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

Designing abiotic foldamer tertiary and quaternary structures

Shuhe Wang

aus

Huludao, Liaoning, China

2024

Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Ivan Huc betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, den 07.11.2024 _____

Shuhe Wang

Dissertation eingereicht am 07.11.2024

1. Gutachter: Prof. Dr. Ivan Huc

2. Gutachter: Prof. Dr. Franz Paintner

Mündliche Prüfung am 09.01.2025

Table of Content

1. List of publications

Published:

S. Wang, B. Wicher, C. Douat, V. Maurizot and I. Huc, Domain Swapping in Abiotic Foldamers. *Angew. Chem. Int. Ed.*, **2024**, 63, e202405091. (doi.org/10.1002/anie.202405091).

(See Chapter 4)

S. Wang, L. Allmendinger and I. Huc, Abiotic foldamer quaternary structures. Angew. *Chem. Int. Ed.*, **2024**, e202413252. (doi.org/10.1002/anie.202413252).

(See Chapter 5)

Accepted:

S. Wang, J. Sigl, L. Allmendinger, V. Maurizot and I. Huc, Design of an abiotic unimolecular three-helix bundle. *Chem. Sci.*, **2025**, Advance Article. (doi.org/10.1039/d4sc07336c).

(See Chapter 6)

2. Abstract

Biopolymers are composed of different folded structural units, allowing them to perform various roles in biological processes. The study of foldamers is based on the concept that backbones chemically distinct from natural biopolymers can lead to different structures, properties, and unique functions. Aromatic backbone foldamers are particularly intriguing due to their unique properties. The folding behavior and resulting structures of these foldamers are more predictable, facilitating design. Over two decades, numerous simple artificial structures, such as macrocycles, helices, sheets, and knots, have been reported. However, the exploration of abiotic foldamers with complex and enriched topologies remains both challenging and highly appealing, as the functions of biopolymers typically emerge at their tertiary and quaternary structural levels.

In this study, we describe the exploration of abiotic quaternary structures based on a previously reported heterochiral helix-turn-helix structure, as well as the design of a unimolecular tertiary fold based on a trimeric helical bundle. First, we modified the heterochiral helix-turn-helix structure by introducing additional hydroxy groups and hoped that two of the modified tertiary folds could dimerize via intermolecular hydrogen bonds to form an abiotic quaternary structure. However, the solid-state structure revealed a novel domain-swapping dimer. Next, we investigated the trimerization of the heterochiral helix-turn-helix structure through a known “shifted dimer” intermolecular interface. The self-assembly of designed sequences was investigated by NMR, CD and molecular dynamic simulations, which revealed that two stable aggregates corresponding to the designed trimer and to a dimer with water molecules bridging the intermolecular hydrogen bonding interface. Finally, we designed and synthesized two new turn units with different flexibilities that were expected to promote the formation of a unimolecular helix-turn-helix-turn-helix tertiary structure, stabilized by the previously reported “parallel trimer” intermolecular hydrogen-bonding pattern. Solution studies and crystal structures of two series of model sequences indicate that both turn units promote the formation of the heterochiral helix-turn-helix structure as anticipated. A fragment condensation strategy was employed in the preparation of the final 6.9 kDa unimolecular three-helix bundle sequences. Solution studies support the formation of the designed tertiary structure in sequences with rigid turn units, suggesting the flexibility in the turn units will bring obstacles in the formation of relatively complex tertiary folds.

Overall, these results explored various complex aggregation modes based on the helix-turn-helix structure, providing a starting point for application development. The relationship between the flexibility of the turn units and tertiary structure formation offers insights for future linker design. Additionally, the use of a fragment condensation strategy paves the way for the synthesis of comparable or even longer sequences in the future.

3. Introduction

Biopolymers, such as proteins (Figure 1b), nucleic acid (Figure 1c) and polysaccharides, serve diverse roles in the biological processes of organisms. Proteins have multiple functions such as catalysis,[1-3] providing structural support to cells and organisms,[4-6] and transporting small molecules or ions through cell membranes.[7-9] The essential functions of nucleic acids are to store and transmit genetic information.[10, 11] The fundamental reason for the significant differences in biopolymer functions lies in the structural diversity of their basic building blocks (Figure 1a). Different building blocks and backbones result in different folding principles and three-dimensional shapes, leading to distinct functions. The study of foldamers arises from the idea that backbones chemically distinct from those of biopolymers might produce novel shapes and functions beyond what natural biopolymers can achieve (Figure 1d, e, f).

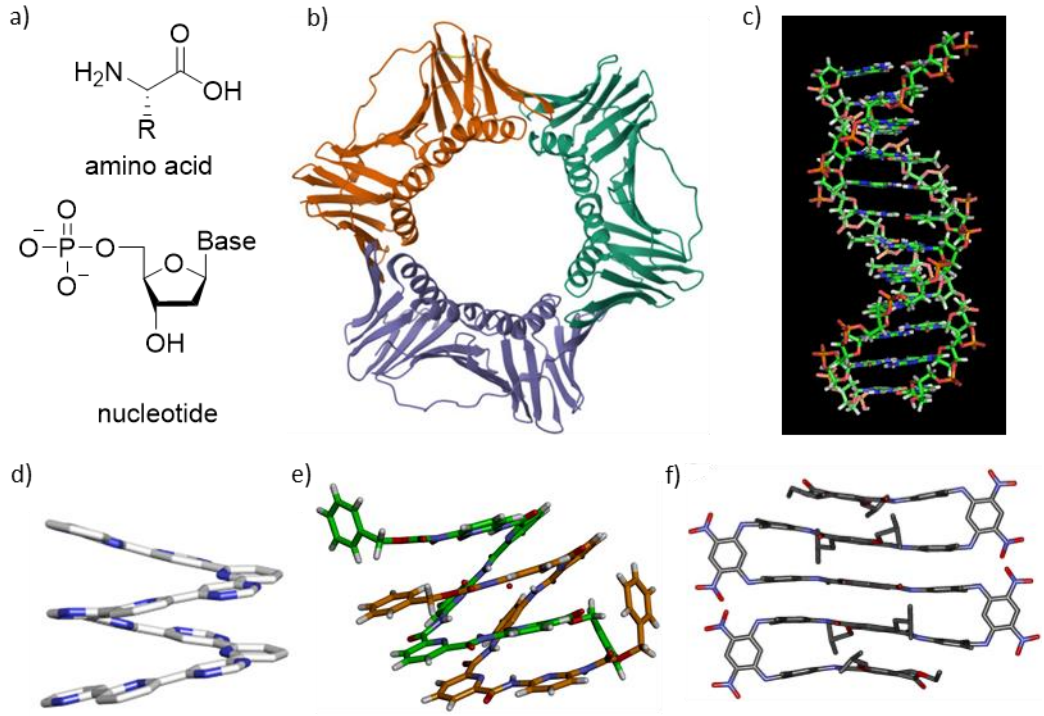


Figure 1. a) Chemical structure of amino acids and nucleotides. b) Structure of human proliferating cell nuclear antigen protein (PDB:1VYM). c) Structure of B-DNA. d) Crystal structure of an aromatic helical foldamer.[12] e) Crystal structure of a double-stranded foldamer.[13] f) Crystal structure of a foldamer sheet structure.[14]

The design of foldamers is rooted in understanding the folding principles of biopolymers. Early studies focused on modified canonical monomers, such as amino acid analogues, whose oligomers are generally known as peptidomimetics.[15-17] They can mimic natural peptides in certain respects and generally exhibit greater stability and strong biocompatibility, making them valuable in drug development. For example, the modification by introducing artificial side chains can often maintain the folding of the main chain while providing different surface properties that can enhance their target affinity and selectivity.[18] The introduction of some side chains will enable the covalent linkage

between two residues that have proximity in the folded state but are distinct from the primary structure, thereby stabilizing the fold.[19]

In addition to peptidomimetics, fully synthetic oligomers offer expanded structural and functional diversity. For example, aromatic oligomers possess predictable folding characteristics and unique properties, enabling them to remain folded even in organic solvents.[20-22] Various structure topologies, such as helices,[23-26] sheets,[14, 27, 28] and knots,[29-31] could be easily accessed. However, the research towards higher-hierarchy structures, such as tertiary or quaternary structure levels, to obtain true artificial protein-like objects remains a challenging yet attractive field, as sophisticated functions of proteins often arise from their tertiary or quaternary structure levels. Quinoline-based foldamer can form stable helical structures through intramolecular hydrogen bonding and π-π stacking across a wide range of solvent environments and temperatures.[32, 33] Their robust and predictable structures facilitate the precise introduction of side-chain residues that protrude from the helix. The first two abiotic helix-turn-helix tertiary structures were achieved by tethering two hydroxy-functionalized quinoline-based helices by turn units. However, the removal of the turn unit leads to the formation of two unexpected aggregates: tilted dimers and a parallel trimer (see section 3.4 below).[34, 35] The research in this thesis focused on the exploration of aggregated tertiary structures to obtain a higher structure level and on the turn unit design to promote the formation of a unimolecular tertiary structure mediated by the parallel trimer pattern. This section provides a general introduction to the field of foldamer science.

3.1. The folding principle, structures and function of biopolymers

The rational design of foldamers relies on understanding the sequence-structure-function relationships of natural biopolymers, particularly nucleic acids and proteins. The double-stranded DNA helix is the most typical structure of nucleic acids, stabilized by base pair recognition between complementary strands and π-stacking between base pairs. The recognition of base pairs via Watson-Crick base pairing is crucial for the functions of nucleic acids. Aptamers[36, 37] are nucleic acid-based artificial folded oligomers, *i.e.* foldamers, with the closest sequence-structure-function relationships. They typically fold into specific shapes, such as stem-loops,[38] hairpins,[39] pseudoknots[40, 41] and G-quadruplexes,[42-46] through intramolecular base-pair recognition, with applications in biomarker detection,[47, 48] therapeutics,[49] drug delivery,[50-52] diagnostics.[53-55] DNA origami technology (Figure 2a) is a recently developed technique that utilizes the properties of DNA base pair recognition to design nucleic acid sequences that can assemble into nanoscale three-dimensional structures, such as nanotubes, and nano-prisms. It is widely used in the field of nano-sensors, drug delivery, and bio-imaging.[56-62]

Proteins have four structure levels: primary, secondary, tertiary and quaternary.[63] The primary structure is the linear sequence of polyamide composed of 20 different amino acids. Notably, the synthesis of the peptide sequences in cells by ribosomes is started from the N-terminus, while solid-phase peptide synthesis (SPPS) progresses in the opposite direction, that is from the C-terminus to the

N-terminus.[64-66] The primary structure of proteins could be determined by sequencing using Mass spectrometry,[67, 68] or by other nanotechnologies, such as those based on nanopore.[69-71]

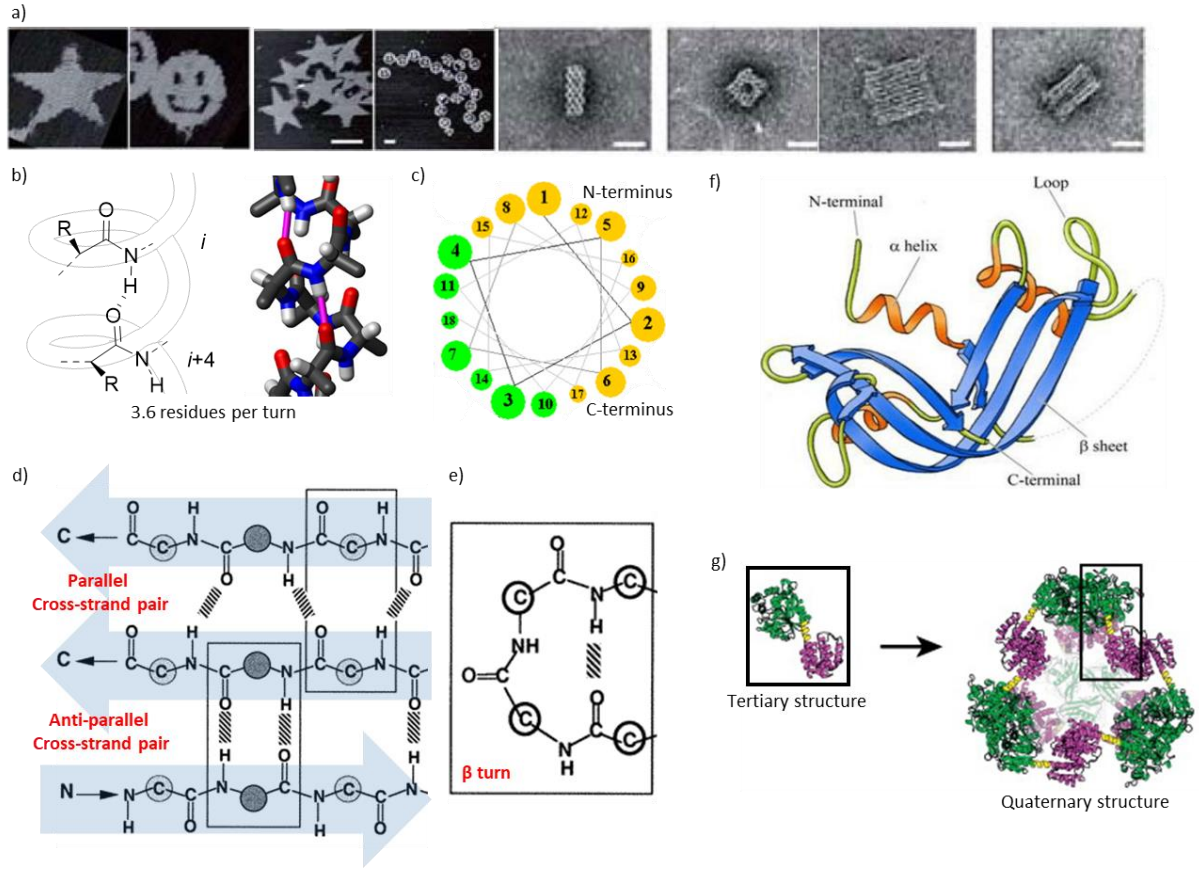


Figure 2. a) Some structures of DNA origami.[59] b) Intramolecular hydrogen bond in α-helix. c) The top view of a α-helix down through the axis shows the positioning of the residues. d) Representation of the patterns of parallel and anti-parallel β-sheet.[91] e) Structural representation of β-turn.[91] f) Schematic representation of the tertiary structure of a protein, illustrating its composition of several secondary structure elements. g) The principle of quaternary structure, which involves the assembly of small protein motifs.[93]

The secondary structures, including α-helix (Figure 2b), β-sheet (Figure 2d), β-turn (Figure 2e), collagen helix, etc., represent the local three-dimensional conformations of peptide chains.[72] Hydrogen bonds usually play a crucial role in the formation of secondary structures. The α-helix is the most common motif, where the peptide chain coils into a right-handed helix with 3.6 amino acid residues per turn (Figure 2b). The helical structure is stabilized by the hydrogen bonds between the carbonyl oxygen and the amide hydrogen of residues that four residues earlier in the sequence (Figure 2b). There are 13 atoms involved in the ring formed by hydrogen bonds, the α-helix is also known as 3.6_{13}-helix.[73-76] The coiled-coil structure, one of the most widespread protein structural elements, consists of two or more α-helices coiled together through hydrophobic effects.[77] Due to their high stability, α-helices and coiled-coil structures play essential structural and functional roles in protein-protein interactions[78, 79] and DNA-protein interactions.[80, 81] Other helical secondary structures include the 3_{10}-helix[82-84] and π-

helix.[85-87] The β-sheet (Figure 2d) is formed by two or three parallel or anti-parallel β-strands, held together by hydrogen bonds to form a stable sheet-like structure.[88-91] β-turns (Figure 2e) are short, flexible regions connecting adjacent β-strands, typically consisting of four amino acids and allowing the chain to reverse direction.[91, 92] β-sheets and β-turns are often located in the core of proteins, contributing to the overall structural stability.

The secondary structures are crucial for the three-dimensional shape and function of proteins, as the combination and arrangement of multiple secondary structures within a single peptide chain form the tertiary structure of the protein (Figure 2f). The tertiary structure is stabilized by hydrophobic effects,[94] hydrogen bonds, salt bridges,[95] disulfide bonds[96] between different secondary structures. The tertiary structure is essential for function. Misfolded proteins, such as amyloid fibrils, are considered to be one of the causes of neurodegenerative diseases.[97] The quaternary structure, the highest level of protein structure classification, refers to the assembly of several protein motifs with tertiary structure into a larger complex (Figure 2g).[93, 98] Of note, not all the proteins have a quaternary structure, as many function as monomers.

Techniques such as X-ray crystallography, NMR, and cryo-electron microscopy help determine the structure of proteins. In recent years, an in-depth understanding of the sequence-structure relationship and machine learning have driven the development of structure prediction software, such as AlphaFold[99], which can predict the protein structure with high accuracy. The targeted modification of natural proteins through mutation also relies on high-resolution structural analysis and precise structure prediction. Enzyme modification is a typical example.[100, 101] Enzymes with low regioselectivity may produce multiple products from a single substrate. By studying enzyme structures and their substrate recognition mechanisms, regioselectivity can be improved through mutations. In recent years, the *de novo* design of proteins has made significant progress. Woolfson and co-workers recently reported the rational design of oligopeptide sequences that can aggregate into stable 3_{10} helices or membrane-less organelles.[102, 103] Baker and co-workers developed a series of software, such as Rosetta which could be utilized in the design of artificial proteins.[104, 105] Additionally, some peptide origami structures with supramolecular self-assembly properties have been reported, capable of forming various nanostructures for applications in drug delivery, molecular cages, and biosensors.[106]

3.2. The design of peptidomimetics and aliphatic foldamers

The limited library of canonical natural building blocks restricts the exploration of new structures and functions. Novel building blocks can endow oligomers with unique properties and functions. Thus, modifications of canonical monomers, such as synthetic L-α-amino acid analogues, have driven the research into peptidomimetics, which could be considered as foldamers with aliphatic backbones.

Peptidomimetics represent an important field in chemistry, pharmacology and material science. Peptidomimetics are oligomers of modified natural amino acid analogues or a class of oligomers that

can mimic the folding of natural peptides but have entirely different backbones. Common modification strategies include amine alkylation, side-chain substitution, structural bond extension, cyclization, carbonyl replacement, and isosteric replacements. The side-chain substitution strategy largely preserves the secondary structure of the peptide while allowing different residue displays. By introducing thiol- or alkenyl-modified amino acids at specific sites, side-chain cross-linking strategies can be used to establish side-chain-to-side-chain bridges for cyclization, thereby stabilizing the folded secondary structure and enhancing binding affinity with the target. In some cases, dithiol- or trithiol-substituted small molecules can serve as a staple to stabilize the relatively complex structure *via* two or three thioether bonds (Figure 3a).[107, 108]

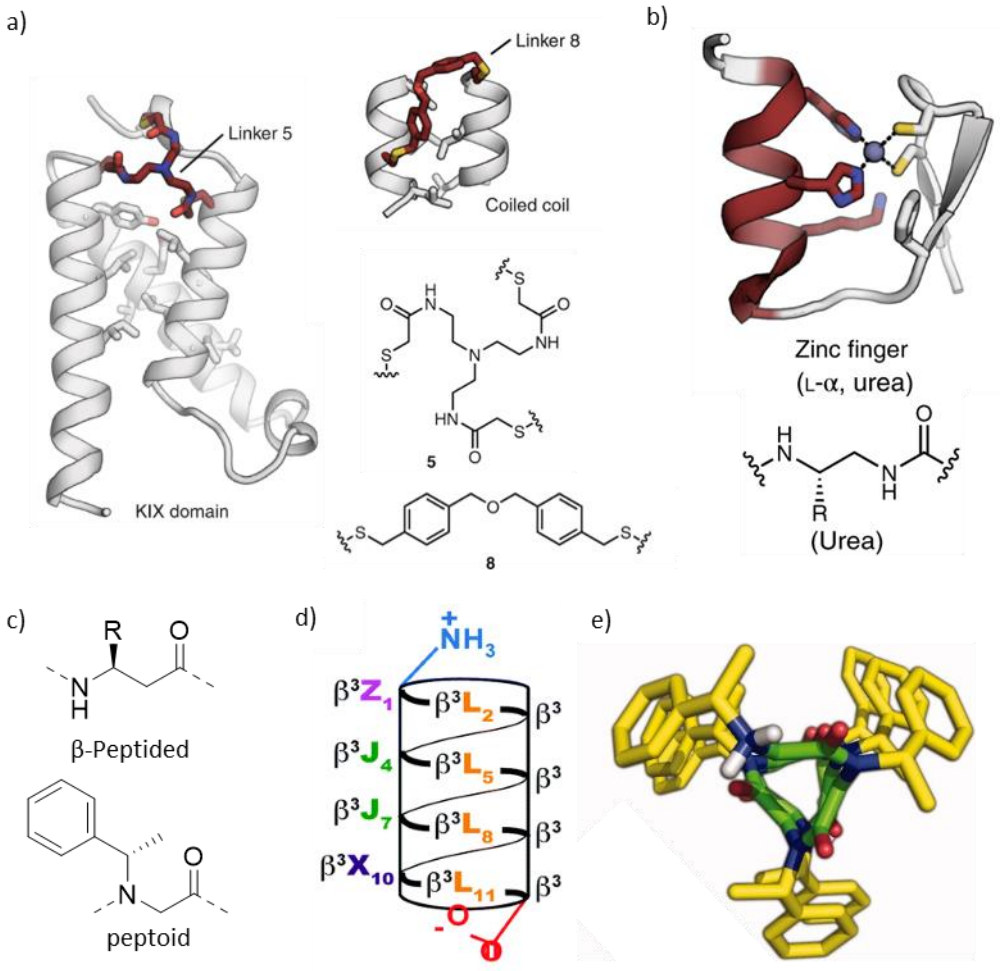


Figure 3. a) The cross-link strategies to stabilize the fold via sidechain-to-sidechain bridges.[15] b) The oligo-urea can mimic the zinc finger of peptides.[15, 111] c) The chemical structure of β-peptide and peptoid. d) The residue positioning of a β-peptide bundle.[110] d) Molecular modelling of a peptoid octamer containing bulky, chiral *S*-1-phenylethyl side chains.[113]

The oligomers with entirely artificial aliphatic backbones can also mimic the surface of peptides but with distinct properties, such as increased stability and resistance to proteolytic degradation. The β-peptide, which has a one carbon chain extension with respect to canonical amino acid, can fold into a

3_{14}-helix (Figure 3c, d).[109, 110] Oligo-ureas, which could be considered as a γ-amino acid analogue, can fold into a helical shape through intramolecular hydrogen bonds. Although its folding behavior differs from that of peptides, it can mimic structures that are similar to natural peptides, such as the zinc finger structure (Figure 3b).[111] The oligomer constituted by N-substituted amino acids will exhibit very different folding properties and residue arrangements, as the folding is not stabilized by intramolecular hydrogen bonds due to the absence of hydrogen bond donors from the amide group. The oligomer of N-substituted glycine (Figure 3c, e), which is known as oligo-peptoid, offers potential as a novel class of peptidomimetics for drug candidates.[112]

3.3. Aromatic foldamers and folding principle

Aromatic foldamers refer to foldamers containing aromatic monomers as the main component of their backbone. The folding of aromatic foldamers is similar to that of biopolymers, mainly driven by non-covalent interactions such as hydrogen bonds, electrostatic repulsion and π-stacking. However, the rigidity of the aromatic monomers significantly enhances the stability and predictability of aromatic foldamer structures.

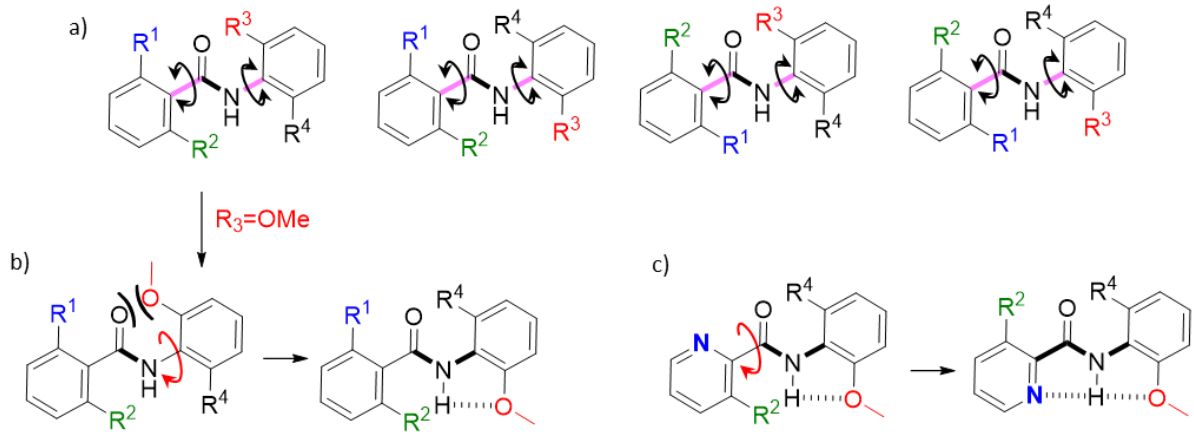


Figure 4. a) The representation showing the four conformations of N-phenylbenzamide. Since there are no restrictions applied on the pink-colored bonds, there is no preference. b) When the R^3 is substituted by a methoxyl group, the intramolecular hydrogen bond and electrostatic repulsion will apply restriction on the (Ar)C−NH(COAr) bond (shown in bold). c) If the left aromatic ring is replaced by pyridine, the three-centred intramolecular hydrogen bond will restrict the rotation of all the bonds.

The most common type of aromatic foldamers is aromatic oligoamides, in which aromatic monomers are connected through amide bonds. Due to the rigidity of the monomers, it is rarely observed that the amide proton forms a hydrogen bond with the carbonyl group of a few units earlier, unlike the behavior observed in peptide helices. Therefore, the (Ar)C−CO(NHAr) and (Ar)C−NH(COAr) bonds (the two bonds connecting the amide group and two adjacent aromatic monomers) can rotate, leading to four possible pseudo-conjugated conformations (Figure 4a). The lack of preference between these conformers hinders the formation of well-defined structures. Introducing a hydrogen bond acceptor at

one of the *ortho*-positions of the amide group will induce the formation of an intramolecular hydrogen bond which restricts the rotation of the amide group. Common hydrogen bond acceptors include oxygen and nitrogen. Effective strategies to introduce hydrogen bond acceptors to the aromatic monomers involve oxygen (ether group) substitution on the aromatic ring and incorporating nitrogen-containing aromatic heterocyclic systems, such as pyridine and quinoline. The electrostatic repulsion and intramolecular hydrogen bonds introduced by these modifications will favor specific amide conformations, thereby constraining the rotation of bonds within the rigid aromatic backbone (Figure 4b,c).

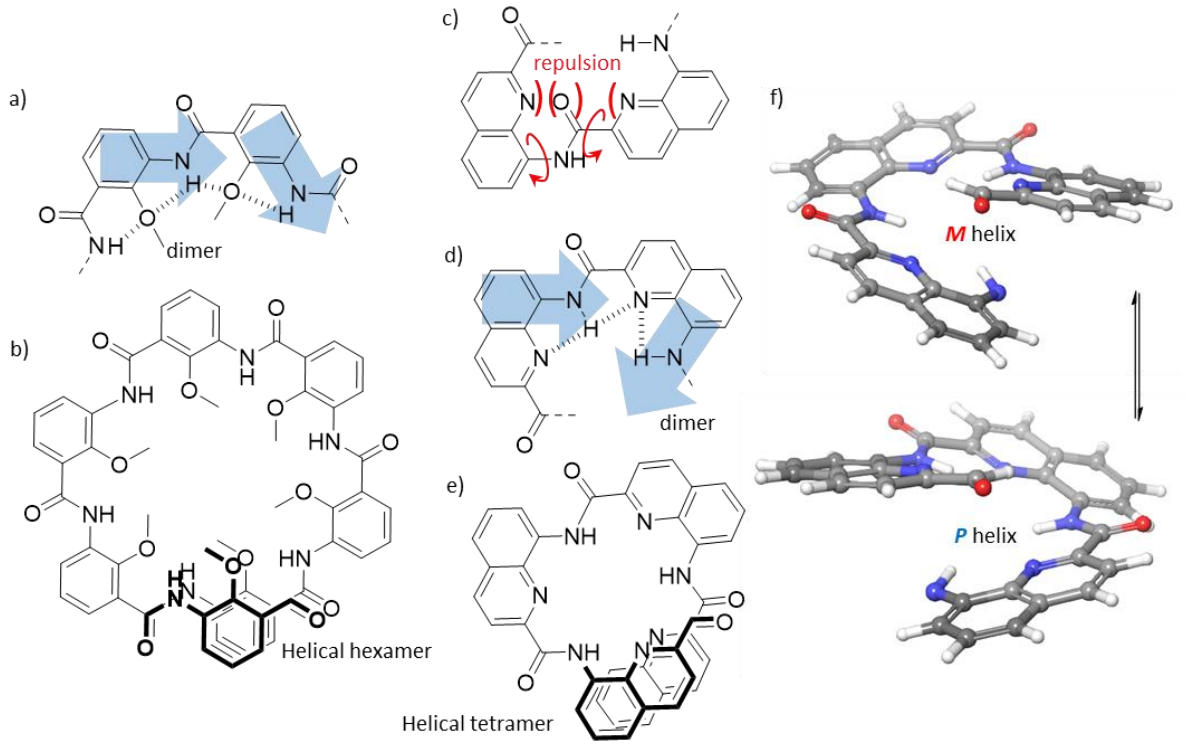


Figure 5. a) The dimer of 3-amino-2-methoxybenzoic acid has a crescent shape. b) The hexamer of 3-amino-2-methoxybenzoic acid will form a helical shape [117] c) The repulsion between the carbonyl group and quinoline nitrogen. d) The dimer of 8-amino-2-quinoline carboxylic acid has a crescent shape. e) The tetramer of 8-amino-2-quinoline carboxylic acid will form a helical shape. f) There is no preference for the *M* helical or *P* helical.

By adjusting the size of the aromatic ring, as well as the position and conformation of the amide group, the distance and defined angle of two consecutive amide bonds can be varied. The angle between two amide bonds is usually less than 180° and the intramolecular hydrogen bond will lead the dimer to form a flat crescent shape (Figure 5a, d). When the oligomer exceeds one turn, the steric clash caused by two termini twists the amide bonds and induces the formation of a helical shape (Figure 5b, e). Notably, since there is no preference for the amide bond twisting either up or down, helical aromatic foldamer helices are generally racemic mixtures of *P* and *M* conformers (Figure 5f), which differs from natural helical peptides. The strategies to bias the handedness of an aromatic foldamer helix include introducing a chiral monomer analogue[114] within the sequence or a chiral group at the terminus.[115] In

the absence of aryl-amide bond conformational restriction, a guest molecule that can form intermolecular hydrogen bonds with the oligomer may still induce the formation of a helical structure.[116]

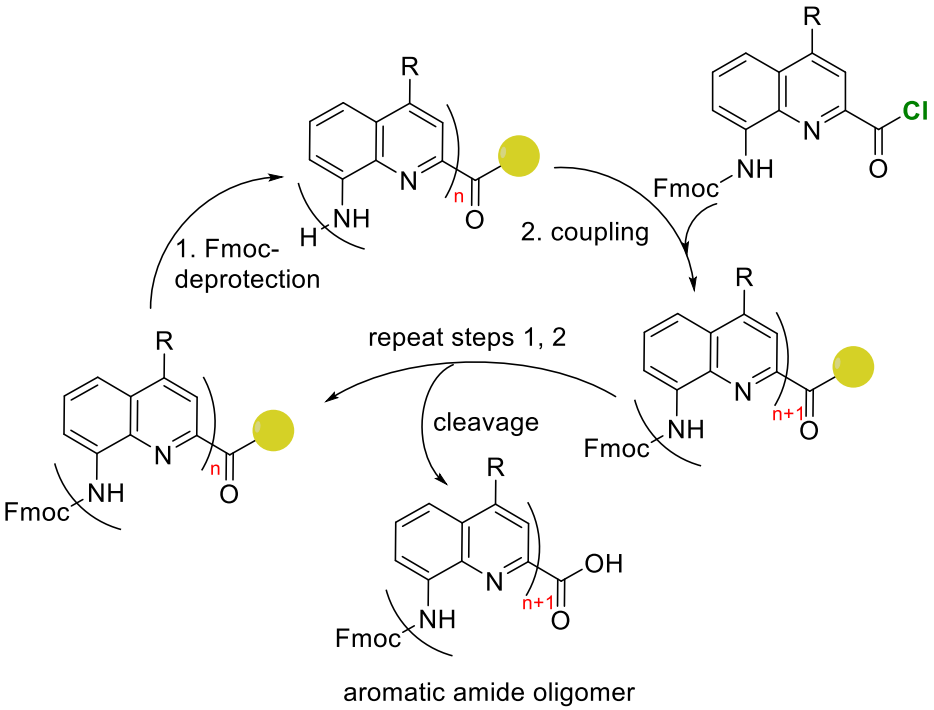


Figure 6. The scheme representation of the cycle of solid phase foldamer synthesis

The 8-amino-2-quinoline carboxylic acid (Q) is an aromatic δ-amino acid.[32] The oligoamide of Q units will fold into a 2.5 units per turn helical structure (Figure 5d, e) stabilized by electrostatic repulsion, intramolecular hydrogen bonding, and π-π stacking interactions. Due to the poor solubility of such an aromatic oligoamide, a side chain is often added to the 4-position to enhance solubility. The synthesis of amide oligomers is usually achieved by coupling an acid with an amine. However, due to steric hindrance caused by the helical conformation, synthesizing relatively long oligomers in solution is challenging: as the length of the oligomer increases, the deprotection and coupling of the N-terminal amine or C-terminal acid becomes more difficult. Relatively long oligoamide of Q units could be synthesized using solid phase foldamer synthesis (SPFS), similar to the method used in solid-phase peptide synthesis (SPPS).[118] Standard coupling reagents such as PyBOP and HBTU, widely used in peptide and some peptidomimetic synthesis, may create challenges in oligo-Q sequence synthesis due to the poor nucleophilicity of the aromatic amine. An *in situ* activation method based on the Appel reaction was developed to pre-activate the acid to form the corresponding acyl chloride. Notably, excess activation reagent does not react with the free amine, unlike other activation reagents such as $(COCl)_2$ or Ghosez's reagent. The coupling reaction was generally performed under microwave conditions, and the corresponding oligomer could be cleaved from the solid support at the end of the synthesis. In this process, different Q units with different side chains could be introduced (Figure 6).

The inner diameter of oligo-Q is relatively small and filled with a hydrogen bond network. Nevertheless, metal ions, such as Cu^{II}, could coordinate by deprotonation of the amide group.[119] The modifications of the helix are generally applied to the outer rim of the helix, especially the 4-, 5- and 6-positions of the quinoline ring. For example, by introducing phosphonate groups at specific positions, the helix can mimic the negatively charged surface of B-DNA.[120, 121] The inner diameter can be adjusted by changing the position of the amine group on quinoline- or naphthalene-based monomers.[122] A larger diameter creates a cavity capable of binding certain guest molecules, such as sugars,[123] and also permits the formation of double helices.[124]

3.4. Helix self-assembly, linker design and abiotic tertiary structure

In nature, the advanced functions of biopolymers typically emerge at the tertiary or quaternary structural levels. Therefore, exploring the self-assembly and higher-order structural hierarchy of foldamers is both challenging and highly intriguing. Foldamers with tertiary structures and beyond are also known as "proteomimetics," as they share many common features with proteins.[15]

Self-assembly is a common phenomenon in nature, where large and complex structures can be formed through the self-assembly of smaller modules. Quaternary structures are essentially self-assembled from already complex subunits (tertiary structures). Examples from the field of peptidomimetics include a β3-residue helical oligomer that forms an octameric bundle [109, 125] and a urea-based helical oligomer that forms a hexameric assembly.[126] It is noteworthy that these self-assemblies are not genuine quaternary structures, as they are assembled secondary (not tertiary) structures.

The quinoline-based helical structures have advantages in exploring complex architectures due to their structural stability in various environments. In proteins, the most common tertiary structure motif is the helix-turn-helix (HTH) structure with a “U” shape. This motif consists of two α-helices connected by a short amino acid chain and typically serves as a DNA-binding domain. Research on higher-order structural levels of abiotic foldamers first investigated the connection of two helical structures via a turn unit (Figure 7).[127] The steric clash causes the two helices to point in opposite directions. Due to the intramolecular hydrogen bonds formed by the turn unit and the terminal amide bond, the two helices are heterochiral, while they do not interact directly with each other, resulting in a structure that deviates from the typical "U" shape of the protein HTH structure.

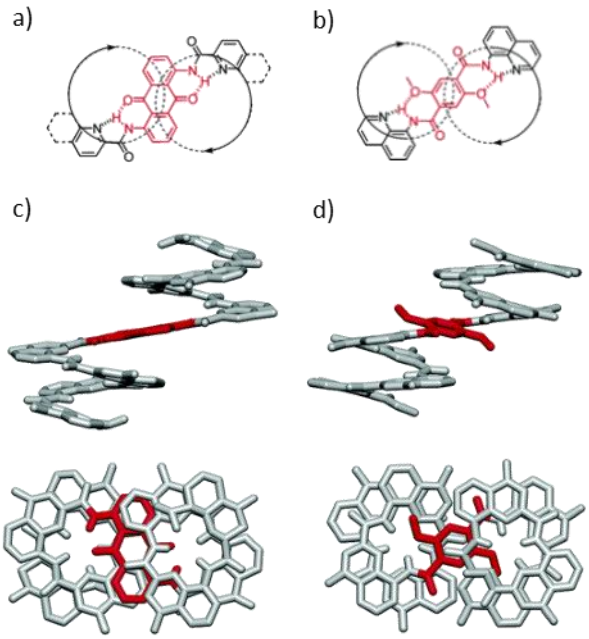


Figure 7. The structure of an aromatic di-amine turn unit (a) and an aromatic di-acid turn unit (b). They can tether the C-terminus or the N-terminus of two helices, respectively. The intramolecular hydrogen bonds are shown in dash lines. The arrow shows the handedness of the structures. The side view and top view of two Q_4-turn-Q_4 structures. The turn units in (c) and (d) are shown in (a) and (b), respectively.[127]

The exploration of the true abiotic tertiary structure was facilitated by using computational tools. A turn unit T1 was picked, and computational modelling was employed to connect two Q_4 oligomers to the hydrazine groups of the turn unit (Figure 8a, 10d).[34] Due to the additional hydrazine groups, the two helices are slightly separated compared to the structure shown in Figure 7d, allowing them to point in the same direction (Figure 8c). The hydrogen bonds between the turn unit and the terminal amide groups provide rigidity to the turn unit and induce homochirality in the two helices (Figure 8a). The resulting molecular model revealed a proximity between the 4-position of one quinoline from one helix and the carbonyl group of the other helix (Figure 8c), which allowed the introduction of hydrogen bond donors by replacing a Q unit with a X unit (Figure 8b) to establish intramolecular but interhelical hydrogen bonds (Figure 8d). modelling indicated that when the oligomers were extended to the second layer, the aromatic ring of the Q unit could cause steric clashes, hindering the formation of the HTH structure (Figure 8e). Therefore, the corresponding X units were replaced with pyridine-based Y units (Figure 8b) to prevent the clash (Figure 8f).

