and Z. Yang, *Anim. Models Exp. Med.*, 2022, **5**, 542-549.
21. Q. Gan, C. Bao, B. Kauffmann, A. Grelard, J. Xiang, S. Liu, I. Huc and H. Jiang, *Angew. Chem., Int. Ed.*, 2008, **47**, 1715-1718.
22. L. Selvaratnam, T. M. Willson and M. Schapira, *J. Med. Chem.*, 2025, **68**, 4022-4039.
23. M. Bossaert, A. Pipier, J.-F. Riou, C. Noirot, L.-T. Nguyen, R.-F. Serre, O. Bouchez, E. Defrancq, P. Calsou, S. Britton*, et al.*, *eLife*, 2021, **10**, e65184.