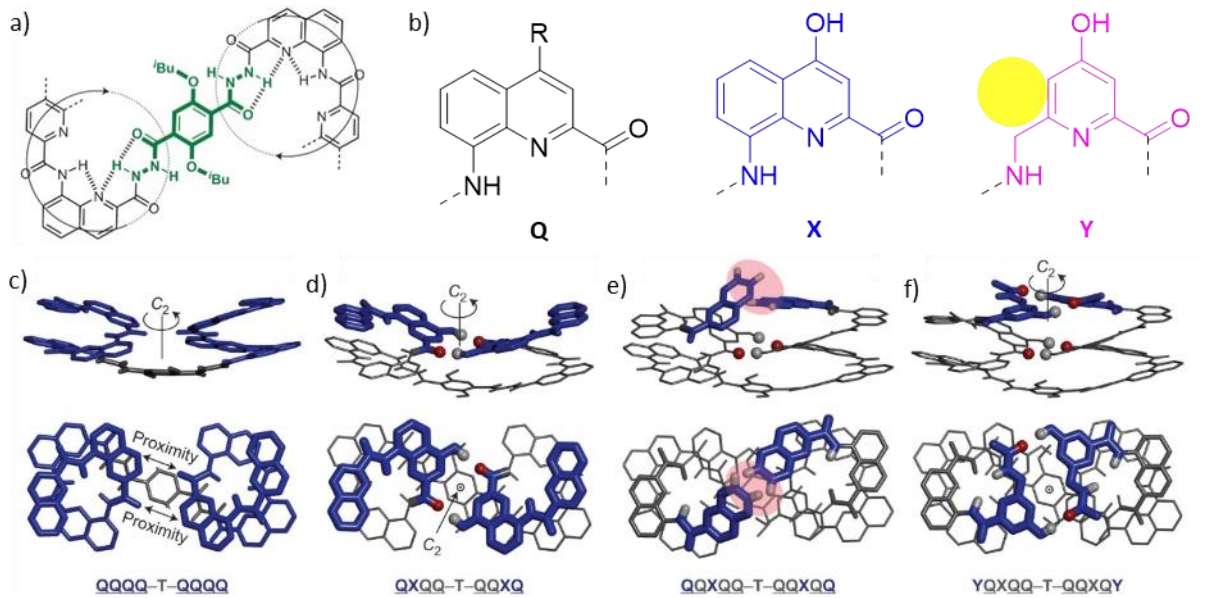


Figure 8. a) The structure of the T1 turn (in green in the structure). The structure was stabilized by hydrogen bonds and induced the handedness and orientation of two helices connected. b) The structure of Q, X and Y units. c) The molecular model of Q_4-T1-Q_4 structure. The positions for the insertion of hydrogen bond donors due to proximity are highlighted. d) The molecular model of QXQQ-T1-QQXQ. The hydroxy proton of the X unit is shown in grey. e) The representation showing the potential steric clash when prolonging the oligomer, which could be prevented by replacing the X monomer with the Y monomer. f) The molecular model of YQXQQ-T1-QQXQY. There are four intramolecular hydrogen bonds.

The crystal structure of the oligomer revealed the formation of the HTH structure stabilized by six intramolecular hydrogen bonds as designed (Figure 9a). Surprisingly, the removal of the turn unit led to the formation of homochiral tilted dimers (Figure 9e) and homochiral parallel trimers (Figure 9f) through intermolecular hydrogen bonds. Notably, the parallel dimer pattern observed in the HTH structure was absent, indicating that this structure is not favored in the absence of T1 (Figure 9c). In later studies the hydrogen bond donors contributed by Y monomers in the middle layer was removed (Figure 9b).[128] The modified oligomer folded into a similar HTH structure without any loss of stability, while the helices adopted their natural curvature (Figure 9d, right). It was concluded that the helices of the original HTH structure underwent some spring torsional load (Figure 9d, left), suggesting the presence of potential conformational frustration. The helices in the tilted dimer and parallel trimer also maintained their natural curvature while forming the same number of hydrogen bonds, which could explain the preference for these two structures in the absence of turn unit.

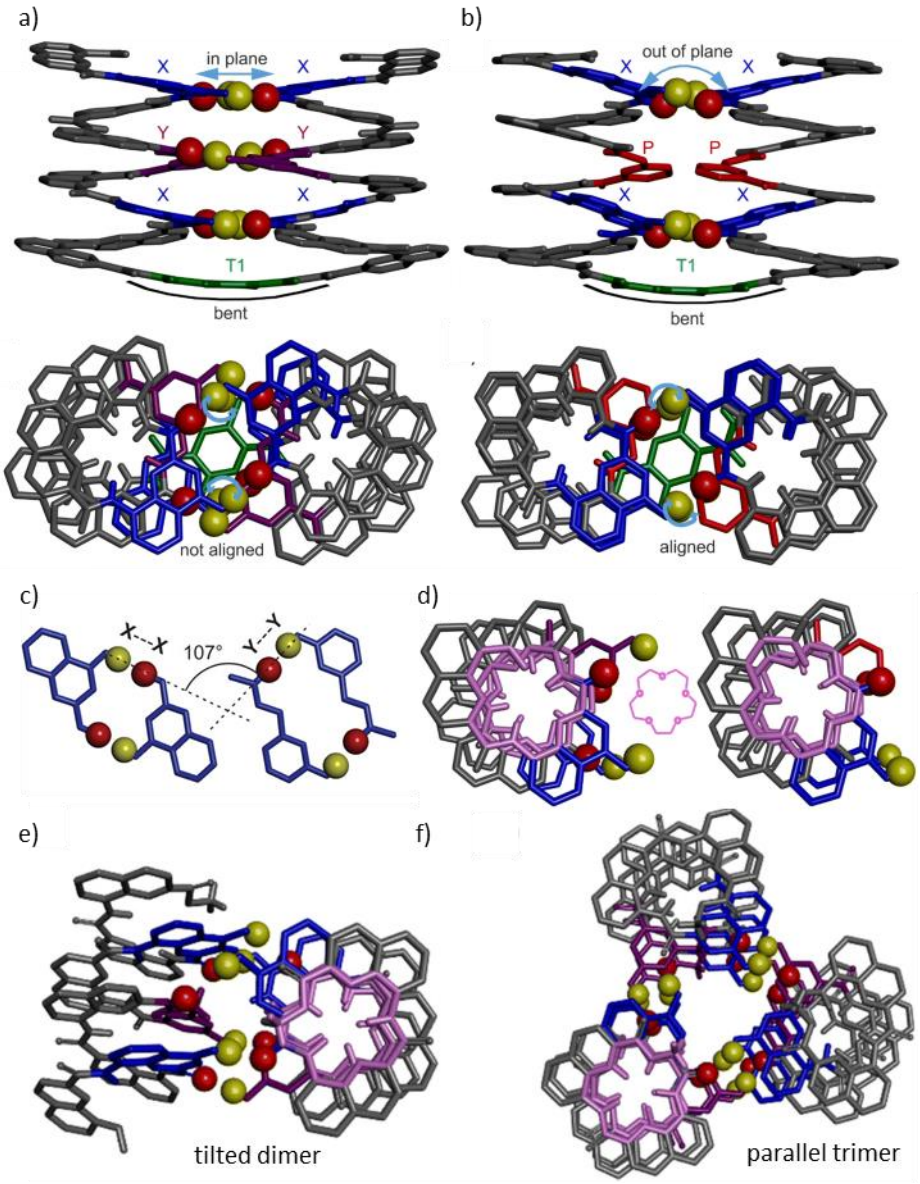


Figure 9. a) The side view and top view of the crystal structure of homochiral HTH structure stabilized by six intramolecular hydrogen bonds. b) The side view and top view of the crystal structure of homochiral HTH structure stabilized by four intramolecular hydrogen bonds. c) The hydrogen bond pattern in (a). d) The top view, down through the axis of the helix, of the helix in (a, left) and in (b, right). The inner rim of the helix is shown in pink showing the potential torsional spring load of the backbone in (a). The crystal structure of tilted dimer (e) and parallel trimer (f).

The helix-T1-helix structure is a symmetrical structure formed by linking two identical helices to a symmetrical T1 unit. When both the handedness and orientation of a helix are inverted, the side chain will align in a similar arrangement (Figure 10c). This transformation is also known as "retro inverso peptide" in the modification of peptides. An unsymmetrical turn unit T2 (Figure 10d) could be introduced between two heterochiral head-to-tail arranged helices to induce the formation of a heterochiral HTH structure[35] while maintaining a similar intramolecular hydrogen bond pattern to that

of the homochiral HTH structure[34] (Figure 9c). The turn unit T2 was prepared as a Fmoc-protected amino acid, facilitating the synthesis of the corresponding sequence based on the SPFS method. This well-defined heterochiral HTH structure represents an abiotic tertiary structure and serves as an excellent candidate for exploring the abiotic self-assemblies, that is, the genuine abiotic quaternary structures. Chapter 5 describes our success at getting such abiotic quaternary structures. However, research on abiotic quaternary structures designed using computational models has also led to the discovery of a novel domain-swapped structure, which will be discussed in Chapter 4.

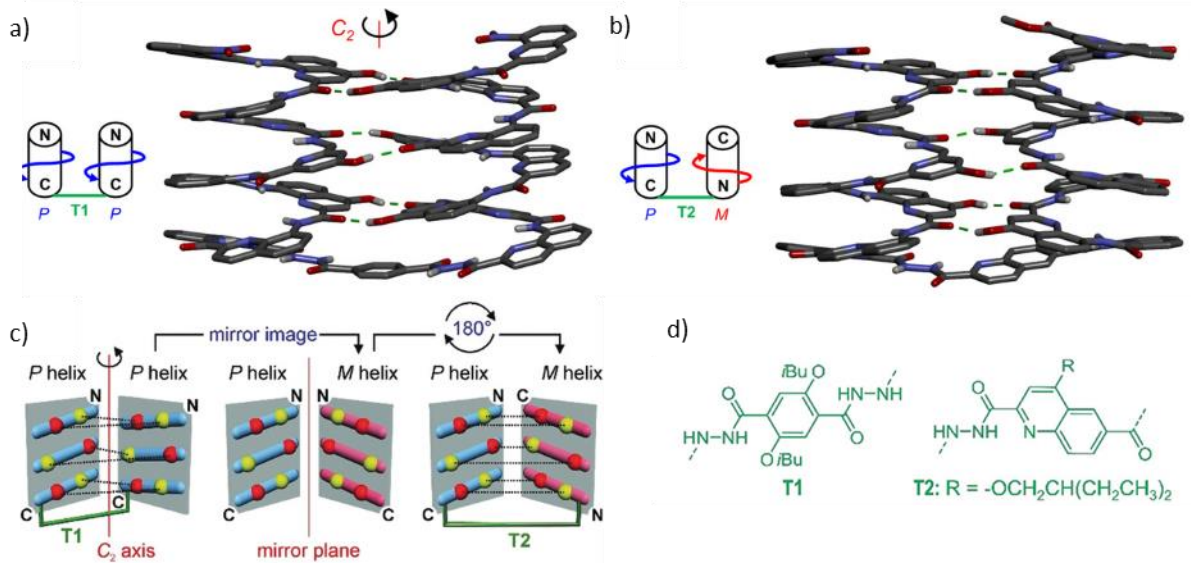


Figure 10. a) The crystal structure and cartoon representation of homochiral helix-turn-helix structure. b) The crystal structure and cartoon representation of heterochiral helix-turn-helix structure. The hydrogen bond is shown with a green dashed line. c) The cartoon representation showing the hydrogen bonding interface. The arrays of six hydrogen bond donors (yellow balls) and acceptors (red balls) are approximated to belong to two planes facing each other. The blue rod and pink rod represent the *P* helix and the *M* helix, respectively. Converting both the handedness and orientation of a helix will remain the arrangement and positioning of hydrogen bond donors and acceptors. d) The chemical structure of T1 and T2.

The subsequent research focused on controlling the self-assembly of hydroxy group-decorated single helices, aiming either to quantitatively achieve one of two unexpected aggregates mentioned above, the tilted dimer and the parallel trimer, or to promote the formation of the parallel dimer observed in the homochiral HTH structure (Figure 9c). The hydrogen bonds contributed by the Y monomer induce a torsional spring load of the main chain within the HTH structure, potentially hindering the formation of the desired parallel dimer in the absence of turn unit. Therefore, Y monomers were replaced with non-hydroxy-substituted pyridine monomer P to reduce conformational strain. The sequence was also extended to favor parallel rather than tilted arrangement (Figure 11a).[129] However, the designed sequence aggregated into yet into new unexpected motifs, a homochiral shifted dimer (Figure 11b, d) and a heterochiral shifted dimer (Figure 11c, e).[129] The social or narcissistic chiral self-sorting occurrence highly depends on the chlorinated solvent that is used: the homochiral shifted dimer

dominated in dichloromethane whereas the heterochiral shifted dimer dominated in chloroform. Since the helices in the helix-T2-helix structure are heterochiral, further research on promoting the self-assembly of the heterochiral HTH structure *via* shifted dimer intermolecular interfaces will be discussed in Chapter 5.

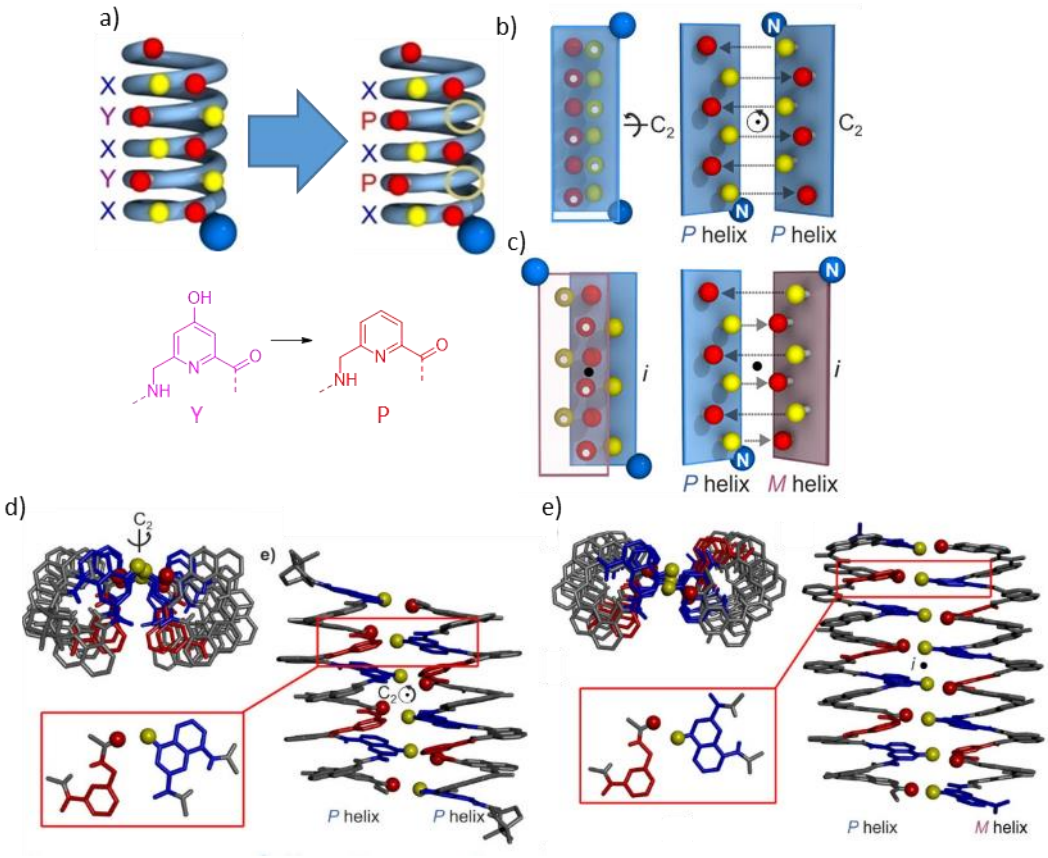


Figure 11. a) Schematic representation highlighting the positions of hydrogen bond donor removal (yellow circles). Schematic representation of the hydrogen bonding arrays of homochiral shifted dimer (b) and heterochiral shifted dimer (c), with overlaid view and open-book view. The N-terminus is shown with a blue ball. The P and M helices are shown with blue and red planes, respectively. The crystal structure of homochiral shifted dimer (d) and heterochiral shifted dimer (e).[129]

Introducing linkers to bridge subdomains in the self-assembly is a common strategy for designing artificial tertiary structures. For example, Woolfson and co-workers started with a coiled-coil tetramer, establishing three connections between the C-terminus and N-terminus of two adjacent helices via short peptide chains to create a single-chain four-helix bundle.[130] As previously mentioned, the tilted dimer is a preferred self-assembly motif, making it a strong candidate for exploring tertiary structures mediated by this pattern. One reason for the particular focus on the tilted dimer pattern is the unique 120° angle between the two helices, which is distinct from a parallel or antiparallel arrangement in other cases. Since the termini of the two helices are not on the same plane, a series of diethylene glycol chains of varying lengths were introduced as flexible turn units (T3, Figure 12a).[131] The sequences fold into

the desired HTH structure with good confidence despite lacking proof from a solid state structure (Figure 12b, c). However, X-ray diffraction of a single crystal of one sequence revealed an unexpected tetrameric, eight-helices bundle (Figure 12d, e). This bundle includes two "U"-shaped HTH sequences (colored in blue and green in Figure 12d, e) and two linear HTH sequences (colored in organ and red in Figure 12d, e), forming a quaternary structure-like object with a molecular weight of 12.9 kDa. Since the parallel trimer is also a preferred self-assembly motif, research on linker design to promote tertiary structures mediated by this pattern, following a similar strategy, will be discussed in Chapter 6.

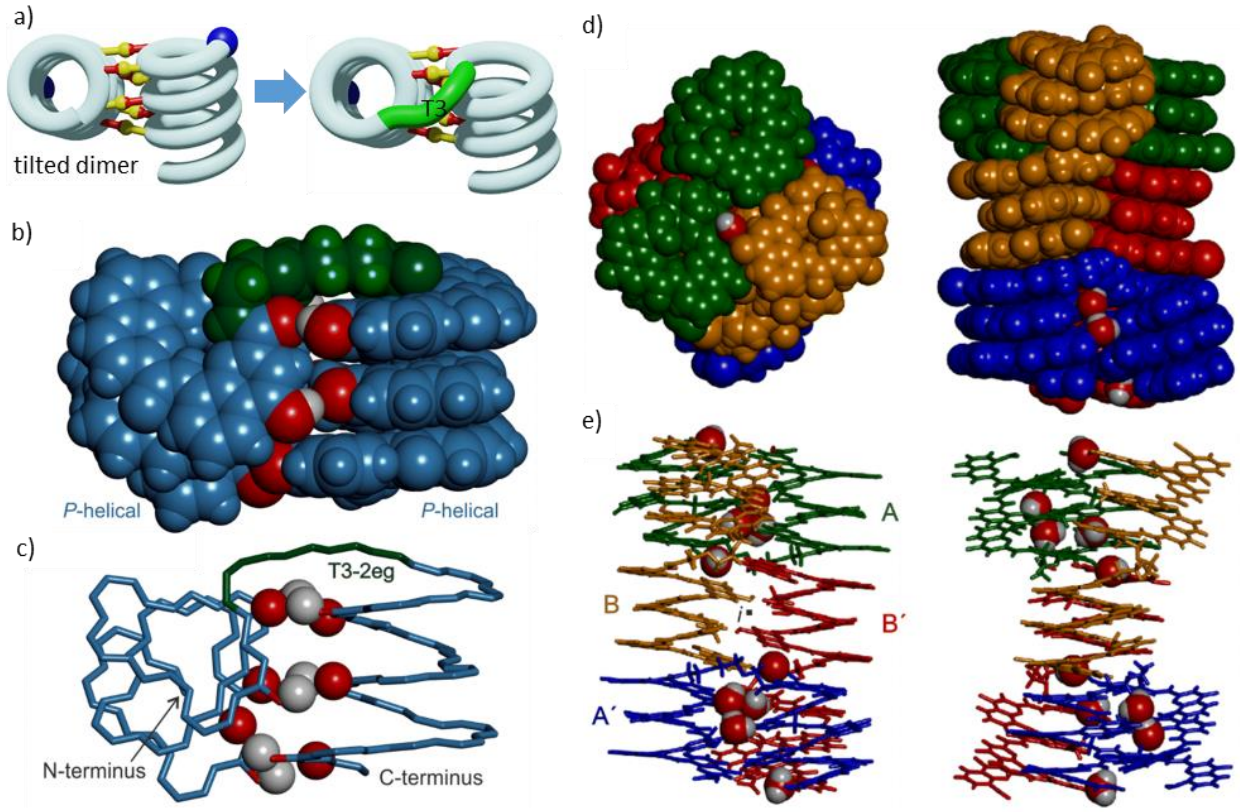


Figure 12. a) The representation of the design strategy. The energy mimized model of a helix-T3-helix structure is shown in a space-filling representation (b) and simplified representation (c, only the outer rim of the backbone is shown). Hydrogen bond donors and acceptors are shown as white and red balls, respectively. The crystal structure of a tetrameric, eight helices bundle is shown in a space-filling representation (d) and tube representation (e). Four identical sequences are colored in different colors.

References:

[1] D. A. Kraut, K. S. Carroll, D. Herschlag, *Annu. Rev. Biochem* **2003**, *72*, 517-571.

[2] J. B. van Beilen, Z. Li, *Curr. Opin. Biotechnol.* **2002**, *13*, 338-344.

[3] O. Kirk, T. V. Borchert, C. C. Fuglsang, *Curr. Opin. Biotechnol.* **2002**, *13*, 345-351.

[4] M. D. Shoulders, R. T. Raines, *Annu. Rev. Biochem* **2009**, *78*, 929-958.

[5] P. Fratzl, in *Collagen: structure and mechanics*, Springer, **2008**, pp. 1-13.

[6] H. Wang, *Polymers* **2021**, *13*, 3868.

[7] T. Zeuthen, *J. Membr. Biol.* **2010**, *234*, 57-73.
[8] A. Ayrton, P. Morgan, *Xenobiotica* **2001**, *31*, 469-497.
[9] M. H. Saier Jr, *Mol. Microbiol.* **2000**, *35*, 699-710.
[10] S. Minchin, J. Lodge, *Essays Biochem.* **2019**, *63*, 433-456.
[11] P. Belmont, J.-F. Constant, M. Demeunynck, *Chem. Soc. Rev.* **2001**, *30*, 70-81.
[12] D. M. Bassani, J. M. Lehn, G. Baum, D. Fenske, *Angew. Chem. Int. Ed.* **1997**, *36*, 1845-1847.
[13] C. Zhan, J. M. Leger, I. Huc, *Angew. Chem. Int. Ed.* **2006**, *45*, 4625-4628.
[14] L. Sebaoun, V. Maurizot, T. Granier, B. Kauffmann, I. Huc, *J. Am. Chem. Soc.* **2014**, *136*, 2168-2174.
[15] W. S. Horne, T. N. Grossmann, *Nat. Chem.* **2020**, *12*, 331-337.
[16] I. Avan, C. D. Hall, A. R. Katritzky, *Chem. Soc. Rev.* **2014**, *43*, 3575-3594.
[17] E. Lenci, A. Trabocchi, *Chem. Soc. Rev.* **2020**, *49*, 3262-3277.
[18] B. Hu, N. Song, Y. Cao, M. Li, X. Liu, Z. Zhou, L. Shi, Z. Yu, *J. Am. Chem. Soc.* **2021**, *143*, 13854-13864.
[19] A. Yeste-Vázquez, F. M. Paulussen, M. Wendt, R. Klintrot, C. Schulte, K. Wallraven, L. van Gijzel, B. Simeonov, M. van der Gaag, A. Gerber, *Angew. Chem. Int. Ed.* **2024**, e202411749.
[20] G. Guichard, I. Huc, *Chem. Commun.* **2011**, *47*, 5933-5941.
[21] I. Huc, *Eur. J. Org. Chem.* **2004**, *2004*, 17-29.
[22] D. W. Zhang, X. Zhao, J. L. Hou, Z. T. Li, *Chem. Rev.* **2012**, *112*, 5271-5316.
[23] Y. Ferrand, I. Huc, *Acc. Chem. Res.* **2018**, *51*, 970-977.
[24] E. R. Gillies, F. Deiss, C. Staedel, J. M. Schmitter, I. Huc, *Angew. Chem. Int. Ed.* **2007**, *46*, 4081-4084.
[25] D. Haldar, H. Jiang, J. M. Leger, I. Huc, *Angew. Chem. Int. Ed.* **2006**, *45*, 5483-5486.
[26] J. Fremaux, L. Mauran, K. Pulka-Ziach, B. Kauffmann, B. Odaert, G. Guichard, *Angew. Chem. Int. Ed.* **2015**, *127*, 9954-9958.
[27] J. Atcher, P. Mateus, B. Kauffmann, F. Rosu, V. Maurizot, I. Huc, *Angew. Chem. Int. Ed.* **2021**, *60*, 2574-2577.
[28] J. Atcher, A. Nagai, P. Mayer, V. Maurizot, A. Tanatani, I. Huc, *Chem. Commun.* **2019**, *55*, 10392-10395.
[29] R. F. Carina, C. Dietrich-Buchecker, J.-P. Sauvage, *J. Am. Chem. Soc.* **1996**, *118*, 9110-9116.
[30] S. D. Fielden, D. A. Leigh, S. L. Woltering, *Angew. Chem. Int. Ed.* **2017**, *56*, 11166-11194.
[31] M. Zhang, R. Nixon, F. Schaufelberger, L. Pirvu, G. De Bo, D. A. Leigh, *Nat. Chem.* **2024**, 1-7.
[32] H. Jiang, J.-M. Léger, I. Huc, *J. Am. Chem. Soc.* **2003**, *125*, 3448–3449.
[33] H. Jiang, J.-M. Léger, C. Dolain, P. Guionneau, I. Huc, *Tetrahedron* **2003**, *59*, 8365-8374.
[34] S. De, B. Chi, T. Granier, T. Qi, V. Maurizot, I. Huc, *Nat. Chem.* **2018**, *10*, 51-57.
[35] D. Mazzier, S. De, B. Wicher, V. Maurizot, I. Huc, *Angew. Chem. Int. Ed.* **2020**, *59*, 1606-1610.
[36] A. D. Ellington, J. W. Szostak, *Nature* **1990**, *346*, 818-822.
[37] C. Tuerk, L. Gold, *Science* **1990**, *249*, 505-510.
[38] R. E. Armstrong, G. F. Strouse, *Bioconjugate Chem.* **2014**, *25*, 1769-1776.
[39] I. Jeddi, L. Saiz, *Sci. Rep.* **2017**, *7*, 1178.
[40] M. Y. Kim, S. Jeong, *Mol. Cell* **2003**, *16*, 413-417.
[41] B. Jiang, F. Li, C. Yang, J. Xie, Y. Xiang, R. Yuan, *Anal. Chem.* **2015**, *87*, 3094-3098.

[42] W. O Tucker, K. T Shum, J. A Tanner, *Curr. Pharm. Des.* **2012**, *18*, 2014-2026.

[43] C. Roxo, W. Kotkowiak, A. Pasternak, *Molecules* **2019**, *24*, 3781.

[44] M. L. Bochman, K. Paeschke, V. A. Zakian, *Nat. Rev. Genet.* **2012**, *13*, 770-780.

[45] A. N. Lane, J. B. Chaires, R. D. Gray, J. O. Trent, *Nucleic Acids Res.* **2008**, *36*, 5482-5515.

[46] R. C. Monsen, J. O. Trent, J. B. Chaires, *Acc. Chem. Res.* **2022**, *55*, 3242-3252.

[47] H. Ulrich, C. Wrenger, *Cytom. Part A* **2009**, *75*, 727-733.

[48] J. Huang, X. Chen, X. Fu, Z. Li, Y. Huang, C. Liang, *Front. Cell Dev. Biol.* **2021**, *9*, 659760.

[49] E. W. Ng, D. T. Shima, P. Calias, E. T. Cunningham Jr, D. R. Guyer, A. P. Adamis, *Nat. Rev. Drug Discov.* **2006**, *5*, 123-132.

[50] F. He, N. Wen, D. Xiao, J. Yan, H. Xiong, S. Cai, Z. Liu, Y. Liu, *Curr. Med. Chem.* **2020**, *27*, 2189-2219.

[51] K. Chen, B. Liu, B. Yu, W. Zhong, Y. Lu, J. Zhang, J. Liao, J. Liu, Y. Pu, L. Qiu, *WIREs. Nanomed. Nanobi.* **2017**, *9*, e1438.

[52] Y. F. Huang, D. Shangguan, H. Liu, J. A. Phillips, X. Zhang, Y. Chen, W. Tan, *ChemBioChem* **2009**, *10*, 862-868.

[53] M. Liu, X. Yu, Z. Chen, T. Yang, D. Yang, Q. Liu, K. Du, B. Li, Z. Wang, S. Li, *J. Nanobiotechnology* **2017**, *15*, 1-16.

[54] W. Zhou, P.-J. J. Huang, J. Ding, J. Liu, *Analyst* **2014**, *139*, 2627-2640.

[55] S. Balamurugan, A. Obubuafo, S. A. Soper, D. A. Spivak, *Anal. Bioanal. Chem.* **2008**, *390*, 1009-1021.

[56] S. Dey, C. Fan, K. V. Gothelf, J. Li, C. Lin, L. Liu, N. Liu, M. A. Nijenhuis, B. Saccà, F. C. Simmel, *Nat. Rev. Method. Prime.* **2021**, *1*, 13.

[57] B. Saccà, C. M. Niemeyer, *Angew. Chem. Int. Ed.* **2012**, *51*, 58-66.

[58] K. F. Wagenbauer, F. A. Engelhardt, E. Stahl, V. K. Hechtl, P. Stömmer, F. Seebacher, L. Meregalli, P. Ketterer, T. Gerling, H. Dietz, *ChemBioChem* **2017**, *18*, 1873-1885.

[59] C. E. Castro, F. Kilchherr, D.-N. Kim, E. L. Shiao, T. Wauer, P. Wortmann, M. Bathe, H. Dietz, *Nat. Methods* **2011**, *8*, 221-229.

[60] F. Hong, F. Zhang, Y. Liu, H. Yan, *Chem. Rev.* **2017**, *117*, 12584-12640.

[61] M. Endo, Y. Yang, H. Sugiyama, *Biomater. Sci.* **2013**, *1*, 347-360.

[62] A. Krissanaprasit, C. M. Key, S. Pontula, T. H. LaBean, *Chem. Rev.* **2021**, *121*, 13797-13868.

[63] G. E. Schulz, R. H. Schirmer, *Principles of protein structure*, Springer Science & Business Media, **2013**.

[64] B. Merrifield, *Science* **1986**, *232*, 341-347.

[65] M. Amblard, J.-A. Fehrentz, J. Martinez, G. Subra, *Mol. Biotechnol.* **2006**, *33*, 239-254.

[66] R. Behrendt, P. White, J. Offer, *J. Pept. Sci.* **2016**, *22*, 4-27.

[67] K. Biemann, *Anal. Chem.* **1986**, *58*, 1288A-1300A.

[68] M. Fountoulakis, H.-W. Lahm, *J. Chromatogr. A* **1998**, *826*, 109-134.

[69] Z. L. Hu, M. Z. Huo, Y. L. Ying, Y. T. Long, *Angew. Chem. Int. Ed.* **2021**, *133*, 14862-14873.

[70] A. Asandei, G. Di Muccio, I. Schiopu, L. Mereuta, I. S. Dragomir, M. Chinappi, T. Luchian, *Small Methods* **2020**, *4*, 1900595.

[71] L. Restrepo-Pérez, C. Joo, C. Dekker, *Nat. Nanotechnol.* **2018**, *13*, 786-796.

[72] L. Pauling, R. B. Corey, H. R. Branson, *Proc.Natl.Acad.Sci.USA* **1951**, *37*, 205-211.
[73] M. Blaber, X.-j. Zhang, B. W. Matthews, *Science* **1993**, *260*, 1637-1640.
[74] R. Aurora, T. P. Creamer, R. Srinivasan, G. D. Rose, *J. Biol. Chem.* **1997**, *272*, 1413-1416.
[75] R. Aurora, R. Srinivasan, G. D. Rose, *Science* **1994**, *264*, 1126-1130.
[76] A. Chakrabartty, R. L. Baldwin, *Adv. Protein Chem.* **1995**, *46*, 141-176.
[77] W. H. Landschulz, P. F. Johnson, S. L. McKnight, *Science* **1988**, *240*, 1759-1764.
[78] B. N. Bullock, A. L. Jochim, P. S. Arora, *J. Am. Chem. Soc.* **2011**, *133*, 14220-14223.
[79] T. A. Edwards, A. J. Wilson, *Amino acids* **2011**, *41*, 743-754.
[80] R. Tan, L. Chen, J. A. Buettner, D. Hudson, A. D. Frankel, *Cell* **1993**, *73*, 1031-1040.
[81] Y. Huang, S. Dey, X. Zhang, F. Sönnichsen, P. Garner, *J. Am. Chem. Soc.* **2004**, *126*, 4626-4640.
[82] L. Pal, G. Basu, P. Chakrabarti, *Proteins* **2002**, *48*, 571-579.
[83] R. S. Vieira-Pires, J. H. Morais-Cabral, *J. Gen. Physiol.* **2010**, *136*, 585-592.
[84] P. Enkhbayar, K. Hikichi, M. Osaki, R. H. Kretsinger, N. Matsushima, *Proteins* **2006**, *64*, 691-699.
[85] M. Fodje, S. Al-Karadaghi, *Protein Eng.* **2002**, *15*, 353-358.
[86] R. P. Riek, R. M. Graham, *J. Struct. Biol.* **2011**, *173*, 153-160.
[87] B. W. Low, R. Baybutt, *J. Am. Chem. Soc.* **1952**, *74*, 5806-5807.
[88] L. Pauling, R. B. Corey, *Proc.Natl.Acad.Sci.USA* **1951**, *37*, 251-256.
[89] C. L. Nesloney, J. W. Kelly, *Biorg. Med. Chem.* **1996**, *4*, 739-766.
[90] J. S. Nowick, *Acc. Chem. Res.* **2008**, *41*, 1319-1330.
[91] C. K. Smith, L. Regan, *Acc. Chem. Res.* **1997**, *30*, 153-161.
[92] P. Y. Chou, G. D. Fasman, *J. Mol. Biol.* **1977**, *115*, 135-175.
[93] Y.-T. Lai, N. P. King, T. O. Yeates, *Trends in cell biology* **2012**, *22*, 653-661.
[94] V. R. Agashe, M. Shastry, J. B. Udgaonkar, *Nature* **1995**, *377*, 754-757.
[95] S. Marqusee, R. T. Sauer, *Protein Sci.* **1994**, *3*, 2217-2225.
[96] W. J. Wedemeyer, E. Welker, M. Narayan, H. A. Scheraga, *Biochemistry* **2000**, *39*, 4207-4216.
[97] J. D. Sipe, A. S. Cohen, *J. Struct. Biol.* **2000**, *130*, 88-98.
[98] B. J. Pieters, M. B. Van Eldijk, R. J. Nolte, J. Mecinović, *Chem. Soc. Rev.* **2016**, *45*, 24-39.
[99] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, *Nature* **2021**, *596*, 583-589.
[100] J. Ikebe, M. Suzuki, A. Komori, K. Kobayashi, T. Kameda, *Sci. Rep.* **2021**, *11*, 19004.
[101] Z. Xu, Y.-K. Cen, S.-P. Zou, Y.-P. Xue, Y.-G. Zheng, *Crit. Rev. Biotechnol.* **2020**, *40*, 83-98.
[102] P. Kumar, N. G. Paterson, J. Clayden, D. N. Woolfson, *Nature* **2022**, *607*, 387-392.
[103] A. T. Hilditch, A. Romanyuk, S. J. Cross, R. Obexer, J. J. McManus, D. N. Woolfson, *Nat. Chem.* **2024**, *16*, 89-97.
[104] P.-S. Huang, S. E. Boyken, D. Baker, *Nature* **2016**, *537*, 320-327.
[105] C. Norn, B. I. Wicky, D. Juergens, S. Liu, D. Kim, D. Tischer, B. Koepnick, I. Anishchenko, F. Players, D. Baker, *Proc.Natl.Acad.Sci.USA* **2021**, *118*, e2017228118.
[106] F. Lapenta, J. Aupič, Ž. Strmšek, R. Jerala, *Chem. Soc. Rev.* **2018**, *47*, 3530-3542.
[107] S. Neubacher, J. M. Saya, A. Amore, T. N. Grossmann, *J. Org. Chem.* **2019**, *85*, 1476-1483.

[108] M. Wendt, R. Bellavita, A. Gerber, N. L. Efrém, T. van Ramshorst, N. M. Pearce, P. R. Davey, I. Everard, M. Vazquez-Chantada, E. Chiarparin, *Angew. Chem. Int. Ed.* **2021**, *60*, 13937-13944.

[109] D. S. Daniels, E. J. Petersson, J. X. Qiu, A. Schepartz, *J. Am. Chem. Soc.* **2007**, *129*, 1532-1533.

[110] P. S. Wang, A. Schepartz, *Chem. Commun.* **2016**, *52*, 7420-7432.

[111] C. M. Lombardo, V. Kumar MV, C. Douat, F. Rosu, J.-L. Mergny, G. F. Salgado, G. Guichard, *J. Am. Chem. Soc.* **2019**, *141*, 2516-2525.

[112] R. J. Simon, R. S. Kania, R. N. Zuckermann, V. D. Huebner, D. A. Jewell, S. Banville, S. Ng, L. Wang, S. Rosenberg, C. K. Marlowe, *Proc.Natl.Acad.Sci.USA* **1992**, *89*, 9367-9371.

[113] R. N. Zuckermann, *Peptide Science* **2011**, *96*, 545-555.

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

[115] A. M. Kendhale, L. Poniman, Z. Dong, K. Laxmi-Reddy, B. Kauffmann, Y. Ferrand, I. Huc, *J. Org. Chem.* **2011**, *76*, 195-200.

[116] Y.-X. Xu, G.-T. Wang, X. Zhao, X.-K. Jiang, Z.-T. Li, *J. Org. Chem.* **2009**, *74*, 7267-7273.

[117] H.-P. Yi, C. Li, J.-L. Hou, X.-K. Jiang, Z.-T. Li, *Tetrahedron* **2005**, *61*, 7974-7980.

[118] V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat, I. Huc, *Chem. Eur. J.* **2023**, *29*, e202300898.

[119] J. Wang, B. Wicher, A. Méndez-Ardoy, X. Li, G. Pecastaings, T. Buffeteau, D. M. Bassani, V. Maurizot, I. Huc, *Angew. Chem. Int. Ed.* **2021**, *133*, 18609-18614.

[120] 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.

[121] V. Corvaglia, J. Wu, D. Deepak, M. Loos, I. Huc, *Chem. Eur. J.* **2024**, *30*, e202303650.

[122] N. Chandramouli, Y. Ferrand, G. Lautrette, B. Kauffmann, C. D. Mackereth, M. Laguerre, D. Dubreuil, I. Huc, *Nat. Chem.* **2015**, *7*, 334-341.

[123] P. Mateus, N. Chandramouli, C. D. Mackereth, B. Kauffmann, Y. Ferrand, I. Huc, *Angew. Chem. Int. Ed.* **2020**, *59*, 5797-5805.

[124] V. Koehler, G. Bruschera, E. Merlet, P. K. Mandal, E. Morvan, F. Rosu, C. Douat, L. Fischer, I. Huc, Y. Ferrand, *Angew. Chem. Int. Ed.* **2023**, *62*, e202311639.

[125] C. J. Craig, J. L. Goodman, A. Schepartz, *ChemBioChem* **2011**, *12*, 1035-1038.

[126] G. W. Collie, K. Pulka-Ziach, C. M. Lombardo, J. Fremaux, F. Rosu, M. Decossas, L. Mauran, O. Lambert, V. Gabelica, C. D. Mackereth, *Nat. Chem.* **2015**, *7*, 871-878.

[127] V. Maurizot, C. Dolain, Y. Leydet, J.-M. Léger, P. Guionneau, I. Huc, *J. Am. Chem. Soc.* **2004**, *126*, 10049-10052.

[128] F. S. Menke, D. Mazzier, B. Wicher, L. Allmendinger, B. Kauffmann, V. Maurizot, I. Huc, *Org. Biomol. Chem.* **2023**, *21*, 1275-1283.

[129] F. S. Menke, B. Wicher, V. Maurizot, I. Huc, *Angew. Chem. Int. Ed.* **2023**, *62*, e202217325.

[130] E. A. Naudin, K. I. Albanese, A. J. Smith, B. Mylemans, E. G. Baker, O. D. Weiner, D. M. Andrews, N. Tigue, N. J. Savery, D. N. Woolfson, *Chem. Sci.* **2022**, *13*, 11330-11340.

[131] F. S. Menke, B. Wicher, L. Allmendinger, V. Maurizot, I. Huc, *Chem. Sci.* **2023**, *14*, 3742-3751.

4. Abiotic domain swapping structure in organic solvents

The exploration of the abiotic helix-turn-helix structure was successfully advanced. By connecting the C-termini of two hydroxy group-decorated quinoline-based oligoamide helices to the di-amine groups of a symmetrical turn unit T1, a homochiral parallel helix-turn-helix structure stabilized by six intramolecular hydrogen bonds was formed. However, in the absence of the T1 unit, the helices tend to aggregate into tilted dimers or parallel trimers, while the parallel dimer arrangement was not observed, suggesting that the hydrogen bond pattern stabilizing the tertiary fold is not inherently preferred. Inverting the handedness and orientation of one helix will lead to a similar side-chain arrangement. This allowed the introduction of an unsymmetrical amino acid turn unit T2 between two heterochiral helices arranged in a head-to-tail manner leading to the formation of a heterochiral helix-turn-helix tertiary structure stabilized by a similar six hydrogen bond pattern. The synthesis of helix-T2-helix sequences benefits from the solid-phase synthesis strategy. Such a well-defined tertiary structure is a good candidate for modification to promote self-assembly, potentially leading to a true abiotic quaternary structure. With the aid of computational tools, three additional hydroxy groups were placed on one side of the helix-T2-helix tertiary structure to create complementary arrays of hydrogen bond donors and acceptors, which may promote the aggregation of tertiary structures.

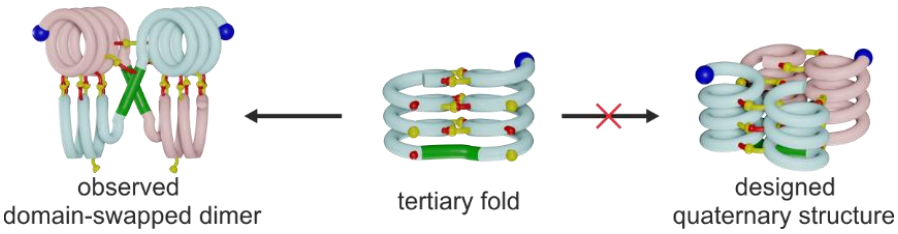


Figure 13. The cartoon representation shows the foldamer initially designed to form a quaternary structure in fact leads to a domain-swapped dimer.

Our findings are summarized in a manuscript which is published by *Angewandte Chemie International Edition*. The single crystal X-ray diffraction and the ^{1}H NMR analysis of the new compounds revealed a novel domain-swapping dimer structure different from the initial design (Figure 13). Domain swapping is commonly seen in proteins, where the subdomains of the protein may bind to each other intramolecularly or may swap with another molecule and bind intermolecularly via the same or similar interaction pattern, leading to multimerization. In this study, intramolecular interactions that stabilize the tertiary fold were not observed; instead, a tilted dimer interface was formed intermolecularly. The diacylhydrazide group of the turn unit T2 underwent a significant conformational change, facilitating the formation of the domain-swapping dimer. The ^{1}H NMR and CD spectra indicated that by slowly adding DMSO as a hydrogen bond disruptor into the solution of the domain-

swapping dimer, the dimer dissociated to form a monomeric tertiary fold, and the tertiary fold was further disrupted at high proportion DMSO.

Contributions: The project was planned by I. Huc. Synthetic monomer precursors have been provided by D. Gill. Monomer synthesis have been performed by me, F. Menke and D. Gill. Foldamer synthesis was performed by me. C. Glas, L. Allmendinger and me carried out NMR measurements. The measurement of CD spectra was performed by me. The LC-MS analysis was performed by C. Douat. Crystallographic studies and structure refinement were performed by B. Wicher and B. Kauffmann. Me, I. Huc, L. Allmendinger, C. Douat, V. Maurizot and B. Wicher contributed to experiment design and interpretation. The research was supervised by I. Huc. The manuscript was written by me and I. Huc. Me, I. Huc, V. Maurizot, L. Allmendinger, B. Wicher and C. Douat proofread and improved the manuscript. This work was supported the China Scholarship Council (CSC, predoctoral fellowship to S. W.).

4.1. Publication

Domain swapping in abiotic foldamers

Authors: Shuhe Wang, Barbara Wicher, Victor Maurizot, and Ivan Huc*
Published: *Angew. Chem. Int. Ed.* **2024**, 63, e202405091. (doi.org/10.1002/anie. 202405091)

How to cite: *Angew. Chem. Int. Ed.* **2024**, *63*, e202405091
doi.org/10.1002/anie.202405091

Foldamers Hot Paper

Domain Swapping in Abiotic Foldamers

*Shuhe Wang, Barbara Wicher, Céline Douat, Victor Maurizot, and Ivan Huc**

Abstract: Foldamer sequences that adopt tertiary helix-turn-helix folds mediated by helix-helix hydrogen bonding in organic solvents have been previously reported. In an attempt to create genuine abiotic quaternary structures, i.e. assemblies of tertiary structures, new sequences were prepared that possess additional hydrogen bond donors at positions that may promote an association between the tertiary folds. However, a solid state structure and extensive solution state investigations by Nuclear Magnetic Resonance (NMR) and Circular Dichroism (CD) show that, instead of forming a quaternary structure, the tertiary folds assemble into stable domain-swapped dimer motifs. Domain swapping entails a complete reorganization of the arrays of hydrogen bonds and changes in relative helix orientation and handedness that can all be rationalized.

Driven by curiosity and creativity, chemists have demonstrated that some folding patterns found in peptides and proteins can be reproduced in completely abiotic oligomers, that is, synthetic molecules chemically remote from aliphatic peptides. Abiotic oligomers also have features of their own, such as folding in organic medium, that can be put to an advantage. For example, aromatic helices with a variable diameter may completely surround a guest,[1–4] while open-ended helices selectively channel molecules or ions through bilayer membranes.[5–9] Early examples of abiotic foldamers comprised isolated secondary structural elements, i.e. helices, linear strands or turns.[10–16] Several such elements were then connected,[17–19] eventually generating the first true abiotic tertiary structures under the form of helix-turn-helix folds.[20–22] These model systems reproduce effects well-known in proteins. For example, abiotic tertiary structures may cooperatively stabilize secondary folds, e.g. helices too flexible to fold well on their own.[23] Conversely, tertiary structures may frustrate secondary folding and accommodate a certain level of strain.[24] Both cooperativity and frustration may promote dynamic changes beneficial to function, as they do in proteins. Here, we present the rare observation in abiotic structures of yet another phenomenon common in proteins, namely domain swapping.[4]

Protein tertiary structures consisting of at least two domains, that is, two subparts whose folds are inherently stable, may undergo multimerization through domain swapping.[25,26] While the domains bind to each other intramolecularly in a protein monomer, they may bind intermolecularly through the same or similar interactions in the domain-swapped multimer. Entropy expectedly favors monomers. Domain swapping is therefore the reflection of additional driving forces that favor multimers. Our discovery of domain swapping in abiotic foldamers stemmed from an attempt to produce the first genuine abiotic quaternary structures, i.e. aggregates of tertiary folds.

In organic solvents, oligoamide sequences of aromatic δ-amino acids of general formula Q and P (Figure 1a) have been shown to fold into stable helices. The amide carbonyl oxygen atoms diverge from these helices and provide hydrogen bond acceptors at their surface (Figure S1).[27–29] Adding complementary hydrogen bond donors such as the 4-hydroxy groups of monomers X and Y (Figure 1a), at precise positions at the surface of the helices promotes tight hydrogen bond-mediated helix-helix associations in chlorinated solvents.[20,22,30] Furthermore, turn units such as monomer T2 (Figure 1a) can be placed between two helix segments within a sequence to induce the formation of a helix-turn-helix tertiary motif.[20–22] Thus, X-, Y- and T2-containing sequence **1** (Figure 1b) folds as depicted in Figure 2a, forming a helix-turn-helix structure stabilized by six inter-helix hydrogen bonds (Figures S2, S3). Three features of the structure of **1** should be highlighted here. First, the axes of the two helices form strictly parallel lines (with head-to-tail orientation). Second, the effect of T2 is such that forming the six hydrogen-bonds requires an inversion of helix handedness (Figure 2a).[21] If the N-terminal helical segment of **1** (before T2) is right-handed (*P*), the C-terminal helix (after T2) must be left-handed (*M*), and reciprocally. Third, and important in the context of this study, the helix-turn-helix structure of **1** is frustrated in that it forces the helices to deviate from their natural curvature to establish the hydrogen bonds.[24] Upon removing T2, different inter-helix hydrogen bonds take place. Specifically, because the six hydrogen-bond donors and the

[*] S. Wang, Dr. C. Douat, Prof. Dr. I. Huc
Department of Pharmacy,
Ludwig-Maximilians-Universität in Munich,
Butenandtstr. 5–13, 81377 München (Germany)
E-mail: ivan.huc@cup.lmu.de

Dr. B. Wicher
Department of Chemical Technology of Drugs,
Poznan University of Medical Sciences,
3 Rokietnicka St., 60-806 Poznan (Poland)

Dr. V. Maurizot
CBMN (UMR 5248), Univ. Bordeaux, CNRS, Bordeaux INP
2, Rue Robert Escarpit, 33600 Pessac (France)

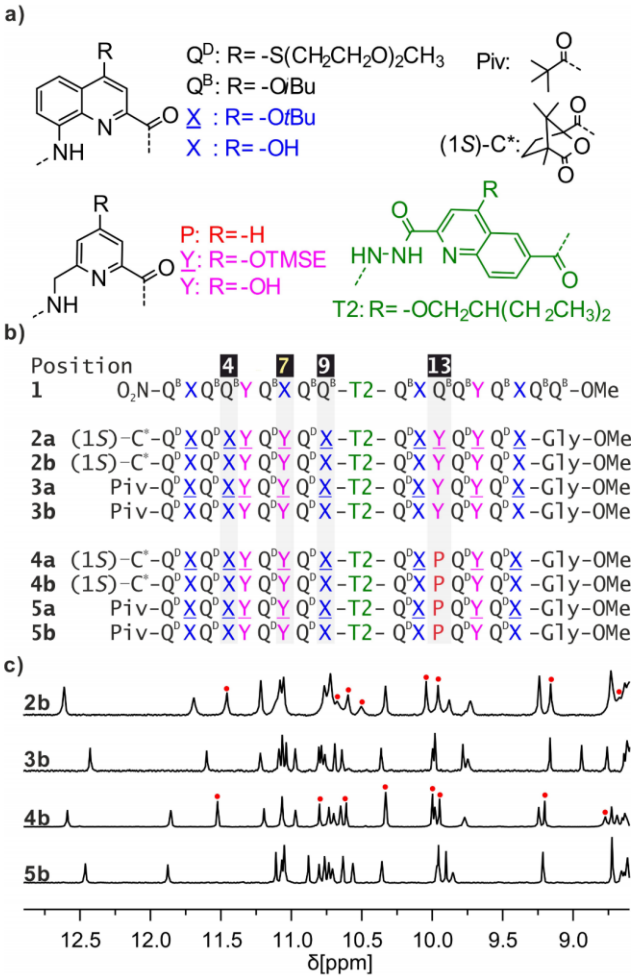


Figure 1. a) Structures of Q^B, Q^D, X, P, Y and T2 amino acid monomers as well as N-terminal Piv and (1*S*)-C* groups. X and Y are the protected precursors of X and Y, respectively. TMSE = 2-trimethylsilylethyl. The side chains of T2, Q^B and Q^D promote solubility in organic solvents. b) Oligoamide foldamer sequences. Gly stands for glycine. In **1**, the nitro group at the N-terminus replaces the NH group. c) Extracts of ^{1}H NMR spectra (500 MHz, 25 °C) of **2b**, **3b**, **4b** and **5b** in $CDCl_3$ showing the amide NH and hydrogen-bonded OH proton resonances. Signals assigned to OH protons are marked with red dots.

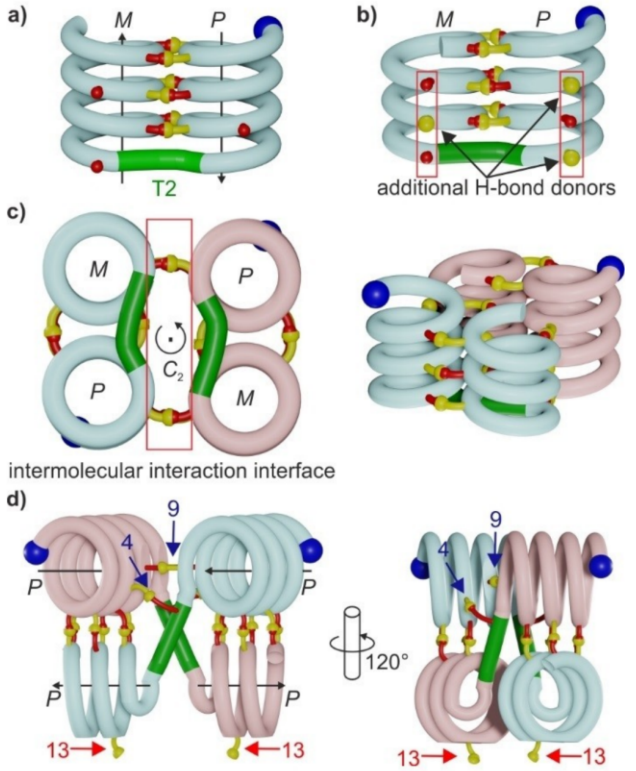


Figure 2. a) Schematic representation of a helix-turn-helix tertiary structure with a reversal of helix handedness. The N-terminus is marked with a blue ball. The T2 linker is shown in green. Red and yellow balls indicate hydrogen bond acceptors (carbonyl oxygen atoms) and donors (hydroxy protons), respectively. b) Modification of the structure in a) so as to create additional complementary arrays of hydrogen bonds (donor-acceptor-donor and acceptor-donor-acceptor shown in red boxes). c) Schematic representations of a hypothetical C_2-symmetrical dimer of the structure in b) preserving the helix-turn-helix fold. d) Representations of the observed domain-swapped dimer of the structure in b), this time without helix handedness reversal. Blue arrows point to unexpected hydrogen bonds. Red arrows point to hydroxy groups not involved in hydrogen bonds. In c) and d), two identical molecules are shown in pink and light blue.

six acceptors of the structure of **1** are arranged in a somewhat distorted hexagon, helices can also associate when their axes have been tilted by 120° in one or the other direction, forming so-called clockwise or counter-clockwise tilted dimers (Figure S4) that are more stable than parallel arrangements.[20,24] In summary, the helix-turn-helix structure of **1** promoted by the T2 turn forms despite the existence of better hydrogen bonding arrangements.

New sequences **2b** and **3b** (Figure 1b) are analogues of **1** with additional hydroxy groups at residues 4, 9, and 13 (Figure 1b). These hydroxy groups were placed on one face of the helix-turn-helix structure so as to create complementary arrays of hydrogen-bond donors and acceptors that, we hoped, would promote the formation of dimers of tertiary structures, i.e. genuine abiotic quaternary structures (Figures 2b,c, S3). An X7Y mutation was also introduced to avoid a possible steric clash in the dimers caused by the benzenic rings of X7. Similarly, Y was introduced in position 13 instead of X to avoid a steric clash (Figure S5a). Sequence **3b** is achiral, whereas an N-terminal (1*S*)-camphanyl group quantitatively biases the N-terminal helix of **2b** to *P* helicity.[31] Sequences **2a** and **3a**, the hydroxy group-protected precursors of **2b** and **3b**, were synthesized on solid phase using reported methods (see Supporting Information).[32,33] Introducing a glycine residue at the C-terminus and the use of 4-(hydroxymethyl)benzoyl-aminomethyl polystyrene (HMBA-AM) resin allowed to directly generate **2a** and **3a** as methyl esters upon sodium methoxide-mediated resin cleavage. The ^{1}H NMR spectra of protected sequences **2a** and **3a** in $CDCl_3$ showed two sets of signals corresponding to the *PM* and *PP* diastereomeric conformers of **2a** and the *PM*/*MP* and *PP*/*MM* conformers of **3a**, respectively, as expected for T2-containing precursors (Figure S6).[21] After side chain deprotection, the ^{1}H NMR

spectra of **2b** and **3b** in $CDCl_3$ showed only one set of sharp signals (Figure 1c), suggesting the formation of discrete species. Heating to 55 °C, changing the solvent to CD_2Cl_2, or diluting to 20 μM did not result in significant changes (Figures S7–S9). Eight (out of nine) O*H* resonances were identified as exchangeable protons showing no correlations to a ^{15}N atom in ^{1}H,^{15}N Heteronuclear Single Quantum Coherence (HSQC) NMR spectra (Figures 1c, S10, S11). Chemical shift values above 8.5 ppm indicated the involvement of these OH groups in hydrogen bonds. This pattern would be hard to fulfill intramolecularly, and it suggests the formation of a symmetrical aggregate too stable to dissociate at NMR concentrations in $CDCl_3$.[34]

A solid state structure[35] of achiral **3b** was obtained that differs from the initially-designed quaternary motif. Instead, it revealed a novel domain-swapped dimer (DSD) with an overall (non-crystallographic, Table S1) C_2-symmetry, in which the individual helix-turn-helix folds have been disrupted (Figures 2d, 3). Unlike **1**, the DSD consists of helical segments all having the same handedness. Furthermore, it contains exclusively intermolecular hydrogen bonds, entailing a massive structural reorganization with respect to the structure of **1** (Figures S12, S13). These hydrogen bonds can be divided into three areas (Table S2), two of which are characteristic of tilted helix-helix interactions (Figure 3c). As mentioned above, this mode of interaction had been previously identified.[20,22] Of note, the structure consists of so-called "counter-clockwise" tilts (Figure S4), whose existence had been hypothesized but had not yet been characterized. Only a "clockwise" tilt had been evidenced.[20,22] A third group of four hydrogen bonds was unexpected. It involves X4 and X9 hydroxy group donors, which are not present in **1** (blue arrows in Figures 2d and 3). The X4 hydroxy protons hydrogen bond to carbonyl groups of T2 units (Figure S2), which themselves undergo a conformation change. Indeed, instead of the flat, conjugated diacylhydrazine conformation that prevails in helix-turn-helix motifs, the diacylhydrazine N–N bonds of T2 units are twisted by around 90° in the DSD.[36] It is this twist at the center of each molecule that orients the helices to enable the formation of the tilted interfaces. Altogether, it appears that the DSD is favored by: (*i*) the more stable tilted helix-helix interface as opposed to the frustrated helix-helix interface with parallel helix axes; (*ii*) intermolecular hydrogen bonds involving donors not present in **1**; (*iii*) a twist of the diacylhydrazine group; (*iv*) the X7Y and the Q13Y mutations without which the DSD would cause steric clashes (Figure S5b,c). However, the DSD formed despite the fact that two Y13 OH groups are not involved in hydrogen bonds (red arrows in Figures 2d and 3, Figure S14).[37]

The ^{1}H NMR spectra of **2b** and **3b** are consistent with the solid state structure of $(\mathbf{3b})_2$. The number of signals aligns with the symmetry of the DSD.[34] The ^{15}N chemical shifts for the diacylhydrazine group are found at 114 and 116 ppm (Figure S10), considerably shifted from their values for compound **1** (at 126 and 128 ppm),[21] which is in agreement with a change of conformation. The count of O*H* signals in solution (Figure 1c) supports that one O*H* is not hydrogen-bonded (that of Y13 in the DSD). Adding 4 vol %

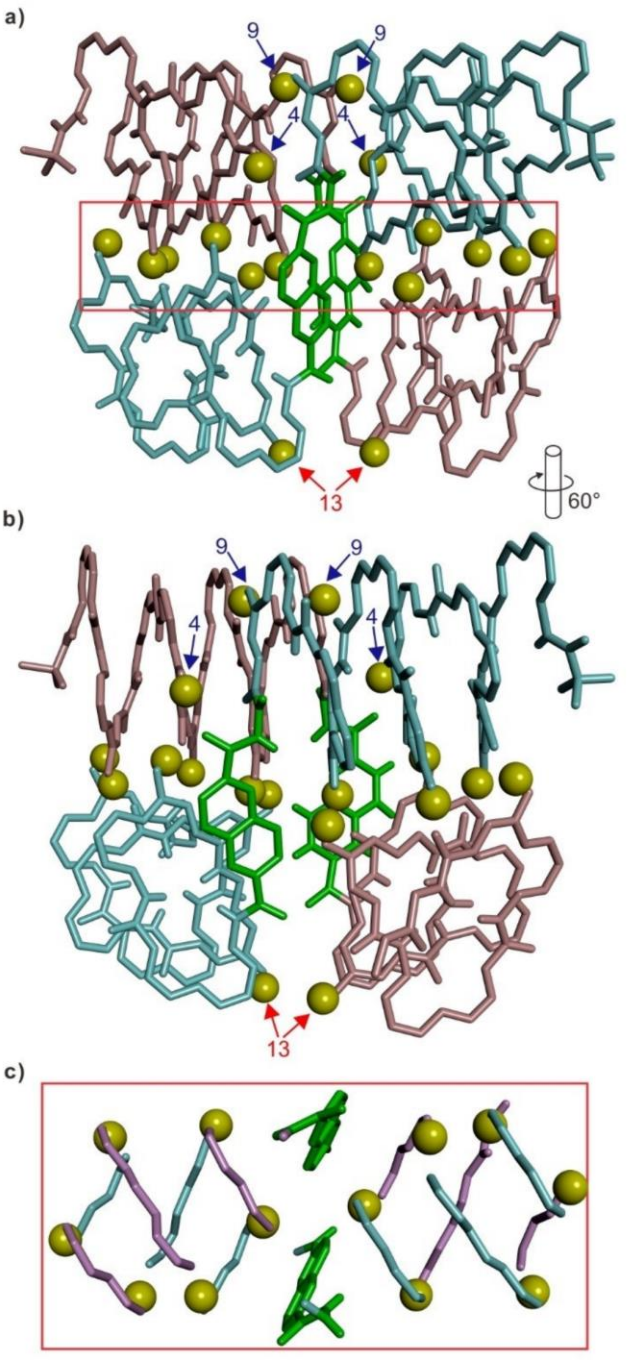


Figure 3. a, b) Simplified views of the solid state structure of $(\mathbf{3b})_2$ as a domain-swapped dimer. Only the all *P* structure is shown. The crystal lattice is centrosymmetrical and also contains the all *M* enantiomer. Individual molecules are shown in light blue and pink tube representation. Hydroxy protons are shown as yellow balls. T2 linkers are shown in green. For clarity, only the outer rim of the helices is shown. The side chains of Q and T2 have been omitted. Blue arrows point to unexpected hydrogen bonds. Red arrows point to hydroxy groups not involved in hydrogen bonds. c) Top view of the intermolecular interface located in the red box in a) showing the two hexagonal arrangement of hydrogen bond donors associated with the tilted helix-helix interaction. Color coding is as in a) and b).

of DMSO-d_6 to a $CDCl_3$ solution of **2b** caused little changes to the ^{1}H NMR spectrum except for the appearance of an additional O*H* signal above 8.5 ppm (Figure S15), consistent with this proton not being hydrogen-bonded within **2b** but shifting downfield upon hydrogen bonding to DMSO. Furthermore, upon varying the concentration of **2b** in DMSO-d_6/$CDCl_3$ (4:96, vol/vol), two sets of signals are observed in varying proportions (Figure 4b). These changes, as well as Diffusion Ordered SpectroscopY (DOSY, Figure 4c), demonstrate that the species that prevails in $CDCl_3$ is an aggregate. Assuming a monomer/dimer equilibrium, we calculated $K_{dim} = 6.3 \cdot 10^{-5}\ L\,mol^{-1}$ in this solvent. The absence of dissociation at NMR concentrations in pure $CDCl_3$ (Figure S7), i.e. a much higher K_{dim}, is consistent with the stability of other hydrogen-bonded dimers we have described.[20,30] Increasing the fraction of DMSO-d_6 above 4 vol% further favored the formation of the monomer (Figures 4b, S16). Above 10 vol%, another set of signals appeared that eventually became major at 20 vol%. This amount of DMSO-d_6 has previously been shown to disrupt the hydrogen bond interface of **1**.[21] A model that would explain the results above is shown in Figure 4a. A small amount of DMSO-d_6 causes the DSD dissociation into a folded monomer having a *P* and an *M* helix. Adding further DMSO-d_6 disrupts the intramolecular helix-helix interface and produces a monomer with two *P* helices.

The effect of DMSO on the CD spectra of **2b** corroborated the NMR data. In pure $CHCl_3$, the CD spectrum of **2b** showed an intense positive band at 400 nm belonging to the quinoline chromophores and indicating the prevalence of *P* helicity (Figure 5a) as expected in the DSD.[31] Adding DMSO caused a drop in CD intensity of about 60%, consistent with the appearance of the *PM* folded monomer. In the *PM* conformation, the CD contribution of the *P* helix and the *M* helix almost entirely cancel each other for the helices have similar numbers of quinoline rings. Further addition of DMSO resulted in the CD band intensity growing again as the *PM* conformer was disrupted to produce some *PP* monomer (Figures 4a, 5a,b). Finally, **4b** and **5b** were produced as analogues of **2b** and **3b**, respectively. These two sequences lack one hydroxy group as the monomer in position 13 is P instead of Y (Figure 1a, b). Nevertheless, **4b** and **5b** behaved like **2b** and **3b** in all respects in solution (Figures 1c, S17–S21). Altogether, these results strongly support that the DSD observed in the solid state structure of $(\mathbf{3b})_2$ corresponds to the species in $CDCl_3$ solutions of **2b**, **3b**, **4b**, and **5b**.

DSD formation reflects the ability of a folded structure to respond to changes. As was recently illustrated, the change may be the presence of a guest molecule.[4] Here, it is

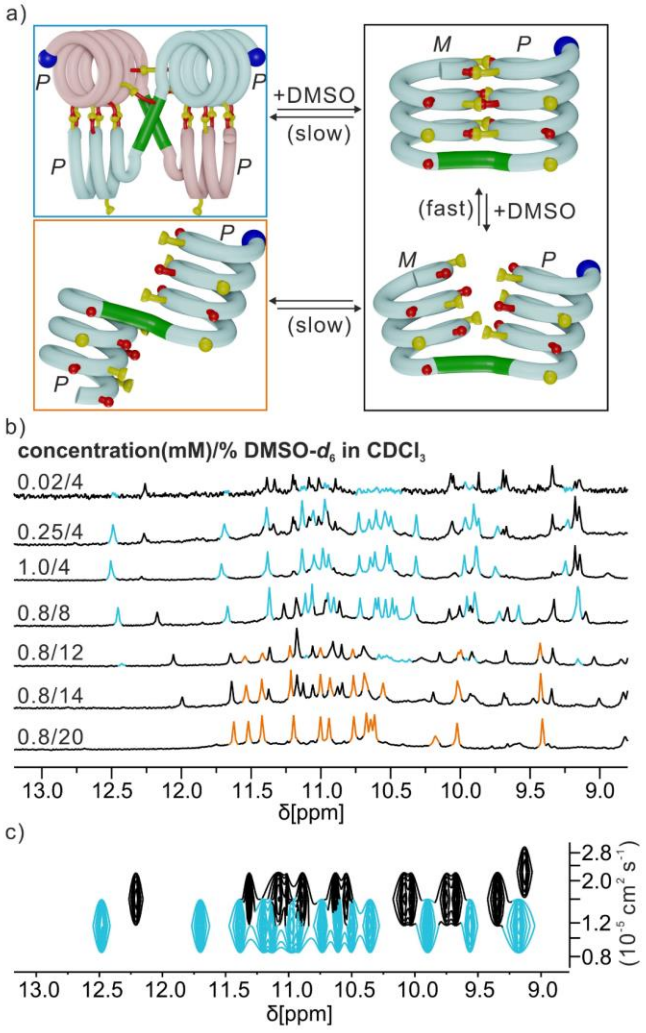


Figure 4. a) Schematic representation of changes induced upon adding DMSO to a solution of DSD. The slow or fast rates are with respect to the NMR time scale. Stoichiometry (a dimer gives two monomers) has been omitted for clarity. b) Excerpts of the ^{1}H NMR spectra (500 MHz, 25 °C) of **2b** in DMSO-d_6/$CDCl_3$ at different concentrations and different vol% of DMSO-d_6. c) 500 MHz DOSY ^{1}H NMR spectrum of **2b** at 25 °C, 1 mM in DMSO-d_6/$CDCl_3$ (10:90 v/v). In b) and c) peaks are colored according to the species they belong to, as coded in the boxes shown in a).

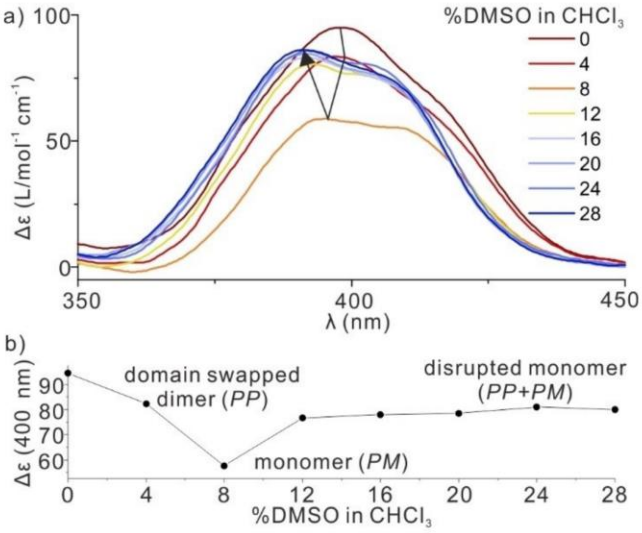


Figure 5. a) CD spectra of **2b**, 0.1 mM in $CHCl_3$ with different proportions of DMSO. b) The Δε values at 400 nm extracted from a) as a function of the vol% of DMSO in $CHCl_3$.

the consequence of a few mutations. Current efforts in our laboratories intend to exploit both the responsiveness and the geometry of DSD's. Meanwhile, progress to genuine quaternary structures has been made and will be reported in due course.

Supporting Information

The authors have cited additional references within the Supporting Information.[38–45]

Acknowledgements

We acknowledge financial support from the China Scholarship Council (CSC, predoctoral fellowship to S.W.). This work has benefited from the facilities and expertise of the Biophysical and Structural Chemistry platform (BPCS) at IECB, CNRS UMS3033, Inserm US001, and Bordeaux University. We thank B. Kauffmann for assistance with X-ray data collection, F. Menke for advice with synthesis, L. Allmendinger for some NMR measurements, and D. Gill for providing some precursors. Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: abiotic foldamer · domain swapping · hydrogen bonding · structure elucidation · tertiary structure

[1] W. Wang, C. Zhang, S. Qi, X. Deng, B. Yang, J. Liu, Z. Dong, *J. Org. Chem.* **2018**, *83*, 1898–1902.
[2] Y. Hua, Y. Liu, C. H. Chen, A. H. Flood, *J. Am. Chem. Soc.* **2013**, *135*, 14401–14412.
[3] N. Chandramouli, Y. Ferrand, G. Lautrette, B. Kauffmann, C. D. Mackereth, M. Laguerre, D. Dubreuil, I. Huc, *Nat. Chem.* **2015**, *7*, 334–341.
[4] G. Song, S. Lee, K.-S. Jeong, *Nat. Commun.* **2024**, *15*, 1501.
[5] J. Shen, Z. Li, H. Oh, H. Behera, H. Joshi, M. Kumar, A. Aksimentiev, H. Zeng, *Angew. Chem. Int. Ed.* **2023**, *62*, e202305623; *Angew. Chem.* **2023**, *135*, e202305623.
[6] J. Shen, R. Ye, Z. Liu, H. Zeng, *Angew. Chem. Int. Ed.* **2022**, *61*, e202200259; *Angew. Chem.* **2022**, *134*, e202200259.
[7] L. Zhang, C. Zhang, X. Dong, Z. Dong, *Angew. Chem. Int. Ed.* **2023**, *62*, e202214194; *Angew. Chem.* **2023**, *135*, e202214194.
[8] Y. Shen, F. Fei, Y. Zhong, C. Fan, J. Sun, J. Hu, B. Gong, D. M. Czajkowsky, Z. Shao, *ACS Cent. Sci.* **2021**, *7*, 2092–2098.
[9] S. Qi, C. Zhang, H. Yu, J. Zhang, T. Yan, Z. Lin, B. Yang, Z. Dong, *J. Am. Chem. Soc.* **2021**, *143*, 3284–3288.
[10] I. Huc, *Eur. J. Org. Chem.* **2004**, *2004*, 7–7.
[11] D.-W. Zhang, X. Zhao, J.-L. Hou, Z.-T. Li, *Chem. Rev.* **2012**, *112*, 5271–5316.
[12] Y. Hamuro, S. J. Geib, A. D. Hamilton, *Angew. Chem. Int. Ed.* **1994**, *33*, 446–448; *Angw. Chem.* **1994**, *106*. 465–467.
[13] J. Zhu, R. D. Parra, H. Zeng, E. Skrzypczak-Jankun, X. C. Zeng, B. Gong, *J. Am. Chem. Soc.* **2000**, *122*, 4219–4220.
[14] B. Gong, H. Zeng, J. Zhu, L. Yua, Y. Han, S. Cheng, M. Furukawa, R. D. Parra, A. Y. Kovalevsky, J. L. Mills, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 11583–11588.
[15] Z.-Q. Wu, X.-K. Jiang, S.-Z. Zhu, Z.-T. Li, *Org. Lett.* **2004**, *6*, 229–232.
[16] J. T. Ernst, J. Becerril, H. S. Park, H. Yin, A. D. Hamilton, *Angew. Chem. Int. Ed.* **2003**, *42*, 535–539; *Angew. Chem.* **2003**, *115*, 553–557.
[17] V. Maurizot, C. Dolain, Y. Leydet, J.-M. Léger, P. Guionneau, I. Huc, *J. Am. Chem. Soc.* **2004**, *126*, 10049–10052.
[18] N. Delsuc, S. Massip, J.-M. Léger, B. Kauffmann, I. Huc, *J. Am. Chem. Soc.* **2011**, *133*, 3165–3172.
[19] C. Dolain, J.-M. Léger, N. Delsuc, H. Gornitzka, I. Huc, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16146–16151.
[20] S. De, B. Chi, T. Granier, T. Qi, V. Maurizot, I. Huc, *Nat. Chem.* **2018**, *10*, 51–57.
[21] D. Mazzier, S. De, B. Wicher, V. Maurizot, I. Huc, *Angew. Chem. Int. Ed.* **2020**, *59*, 1606–1610; *Angew. Chem.* **2020**, *132*, 1623–1627.
[22] F. S. Menke, B. Wicher, L. Allmendinger, V. Maurizot, I. Huc, *Chem. Sci.* **2023**, *14*, 3742–3751.
[23] D. Mazzier, S. De, B. Wicher, V. Maurizot, I. Huc, *Chem. Sci.* **2019**, *10*, 6984–6991.
[24] F. S. Menke, D. Mazzier, B. Wicher, L. Allmendinger, B. Kauffmann, V. Maurizot, I. Huc, *Org. Biomol. Chem.* **2023**, *21*, 1275–1283.
[25] Y. Huang, H. Cao, Z. Liu, *Proteins* **2012**, *80*, 1610–1619.
[26] N. Nandwani, P. Surana, H. Negi, N. M. Mascarenhas, J. B. Udgaonkar, R. Das, S. Gosavi, *Nat. Commun.* **2019**, *10*, 452.
[27] D. Sanchez-Garcia, B. Kauffmann, T. Kawanami, H. Ihara, M. Takafuji, M.-H. Delville, I. Huc, *J. Am. Chem. Soc.* **2009**, *131*, 8642–8648.
[28] C. Dolain, A. Grélard, M. Laguerre, H. Jiang, V. Maurizot, I. Huc, *Chem. Eur. J.* **2005**, *11*, 6135–6144.
[29] H. Jiang, J.-M. Léger, I. Huc, *J. Am. Chem. Soc.* **2003**, *125*, 3448–3449.
[30] F. S. Menke, B. Wicher, V. Maurizot, I. Huc, *Angew. Chem. Int. Ed.* **2023**, *62*, e202217325; *Angew. Chem.* **2023**, *135*, e202217325.
[31] A. M. Kendhale, L. Poniman, Z. Dong, K. Laxmi-Reddy, B. Kauffmann, Y. Ferrand, I. Huc, *J. Org. Chem.* **2011**, *76*, 195–200.
[32] B. Baptiste, C. Douat-Casassus, K. Laxmi-Reddy, F. Godde, I. Huc, *J. Org. Chem.* **2010**, *75*, 7175–7185.
[33] V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat, I. Huc, *Chem. Eur. J.* **2023**, *29*, e202300898.
[34] Earlier systems showed that exchange between different aggregates or between aggregate and monomer is slow on the NMR times scale. If an aggregate was not symmetrical, different resonances would be seen for the different subunits. See Ref. [20].
[35] Deposition Number 2337052 (for **3b**) contain(s) the supplementary crystallographic data for this paper. These data are provided free of charge by the joint Cambridge Crystallographic Data Centre and Fachinformationszentrum Karlsruhe Access Structures service.
[36] Such a twisted diacylhydrazine has been observed before. See Ref. [3].
[37] The preference for the tilted hydrogen-bonded helix-helix interface has already been observed to be strong enough to compensate for some orphan OH groups. See Ref. [20].
[38] Rigaku-Oxford-Diffraction, *CrysAlisPro Software System, Version 171.42.79a*, **2022**, Rigaku Corporation: Yarton, England.
[39] G. M. Sheldrick, *Acta Crystallogr.* **2015**, *A71*, 3–8.
[40] G. M. Sheldrick, *Acta Crystallogr.* **2015**, *C71*, 3–8.
[41] O. V. Dolomanov, L. J. Bourhis, R. J. Gildea, J. A. K. Howard, H. Puschmann, *J. Appl. Crystallogr.* **2009**, *42*, 339–341.

[42] V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat, I. Huc, *Chem. Eur. J.* **2023**, *29*, e202300898.
[43] M. Vallade, P. Sai Reddy, L. Fischer, I. Huc, *Eur. J. Org. Chem.* **2018**, *2018*, 5489–5498.
[44] O. Al Musaimi, A. Basso, B. G. de la Torre, F. Albericio, *ACS Comb. Sci.* **2019**, *21*, 717–721.
[45] J. Hansen, F. Diness, M. Meldal, *Org. Biomol. Chem.* **2016**, *14*, 3238–3245.

Manuscript received: March 14, 2024
Accepted manuscript online: April 25, 2024
Version of record online: June 4, 2024

4.2. Supplementary Information

For:

Domain swapping in abiotic foldamers

Shuhe Wang, Barbara Wicher, Victor Maurizot, and Ivan Huc*

Table of contents

1 List of Abbreviations

CD	circular dichroism
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DCM	dichloromethane
DIPEA	*N*,*N*-diisopropylethylamine
DMF	*N*,*N*-dimethylformamide
DMSO	dimethyl sulfoxide
DOSY	diffusion ordered spectroscopy
HR-ESI	high resolution electrospray ionization
eq.	equivalent
Fmoc	fluorenylmethoxycarbonyl
HBTU	2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HSQC	heteronuclear single quantum correlation
Me	methyl
MeOH	methanol
min	minutes
MS	mass spectrometry
MW	microwave
NMP	*N*-methyl-2-pyrrolidone
NMR	nuclear magnetic resonance
r. t.	room temperature
SPS	solid phase synthesis
***t*Bu**	*tert*-butyl
TFA	trifluoroacetic acid
THF	tetrahydrofuran
UV/Vis	ultraviolet–visible

2 Supplementary figures

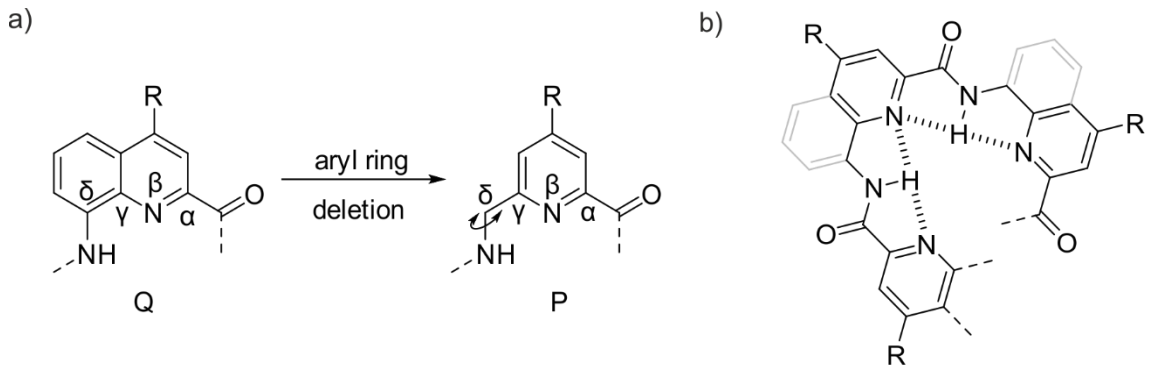


Figure S1. Chemical structure of Q and P units and the folding principle of their oligomers. a) Chemical structure of Q (or X with R = OH) and P (or Y with R = OH). b) Intramolecular H-bonding and helical folding principle of P/Q oligomers. Note that the amide carbonyl groups diverge from the folded structures and thus provide hydrogen bond acceptors.

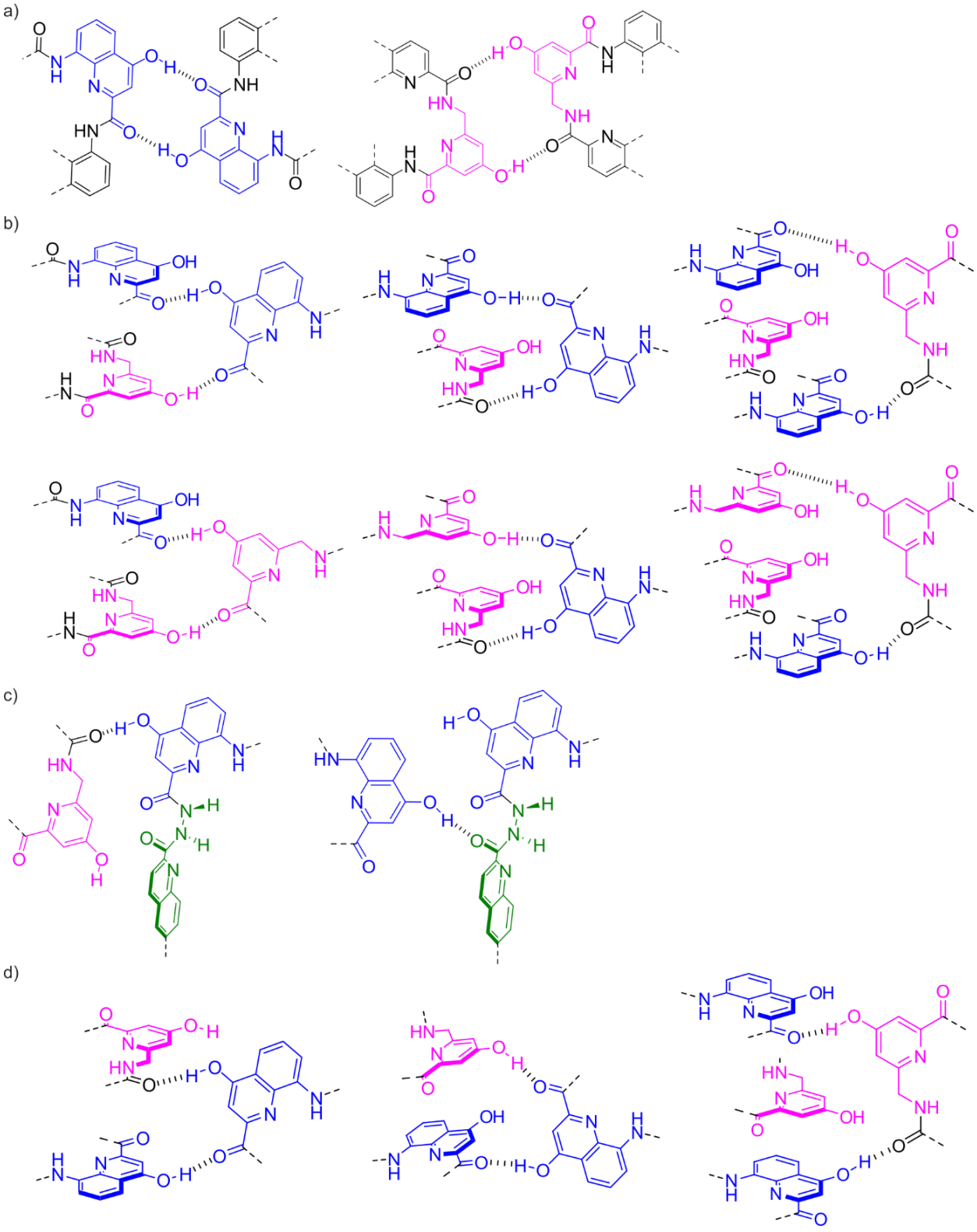


Figure S2. H-bonding patterns involving X and Y units: a) as observed in the helix-turn-helix structure of **1**;[21] b) and c) as observed in the solid state structure of (**3b**)$_2$; d) as observed in clockwise tilted-dimer interfaces.[20]

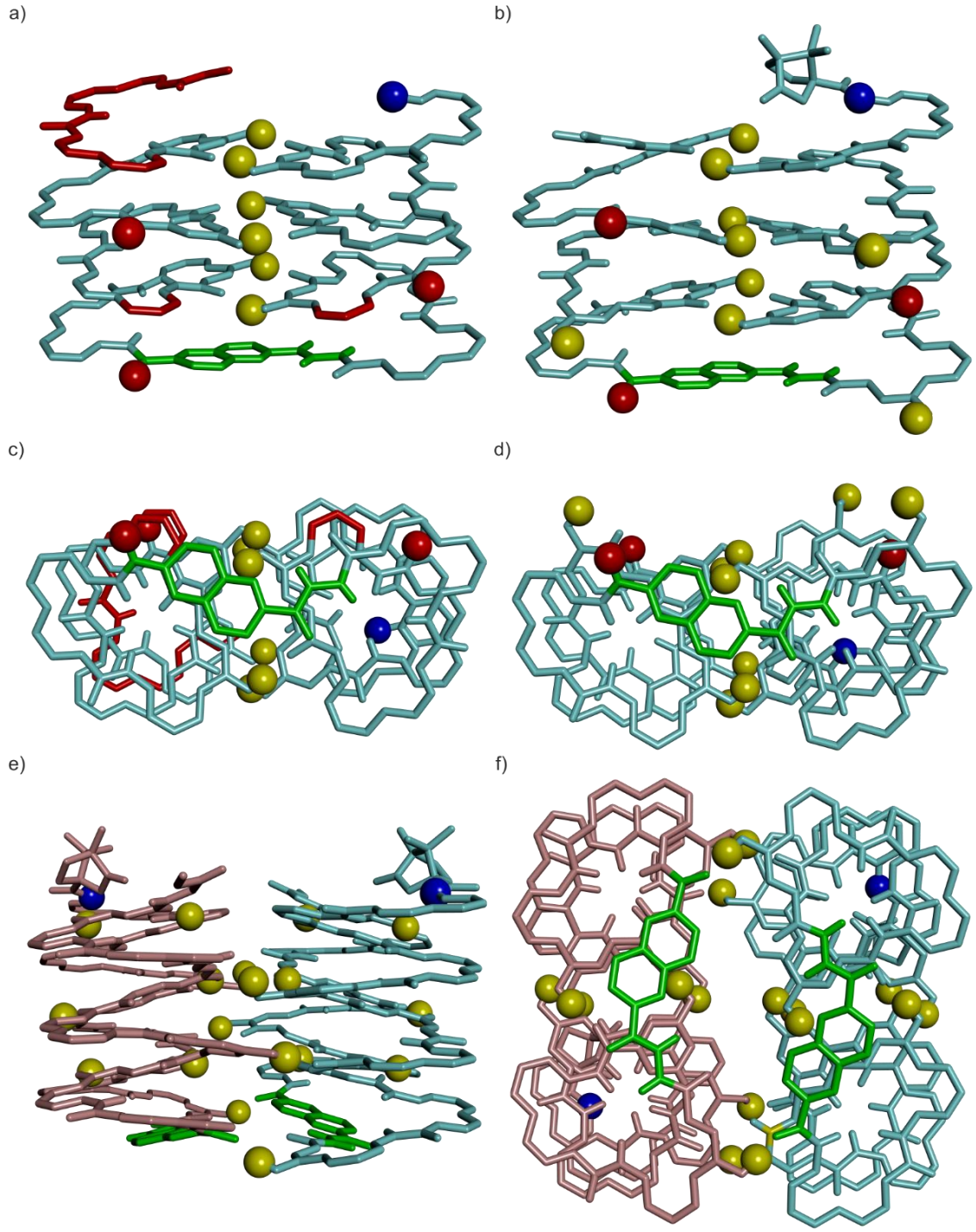


Figure S3. Energy minimized models of dimerized helix-T2-helix structure. The side view (a) and bottom view (c) of the crystal structure of 1.[21] Parts shown as red tubes highlight aromatic rings that were removed to prevent steric hindrance in the design of 2b and 3b. The side view (b) and bottom view (d) of an energy-minimized model of monomeric 2b in the same (hypothetical) conformation as 1. The red balls represent H-bond acceptors intended for intermolecular interactions. The yellow balls represent H-bond donors. The side (e) and bottom (f) views of an energy-minimized model of dimerized 2b maintaining the helix-T2-helix structure of 1. The T2 linkers are colored in green. The N-termini are shown as blue balls. For clarity, only the outer rim of the helical backbone is shown in tube representation. Side chains of Q and T2 are omitted.

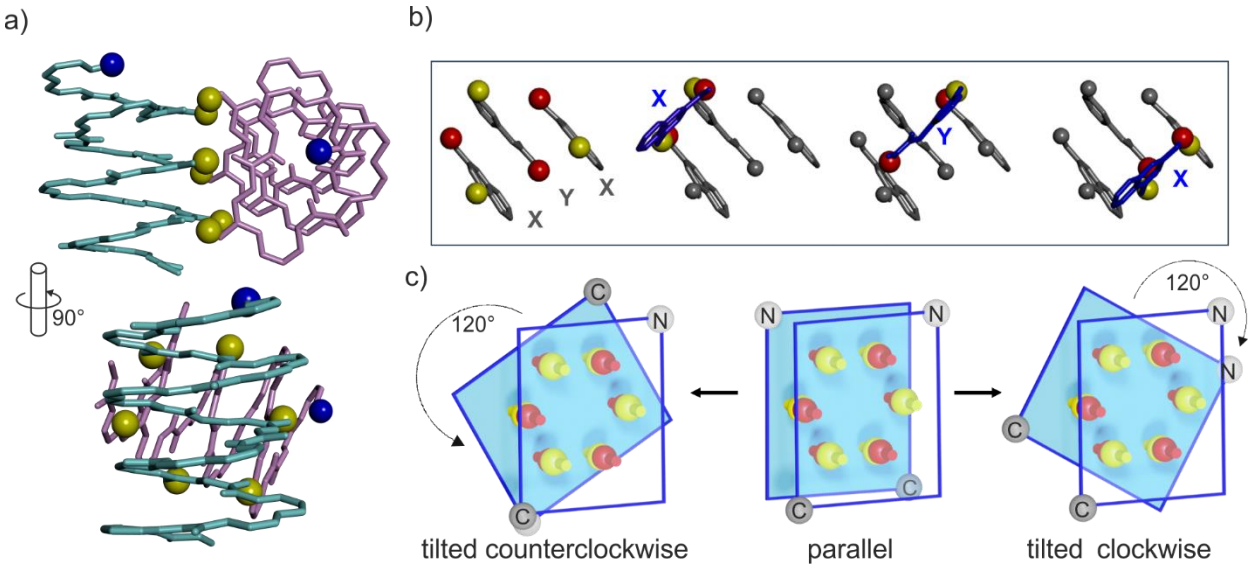


Figure S4. Crystal structure of a clockwise tilted dimer.[20] a) Two views of a clockwise tilted dimer as seen in the crystal. Yellow balls represent the hydrogen bond donors involved in helix-helix interactions. Blue balls highlight the N-terminus of the helix. For clarity, only the outer rim of the helical backbone is shown in tube representation. Side chains are omitted. b) Views of the H-bond patterns of the structure shown in a). c) Schematic representations in stacked view of a hydrogen-bonded parallel head-to-head arrangement and of the related clockwise and counter-clockwise tilted dimers. The top plane is transparent so that one can see the six hydrogen bonds (knobs into cups) and the plane behind.

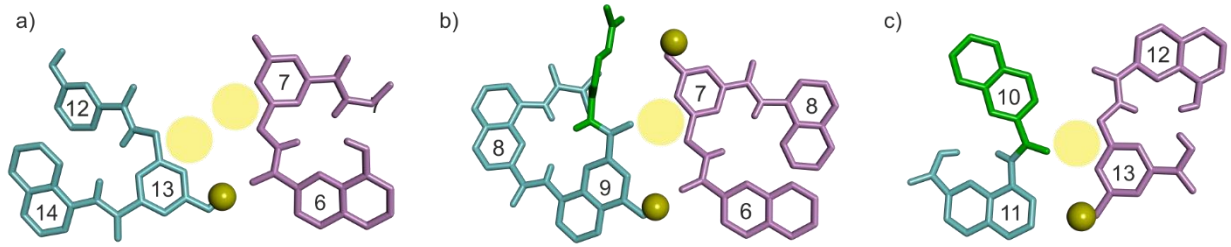


Figure S5. The steric clash avoided by X7Y and Q13Y mutations in the DSD structure. Removing the benzenic rings at positions 7 and 13 was intended to prevent the steric clash that could occur in the dimerized helix-turn-helix structures according to the molecular model in the initial design (a). The crystal structure of DSD shows the X7Y (b) and Q13Y (c) mutations avoid steric clashes.

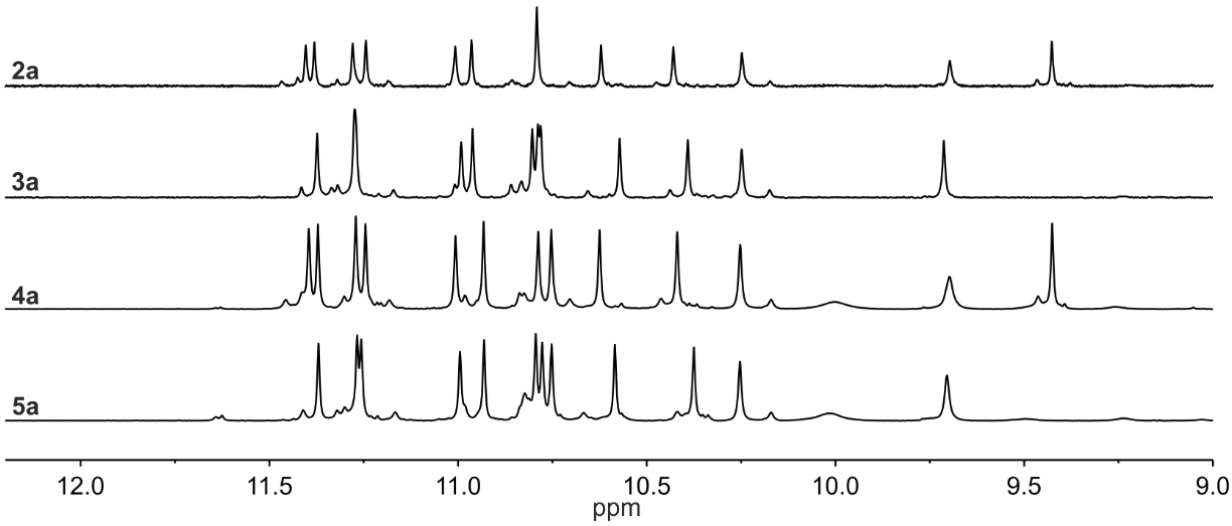


Figure S6. ^{1}H NMR spectra of hydroxy-group-protected precursors. Extracts of the ^{1}H NMR spectra (500 MHz, $CDCl_3$, 25 °C) of 2a, 3a, 4a and 5a. A major and a minor species coexist in all cases corresponding to two different conformers.

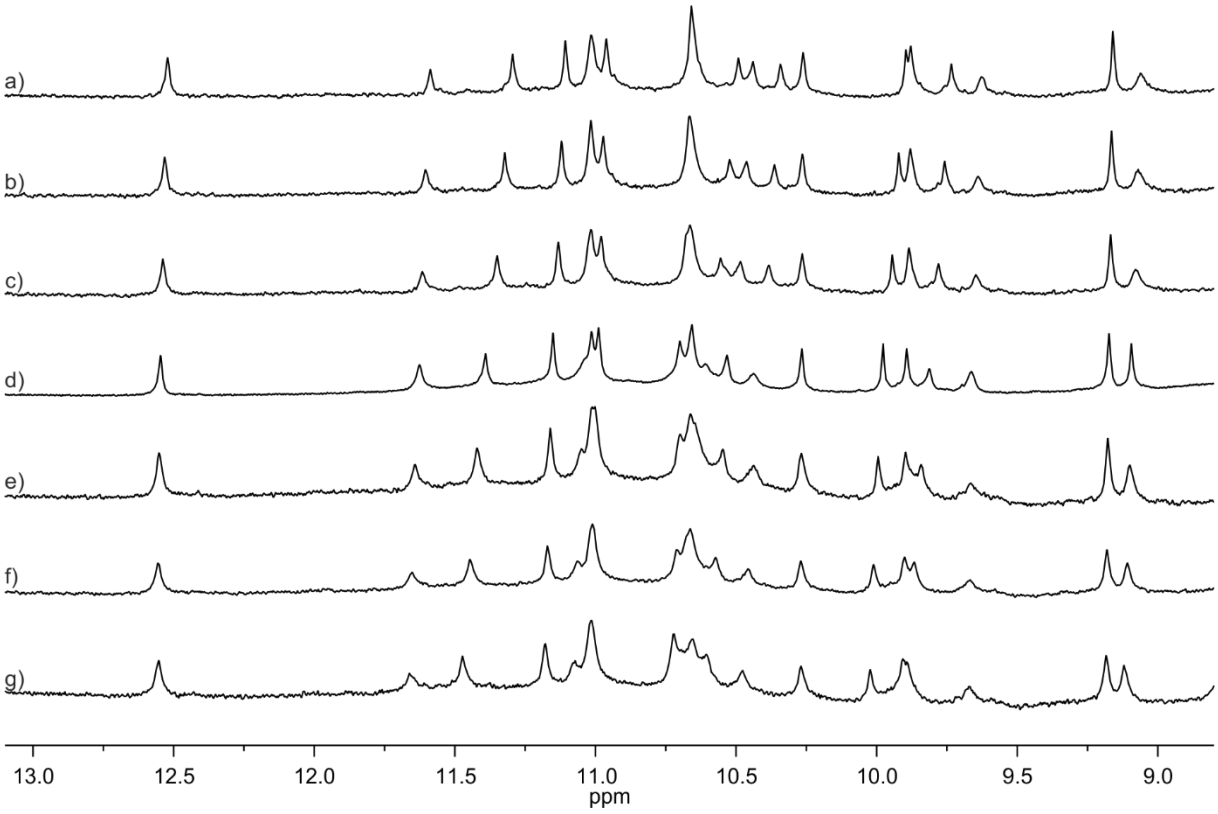


Figure S7. ^{1}H NMR spectra of 2b at different temperatures. Extracts of the ^{1}H NMR spectra (400 MHz, $CDCl_3$) of 2b at 328 K (a), 318 K (b), 308 K (c), 298 K (d), 283 K (e), 273 K (f) and 263 K (g).

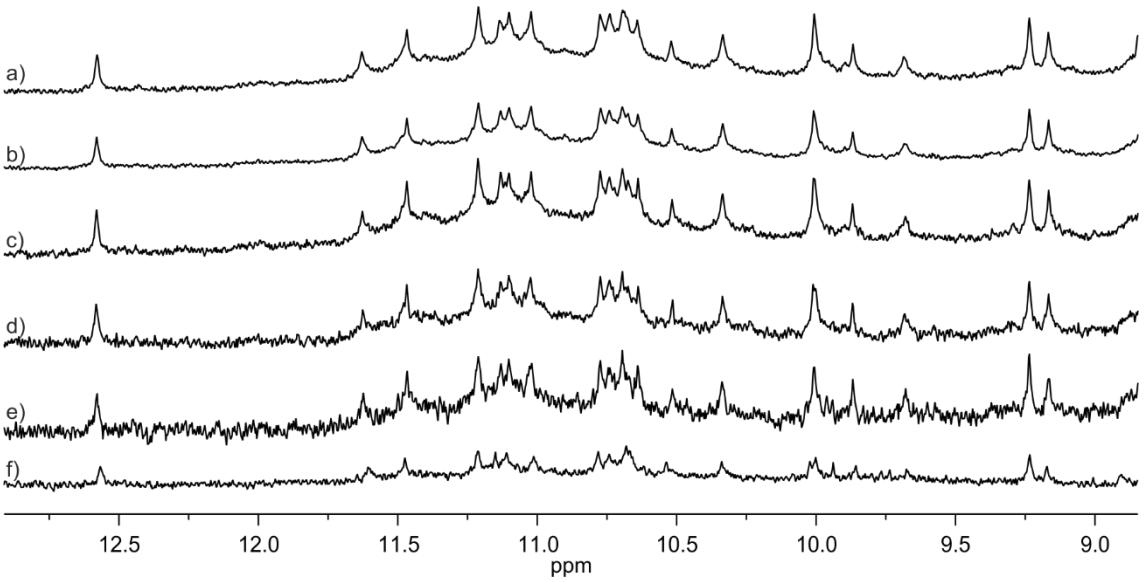


Figure S8. ^{1}H NMR spectra of 2b at different concentrations. Extracts of the ^{1}H NMR spectra (500 MHz, $CDCl_3$, 25 °C) of **2b** at 1.0 mM (a), 0.5 mM (b), 0.25 mM (c), 0.1 mM (d), 0.05 mM (e) and 0.02 mM (f).

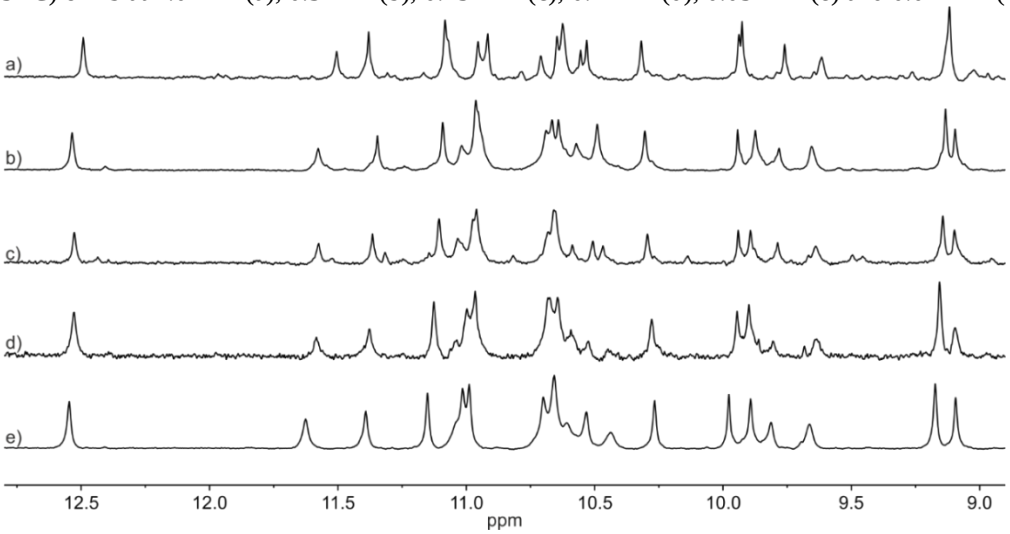


Figure S9. ^{1}H NMR spectra of 2b in varying $CDCl_3$/CD_2Cl_2 mixtures. Extracts of the ^{1}H NMR spectra (500 MHz, 25 °C) of 2b in $CDCl_3$/CD_2Cl_2 mixtures showing the amide and hydroxy proton resonances. The vol% of CD_2Cl_2 are 100 (a), 75 (b), 50 (c), 25 (d), and 0 (e). With other sequences, aggregation behavior or conformation has been observed to change radically between $CDCl_3$ and CD_2Cl_2. [20,30]

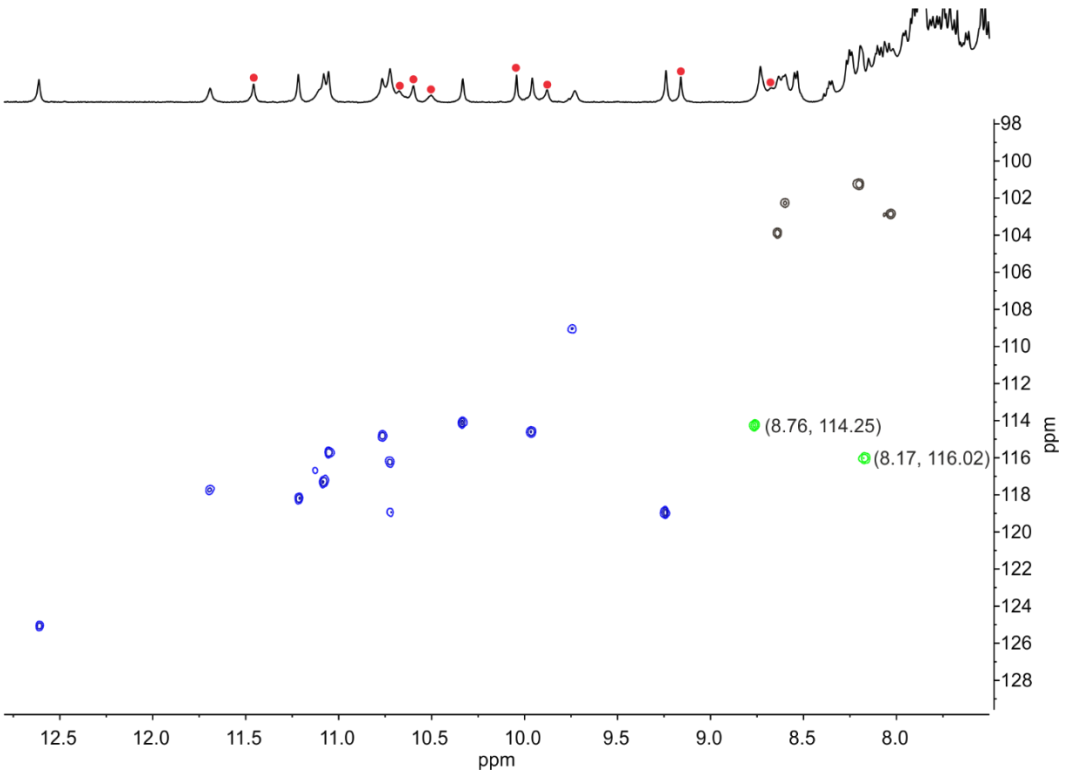


Figure S10. Identification of the signals of hydrogen-bonded OH protons of 2b in $CDCl_3$. Extract of the $^{1}H^{15}N$ HSQC NMR spectrum (500 MHz, $CDCl_3$, 25 °C) of 2b. Only NH resonances correlate, red dots indicate the signals of H-bonded O<u>H</u> protons. The grey correlations are assigned to NH signals of the pyridine-based (Y) monomers. The green correlations are assigned to the NH-NH signals of T2 unit.

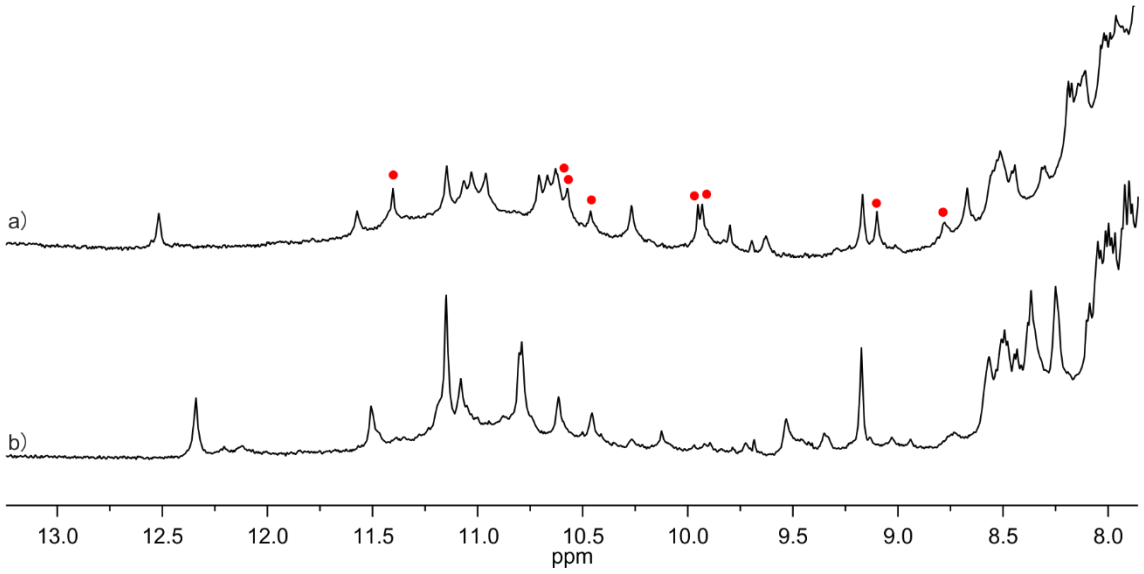


Figure S11 Deuterium exchange experiment of 2b. a) Extract of the ^{1}H NMR of **2b** (500 MHz, $CDCl_3$, 25 °C). The signals assigned to OH protons are marked with a red dot. b) Extract of the ^{1}H NMR of **2b** (500 MHz, 5% $CD_3OD/CDCl_3$, 25 °C).

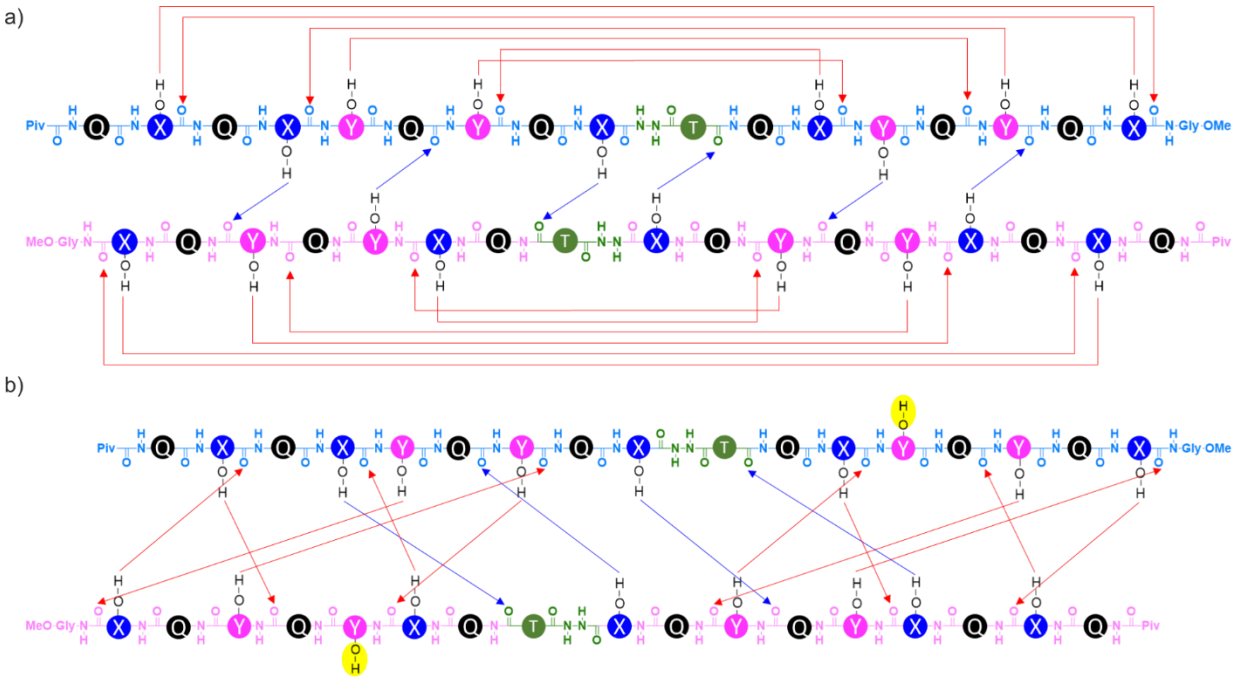


Figure S12. Hydrogen bonding patterns: a) in the initially designed dimer of 3b based on the structure of 1 (see Figure S3 e, f). Red and blue arrows indicate intramolecular and intermolecular H-bonds, respectively. b) in the crystal structure of $(3b)_2$. Red and blue arrows involve the same H-bond donors as in a). Two orphan H-bond donors are highlighted in yellow.

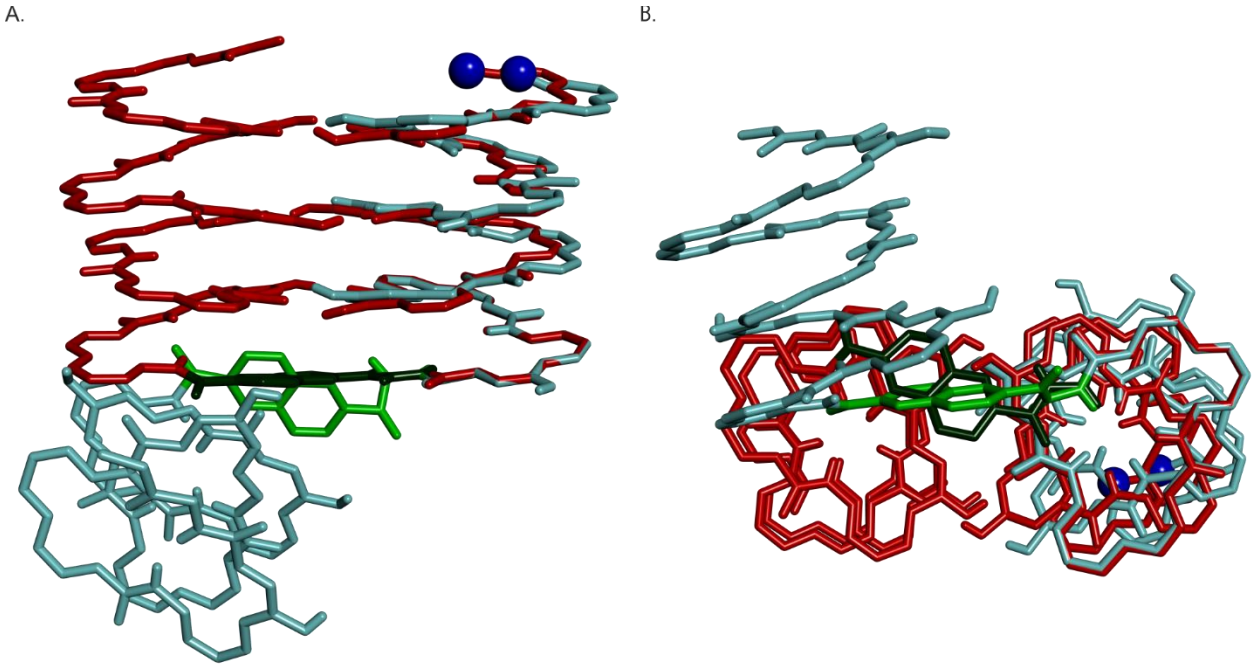


Figure S13. Overlay of the structures of 1 and 3b in the solid state. Compound 1 is shown in red tube representation with a dark green T2 linker. Compound 3b is shown in light blue tube representation with a light green T2 linker. The N-termini are shown as blue balls. Only the outer rim of the helical backbone is shown in tube representation. Side chains are omitted for clarity.

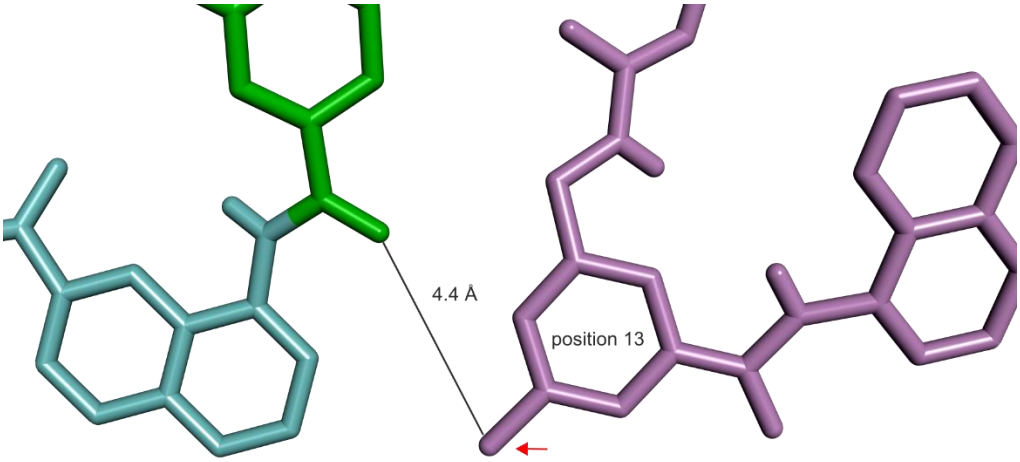


Figure S14. Orphan OH proton in the solid state structure of (3b)$_2$. The distance between the hydroxy group of unit 13 and the nearest carbonyl group is 4.4 Å, which is too far for H-bonding.

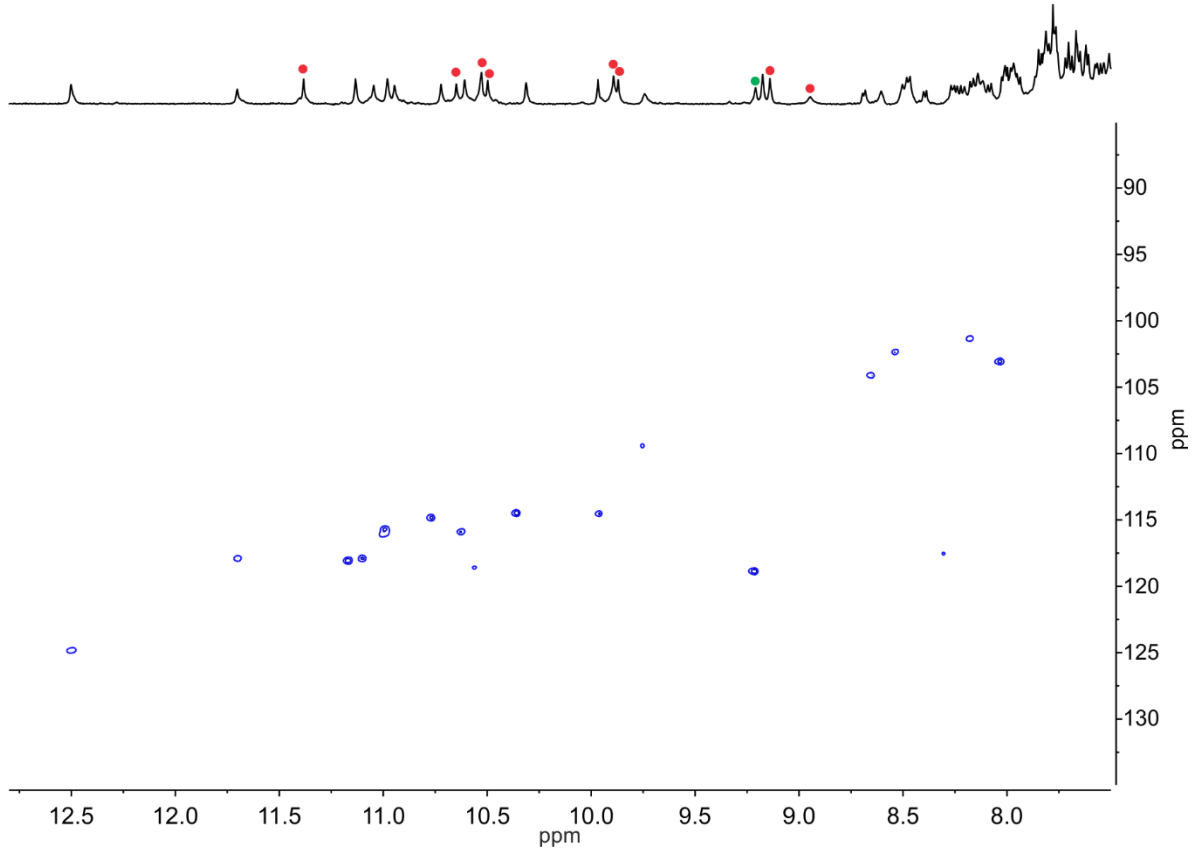


Figure S15. Identification of free OH that is hydrogen-bonded to DMSO. Extract of the $^{1}H^{15}N$ HSQC NMR spectrum (500 MHz, in 4% DMSO-d_6/$CDCl_3$, 25 °C) of 2b. Only NH resonances correlate. Red dots on the ^{1}H trace assign the signals of O<u>H</u> protons hydrogen bonded to carbonyl groups found at similar chemical shift values in the absence of DMSO-d_6. The green dot indicates an O<u>H</u> resonance not visible in this region in the absence of DMSO-d_6 that may be assigned to the O<u>H</u> of Y13.

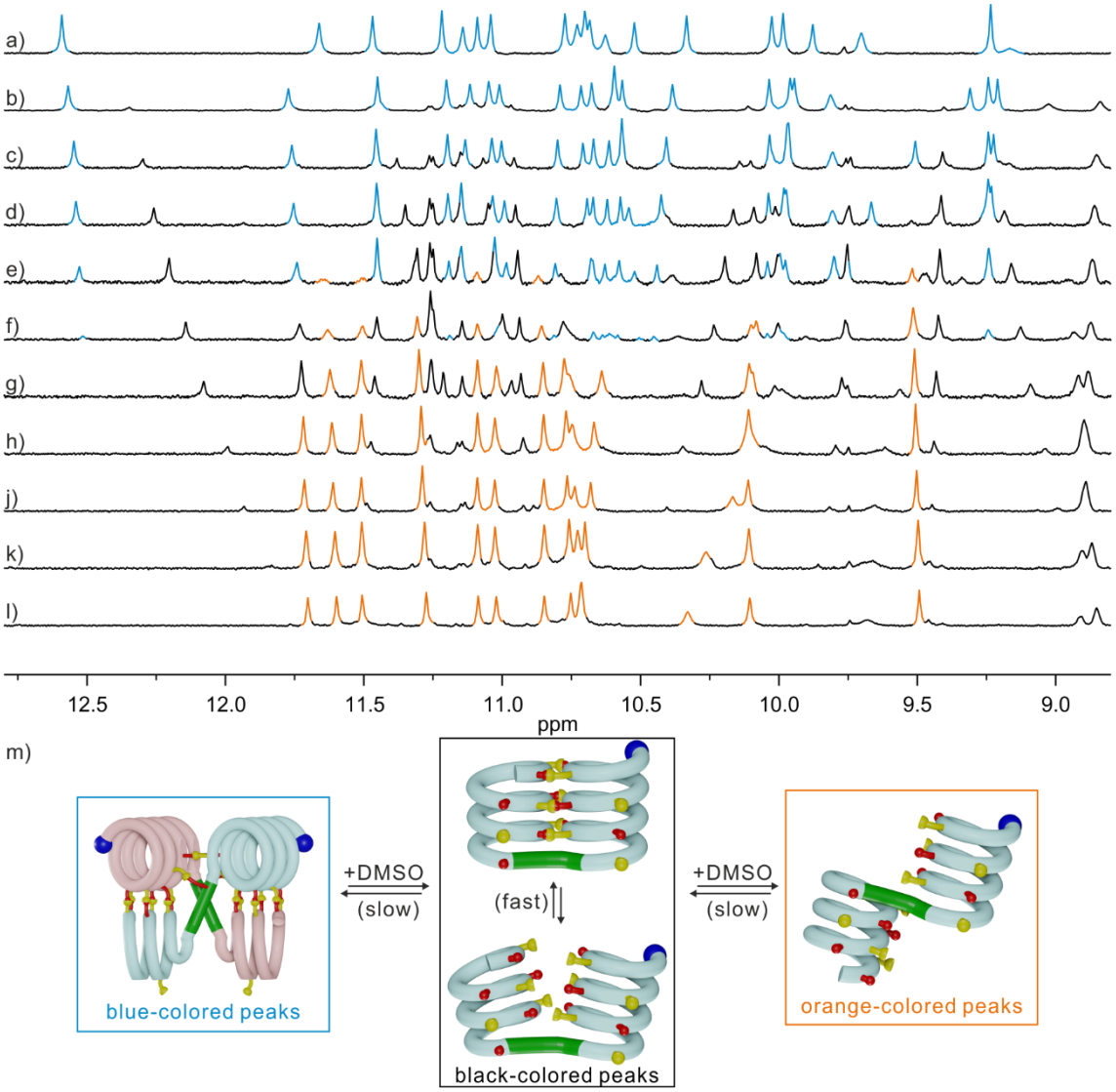


Figure S16. ^{1}H NMR observation of the DMSO-induced dissociation of (2b)$_2$. Extracts of the ^{1}H NMR spectra (500 MHz, 25 °C) showing amide resonances of **2b** in $CDCl_3$/DMSO-d_6 mixture. The vol% of DMSO-d_6 are 0 (a), 4 (b), 6 (c), 8 (d), 10 (e), 12 (f), 14 (g), 16 (h), 18 (j), 20 (k) and 22 (l), respectively. The signals of DSD, helix-T2-helix tertiary structure (*P*-T2-*M*) and disrupted helix-T2-helix tertiary structure (*P*-T2-*P*) are colored in light blue, black and orange, respectively. m) Schematic representation of changes induced by the addition of DMSO to a solution of DSD.

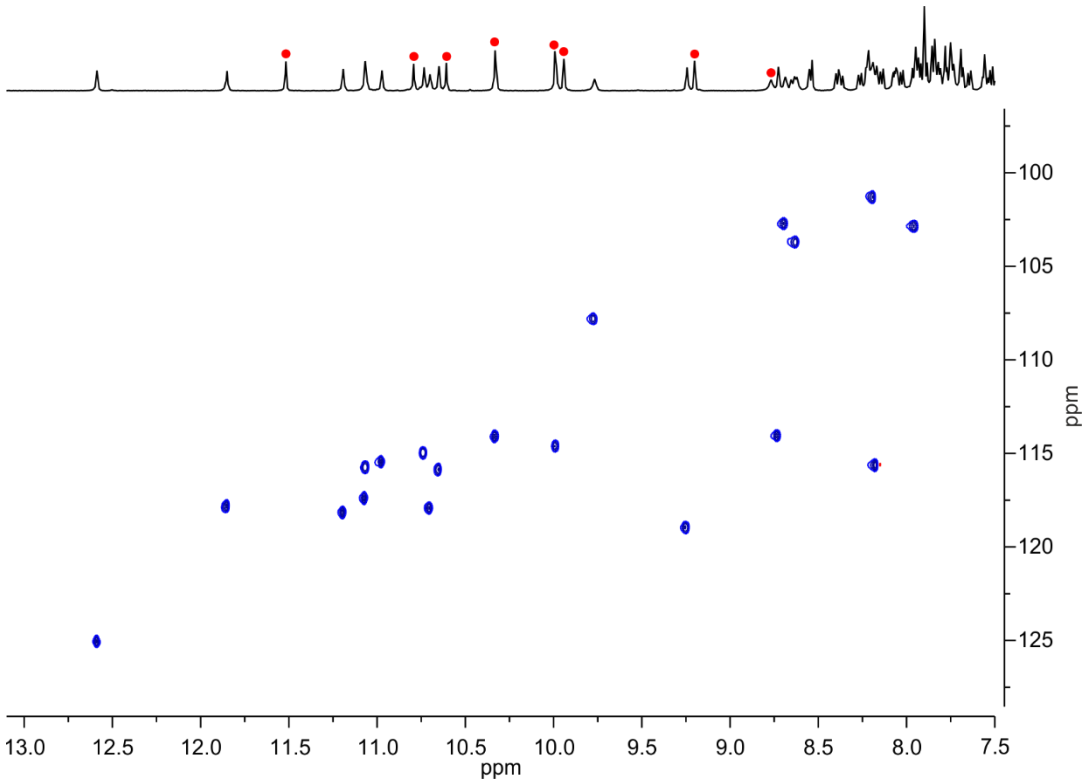


Figure S17. Identification of the signals of hydrogen-bonded OH protons of 4b in $CDCl_3$. Extract of the $^{1}H^{15}N$ HSQC NMR spectrum of 4b (500 MHz, $CDCl_3$, 25 °C). Only NH resonances correlate, red dots indicate the signals of H-bonded O<u>H</u> protons.

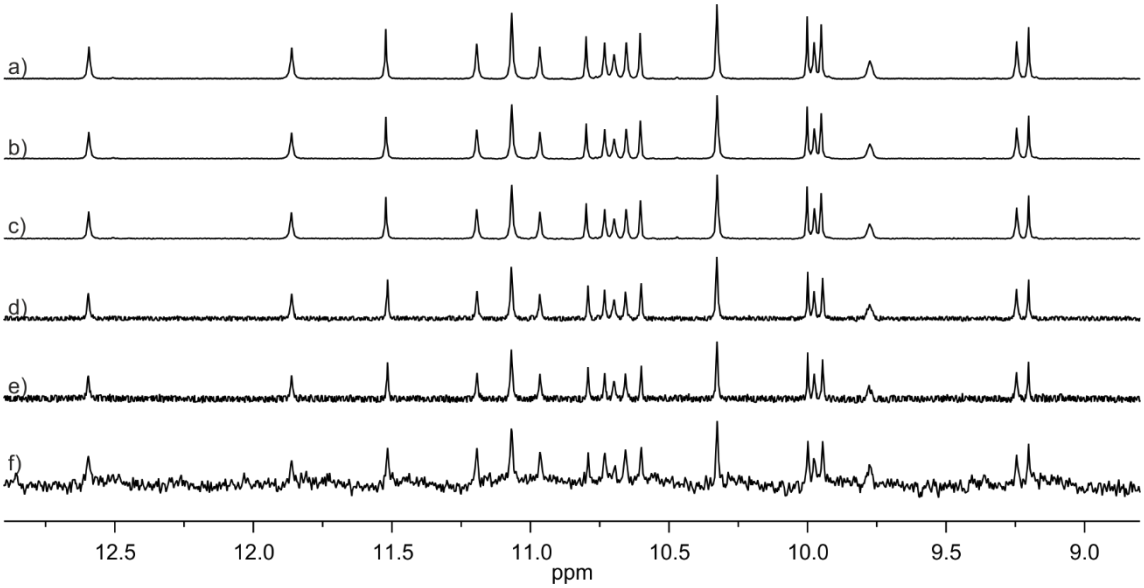


Figure S18. ^{1}H NMR spectra of 4b at different concentrations. a) Extracts of the ^{1}H NMR spectra of **4b** (500 MHz, $CDCl_3$, 25 °C) at 0.8 mM (a), 0.5 mM (b), 0.25 mM (c), 0.1 mM (d), 0.05 mM (e) and 0.02 mM (f).

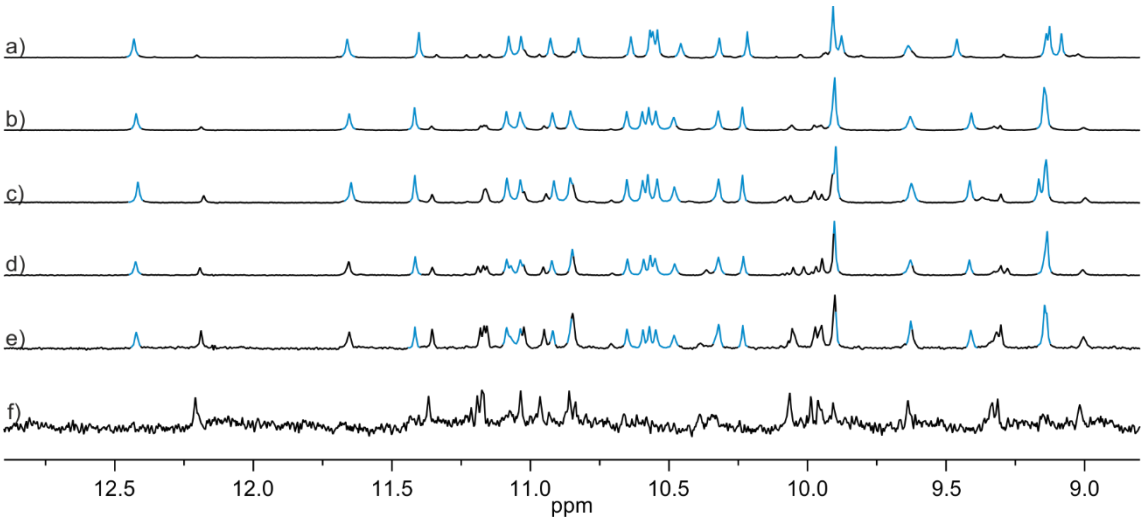


Figure S19. ^{1}H NMR spectra of 4b in 6% DMSO-d_6/$CDCl_3$ at different concentrations. a) Extracts of the ^{1}H NMR spectra of **4b** (500 MHz, 6% DMSO-d_6/$CDCl_3$, 25 °C) at 0.8 mM (a), 0.5 mM (b), 0.25 mM (c), 0.1 mM (d), 0.05 mM (e) and 0.02 mM (f).

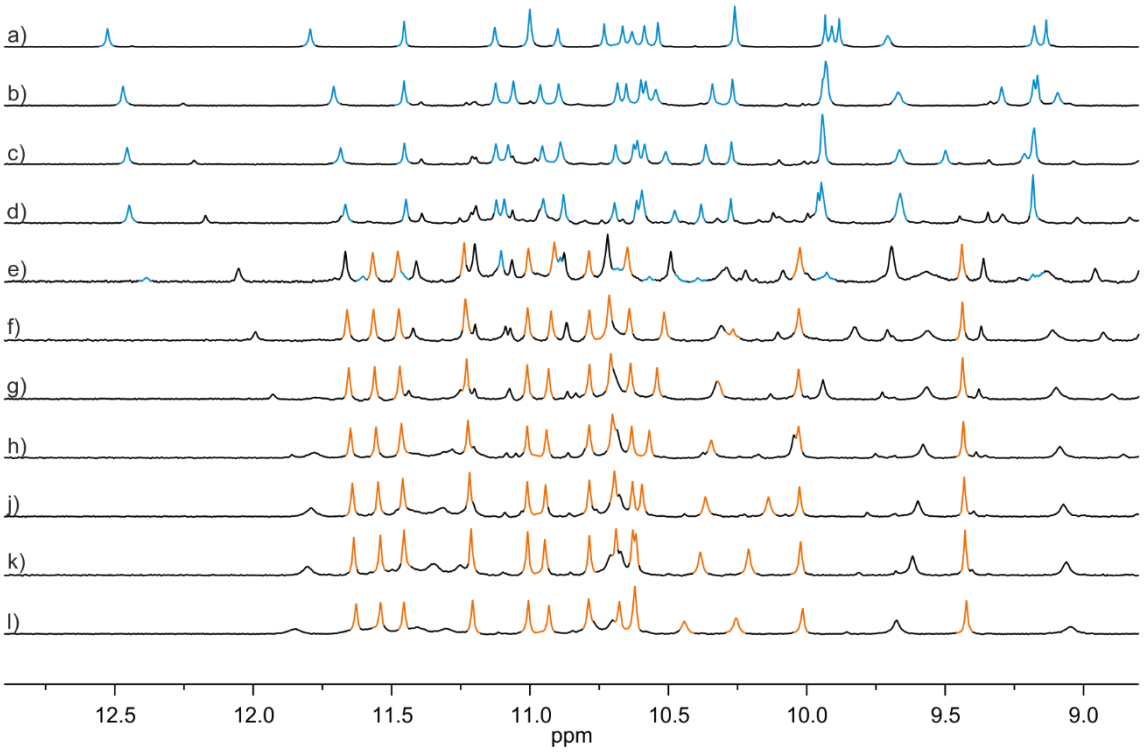


Figure S20. ^{1}H NMR observation of the DMSO-induced dissociation of $(4b)_2$. Extracts of the ^{1}H NMR spectra (500 MHz, 25 °C) showing amide resonances of **4b** in $CDCl_3$/DMSO-d_6 mixture. The vol% of DMSO-d_6 are 0 (a), 4 (b), 6 (c), 8 (d), 10 (e), 12 (f), 14 (g), 16 (h), 18 (j), 20 (k) and 22 (l), respectively. The signals of DSD, helix-T2-helix tertiary structure (*P*-T2-*M*) and disrupted helix-T2-helix tertiary structure (*P*-T2-*P*) are colored in light blue, black and orange, respectively.

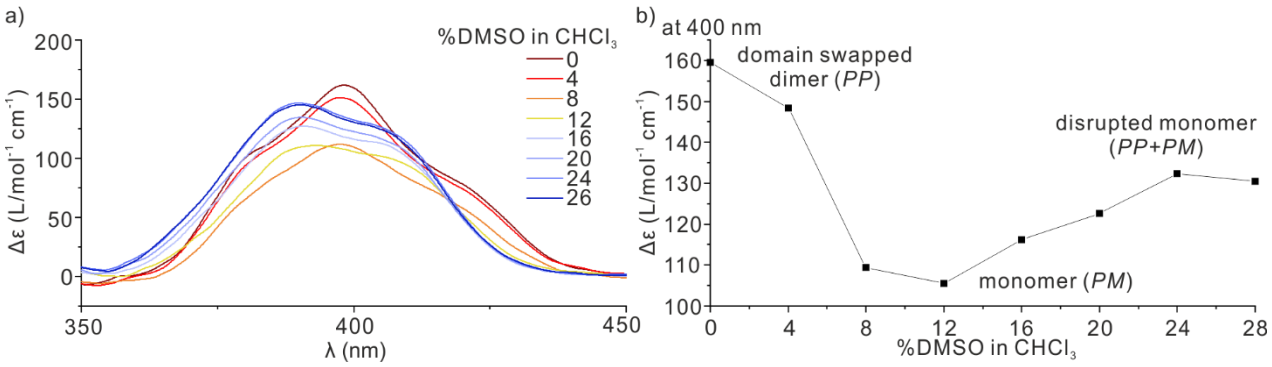


Figure S21. **CD spectra observation of the DMSO-induced dissociation of $(4b)_2$.** a) The CD spectra of **4b**, 0.1 mM in $CHCl_3$ with different proportions of DMSO. b) The Δε values at 400 nm extracted from a) as a function of the vol% of DMSO in $CHCl_3$.

3 Supplementary methods

3.1 LC-MS analyses

LC-MS spectra were recorded on a Bruker microTOF II in positive ionization mode. The instrument was calibrated in positive mode by direct infusion of a calibration solution (Agilent Technologies ESI-L Low Concentration Tuning Mix). The HPLC line was an Ultimate 3000 RP-HPLC system (ThermoFisher Scientific) equipped with a Aeris™ Widepore C4 column (2.1 x 150 mm, 3.6 µm) at a flow rate of 0.25 mL/min. 0.1 % formic acid and 0.025% TFA was added to the aqueous mobile phase (solvent A) and to acetonitrile (solvent B). The gradient is: 0-5 min, 50% to 100% solvent B; 5-14 min, 100% solvent B at 50°C. The column eluent was monitored by UV detection at 214, 254, and 300 nm with a diode array detector. The sample was prepared by adding 10 µL of a solution of the sample in DCM (0.1 mg/mL) to 1 mL acetonitrile containing 0.05-0.1% formic acid.

3.2 Molecular modeling

Models were simulated by using Maestro version 11.5 (Schrödinger Inc.). Energy minimized structures were obtained using MacroModel energy minimization with the following parameters: force field: MMFFs; solvent: none; electrostatic treatment: constant dielectric; dielectric constant: 1.0; charges from: force field; cutoff: normal; Van der Waals: 7.0; electrostatic: 12.0; H-bond: 4.0; mini method: TNCG; maximum iterations: 2500; converge on gradient; convergence threshold: 0.05; constraints: distances. As a starting point, the coordinates of the crystal structure of **1** (CCDC entry # 1955168) was used. Some hydroxy groups were inserted and some aromatic rings were removed. The modified sequence was first energy-minimized. In a second round, two identical modified helix-turn-helix structures were placed in a plausible arrangement, and distance constraints between plausible hydrogen-bonding partners were set to 1.8 on purpose. While setting the constraints, it was important to match the hydroxy group to their correct hydrogen-bonding carbonyl partner. Then all constraints were removed, and energy minimization was repeated.

3.3 Nuclear magnetic resonance spectroscopy

NMR spectra were recorded on different NMR spectrometers: (I) an Avance III HD NMR spectrometer 500 MHz (Bruker BioSpin) with CryoProbe™ Prodigy for ^{1}H NMR, $^{1}H^{15}N$-HSQC, and DOSY spectra of foldamers. (II) a Bruker HD NMR spectrometer 400 MHz (Bruker BioSpin) for variable temperature measurements. Chemical shifts are described in part per million (ppm, δ) relative to the ^{1}H residual signal of the deuterated solvent used – meaning DMSO-d_6 (δ 2.50 ppm), CD_2Cl_2 (δ 5.32 ppm) and $CDCl_3$ (δ 7.26 ppm). For the DMSO-d_6 and $CDCl_3$ solvent mixture, the chemical shifts were calibrated according to DMSO-d_6 (δ 2.50 ppm). For the CD_2Cl_2 and $CDCl_3$ solvent mixture, the chemical shifts were calibrated according to internal standard tetramethylsilane (δ 0.00 ppm).^{1}H NMR splitting patterns with observed first-order coupling are entitled as singlet (s), doublet (d), triplet (t), quartet (q) or multiplet (m). Coupling constants (*J*) are ported in Hz.

The ^{1}H NMR spectra of each sample were measured at different times respectively until no further change was observed within a week. We generally consider that at this point the compound reached equilibrium. When the sample reached equilibrium, re-dissolving the compound solid results in the equilibrated spectrum immediately without going through the equilibration process again. Complete disruption of the hydrogen bonds was achieved by dissolving the sample in polar solvents (such as DMSO, pyridine or MeOH/chloroform mixture), followed by removal of the solvent. When all of the hydrogen bonds were completely disrupted, it has to go through the equilibrium process again to reach the equilibrium. The equilibrium time (the measurement time gap between two different conditions) of ^{1}H NMR spectra at different temperatures and in different proportions of DMSO-d_6/$CDCl_3$ solvent was usually several minutes. Due to similar properties of $CDCl_3$ and CD_2Cl_2, the individual samples in different proportions of $CDCl_3$/CD_2Cl_2 mixture were prepared and the ^{1}H NMR spectra of all of the samples were measured over time whereas no change was observed.

$^{1}H^{15}N$ HSQC spectra were recorded with a phase-sensitive pulse sequence with sensitivity enhancement using trim pulses in inept transfer (hsqcetgpsi2) from the Bruker pulse program library. Data acquisition was performed utilizing non-uniform sampling (NUS; NUS amount: 50% with an automatically created NUSList) yielding 1024 (F2) x 128 (F1) data points in Echo/Antiecho gradient selection mode. The recycling delay was 2.0 s and 64 transients per increment were applied at a sweep width of 2.5 kHz in F2 and 7 kHz in F1 resulting in an acquisition time of 0.1462 s. NUS processing was performed using the fully automated NUS processing tool provided by MestReNova. Zero filling in F1 has been used to yield a final matrix of 1K x 1K real points.

The DOSY spectrum was recorded applying a pulse sequence with stimulated echo using stimulated echo for diffusion from the Bruker pulse program library (stegp1s). The diffusion delay Δ (big delta) was set to 220 ms and the diffusion gradient pulse length δ (little delta) was set to 1.0 ms. The number of gradient steps were set to 32 with linear spacing starting from 2% reaching 95% of the full gradient strength in the final step. For each of the 32 gradient amplitudes, 256 transients of 65K complex data points were acquired. DOSY processing was performed with the DOSY processing tool from MestReNova (v.12.x64) employing the “Peak Heights Fit” algorithm including the “overlapped peaks analysis” with 128 points in diffusion dimension and a window of $1.00 \cdot 10^{-16}$ to $1.00 \cdot 10^{+03}$ cm^2 s^{-1}.

3.4 CD studies

The CD spectra of **2b** were recorded on a Jasco J-1500 spectrometer with 1 mm quartz cuvette. The following parameters were used: wavelength range from 460 to 280 nm. Scan speed: 100 nm/min; accumulation:

2; response time: 1.0 s; bandwidth: 2; temperature: 20 °C; sensitivity: standard (100 mdeg); data pitch: 0.5 nm; nitrogen gas flow rate: 500 L/h. The CD spectra of **4b** were recorded on a Jasco J-810 spectrometer with 1 mm quartz cuvette. The following parameters were used: wavelength range from 460 to 280 nm. Scan speed: 200 nm/min; accumulation: 3; response time: 1.0 s; bandwidth: 2; temperature: 20 °C; sensitivity: standard (100 mdeg); data pitch: 0.1 nm; nitrogen gas flow rate: 500 L/h. The sample solution of **2b** and **4b** was prepared in different proportions of DMSO/$CHCl_3$ solvents. The concentration is 0.1 mM. Δε values (in $cm^2 \cdot mol^{-1}$) were obtained by using the formula: Δε = m°/(C.l.32980) where m = CD value in millidegrees; l = cuvette pathlength in cm; C = sample concentration in mol/L.

3.5 X-ray crystallography

A single crystal of $(\mathbf{3b})_2$ used for the data diffraction experiment was obtained by slow evaporation of the undried chloroform/acetonitrile mixture. Data were collected at the IECB x-ray facility (CNRS UMS 3033 – INSERM US001) with a Rigaku FRX rotating anode (2.9 kW) diffractometer. CuKα radiation is monochromated with high flux Osmic Varimax HF mirrors. The x-ray source has a Dectris Pilatus 200K detector and partial chi goniometer. Crystal was kept at 100(2) K during data collection. The data were processed with the CrysAlis PRO software[38] with a multiscan absorption correction. Structure was solved using a dual-space algorithm with the ShelXT[39] structure solution program. Crystal model refinement was performed with the ShelXL[40] package using the Least Squares minimization. Both programs are implemented in Olex2.[41]

For some side chains, not all C or O atoms were found. During refinement, anisotropic displacement parameters were used for most atoms of backbone and S atoms of side chains. The C- and N-bound hydrogen atoms were placed in an idealized position. The positions of hydrogen atoms of O-H groups and one H atom of water molecule were found based on possible hydrogen bonds. All H atoms were refined in the riding-model approximation, with $U_{iso}(H)=1.2U_{eq}$(CH, CH_2, NH) and $U_{iso}(H)=1.5U_{eq}$ (OH). EADP, DELU, SIMU and RIGU instructions were employed to model temperature parameters. The geometry of the molecules was improved with DFIX, FLAT or AFIX commands.

Wide channels occupying about 33% of the unit cell volume are formed in the structure. These channels are filled with severely disordered solvent molecules removed using the solvent masking procedure implemented in Olex2. The solvent radius was set to 1.2 Å, and the calculated total potential solvent-accessible void volume and electron counts per unit cell were 8674 Å^3 and 2737. Considering the high number of electrons calculated for the channels and the variety of solvents used for crystallization (acetonitrile, water, chloroform), it is impossible to determine the solvent composition reliably. However, structure factors include contributions from the .fab file.

The final cif files were checked using IUCR's checkcif algorithm. Due to the characteristics of the crystal, *i.e.* large volume fractions of disordered solvent molecules, weak diffraction intensity, incompleteness of the data and low resolution, many A - level and B - level alerts remain in the check cif file. These alerts are inherent to the data and refinement procedures and do not reflect errors. They are explicitly listed below and have been divided into two groups. The first group illustrates the poor quality of the data and refinement statistics compared to that expected for small molecule structures from highly diffracting crystals. The second group is connected to decisions made during

GROUP 1

THETM01_ALERT_3_A The value of sine(theta_max)/wavelength is less than 0.550

PLAT023_ALERT_3_A, B Resolution (too) Low [sin(theta)/Lambda < 0.6].

PLAT084_ALERT_3_A, B High wR2 Value (i.e. > 0.25)

PLAT934_ALERT_3_A Number of (Iobs-Icalc)/Sigma(W) > 10 Outliers

PLAT082_ALERT_2_B High R1 Value

PLAT088_ALERT_3_B Poor Data / Parameter Ratio

PLAT241_ALERT_2_B High 'MainMol' Ueq as Compared to Neighbors

PLAT242_ALERT_2_B Low 'MainMol' Ueq as Compared to Neighbors

PLAT340_ALERT_3_B Low Bond Precision on C-C Bonds

GROUP 2

PLAT201_ALERT_2_A Isotropic non-H Atoms in Main Residue(s)

PLAT202_ALERT_3_A Isotropic non-H Atoms in Anion/Solvent

As mentioned above, not all atoms were refined with ADPs

PLAT306_ALERT_2_B Isolated Oxygen Atom (H-atoms Missing)

Unrecognized electron density was introduced to the refinement as dummy oxygen atoms.

PLAT315_ALERT_2_B Singly Bonded Carbon Detected (H-atoms Missing)

Not all H-atoms were localized, but they were used in SFAC calculation[40]

Table S1. Crystal data and refinement details.

Identification code	**(3b)$_2$**
Chemical formula	$2(C_{208}H_{200}N_{36}O_{45}S_7)\cdot H_2O\cdot 2(O)$**, [+solvents]*
Formula weight	8346.91
Crystal system	Triclinic
Space group	*P*-1
Unit cell dimensions (Å)	a= 26.866 (2), b=29.677 (2) c= 31.985 (2)
Volume (Å^3)	24326 (3)
Z	2
Radiation type	Cu *K*□
Density (calculated) (Mg m^{-3})	1.140
Absorption coefficient (mm^{-1})	1.22
Crystal size (mm)	0.15 × 0.15 × 0.03
Completeness	98.6 (up to 35.42°)

Reflections collected	68886
Reflections observed [$I > 2\sigma(I)$]	11712
R_{int}	0.064
Data/parameters/restrains	21353/2917/ 2518
Goodness-of-fit on F^2	2.17
Final R indices [$I > 2\sigma(I)$]	0.1995, 0.5136
R indices (all data)	0.2558, 0.5551
Largest diff. peak and hole	0.79, -0.45
CCDC #	2337052

* A solvent mask was used to remove severely disordered solvent molecules.

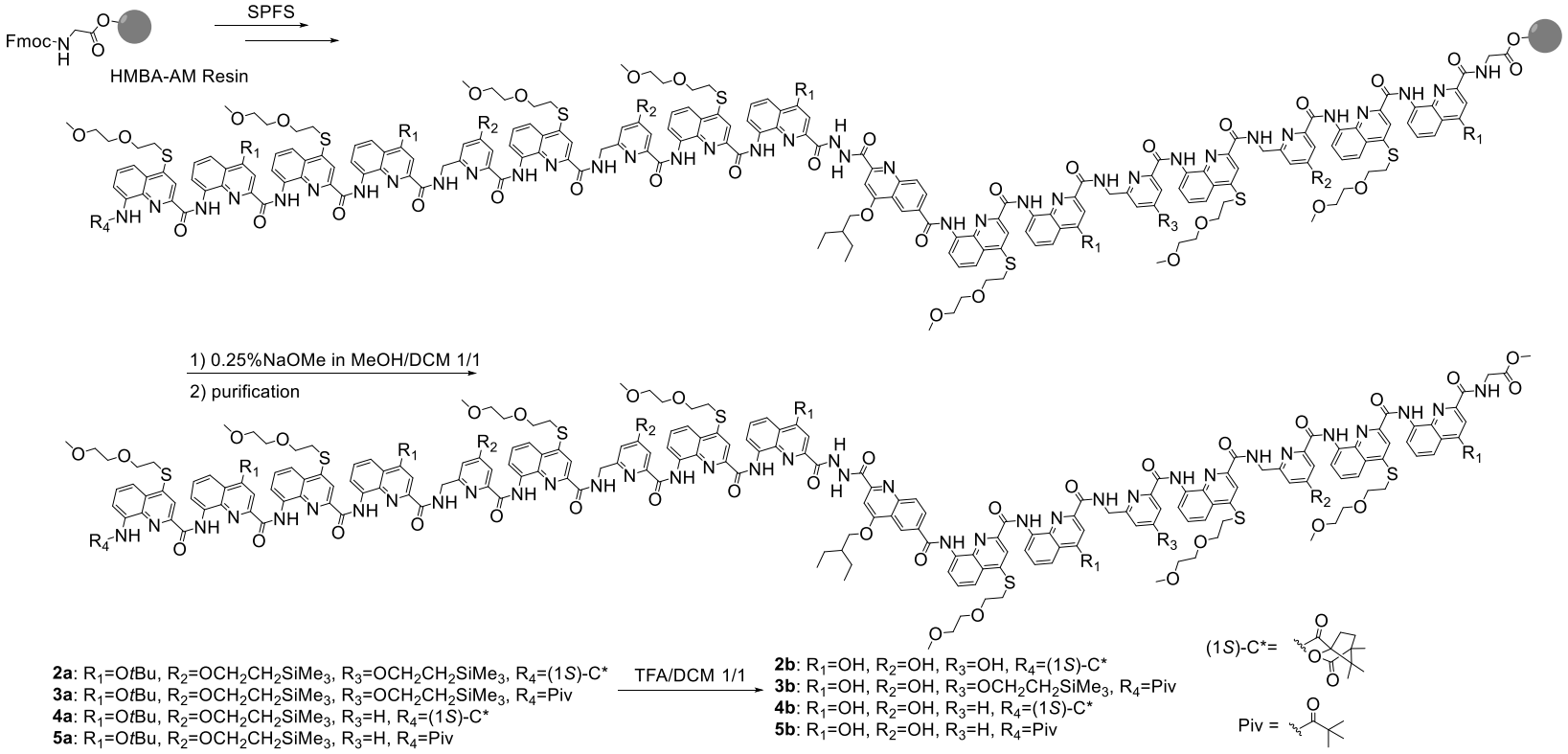
** Unrecognized electron density was introduced to the refinement as dummy oxygen.

Table S2. Hydrogen-bond geometry. Atoms are named as in the cif.

D—H···*A*	*D*—H (Å)	H···*A* (Å)	*D*···*A* (Å)	*D*—H···*A* (°)
First group				
O18D—H18D···O3C	0.84	1.74	2.57 (2)	169
O1D—H1D···O16G	0.84	1.82	2.62 (2)	161
O3D—H3D···O19G	0.84	1.82	2.65 (3)	175
O15D—H15D···O5C	0.84	1.73	2.57 (2)	171
O17D—H17D···O8C	0.84	1.81	2.64 (3)	168
O4D—H4D···O14G	0.84	1.86	2.69 (2)	168
Second group				
O9D—H9D···O3G	0.84	1.84	2.66 (3)	165
O10D—H10D···O16C	0.84	1.88	2.72 (3)	177
O12D—H12D···O19C	0.84	1.78	2.62 (2)	174
O6D—H6D···O5G	0.84	1.90	2.74 (3)	173
O8D—H8D···O8G	0.84	1.80	2.63 (3)	176
O13D—H13D···O14C	0.84	2.01	2.80 (2)	156
Third group				
O14D—H14D···O7C	0.84	1.78	2.63 (3)	174
O5D—H5D···O7G	0.84	1.82	2.65 (3)	168
O12D—H12D···O19C	0.84	1.78	2.62 (2)	174
O1D—H1D···O16G	0.84	1.82	2.62 (2)	161

Symmetry code(s): (i) -*x*+2, -*y*+2, -*z*; (ii) -*x*+2, -*y*+2, -*z*+1

1. Synthetic Scheme of foldamer synthesis



Scheme 1. Solid phase synthesis of aromatic oligomers.

4 Experimental Procedures

4.1 General methods

Commercially available reagents were purchased from Sigma-Aldrich, Alfa-Aesar or TCI and were used without further purification unless specified. HMBA-AM resin (200-400 mesh, loading 0.8–1.2 mmol/g) was purchased from Iris-biotech. THF, DCM and toluene were dried over alumina columns (MBRAUN SPS-800 solvent purification system). *N*,*N*-Diisopropylethylamine and chloroform were distilled over CaH_2 prior to use. Extra dry DMF was purchased from Sigma-Aldrich. Ultrapure water was obtained via a Stakpure OmniaPure-T UV-TOC ultrapure water system.

Analytical reversed phase (RP) high performance liquid chromatography (HPLC) was performed on a Thermo Fisher Scientific Ultimate 3000 HPLC System using Macherey-Nagel Nucleodur C8 Gravity columns (4 × 50 mm, 5 μm). UV absorbance was monitored at 300 nm and 254 nm, if not stated otherwise. The semi-preparative HPLC was performed on a Waters system equipped with a 2545 Quaternary Gradient Module with automated fraction collector system on a XBridge® Prep C8 OBD™ column (19 × 150 mm, 5 μm) at a flow rate of 25 mL/min. 0.1 % TFA was added to the aqueous mobile phase (referred to as mobile phase A) and to acetonitrile (referred to as mobile phase B). The gradient is: 0-5 min, 90% to 100% solvent B; 5-25 min, 100% solvent B at r.t.. The column eluent was monitored by UV detection at 254 and 300 nm with a diode array detector.

The ultraviolet–visible (UV/Vis) absorbance measurements were done with a Thermo Fisher Scientific Nanodrop One instrument using a 1 cm path length quartz cuvette. Circular dichroism (CD) spectra were measured on Jasco J-810 or Jasco J-1500 spectrometers. Measurements were performed at 20 °C if not stated otherwise.

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.[42]

4.2 Synthesis of monomers

The Fmoc-<u>Y</u>-OH,[21] Fmoc-T2-OH,[21] Fmoc-Q^{Deg}-OH,[30] Fmoc-<u>X</u>-OH[30] and Fmoc-P-OH[43] were synthesized according to literature. All of the Fmoc-protected monomers was ≥ 98% pure before used in the solid phase synthesis.

1.1 Solid phase synthesis general methods

1.1.1 Loading of the resin via HBTU activation

HMBA-AM resin (500 mg, 0.8-1.2 mmol, 1 eq.) was swollen in 5 mL DCM for 1 h, transferred to the microwave vessel and washed 3 times with extra dry DMF. DIPEA (170 µL, 1.0 mmol, 2 eq.) was added to a mixture of Fmoc-Gly-OH (134 mg; 0.45 mmol, 0.9 eq.) and HBTU (228 mg, 1.2 eq.) in extra dry DMF (5 mL) and the resulting solution was shaken for 30 s before to be poured to the resin-containing reaction vessel. The reaction mixture was subjected to treatment in a microwave oven (50 °C, 20 min, 25 W). The resin was filtered and washed with DMF (5 x 2 mL) and DCM (10 x 2 mL). Capping was performed by adding a mixture of DCM/pyridine/benzoyl chloride (3:1:1 (v/v/v), 5 mL) to the resin followed by shaking for 30 min at r.t., and subsequent washing with DCM (20 x 2 mL). For monitoring the efficiency of the 1st loading, small amount of resin (around 2 mg) was taken and dried under vacuum. The loading was estimated at this scale.

1.1.2 Estimation of the loading

To a small amount of resin-bound Fmoc-Gly (1–2 mg), a freshly prepared solution of DMF/piperidine (8:2 (v/v), 3.0 mL) was added. The mixture was shaken and incubated for 5 min. Then the absorption was measured at 301 nm using a NanoDrop One Microvolume UV-Vis Spectrophotometer and a Hellma quartz glass cuvette 104 (path length 10 mm). Three replicates were measured, then the loading was calculated with the following equation:

$$\text{loading}\left(\text{in } \frac{\text{mmol}}{\text{g}}\right) = \frac{\text{Abs}_{301\,\text{nm}} \times \text{V}}{\varepsilon_{301\text{nm}} \times l \times \text{m}}$$

$$\varepsilon_{301\text{nm}} = 7800\ \text{L/mol/cm}^{[44]}$$

1.1.3 Solid Phase Synthesis via in-situ-activation[42]

Fmoc-Gly-HMBA-AM resin (100 mg, loading 0.3 mmol/g, 30 µmol) was first swollen in DCM (3 mL) for 1 h, the resin was transferred into the microwave vessel and washed 3 times with DMF and 3 times with NMP.

The deprotection of the Fmoc group was performed by adding a solution of 2% DBU in NMP (3 mL) to the resin and incubation for 3 min. The resin was next filtered off and the deprotection step was repeated once. After filtration, the resin was washed with DCM (3 x 2 mL) and then with anhydrous THF (5 x 2 mL). This deprotection step was performed after each aromatic monomer coupling.

The resin was next suspended in anhydrous THF (1 mL) and 2,3,5-collidine (5 eq. with respect to the resin-loading) was added to the resin supernatant. The Fmoc-protected monomer (2 eq. with respect to the resin-loading) and PPh_3 (4 eq. with respects to the resin-loading) were successively added in a vial to be solubilized in freshly distilled $CHCl_3$ (1 mL).

Trichloroacetonitrile (4.5 eq. with respect to the resin loading) was next added to the vial and the resulting acid chloride solution was shaken for 30 s before to be poured to the resin-containing reaction vessel. The reaction vessel was then placed in the microwave oven and subjected to MW irradiation for 15 min (50°C, 50 W). The resin was then washed 3 times with anhydrous THF. This entire coupling step was then repeated once more. For the coupling of Fmoc-T2-OH, the same acid chloride activation process was followed but the coupling was, this time, performed at r.t. by shaking the resin for 2 h. Then the resin was washed with anhydrous THF (3 x 2 mL).

For the final coupling of pivaloyl- (Piv-) or (1*S*)-camphanic ((1*S*)-C*-) amides, the resin was suspended in anhydrous THF (1 mL) and 2,3,5-collidine (5 eq. with respect to the resin-loading) was added to the resin suspensions. A solution of pivaloyl chloride (2 eq. with respect to the resin-loading) or (1*S*)-camphanic chloride (2 eq. with respect to the resin-loading, purchased from Sigma-Aldrich, 98%, ee: 99%) in freshly distilled $CHCl_3$ (1 mL) was added to the supernatant and the resin was shaken at r.t. for 2 h. The resin was filtered off, washed 3 times with dry THF, and the same process was repeated once. After coupling, the resin was vigorously washed 3 times with DMF and 3 times with DCM.

1.1.4 Mini-Cleavage

To perform a mini cleavage, the resin (1–2 mg) was swollen in 1 mL MeOH/DCM (1:1, v/v) solution followed by addition of 10 µL NaOMe (25% (m/m)) and incubated at r.t. for 10 min. The cleavage solution was diluted with DCM, washed with aqueous citric acid solution (5%), dried over $MgSO_4$, filtered and the solvent was finally removed under reduced pressure.

1.1.5 Full Cleavage[45]

Preparation of cleavage solution: 200 mL dry MeOH was added to 200 mL dry DCM under N_2 atmosphere. 2 mL NaOMe (25% (m/m)) in methanol was added and the mixture was well-mixed by magnetic stirring. Sufficient amount of cleavage solution (at least 400 mL cleavage solution for 100 mg resin) was important to avoid the formation of oligomer acid as the side-product.

The resin (around 100 mg) was dried under vacuum and slowly added to 400 mL cleavage solution under N_2 atmosphere. The mixture was stirred under N_2 atmosphere for 2 h before it was added to 100 mL aqueous citric acid solution (5%). The aqueous layer was extracted with DCM(3 x 50 mL). The combined organic phases were washed with brine, dried over $MgSO_4$, filtered and the solvent was evaporated under reduced pressure. The crude was recovered as solid.

1.2 Synthesis of oligomers

(1*S*)-Camph-Q^D<u>X</u>Q^D<u>XY</u>Q^D<u>Y</u>Q^D<u>X</u>-T2-Q^D<u>XY</u>Q^D<u>Y</u>Q^D<u>X</u>-Gly-OMe (**2a**): Compound **2a** was synthesized using the SPS procedures reported in paragraph 5.3 on Fmoc-Gly-HMBA AM resin. The crude product was obtained after full cleavage and purification by RP-HPLC. (15 mg, 12%). **^{1}H NMR** (500 MHz, chloroform-*d*) δ 11.40 (s, 1H), 11.38 (s, 1H), 11.28 (s, 1H), 11.24 (s, 1H), 11.01 (s, 1H), 10.96 (s, 1H), 10.79 (s, 2H), 10.62 (s, 1H), 10.43 (s, 1H), 10.25 (s, 1H), 9.70 (s, 1H), 9.43 (s, 1H), 8.52 (t, *J* = 3.3 Hz, 1H), 8.27 (s, 1H), 8.26 (s, 1H), 8.20 (d, *J* = 7.3 Hz, 1H), 8.14–8.11 (m, 2H), 8.05 (t, *J* = 3.3 Hz, 1H), 8.02 (s, 1H), 7.78 (s, 2H), 7.76 (s, 2H), 7.74 (s, 2H), 7.73–7.70 (m, 6H), 7.66–7.61 (m, 4H), 7.59 (s, 1H), 7.56 (s, 2H), 7.54 (s, 3H), 7.47 (s, 2H), 7.45 (s, 2H), 7.42–7.40 (m, 2H), 7.39 (s, 3H), 7.32–7.29 (m, 4H), 7.15 (s, 1H), 7.12 (s, 2H), 7.11 (s, 2H), 7.10 (s, 1H), 7.08 (s, 1H), 7.07 (s, 2H), 7.05 (s, 1H), 6.93 (t, *J* = 7.8 Hz, 1H), 6.69 (s, 1H), 6.67 (s, 1H), 6.61 (s, 1H), 6.50 (s, 1H), 6.41 (s, 1H), 6.40 (s, 1H), 6.37 (s, 1H), 6.33 (s, 1H), 6.29 (s, 1H), 6.26 (s, 1H), 6.18 (s, 1H), 5.93 (s, 1H). 4.33–4.27 (m, 1H), 4.26–4.13 (m, 4H), 4.08–3.95 (m, 9H), 3.93–3.90 (m, 4H), 3.89–3.77 (m, 12H), 3.76–3.66 (m, 13H), 3.65–3.58 (m, 8H), 3.57 (s, 3H), 3.54 (s, 3H), 3.52–3.49 (m, 4H), 3.47 (s, 3H), 3.46–3.44 (m, 6H), 3.43 (s, 3H), 3.42 (s, 4H), 3.37 (t, *J* = 6.5 Hz, 3H), 3.33–3.18 (m, 12H), 3.16 (s, 3H), 3.01 (d, *J* = 16.0 Hz, 1H), 2.90 (t, *J* = 8.4 Hz, 1H), 2.80 (d, *J* = 15.3 Hz, 1H), 2.58 (d, *J* = 14.6 Hz, 1H), 2.33 (d, *J* = 15.0 Hz, 1H), 2.03–1.99 (m, 3H), 1.95–1.85 (m, 2H), 1.67 (s, 9H), 1.66 (s, 9H), 1.64 (s, 9H), 1.47 (s, 9H), 1.32–1.24 (m, 5H), 1.09 (s, 10H), 1.05–0.99 (m, 4H), 0.91–0.82 (m, 2H), 0.70 (d, *J* = 7.2 Hz, 6H), 0.57–0.49 (m, 6H), 0.33 (s, 9H), 0.26 (s, 9H), 0.17 (s, 9H), 0.13–0.04 (m, 2H), -0.03 (s, 9H). (mixture of two conformers in a ratio of 1:0.3, only the major peaks are reported.) **HRMS** (ESI+) calcd. for $C_{253}H_{292}N_{36}O_{47}S_7Si_4$ $[M+2H]^{2+}$ 2461.9417, found 2461.8227.

(1*S*)-Camph-Q^{D}XQ^{D}XYQ^{D}YQ^{D}X-T2-Q^{D}XYQ^{D}YQ^{D}X-Gly-OMe (**2b**): Compound **2a** was treated with a solution of TFA/DCM (1:1 (v/v), 2 mL) at r.t. overnight. The solvent was removed under vacuum. The solid was precipitated from MeOH, was subsequently filtered and washed 2 times with MeOH to yield the desired product as a yellow solid. (13 mg, quant.) **^{1}H NMR** (500 MHz, chloroform-*d*) δ 12.61 (s, 1H), 11.69 (s, 1H), 11.46 (s,

1H), 11.22 (s, 1H), 11.11 (s, 1H), 11.08 (s, 1H), 11.05 (s, 2H), 10.77 (s, 1H), 10.72 (s, 2H), 10.68 (s, 1H), 10.60 (s, 1H), 10.50 (s, 1H), 10.33 (s, 1H), 10.04 (s, 1H), 9.96 (s, 1H), 9.88 (s, 1H), 9.74 (s, 1H), 9.24 (s, 1H), 9.16 (s, 1H), 8.73 (s, 1H), 8.68 (s, 1H), 8.64–8.59 (m, 2H), 8.54 (d, *J* = 8.0 Hz, 1H), 8.39–8.34 (m, 1H), 8.27–8.23 (m, 3H), 8.20–8.18 (m, 2H), 8.15 (s, 1H), 8.11 (s, 1H), 8.10 (d, *J* = 2.9 Hz, 1H), 8.08 (s, 1H), 8.07–8.05 (m, 1H), 8.04 (s, 1H), 8.02 (s, 1H), 7.96 (d, *J* = 7.0 Hz, 2H), 7.93–7.89 (m, 3H), 7.88 (s, 1H), 7.86 (s, 2H), 7.84 (s, 2H), 7.81 (d, *J* = 7.3 Hz, 2H), 7.78 (d, *J* = 6.7 Hz, 2H), 7.75 (s, 2H), 7.73 (s, 1H), 7.71 (s, 2H), 7.69 (s, 1H), 7.67 (s, 1H), 7.65 (d, *J* = 8.0 Hz, 1H), 7.62 (d, *J* = 7.7 Hz, 2H), 7.56 (s, 1H), 7.54 (s, 2H), 7.52 (s, 1H), 7.51 (s, 1H), 7.47 (s, 2H), 7.45 (s, 1H), 7.43 (s, 1H), 7.37–7.30 (m, 3H), 7.25–7.20 (m, 2H), 7.14–7.11 (m, 1H), 7.09 (s, 1H), 6.98 (s, 1H), 6.94 (s, 1H), 6.90 (s, 1H), 6.83 (s, 1H), 6.76 (s, 1H), 6.54–6.51 (m, 2H), 6.31 (s, 1H), 6.15 (s, 1H), 4.31–4.18 (m, 7H), 4.16–4.06 (m, 10H), 4.05–3.98 (m, 8H), 3.97–3.95 (m, 5H), 3.94–3.90 (m, 4H), 3.85–3.78 (m, 10H), 3.78–3.71 (m, 4H), 3.70 (s, 3H), 3.68–3.63 (m, 6H), 3.61 (s, 3H), 3.56–3.48 (m, 6H), 3.46 (s, 6H), 3.42–3.32 (m, 5H), 3.30 (s, 6H), 3.24 (d, *J* = 6.2 Hz, 3H), 3.15–3.00 (m, 4H), 2.85–2.80 (m, 2H), 1.53 (m, 2H), 1.27 (d, *J* = 15.6 Hz, 12H), 0.87 (dd, *J* = 16.1, 9.4 Hz, 10H). **HRMS** (ESI+) calcd. for $C_{213}H_{204}N_{36}O_{47}S_7$ $[M+2K]^{2+}$ 2159.5994, found 2159.5140.

Piv-Q^{D}<u>X</u>Q^{D}<u>XY</u>Q^{D}<u>Y</u>Q^{D}<u>X</u>-T2-Q^{D}<u>XY</u>Q^{D}<u>Y</u>Q^{D}<u>X</u>-Gly-OMe (**3a**): Compound **3a** was synthesized using the SPS procedures reported in 5.3 on Fmoc-Gly-HMBA AM resin. The crude product was obtained after full cleavage and purified by RP-HPLC. (11 mg, 9%). **^{1}H NMR** (500 MHz, chloroform-*d*) δ 11.37 (s, 1H), 11.28 (s, 2H), 10.99 (s, 1H), 10.96 (s, 1H), 10.80 (s, 1H), 10.79 (s, 1H), 10.78 (s, 1H), 10.57 (s, 1H), 10.39 (s, 1H), 10.25 (s, 1H), 9.71 (s, 1H), 8.69 (s, 1H), 8.54 (t, *J* = 3.3 Hz, 1H), 8.35 (d, *J* = 3.3 Hz, 1H), 8.27 (s, 1H), 8.25 (s, 1H), 8.21 (d, *J* = 3.3 Hz, 1H), 8.16–8.11 (m, 4H), 8.04–8.01 (m, 3H), 7.93–7.88 (m, 5H), 7.82–7.77 (m, 3H), 7.76–7.75 (m, 2H), 7.75–7.73 (m, 2H), 7.71 (dd, *J* = 7.6, 4.8 Hz, 2H), 7.63 (s, 1H), 7.62–7.59 (m, 3H), 7.58 (s, 1H), 7.55 (d, *J* = 8.1 Hz, 3H), 7.46–7.43 (m, 2H), 7.42–7.37 (m, 4H), 7.32 (d, *J* = 7.5 Hz, 2H), 7.30 (s, 2H), 7.21 (s, 1H), 7.20 (s, 1H), 7.19–7.16 (m, 2H), 7.15 (s, 1H), 7.13 (s, 1H), 7.12 (s, 1H), 7.10–7.09 (m, 2H), 7.06 (s, 1H), 6.93 (t, *J* = 7.8 Hz, 1H), 6.68–6.67 (m, 1H), 6.62 (s, 1H), 6.60 (s, 1H), 6.48 (s, 1H), 6.46 (s, 1H), 6.42 (s, 1H), 6.36 (s, 1H), 6.34 (s, 1H), 6.31 (d, *J* = 2.4 Hz, 1H), 6.26 (s, 1H), 6.19 (s, 1H), 5.93–5.91 (m, 1H). 4.32–4.14 (m, 5H), 4.08–3.91 (m, 9H), 3.89–3.77 (m, 18H), 3.76–3.66 (m, 12H), 3.65–3.59 (m, 16H), 3.57 (s, 3H), 3.54 (s, 3H), 3.50 (s, 4H), 3.47 (s, 3H), 3.44 (s, 5H), 3.43 (s, 3H), 3.42 (s, 3H), 3.40–3.17 (m, 7H), 3.16 (s, 3H), 2.97 (s, 2H), 2.90 (s, 2H), 2.80 (d, *J* = 14.6 Hz, 1H), 2.54 (d, *J* = 15.5 Hz, 1H), 1.72 (s, 9H), 1.66 (s, 9H), 1.65 (s, 9H), 1.63 (d, *J* = 2.4 Hz, 2H), 1.60 (s, 2H), 1.50 (s, 2H), 1.47 (s, 9H), 1.38–1.20 (m, 4H), 1.12 (s, 9H), 1.07 (s, 9H), 1.05–0.97 (m, 1H), 0.86–

0.81 (m, 1H), 0.71–0.67 (m, 1H), 0.57 (s, 9H), 0.50 (t, *J* = 7.4 Hz, 2H), 0.32 (d, *J* = 2.3 Hz, 9H), 0.25 (s, 9H), 0.17 (s, 9H) (mixture of two conformers in a ratio of 1:0.3, only the major peaks are reported). **HRMS** (ESI+) calcd. for $C_{248}H_{288}N_{36}O_{45}S_7Si_4$ $[M+2H]^{2+}$ 2413.9311, found 2413.8372.

Piv-Q^{D}XQ^{D}XYQ^{D}YQ^{D}X-T2-Q^{D}XYQ^{D}YQ^{D}X-Gly-OMe (**3b**): Compound **3a** was treated with a solution of TFA/DCM (1:1 (v/v), 2 mL) at r.t. overnight. The solvent was removed under vacuum. The solid was precipitate from MeOH and recovered by filtration. The solid was washed with MeOH for another 2 times. The compound was obtained as a yellow solid. (9.5 mg, quant.) **^{1}H NMR** (500 MHz, chloroform-*d*) δ 12.43 (s, 1H), 11.60 (s, 1H), 11.22 (s, 1H), 11.09 (s, 1H), 11.06 (s, 1H), 11.04 (s, 1H), 10.97 (s, 1H), 10.80 (s, 1H), 10.79 (s, 1H), 10.76 (s, 0H), 10.69 (s, 1H), 10.64 (s, 1H), 10.36 (s, 1H), 9.99 (s, 1H), 9.98 (s, 1H), 9.78 (s, 1H), 9.75 (s, 1H), 9.16 (s, 1H), 8.93 (s, 1H), 8.76 (s, 1H), 8.64 (s, 1H), 8.63–8.60 (m, 3H), 8.60 (s, 2H), 8.58–8.57 (m, 1H), 8.52–8.49 (m, 2H), 8.39 (d, *J* = 8.0 Hz, 1H), 8.31 (d, *J* = 8.0 Hz, 1H), 8.28–8.25 (m, 2H), 8.22 (s, 1H), 8.18 (s, 1H), 8.14 (d, *J* = 8.2 Hz, 1H), 8.09–8.05 (m, 3H), 8.04 (s, 1H), 8.02 (s, 1H), 8.00 (s, 1H), 7.96 (s, 2H), 7.94 (s, 1H), 7.92 (s, 1H), 7.91–7.90 (m, 2H), 7.90 (s, 1H), 7.89 (s, 2H), 7.87 (s, 1H), 7.87–7.86 (m, 2H), 7.85–7.83 (m, 2H), 7.82 (s, 1H), 7.79–7.77 (m, 2H), 7.75–7.73 (m, 2H), 7.72–7.71 (m, 2H), 7.70 (d, *J* = 7.3 Hz, 1H), 7.67 (s, 1H), 7.66 (s, 1H), 7.56 (d, *J* = 2.2 Hz, 1H), 7.54 (s, 1H), 7.52 (s, 1H), 7.48 (s, 1H), 7.45 (d, *J* = 2.7 Hz, 1H), 7.44–7.42 (m, 1H), 7.36–7.34 (m, 2H), 7.19 (s, 1H), 7.18 (s, 1H), 7.08 (s, 1H), 6.97 (s, 1H), 6.94 (s, 1H), 6.91 (s, 1H), 6.85 (s, 1H), 6.72 (s, 1H), 6.55 (d, *J* = 7.2 Hz, 2H), 6.28 (d, *J* = 1.3 Hz, 2H), 6.04 (s, 1H), 5.77 (s, 1H), 5.49 (s, 1H), 4 4.30–4.18 (m, 2H), 4.21–4.14 (m, 3H), 4.13–4.05 (m, 3H), 4.06–3.88 (m, 18H), 3.88–3.77 (m, 12H), 3.79–3.71 (m, 6H), 3.69 (s, 3H), 3.69–3.60 (m, 10H), 3.61 (s, 3H), 3.60 (s, 3H), 3.54 (s, 3H), 3.54–3.47 (m, 3H), 3.46 (s, 3H), 3.46 (s, 3H), 3.29 (s, 3H), 3.28–3.22 (m, 3H), 3.13 (s, 3H), 3.08–3.03 (m, 3H), 3.00–2.94 (m, 2H), 2.89–2.82 (m, 1H), 2.62–2.57 (m, 1H), 2.38–2.33 (m, 1H), 2.02–1.99 (m, 1H), 1.28 (s, 9H), 1.26–1.22 (m, 10H). **HRMS** (ESI+) calcd. for $C_{208}H_{200}N_{36}O_{45}S_7$ $[M+2H]^{2+}$ 2073.6329, found 2073.6093.

(1*S*)-Camph-Q^{D}<u>X</u>Q^{D}<u>XY</u>Q^{D}<u>Y</u>Q^{D}<u>X</u>-T2-Q^{D}<u>X</u>PQ^{D}<u>Y</u>Q^{D}<u>X</u>-Gly-OMe (**4a**): Compound **4a** was synthesized using the SPS procedures reported in 5.3 on Fmoc-Gly-HMBA AM resin. The crude product was obtained after full cleavage and purified by RP-HPLC. (18 mg, 15%). **^{1}H NMR** (500 MHz, chloroform-*d*) δ 11.40 (s, 1H), 11.37 (s, 1H), 11.27 (s, 1H), 11.25 (s, 1H), 11.01 (s, 1H), 10.93 (s, 1H), 10.79 (s, 1H), 10.75 (s, 1H), 10.63 (s, 1H), 10.42 (s, 1H), 10.25 (s, 1H), 9.70 (s, 1H), 9.43 (s, 1H), 8.48 (t, *J* = 3.3 Hz, 1H), 8.28–8.24 (m, 2H), 8.22–8.19 (m, 2H), 8.14 (d, *J* = 7.4 Hz, 1H), 8.04–8.01 (m, 3H), 7.96 (d, *J* = 7.3 Hz, 1H), 7.88 (s, 2H), 7.86 (s, 1H), 7.82 (d, *J* = 8.1 Hz, 1H), 7.77–7.74 (m, 3H), 7.73–7.72 (m, 2H), 7.71–7.70 (m, 2H), 7.69–7.67 (m, 2H), 7.65 (s, 1H), 7.63 (s, 2H),

7.61 (s, 1H), 7.60 (s, 1H), 7.58 (s, 1H), 7.57–7.55 (m, 1H), 7.54 (s, 2H), 7.49 (dd, *J* = 12.7, 7.7 Hz, 2H), 7.45 (s, 1H), 7.39 (d, *J* = 8.0 Hz, 2H), 7.34 (s, 3H), 7.30 (s, 3H), 7.28 (s, 2H), 7.25 (s, 1H), 7.16–7.14 (m, 2H), 7.14–7.13 (m, 1H), 7.12–7.10 (m, 2H), 7.10 (s, 1H), 7.08 (s, 1H), 7.07–7.06 (m, 2H), 6.95–6.91 (m, 2H), 6.68 (s, 1H), 6.63 (s, 1H), 6.49 (s, 1H), 6.41 (s, 1H), 6.39 (d, *J* = 4.8 Hz, 2H), 6.33 (s, 1H), 6.27 (s, 1H), 6.17 (s, 1H), 5.93 (s, 1H), 5.87 (d, *J* = 7.8 Hz, 1H), 4.29 (d, *J* = 8.3 Hz, 1H), 4.23–4.13 (m, 4H), 4.06–4.02 (m, 4H), 4.01–3.96 (m, 6H), 3.93–3.90 (m, 4H), 3.89–3.76 (m, 16H), 3.76–3.73 (m, 2H), 3.73–3.66 (m, 8H), 3.64–3.58 (m, 9H), 3.57 (s, 3H), 3.54 (s, 2H), 3.52 (s, 1H), 3.50 (s, 3H), 3.47 (s, 2H), 3.44 (s, 5H), 3.43 (s, 3H), 3.42 (s, 3H), 3.38–3.18 (m, 14H), 3.16 (s, 2H), 3.14 (s, 1H), 3.00 (d, *J* = 16.4 Hz, 1H), 2.90 (t, *J* = 8.4 Hz, 1H), 2.77 (d, *J* = 14.9 Hz, 1H), 2.61 (d, *J* = 14.9 Hz, 1H), 2.32 (d, *J* = 13.8 Hz, 1H), 1.86–1.75 (m, 1H), 1.66 (s, 9H), 1.66 (s, 9H), 1.63 (s, 9H), 1.46 (s, 9H), 1.35–1.22 (m, 9H), 1.09 (s, 9H), 1.05–0.99 (m, 1H), 0.86–0.78 (m, 2H), 0.70 (t, *J* = 7.2 Hz, 2H), 0.55 (s, 3H), 0.52 (s, 3H), 0.32 (t, *J* = 7.4 Hz, 4H), 0.25 (s, 9H), 0.17 (s, 9H), -0.03 (s, 9H). (mixture of two conformers in a ratio of 1:0.3, only the major peaks are reported.) **HRMS** (ESI+) calcd. for $C_{248}H_{280}N_{36}O_{46}S_7Si_3$ $[M+2H]^{2+}$ 2403.9088, found 2403.8301.

(1*S*)-Camph-Q^{D}XQ^{D}XYQ^{D}YQ^{D}X-T2-Q^{D}XPQ^{D}YQ^{D}X-Gly-OMe (**4b**): Compound **4a** was treated with a solution of TFA/DCM (1:1 (v/v), 2 mL) at r.t. overnight. The solvent was removed under vacuum. The solid was precipitate from MeOH and recovered by filtration. The solid was washed by MeOH for another 2 times. The compound was obtained as a yellow solid. (15.8 mg, quant.) **^{1}H NMR** (500 MHz, chloroform-*d*) δ 12.59 (s, 1H), 11.86 (s, 1H), 11.52 (s, 1H), 11.19 (s, 1H), 11.07 (s, 2H), 10.97 (s, 1H), 10.80 (s, 1H), 10.73 (s, 1H), 10.70 (s, 1H), 10.65 (s, 1H), 10.61 (s, 1H), 10.33 (s, 2H), 10.00 (s, 1H), 9.98 (s, 1H), 9.95 (s, 1H), 9.77 (s, 1H), 9.24 (s, 1H), 9.20 (s, 1H), 8.77 (s, 1H), 8.73 (s, 1H), 8.69 (s, 1H), 8.66–8.61 (m, 2H), 8.56–8.53 (m, 2H), 8.41–8.35 (m, 2H), 8.26 (d, *J* = 8.0 Hz, 1H), 8.23–8.18 (m, 4H), 8.17 (s, 1H), 8.14 (d, *J* = 7.8 Hz, 1H), 8.08–8.05 (m, 3H), 8.03 (d, *J* = 7.0 Hz, 1H), 7.97–7.93 (m, 3H), 7.92–7.88 (m, 2H), 7.86–7.81 (m, 5H), 7.79–7.77 (m, 2H), 7.76–7.74 (m, 2H), 7.73 (s, 1H), 7.69 (t, *J* = 4.1 Hz, 2H), 7.64 (d, *J* = 7.8 Hz, 1H), 7.57–7.54 (m, 3H), 7.52 (d, *J* = 7.6 Hz, 1H), 7.50–7.44 (m, 5H), 7.42–7.35 (m, 4H), 7.33–7.27 (m, 2H), 7.25–7.20 (m, 3H), 7.14 (s, 1H), 7.07 (d, *J* = 8.4 Hz, 2H), 6.94 (s, 1H), 6.87 (d, *J* = 13.8 Hz, 2H), 6.80 (s, 1H), 6.58 (d, *J* = 8.1 Hz, 1H), 6.51 (s, 1H), 6.48 (s, 1H), 6.31 (d, *J* = 1.6 Hz, 1H), 6.17 (s, 1H), 4.33–4.28 (m, 1H), 4.26–4.20 (m, 4H), 4.14 (dd, *J* = 10.1, 4.4 Hz, 8H), 4.08–3.90 (m, 16H), 3.85–3.82 (m, 3H), 3.81–3.76 (m, 9H), 3.74–3.71 (m, 5H), 3.69–3.64 (m, 10H), 3.62 (s, 3H), 3.57 (s, 3H), 3.54–3.51 (m, 3H), 3.48 (s, 3H), 3.46 (s, 3H), 3.46 (s, 3H), 3.43–3.35 (m, 6H), 3.29 (s, 3H), 3.24 (s, 3H), 3.09 (dd, *J* = 18.5, 4.2 Hz, 2H), 3.05–3.00 (m, 1H), 2.89–2.80 (m, 2H), 2.42–2.38 (m, 1H), 1.63–1.44 (m, 1H),

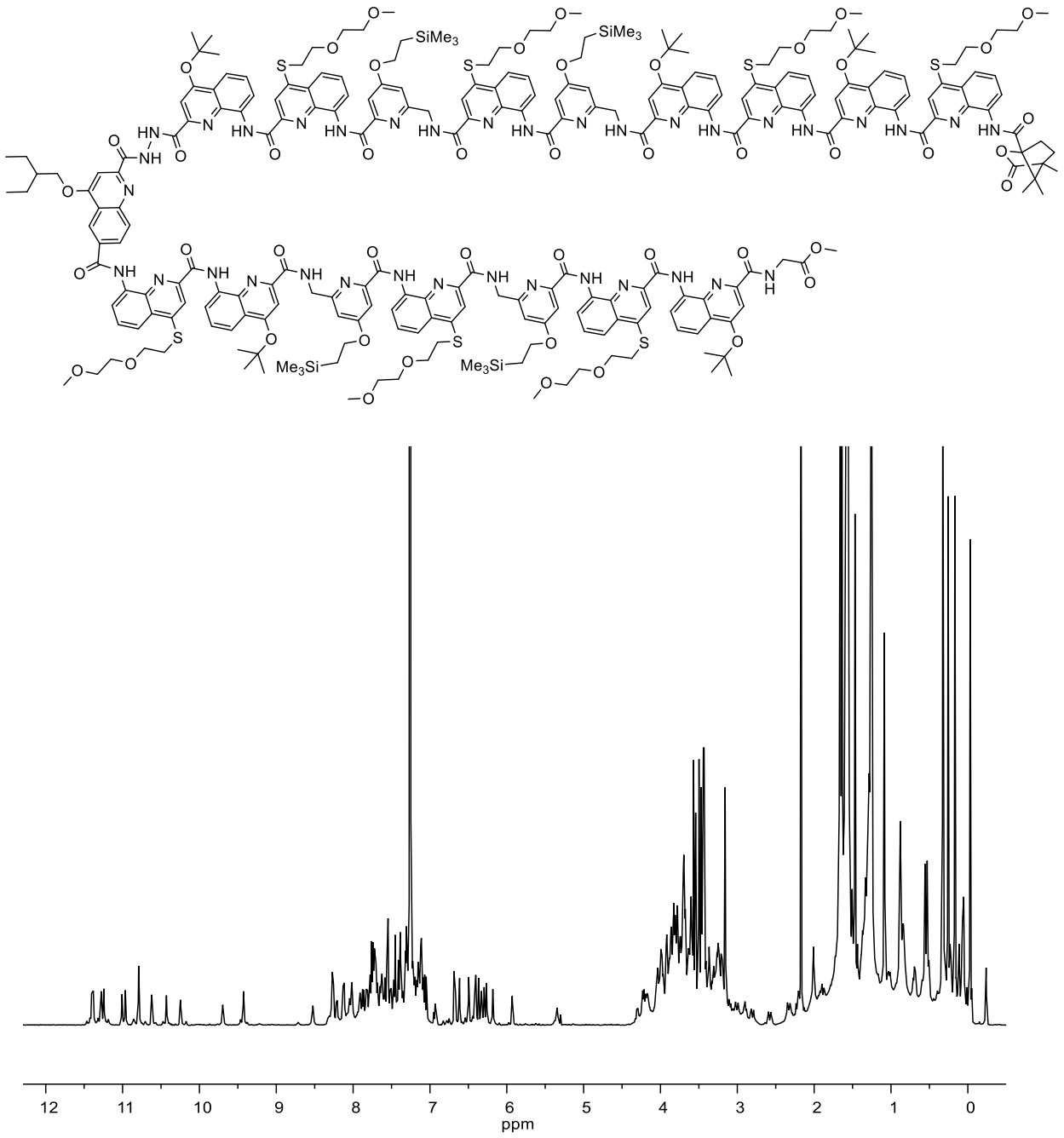
1.25 (s, 10H, overlap with impurities), 0.92–0.79 (m, 8H, overlap with impurities), 0.67 (s, 3H), 0.60 (s, 2H). **HRMS** (ESI+) calcd. for $C_{213}H_{204}N_{36}O_{46}S_7$ $[M+2H]^{2+}$ 2113.6460, found 2113.6315.

Piv-Q^{D}<u>X</u>Q^{D}<u>XY</u>Q^{D}<u>Y</u>Q^{D}<u>X</u>-T2-Q^{D}<u>X</u>PQ^{D}<u>Y</u>Q^{D}<u>X</u>-Gly-OMe (**5a**): Compound **5a** was synthesized using the SPS procedures reported in 5.3 on Fmoc-Gly-HMBA AM resin. The crude product was obtained after full cleavage and purified by RP-HPLC. (16 mg, 14%). **^{1}H NMR** (500 MHz, chloroform-*d*) δ 11.37 (s, 1H), 11.27 (s, 1H), 11.26 (s, 1H), 10.99 (s, 1H), 10.93 (s, 1H), 10.79 (s, 1H), 10.78 (s, 1H), 10.75 (s, 1H), 10.58 (s, 1H), 10.38 (s, 1H), 10.25 (s, 1H), 9.70 (s, 1H), 8.69 (s, 1H), 8.48 (t, *J* = 3.3 Hz, 2H), 8.26 (t, *J* = 3.3 Hz, 2H), 8.26–8.24 (m, 1H), 8.20 (t, *J* = 3.3 Hz, 1H), 8.17–8.12 (m, 2H), 8.04–8.00 (m, 3H), 7.96 (d, *J* = 7.2 Hz, 1H), 7.92–7.89 (m, 1H), 7.88 (s, 2H), 7.86 (s, 1H), 7.80 (d, *J* = 8.3 Hz, 1H), 7.78–7.75 (m, 2H), 7.75 (s, 2H), 7.73 (s, 1H), 7.71–7.67 (m, 2H), 7.64 (d, *J* = 3.5 Hz, 1H), 7.63 (s, 2H), 7.61 (s, 1H), 7.60–7.59 (m, 2H), 7.58 (s, 1H), 7.56 (s, 2H), 7.55 (d, *J* = 1.8 Hz, 1H), 7.48 (d, *J* = 8.1 Hz, 1H), 7.44 (s, 1H), 7.39 (d, *J* = 7.9 Hz, 2H), 7.35 (s, 3H), 7.32 (d, *J* = 6.3 Hz, 2H), 7.30 (s, 2H), 7.24 (s, 1H), 7.19–7.17 (m, 1H), 7.16 (s, 2H), 7.15 (s, 2H), 7.14–7.13 (m, 1H), 7.11 (s, 2H), 7.07–7.06 (m, 2H), 6.95–6.91 (m, 2H), 6.65 (s, 1H), 6.63 (s, 1H), 6.49 (s, 1H), 6.42 (s, 2H), 6.38 (s, 1H), 6.33 (s, 1H), 6.27 (s, 1H), 6.17 (s, 1H), 5.92 (d, *J* = 1.3 Hz, 1H), 4.29 (q, *J* = 8.0 Hz, 1H), 4.17 (dd, *J* = 19.6, 8.8 Hz, 4H), 4.07–4.02 (m, 4H), 4.01–3.94 (m, 5H), 3.93–3.90 (m, 4H), 3.88–3.83 (m, 9H), 3.81–3.77 (m, 9H), 3.76–3.73 (m, 3H), 3.72–3.67 (m, 10H), 3.63–3.59 (m, 7H), 3.57 (s, 3H), 3.54 (s, 3H), 3.52–3.48 (m, 4H), 3.47 (s, 3H), 3.44 (s, 3H), 3.43 (s, 3H), 3.42 (s, 3H), 3.38–3.35 (m, 2H), 3.33–3.29 (m, 2H), 3.28–3.18 (m, 8H), 3.16 (s, 1H), 2.98 (d, *J* = 15.9 Hz, 1H), 2.90 (t, *J* = 8.4 Hz, 1H), 2.77 (d, *J* = 14.7 Hz, 1H), 2.60 (d, *J* = 15.3 Hz, 1H), 2.30 (d, *J* = 13.6 Hz, 1H), 1.83 (d, *J* = 14.0 Hz, 1H), 1.71 (s, 9H), 1.66 (s, 9H), 1.66 (s, 9H), 1.59 (s, 3H), 1.50 (s, 3H), 1.46 (s, 9H), 1.37–1.29 (m, 5H), 1.09 (s, 9H), 0.72–0.68 (m, 2H), 0.56 (s, 9H), 0.52 (d, *J* = 7.4 Hz, 2H), 0.32 (t, *J* = 7.4 Hz, 2H), 0.25 (s, 9H), 0.16 (s, 9H), 0.06 (s, 9H). (mixture of two conformers in a ratio of 1:0.3, only the major peaks are reported.) **HRMS** (ESI+) calcd. for $C_{243}H_{276}N_{36}O_{44}S_7Si_3$ $[M+2H]^{2+}$ 2355.8982, found 2355.8719.

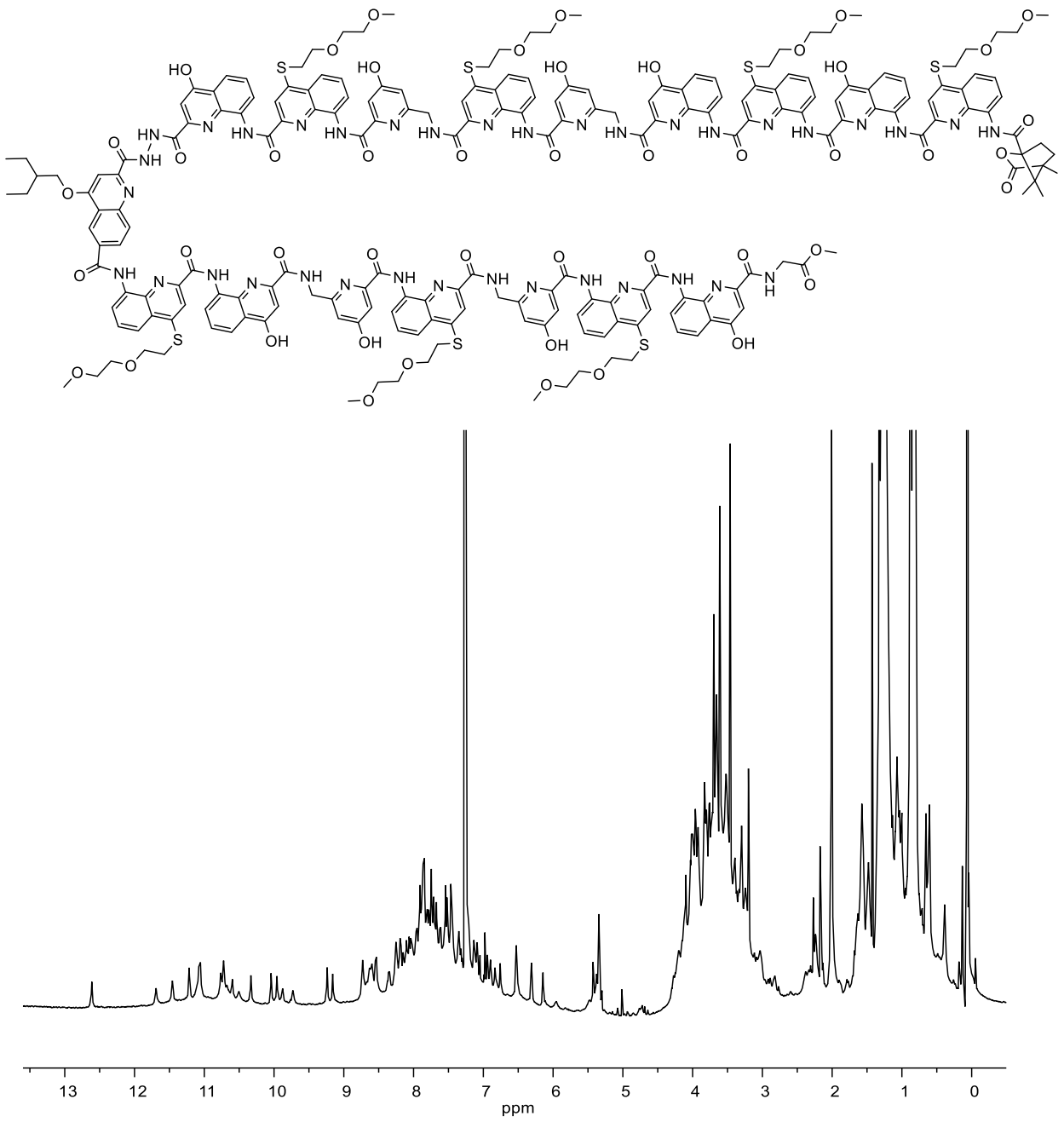
Piv-Q^{D}XQ^{D}XYQ^{D}YQ^{D}X-T2-Q^{D}XPQ^{D}YQ^{D}X-Gly-OMe (**5b**): Compound **5a** was treated with a solution of 50% TFA/DCM (2 mL) at r.t. overnight. The solvent was removed under vacuum. The solid was precipitated from MeOH and recovered by filtration. The solid was washed by MeOH for another 2 times. The compound was obtained as a yellow solid. (14 mg, quant.) **^{1}H NMR** (500 MHz, chloroform-*d*) δ 12.46 (s, 1H), 11.88 (s, 1H), 11.11 (s, 1H), 11.07 (s, 1H), 11.05 (s, 2H), 10.88 (s, 1H), 10.80 (s, 1H), 10.76 (s, 1H), 10.73 (s, 1H), 10.71 (s, 1H), 10.63 (s, 1H), 10.56 (s, 1H), 10.36 (s, 1H), 9.96 (s, 1H), 9.96 (s, 1H), 9.90 (s, 1H), 9.85 (s, 1H), 9.21 (s, 1H), 8.72 (s, 2H), 8.66 (s, 1H), 8.64–8.61 (m, 2H), 8.59 (s, 1H), 8.57 (d, *J* = 7.3 Hz, 1H), 8.54 (d, *J* = 8.4 Hz, 1H), 8.44

(d, *J* = 7.1 Hz, 1H), 8.35 (d, *J* = 8.1 Hz, 1H), 8.29 (d, *J* = 8.2 Hz, 2H), 8.21 (s, 1H), 8.17 (d, *J* = 7.0 Hz, 3H), 8.13 (s, 1H), 8.12–8.06 (m, 3H), 7.97–7.93 (m, 4H), 7.92 (s, 1H), 7.91–7.88 (m, 3H), 7.87–7.84 (m, 2H), 7.84–7.83 (m, 1H), 7.83–7.80 (m, 3H), 7.78 (d, *J* = 7.8 Hz, 2H), 7.76 (s, 2H), 7.71 (s, 1H), 7.67 (d, *J* = 8.8 Hz, 1H), 7.58 (s, 1H), 7.56 (d, *J* = 8.3 Hz, 1H), 7.53 (d, *J* = 7.7 Hz, 1H), 7.50–7.44 (m, 5H), 7.42–7.40 (m, 1H), 7.39–7.37 (m, 1H), 7.36 (d, *J* = 2.0 Hz, 1H), 7.34 (s, 1H), 7.33–7.29 (m, 3H), 7.28 (s, 1H), 7.23 (d, *J* = 9.1 Hz, 1H), 7.09–7.05 (m, 2H), 6.95 (s, 1H), 6.90 (s, 2H), 6.79 (s, 1H), 6.53 (s, 1H), 6.48 (s, 1H), 6.45 (d, *J* = 8.0 Hz, 1H), 6.29 (d, *J* = 1.4 Hz, 1H), 6.06 (s, 1H), 5.45 (s, 1H), 4.24 (s, 3H), 4.20–4.09 (m, 4H), 4.07–4.04 (m, 1H), 4.03–3.98 (m, 6H), 3.97–3.95 (m, 4H), 3.94–3.91 (m, 3H), 3.85–3.82 (m, 3H), 3.81–3.76 (m, 8H), 3.73 (d, *J* = 4.4 Hz, 3H), 3.71 (s, 3H), 3.69 (s, 1H), 3.68–3.67 (m, 2H), 3.67 (s, 2H), 3.66 (s, 1H), 3.66 (s, 3H), 3.65–3.63 (m, 3H), 3.62 (s, 3H), 3.61–3.58 (m, 2H), 3.57 (s, 3H), 3.56–3.49 (m, 8H), 3.49 (s, 3H), 3.47 (s, 3H), 3.45 (s, 3H), 3.42 (d, *J* = 4.0 Hz, 2H), 3.41–3.38 (m, 3H), 3.38–3.35 (m, 1H), 3.29 (s, 3H), 3.20 (s, 3H), 3.19–3.17 (m, 1H), 3.09 (d, *J* = 4.4 Hz, 1H), 3.07–2.99 (m, 2H), 2.87–2.83 (m, 1H), 1.65–1.54 (m, 1H), 1.28 (s, 9H), 1.27–1.22 (m, 10H). **HRMS** (ESI+) calcd. for $C_{208}H_{200}N_{36}O_{44}S_7$ $[M+2H]^{2+}$ 2065.6355, found 2065.6050.

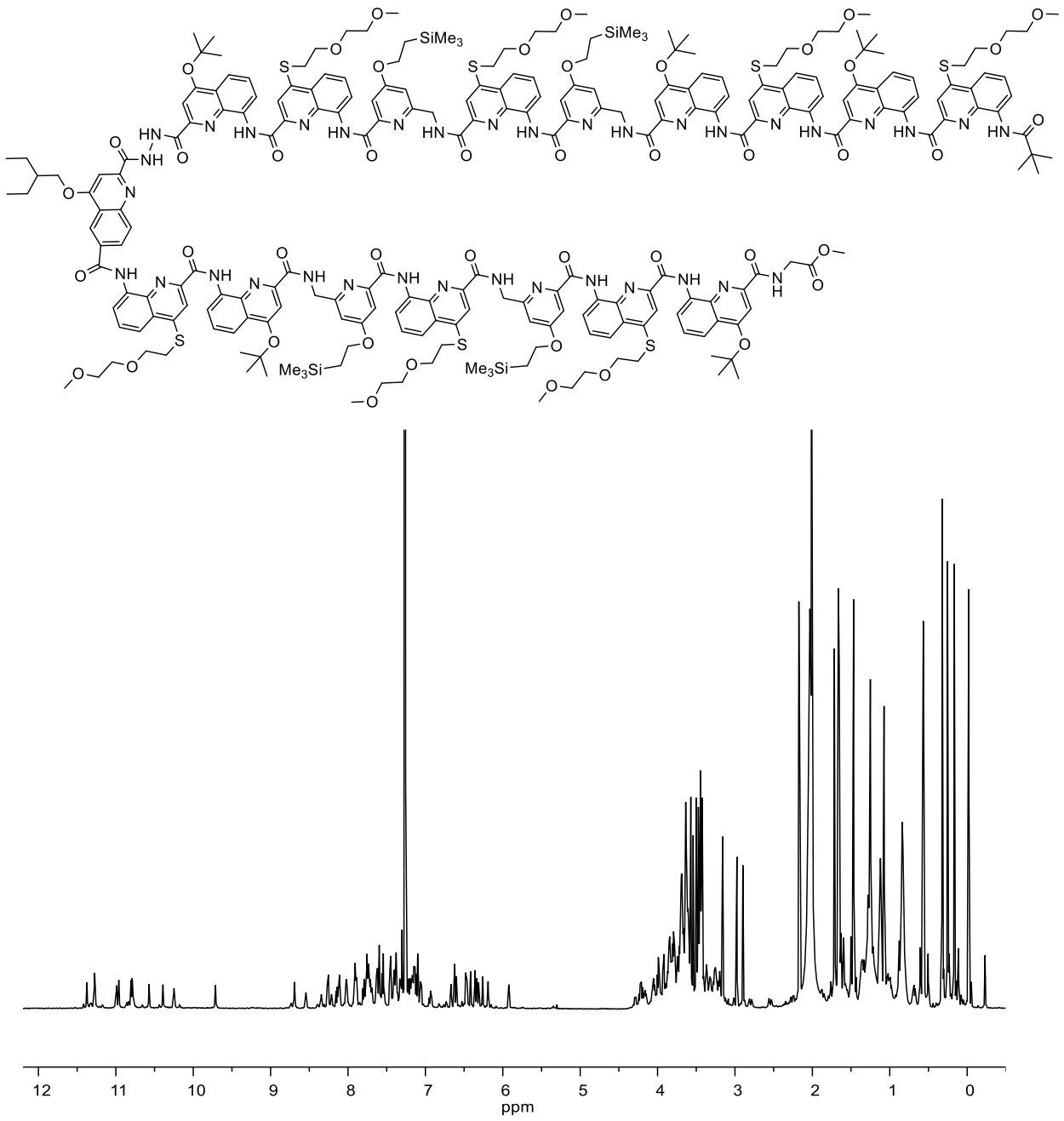
5 NMR spectra of new compounds



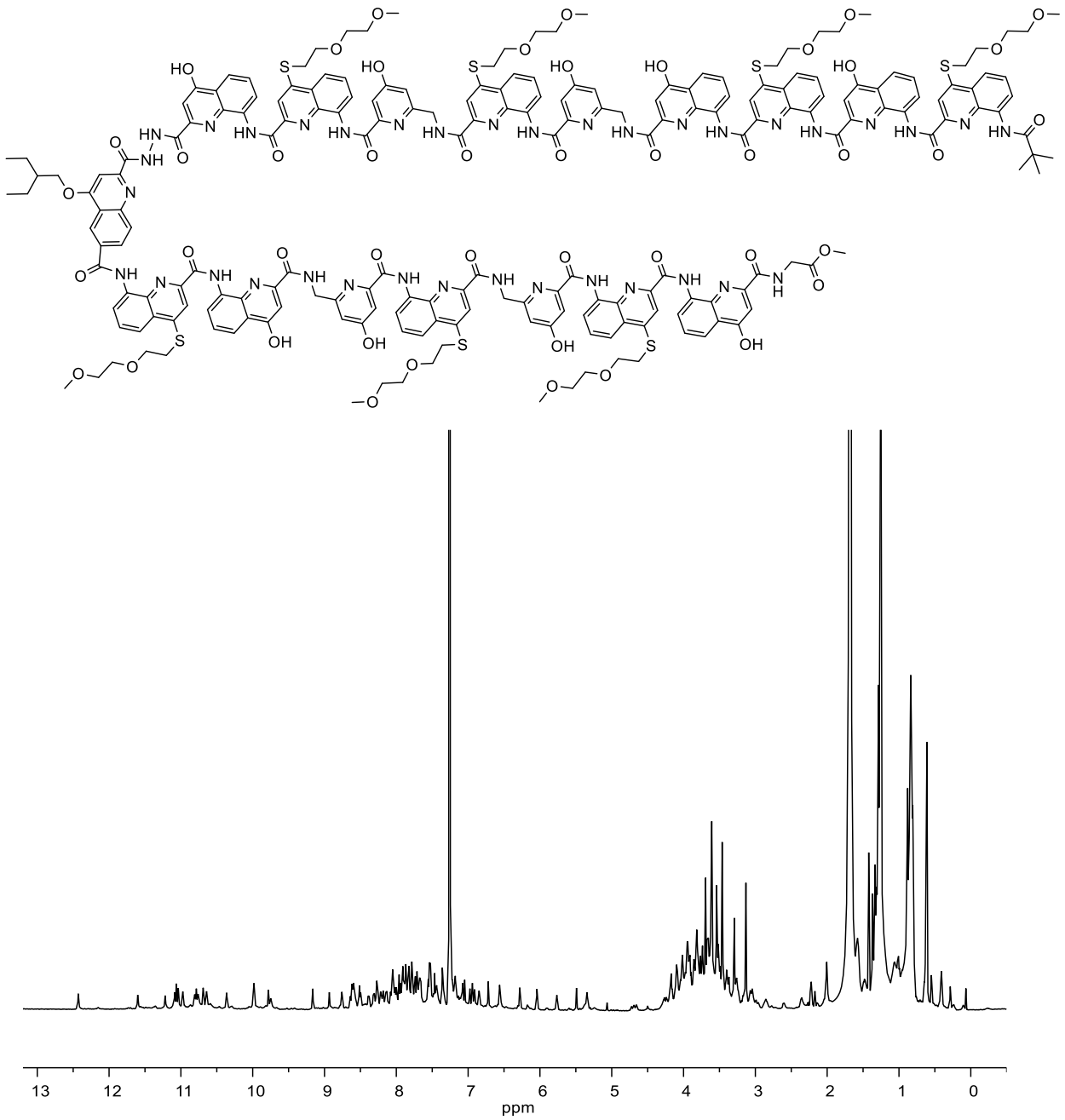
Chemical structure and ^{1}H NMR spectrum (500 MHz, $CDCl_3$, 25 °C) of **2a**.



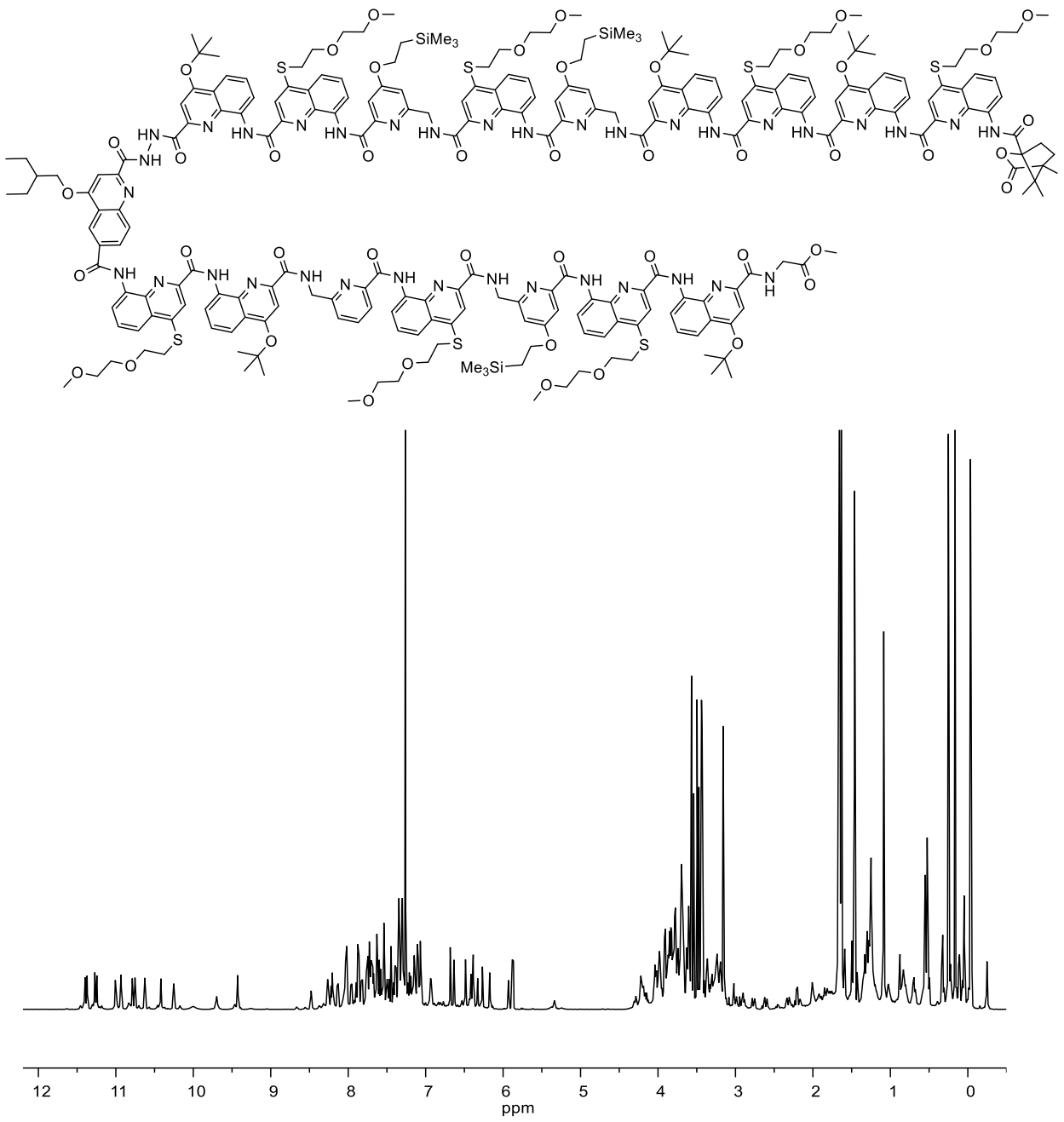
Chemical structure and ^{1}H NMR spectrum (500 MHz, $CDCl_3$, 25 °C) of **2b**.



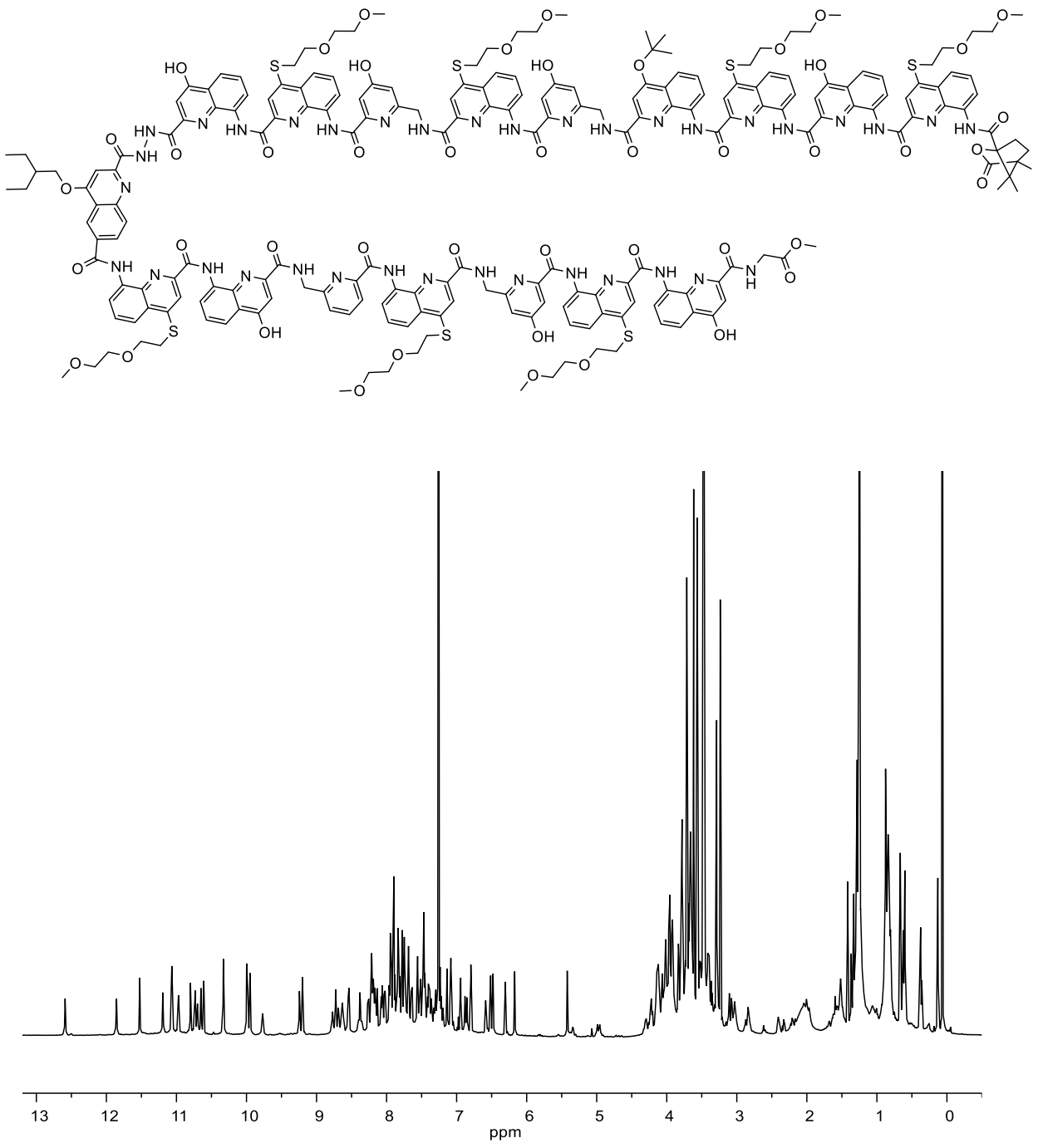
Chemical structure and ^{1}H NMR spectrum (500 MHz, $CDCl_3$, 25 °C) of **3a**.



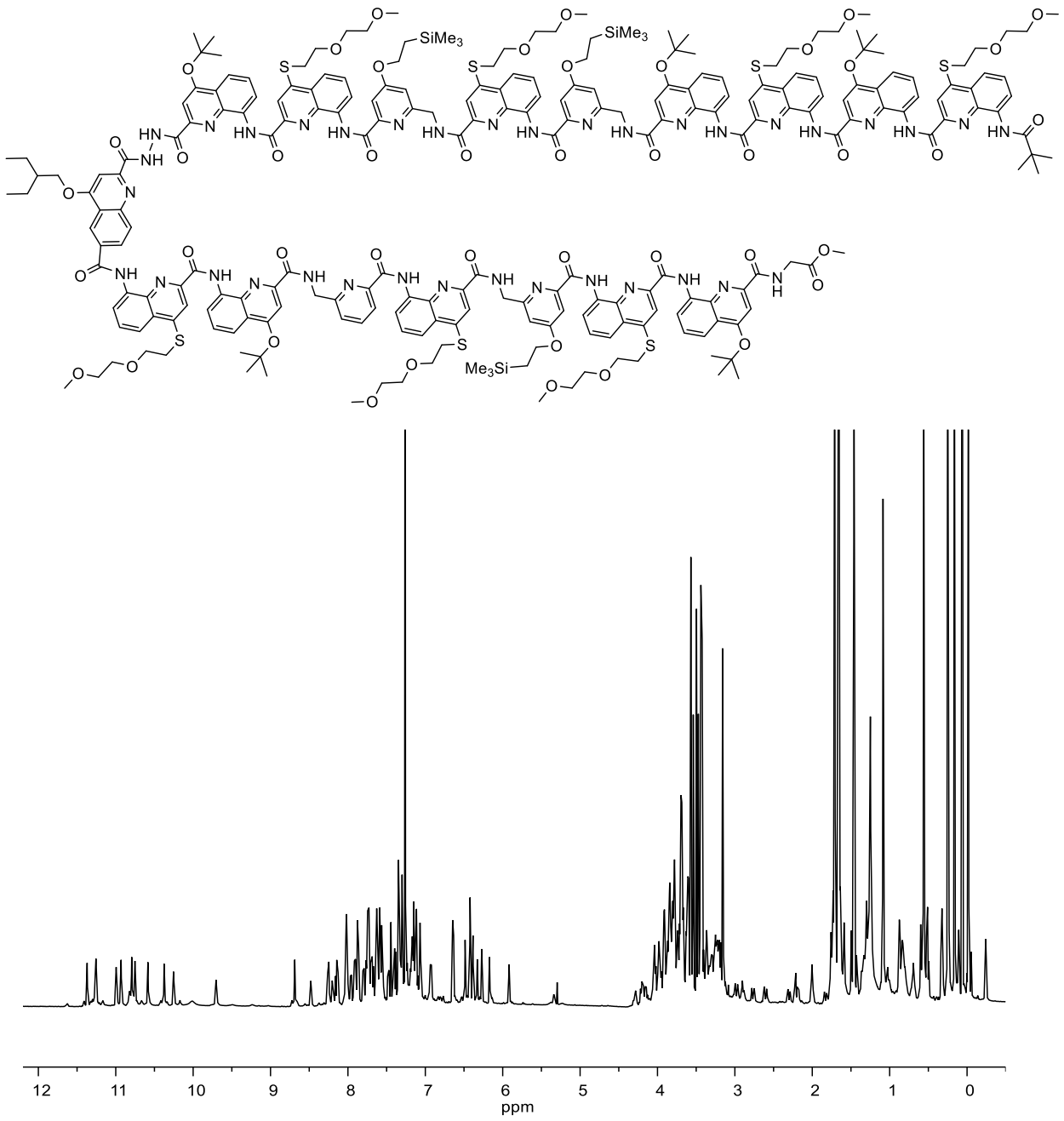
Chemical structure and ^{1}H NMR spectrum (500 MHz, $CDCl_3$, 25 °C) of **3b**.



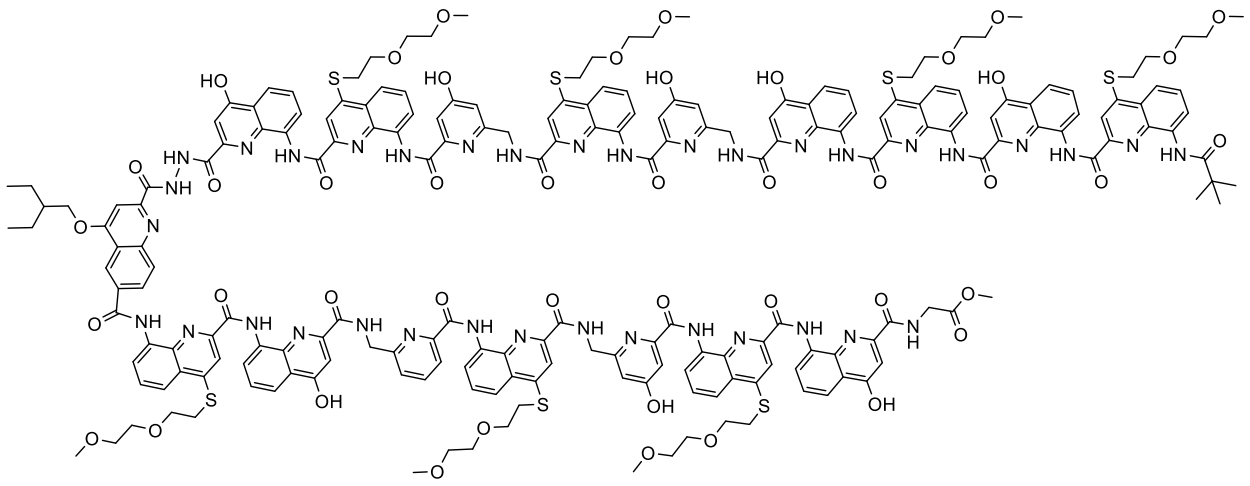
Chemical structure and ^{1}H NMR spectrum (500 MHz, $CDCl_3$, 25 °C) of **4a**.

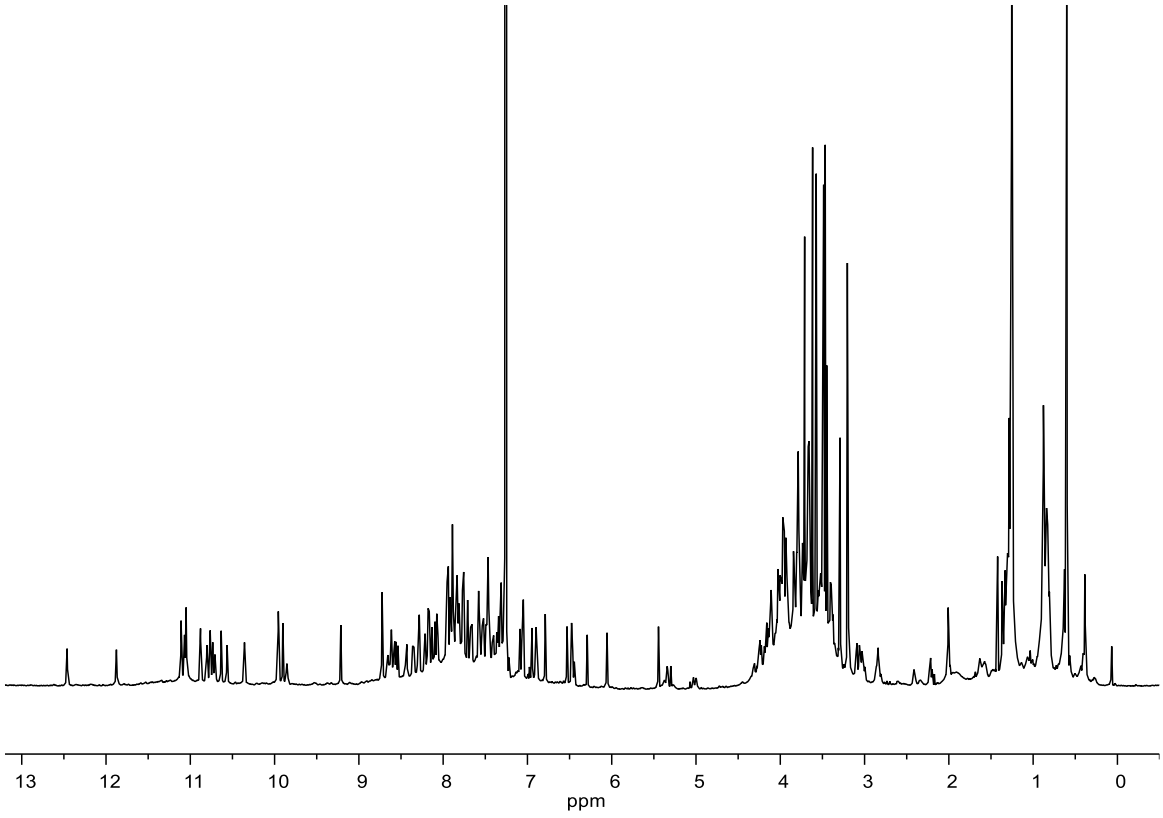


Chemical structure and ^{1}H NMR spectrum (500 MHz, $CDCl_3$, 25 °C) of **4b**.



Chemical structure and ^{1}H NMR spectrum (500 MHz, $CDCl_3$, 25 °C) of **5a**.





Chemical structure and ^{1}H NMR spectrum (500 MHz, $CDCl_3$, 25 °C) of **5b**.

5. Abiotic foldamer quaternary structures

As mentioned in Chapter 3, the absence of the T1 turn unit in the helix-T1-helix structure leads to the formation of unexpected aggregates, including tilted dimers and parallel trimers. These two distinct aggregates typically coexist in solution. Inspired by the critical role of the Y unit in the helix-T1-helix structure and considering that longer helices favor parallel arrangements to reduce the number of un-bonded hydroxy groups, several sequences were designed by removing the hydroxy groups on the Y monomer and extending its length to better control aggregate formation. The single crystal X-ray diffraction and ^{1}H NMR studies on the new sequences revealed a series of novel shifted dimer structures, which can be used for quaternary structure design. In our first attempt to drive aggregation of abiotic tertiary folds, an unexpected domain-swapped dimer structure was discovered (Chapter 4). Due to the absence of a true tertiary structure (*i.e.*, no interactions exist between two secondary subdomains), this domain-swapped dimer, despite its structural complexity, is not defined as a genuine quaternary structure. Here, we attempted to introduce hydroxy groups as H-bond donors at other position of the helix-turn-helix than the positions that led to the formation of the domain swapped dimer structure. Specifically, we intended to promote the aggregation of three tertiary folds into a C_3-symmetrical true abiotic quaternary structure through intermolecular shifted dimer interfaces.

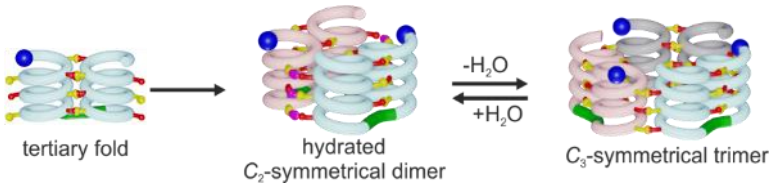


Figure 14. The cartoon representation shows the modified helix-turn-helix tertiary fold could aggregate in chloroform to produce two genuine quaternary structures which were assigned to initially designed trimer and unexpected hydrated dimer.

Our findings are summarized in a manuscript which is accepted by *Angewandte Chemie International Edition*. The results of the ^{1}H NMR analysis of the new compounds revealed two sets of signals, with their ratios changing according to temperature or water content in the chlorinated solvent. The DOSY, CD and 1D-NOESY spectra suggested that the two species are both aggregated heterochiral helix-turn-helix tertiary structures. However, the larger aggregate is favored at higher temperatures or in dried chloroform, which seems to conflict with the expected effect of entropy. A molecular model of

a hydrated dimer could be built by inserting several water molecules between the intermolecular interface to serve as a bridge. Our results suggest that the larger aggregate is the designed C_3-symmetrical trimer, while the smaller aggregate is a hydrated dimer (Figure 14). Molecular dynamics simulation of the hydrated dimer showed that water molecules have a high tendency to leave the intermolecular interface at a high temperature, leading to the formation of a shifted dimer interface and causing the angle between the two tertiary structures to approach 60°. This phenomenon was not observed at a lower temperature. The 2D-NMRs also helped us identify specific NOE correlations between CH_2/OH within both aggregates, which highly supports our conclusion.

Contributions: The project was planned by I. Huc. Synthetic monomer precursors have been provided by D. Gill. Monomer synthesis have been performed by me, F. Menke and D. Gill. Foldamer synthesis was performed by me. C. Glas, L. Allmendinger and me carried out NMR measurements. The measurement of CD spectra was performed by me. The LC-MS analysis was performed by C. Douat. Me, I. Huc and L. Allmendinger contributed to experiment design and interpretation. The research was supervised by I. Huc. The manuscript was written by me and I. Huc. Me, I. Huc and L. Allmendinger proofread and improved the manuscript. This work was supported the China Scholarship Council (CSC, predoctoral fellowship to S. W.).

5.1. Publication

Abiotic foldamer quaternary structures

Authors: Shuhe Wang, Lars Allmendinger and Ivan Huc*

Published: *Angew. Chem. Int. Ed.* **2024**, e202413252. (doi/10.1002/anie.202413252)

How to cite: *Angew. Chem. Int. Ed.* **2024**, e202413252
doi.org/10.1002/anie.202413252

Supramolecular Chemistry

Abiotic Foldamer Quaternary Structures

*Shuhe Wang, Lars Allmendinger, and Ivan Huc**

Abstract: Abiotic aromatic foldamer sequences have been previously shown to fold in helix-turn-helix motifs in organic solvents. Using simple computational tools, a new helix-turn-helix motif was designed that bears additional hydrogen bond donor OH groups to promote its aggregation into a genuine, trimeric, abiotic quaternary structure. This sequence was synthesized and its self-assembly in solution was investigated by Nuclear Magnetic Resonance (NMR), Circular Dichroism (CD) and Molecular Dynamics (MD) simulations. The existence of two stable discrete aggregates was evidenced, one assigned to the initially designed trimer, the other to a dimer including multiple water molecules. The two species may be quantitatively interconverted upon changing the water content of the solution or the temperature. These results represent important steps in the design of protein-like abiotic architectures.

A current challenge in foldamer research is to move beyond isolated non-natural molecular helices, turns or linear strands and combine several of these secondary elements into tertiary and quaternary structures. This challenge stems from the hope that sophisticated functions will emerge from more complex objects, as it is the case for proteins. In this regard, computational tools[1,2] and other rational[3–7] or evolutionary approaches[8–10] to produce non-natural proteins have made great progress. Including non-natural units within α-peptidic structures has also been demonstrated.[11–13] However, complex folded objects at the exclusion of natural monomers remain rare. Some helix bundles based on β-peptides,[14,15] oligoureas,[16,17] or abiotic aromatic amino acids[18,19] have been described. Related to the latter, abiotic helix-turn-helix tertiary structures such as that of sequence **1** (Figure 1b) have been designed.[18,20–23] To the best of our knowledge, genuine quaternary structures, i.e. assemblies of tertiary structures are lacking.[24] For instance, a recent attempt to design a quaternary structure yielded instead an unexpected domain-swapped dimer (DSD) in which the initial tertiary fold was disrupted.[25] Here, we report what we believe to be the first two examples of abiotic quaternary structures, one the result of rational design and the other of discovery.

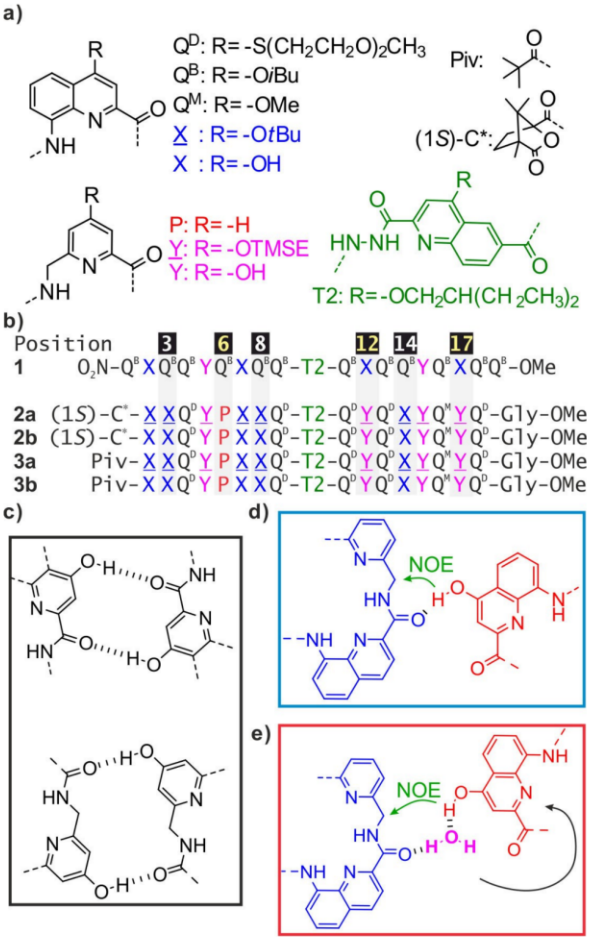


Figure 1. a) Formulas of Q^B, Q^D, Q^M, X, P, Y and T2 amino acid monomers as well as N-terminal Piv and (1*S*)-C* groups. X and Y are the protected precursors of X and Y, respectively. TMSE = 2-trimethylsilylethyl. Q^B, Q^D, Q^M and T2 carry organic solubilizing side chains. b) Oligoamide foldamer sequences. Gly stands for glycine. In **1**, the nitro group at the N-terminus replaces the NH group. The unit numbering of **1** was adopted for all sequences, i.e. **2b** starts with residue 2. c) Hydrogen bonding patterns involving X and Y units within the intramolecular helix-helix interface of e.g. **1**. d) Intermolecular hydrogen bonding pattern mediating the shifted-dimer helix-helix interface used to program the assembly of $(\mathbf{2b})_3$. e) Hypothetical insertion of water in the pattern shown in d). The black arrow highlights the change of position of the X unit with respect to its position in d). Green arrows highlight observed intermolecular NOE's.

[*] S. Wang, Dr. L. Allmendinger, Prof. Dr. I. Huc
Department of Pharmacy
Ludwig-Maximilians-Universität München
Butenandtstr. 5–13, 81377 München (Germany)
E-mail: ivan.huc@cup.lmu.de

Sequence **1**[21] comprises two δ-peptidic segments consisting of 8-amino-2-quinolinecarboxylic acid units (Q, X in Figure 1a) and 6-aminomethyl-2-pyridinecarboxylic acid units (P, Y in Figure 1a). Both segments adopt stable helical conformations in organic solvents following the well-established folding rules of aromatic oligoamide foldamers (Figure S1 in the Supporting Information).[26,27] Q and X units are more rigid and template the folding of the more flexible Y units.[28] Essential for tertiary structure design, X and Y bear 4-hydroxy substituents that can hydrogen bond to amide carbonyl groups and mediate helix-helix associations (Figure 1c). The two helical segments are connected by a T2 turn unit that has been designed to promote a stable helix-turn-helix tertiary fold stabilized by six inter-helix intramolecular hydrogen bonds. T2 also mediates a reversal of helix handedness. When the N-terminal helix of **1** is right-handed (*P*), the C-terminal helix is left-handed (*M*), and vice versa. A cartoon representation of the *PM* structure of **1** is shown in Figure 2a and its crystal structure in Figure S2.

The tertiary fold of **1** appears to be a prime candidate for the introduction of additional hydroxy groups to promote intermolecular interactions and quaternary structure formation. In a first attempt, we introduced hydroxy groups at positions 4, 9 and 13 of the sequence of **1** to create a new intermolecular hydrogen-bonded interface. However, this *ab initio* design based on molecular modelling was not observed experimentally. As mentioned above, a DSD formed instead, leading to the disruption of the tertiary fold.[25] This dimerization entailed a *gauche* ~90° torsion about the N–N bond of T2 instead of the flat *anti* conformation observed in the structure of **1** (Figure S2).[21] The two conformations resulted in characteristic ^{15}N NMR chemical shift values.

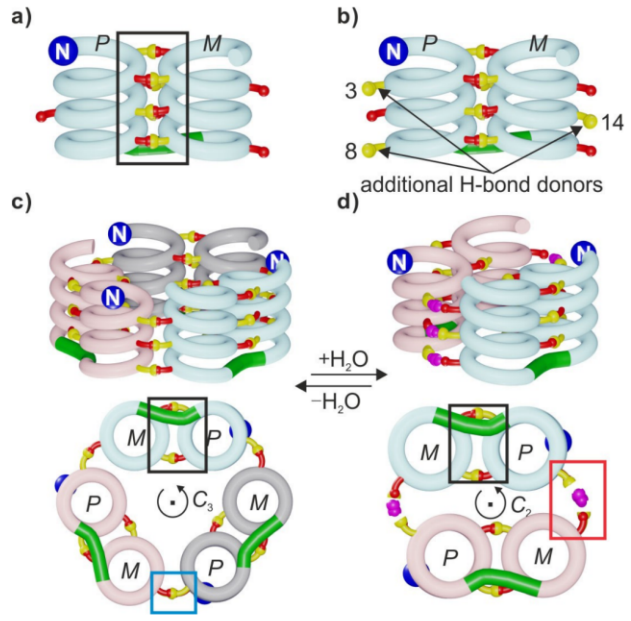


Figure 2. a) Schematic representation of a helix-turn-helix tertiary structure with a reversal of helix handedness. b) Modification of the structure in a) so as to create additional complementary arrays of hydrogen bonds. Schematic representations of: c) a C_3-symmetrical trimer, and d) a C_2-symmetrical dimer with six water molecules (colored in purple) bridging the hydrogen bonding interface. The N-terminus is marked with a blue ball. The T2 linker is shown in green. Red and yellow balls indicate hydrogen bond acceptors (carbonyl oxygen atoms) and donors (hydroxy protons), respectively. Intramolecular interfaces are highlighted with a black box. Intermolecular interfaces within the trimer and the hypothetical water-mediated dimer are highlighted with cyan and red boxes, respectively.

In a new attempt, we planned to exploit an already characterized hydrogen bond-mediated intermolecular helix-helix association termed "shifted-dimerization".[19] In $CDCl_3$, hydrogen bonding between hydroxy groups borne by X units and carbonyl groups linked to P units (Figure 1d) promotes the head-to-tail association of a *P* helix and an *M* helix (see Figure S3 for details).[19] Since the structure of **1** already consists of a *P* helix and an *M* helix arranged head-to-tail, albeit within a tertiary fold, intermolecular shifted dimerization was predicted to promote the discrete aggregation of **1**. For that purpose, hydroxy groups were introduced at positions 3, 8 and 14 of **1**, that is, on both its N-terminal and C-terminal helices, so as to create complementary arrays of hydrogen-bonding donors and acceptors at the periphery of the tertiary structure (Figure 2b). Shifted dimerization also required the mutation of a Q unit into P and of X units into Y at positions 6, 12 and 17. New sequences **2b** and **3b** bearing all these features were synthesized (see Supporting Information for experimental details). They differ only through their N terminus. Sequence **2b** bears a (1*S*)-camphanyl group that quantitatively biases the handedness of its N-terminal helix to *P*, and thus of its C-terminal helix to *M*.[21,29] In contrast, achiral **3b** exists as a racemic mixture of *PM* and *MP* conformers.

A molecular model of two *PM* conformers of **2b** associated by a shifted dimerization interface (i.e. with the *P* helix of one molecule and the *M* helix of the other molecule) was energy-minimized using Maestro.[30] It showed that the two molecules were at an angle of 54° (Figure S4). This value is close to the 60° that would be found in a cyclic trimer (*PM*-**2b**)$_3$ involving three shifted intermolecular interfaces. Indeed, such a C_3-symmetrical trimer could be built with minimal distortion with respect to the dimer structure, i.e. with presumably minimal energy penalty (Figures 2c, 3a and S4). Of note, the *PM*-**3b** and *MP*-**3b** conformational enantiomers of achiral sequence **3b** may have more ways to assemble. Solution studies were thus started with chiral sequence **2b** which is exclusively *PM*.

The ^{1}H NMR spectra of hydroxy group-protected precursors **2a** and **3a** in $CDCl_3$ were similar to that of the precursor of **1**,[21] showing a mixture of a minor and a major diastereomeric conformer e.g. *PM* and *PP* for **2a** (Figures 4a, S5). After side chain deprotection, the spectrum of chiral sequence **2b** in $CDCl_3$ "from the bottle" showed two sets of sharp signals indicating the presence of two well-defined species (Figure 2c). In carefully dried $CDCl_3$ (over Al_2O_3, note that "dried" does not imply rigorously dry) only one species was observed, whereas in wet $CDCl_3$ (after shaking with water) only the other species was observed (Figures 2b, 2d). Intermediate proportions were obtained upon using mixtures of wet and dried chloroform (Fig-

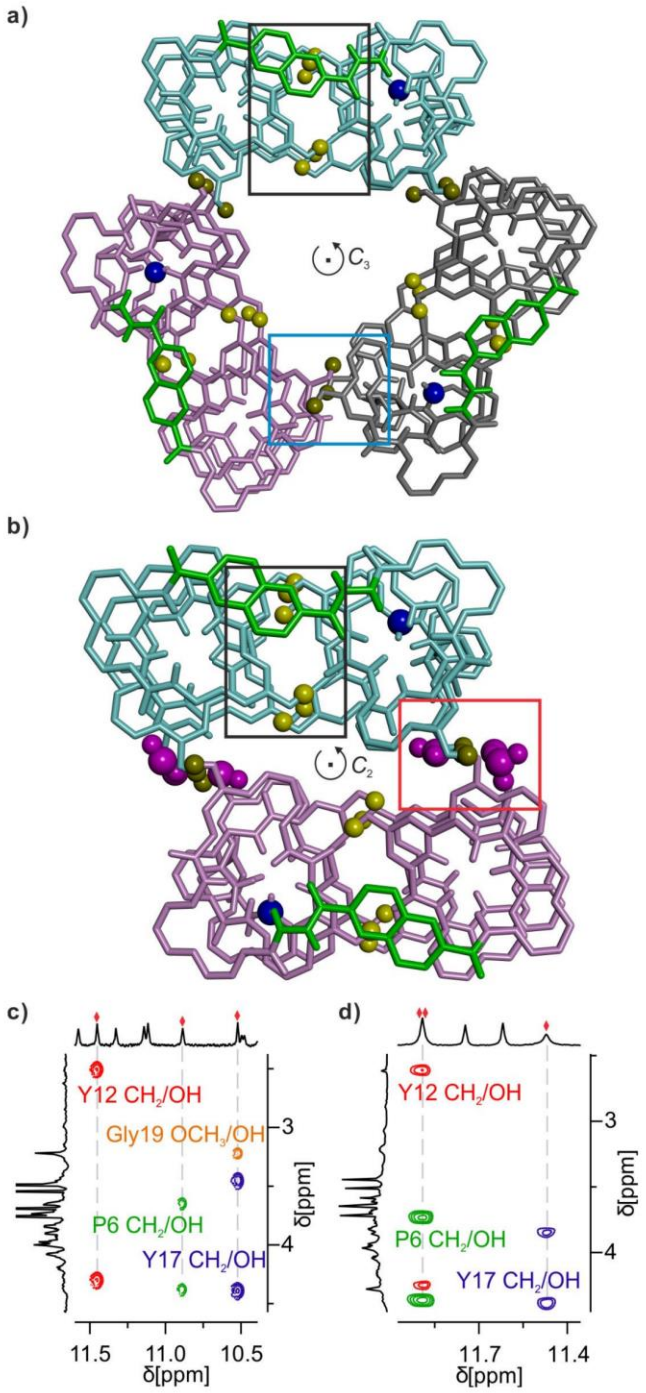


Figure 3. Simplified bottom views of energy-minimized models of: a) (*PM*-**2b**)$_3$; and b) (*PM*-**2b**)$_2$(H_2O)$_6$. Water molecules bridging the intermolecular interface are colored in purple. Individual sequences are shown in light blue, pink and grey tube representation. Hydroxy protons are shown as yellow balls. T2 linkers are shown in green. For clarity, only the outer rim of the helices is shown. The side chains of Q and T2 have been omitted. The black, cyan and red boxes outline the same areas as in Figure 2. c) Excerpt from the ^{1}H,^{1}H NOESY spectrum (500 MHz, 298 K) of **2b** in dried $CDCl_3$ showing intermolecular NOE's between O<u>H</u> protons (red diamonds) and diastereomeric pairs of C<u>H</u>$_2$ protons of P6, Y12 and Y17 as well as the C-terminal OC<u>H</u>$_3$. d) Same as (c) in wet $CDCl_3$.

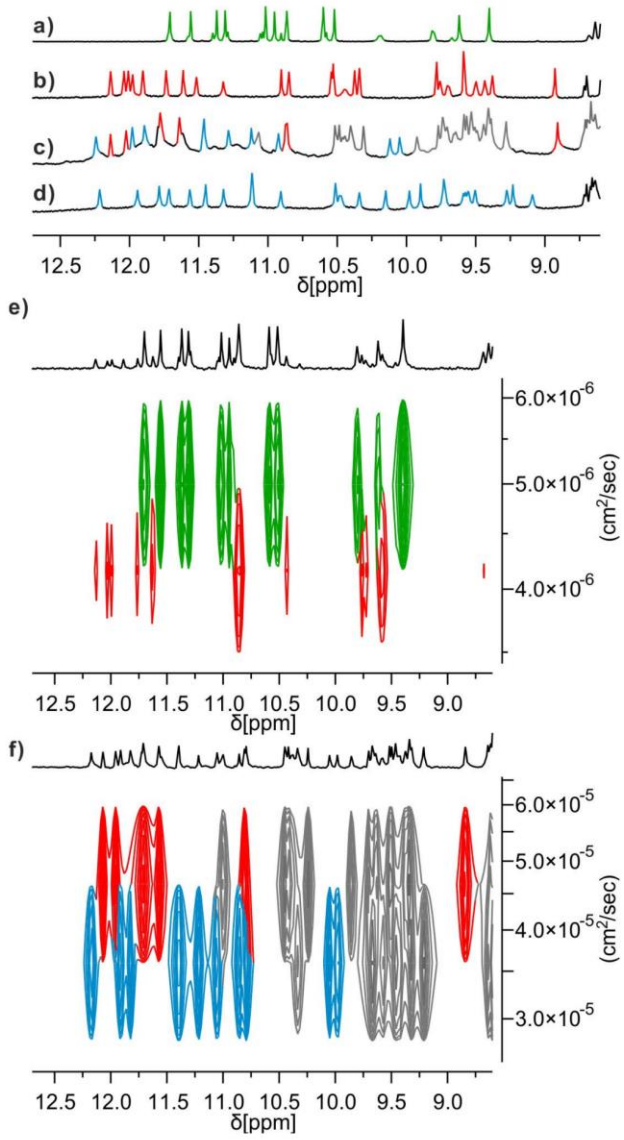


Figure 4. Extract of ^{1}H NMR spectra (500 MHz, 298 K) of **2a** (a), **2b** in wet $CDCl_3$ (b), **2b** in dry $CDCl_3$ (d) and "normal" $CDCl_3$ (c). e) DOSY spectrum of a mixture of **2a** and **2b** in wet $CDCl_3$. Correlations corresponding to compound **2a** are highlighted in green and those to compound **2b** are highlighted in red. f) DOSY spectrum of **2b** in which two different species coexist. Signals that correspond to two different species are highlighted in red and blue, respectively.

ure S6). In contrast, adding comparable amounts of dry methanol marginally altered the equilibrium. The ^{1}H,^{15}N Heteronuclear Single Quantum Coherence (HSQC) spectra allowed for the assignment of OH resonances as exchangeable protons that do not correlate with ^{15}N (Figures S7, S8, S9). The OH signals resonate between 7.9 ppm and 12.0 ppm, indicating their involvement in hydrogen bonds of variable strength, and remain sharp even in the wet sample, indicating slow exchange with water. Since all hydrogen bonds are unlikely to be fulfilled intramolecularly, one can presume that both species are aggregates. In agreement, Diffusion Ordered Spectroscopy (DOSY) experiments showed that the species formed by **2b** in wet $CDCl_3$ is

significantly larger than its monomeric precursor **2a**, i.e. that it must be dimeric or larger. Conversely, DOSY showed that the aggregate of **2b** in dried $CDCl_3$ is significantly larger than its aggregate in wet $CDCl_3$, i.e. that it must be at least trimeric (Figure 4e, f, see also Figure S10). Diluting a solution of the trimeric aggregate in dried $CDCl_3$ down to 20 μM showed no sign of dissociation indicating the robustness of this species in the absence of water.

The effect of temperature on the aggregation of **2b** corroborated the hypothesis that water is involved in the smaller aggregate. In chloroform from the bottle, the larger aggregate was quantitatively produced at higher temperatures ($\geq$308 K) and the smaller aggregate was quantitatively produced at lower temperature ($\leq$253 K) (Figure S11). The temperature at which the transition occurs (at which signals for both species have the same intensity) can be modulated from above 318 K to below 223 K upon varying the water content of the solution (Figures S12, S13). Such a dependence on temperature seemingly conflicts with the expected effect of entropy, but it can be explained if the smaller aggregate (the one that has the smaller diffusion coefficient according to DOSY) in fact has the higher molecularity because it contains both **2b** and water molecules. The exact number of water molecules is not known but it must be high enough for temperature dependence to be so pronounced. That an almost quantitative interconversion of dimer and trimer occurs over a range of only ca. 35 K indicates that the associated entropic term must be quite large (Figure S11).

Next we gathered information about the nature and structure of the aggregates. With the instruments at our disposal, we could not obtain evidence about molecularity by mass spectrometry. The multiplicity of the NMR signals clearly demonstrates that, within each aggregate, all **2b** molecules are equivalent. The aggregates are thus symmetrical. While the sequence of **2b** was designed to produce a C_3-symmetrical trimer (Figures 2c, 3a), we were unable to produce a reasonable model for a symmetrical aggregate with four or more molecules. We thus think it reasonable to propose that the larger aggregate of **2b** is the designed C_3-symmetrical trimer and that the smaller aggregate is a dimer $(\mathbf{2b})_2\cdot(H_2O)_n$. We can infer that $n>1$ but its value could not be determined experimentally. Even at 223 K, water seems to undergo rapid tumbling within the aggregate or exchange between the solution and the aggregate.

A plausible model of a hydrated C_2-symmetrical dimer could be proposed upon inserting water molecule bridges within each hydrogen bonds of the shifted interfaces (Figures 1e, 2d, 3b). This amounts to using the array of hydrogen bond as a hinge and reducing the angle between the molecules of **2b** to 0° in $(\mathbf{2b})_2\cdot(H_2O)_6$, compared to 60° in $(\mathbf{2b})_3$. Such insertion of several water molecules have been observed previously in related foldamer assemblies,[23] as well as in self-assembled calix[4]resorcinarene-based capsules.[31] To assess its stability, the energy-minimized model of $(\mathbf{2b})_2\cdot(H_2O)_6$, was subjected to 1 ns long MD simulations at 200, 300, 400 and 500 K using $CHCl_3$ as an implicit solvent (see Supporting Information). Increasing temperature was found to result in the dissociation of water molecules (all of them dissociate at 500 K) but the dimer **2b** kept together even if all intermolecular hydrogen bonds could not be established in the absence of bridging water molecules (Figure S14). Another simulation was run at 400 K after having removed all three OH groups and the three water molecules of one of the two intermolecular hydrogen-bonded interfaces of $(\mathbf{2b})_2\cdot(H_2O)_6$. Upon the release of the water molecules from the other interface during the MD run, the dimer was found to open like a Jackknife, the angle between the two molecules of **2b** finally adjusting back to 54° (Figure S15) and supporting the role of a hinge of the shifted intermolecular interface depending on whether or not it involves bridging water molecules (Figure S15).

Crystals of **2b** were obtained but did not diffract sufficiently for a crystallographic structure elucidation so we relied on NMR and CD to assess the structure of its aggregates in solution. The CD spectra of **2b** showed relatively weak intensity and low temperature dependence (Figures S16, S17) in agreement with the helix handedness reversal mediated by T2 which makes the contributions of the two helices largely cancel each other. The characteristic ^1H and ^{15}N NMR chemical shift values of the diacyl hydrazine group of T2 (9.0–10.0 ppm for ^1H, 125–130 ppm for ^{15}N) indicate that it adopts the flat *anti* conformation found in **1** in both the small and large aggregates (Figures S7, S8). For comparison, these resonances were found at 8.0–8.5 ppm for ^1H and 110–120 ppm for ^{15}N in the *gauche* conformer of the DSD.[25] Also consistent with the *anti* conformation, correlations between these two N<u>H</u> protons are seen by $^1H,^1H$ Correlation Spectroscopy (COSY) but not by $^1H,^1H$ Nuclear Overhauser Effect Spectroscopy (NOESY). Overall, this suggests that **2b** adopts its U-shape helix-turn-helix conformation in its aggregates.

The $^1H,^1H$ NOESY spectrum of **2b** allowed for the unambiguous assignment of all amide N<u>H</u> signals in both aggregates (Figures S19, S22) and corroborated the conformation of T2 (Figure S23). Scalar couplings between N<u>H</u> protons and neighbor C<u>H</u>$_2$ protons in Y and P then allowed for the assignment of the later as diastereomeric pairs (Figures S18, S21). NOESY spectra showed that the C<u>H</u>$_2$ pairs of protons at positions 6, 12 and 17, each correlate with an O<u>H</u> proton in both aggregates (Figure 3c, d). Distances within the helix-turn-helix motifs indicate that such correlations cannot be achieved intramolecularly. However, they are fully consistent with the shifted dimer intermolecular interface whether or not a water molecule has been inserted (Figure 1d, e) and strongly support the models shown in Figures 2 and 3.

Further experiments involved varying the solvent and using **3b**, the achiral analogue of **2b**. The ^1H NMR spectra of **2b** in wet and dried CD_2Cl_2 showed single species. Using solvent mixtures, these species were shown to be the same as the species observed in wet and dried $CDCl_3$, respectively (Figures S24, S25). This came in part as a surprise because the nature of the chlorinated solvent in principle alters the shifted dimerization. CD_2Cl_2 may favor a shifted dimer between helices having the same handedness, instead of the

P helix/*M* helix favored in $CDCl_3$ and used for the design of $(\mathbf{2b})_3$ (Figure S3).[19] The absence of perturbation of the quaternary structures in CD_2Cl_2 may be the result of their inherent stability and of the fact that the effect of CD_2Cl_2 is weaker on shifted interfaces with fewer hydrogen bonds as in **2b**.[19] Next, and consistent with earlier experiments of the same kind,[21,25] the progressive addition of DMSO-d_6, a solvent that competes with hydrogen bonding, to a solution of **2b** in $CDCl_3$ led to the successive appearance of new species that could be assigned to the monomeric tertiary fold of **2b** (at 4–12 vol % of DMSO-d_6) and then to the disrupted (i.e. partially unfolded) monomeric **2b** (at 20–36 vol % of DMSO-d_6) (Figure S26).

Finally, in wet $CDCl_3$, the ^{1}H NMR spectrum of **3b** is very similar to that of **2b** indicating the formation of a single aggregate and suggesting that *PM* and *MP* conformational enantiomers of **3b** undergo narcissistic self-sorting into $(PM\text{-}\mathbf{3b})_2\cdot(H_2O)_n$ and $(MP\text{-}\mathbf{3b})_2\cdot(H_2O)_n$ (Figures S27, S28). In contrast, the spectra of **3b** in dried $CDCl_3$ and CD_2Cl_2 are broad to the extent that no signals can be distinguished (Figures S29, S30), suggesting some sort of polymerization. In the absence of self-sorting, shifted dimerization may occur between each type of *P* helix (from either *PM* and *MP* conformer) with each type of *M* helix leading to multiple aggregates (Figure S31). Polymerization also occurs in wet $CDCl_3$ upon heating (Figure S32).

In summary, we have promoted the aggregation of the helix-turn-helix tertiary fold of **1** upon adding hydrogen bond donors at precise locations at its periphery and without any sign of perturbation of the conformation it adopts as a monomer. The structures of the resulting well-defined dimeric and trimeric quaternary structures are strongly supported by CD and NMR data and molecular modelling, and their proportions can be quantitatively controlled by varying temperature or the amount of water in $CDCl_3$. Further elaboration aiming at preventing dimer formation, at promoting V-shaped heteromeric dimers (trimers in which one unit is missing), at increasing the number of hydrogen bonds involved in the intermolecular contacts, and at exploiting the large cavity of the trimer is in progress and will be reported on in due course.

Supporting Information

The authors have cited additional references within the Supporting Information.[32–34]

Acknowledgements

We acknowledge financial support from the China Scholarship Council (CSC, predoctoral fellowship to S. W.) and thank Victor Maurizot and Alexandru Grozavu for useful discussions. Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

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

Keywords: abiotic foldamer · hydrogen bonding · quaternary structure · tertiary structure · self-assembly

[1] L. An, G. R. Lee, *Curr. Protoc. Protein Sci.* **2020**, *102*, e116.
[2] C. Norn, B. I. Wicky, D. Juergens, S. Liu, D. Kim, D. Tischer, B. Koepnick, I. Anishchenko, F. Players, D. Baker, *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2017228118.
[3] A. Ljubetič, F. Lapenta, H. Gradišar, I. Drobnak, J. Aupič, Ž. Strmšek, D. Lainšček, I. Hafner-Bratkovič, A. Majerle, N. Krivec, *Nat. Biotechnol.* **2017**, *35*, 1094–1101.
[4] F. Lapenta, J. Aupič, Ž. Strmšek, R. Jerala, *Chem. Soc. Rev.* **2018**, *47*, 3530–3542.
[5] E. A. Naudin, K. I. Albanese, A. J. Smith, B. Mylemans, E. G. Baker, O. D. Weiner, D. M. Andrews, N. Tigue, N. J. Savery, D. N. Woolfson, *Chem. Sci.* **2022**, *13*, 11330–11340.
[6] W. M. Dawson, K. L. Shelley, J. M. Fletcher, D. A. Scott, L. Lombardi, G. G. Rhys, T. J. LaGambina, U. Obst, A. J. Burton, J. A. Cross, *Nat. Commun.* **2023**, *14*, 383.
[7] A. T. Hilditch, A. Romanyuk, S. J. Cross, R. Obexer, J. J. McManus, D. N. Woolfson, *Nat. Chem.* **2024**, *16*, 89–97.
[8] I. Anishchenko, S. J. Pellock, T. M. Chidyausiku, T. A. Ramelot, S. Ovchinnikov, J. Hao, K. Bafna, C. Norn, A. Kang, A. K. Bera, *Nature* **2021**, *600*, 547–552.
[9] L. Cao, B. Coventry, I. Goreshnik, B. Huang, W. Sheffler, J. S. Park, K. M. Jude, I. Marković, R. U. Kadam, K. H. Verschueren, *Nature* **2022**, *605*, 551–560.
[10] N. Ferruz, S. Schmidt, B. Höcker, *Nat. Commun.* **2022**, *13*, 4348.
[11] M. Bejger, P. Fortuna, M. Drewniak-Świtalska, J. Plewka, W. Rypniewski, Ł. Berlicki, *Chem. Commun.* **2021**, *57*, 6015–6018.
[12] M. W. Giuliano, S. J. Maynard, A. M. Almeida, L. Guo, I. A. Guzei, L. C. Spencer, S. H. Gellman, *J. Am. Chem. Soc.* **2014**, *136*, 15046–15053.
[13] a) J. R. Santhouse, J. M. Leung, L. T. Chong, W. S. Horne, *Chem. Sci.* **2022**, *13*, 11798–11806; b) Z. E. Reinert, G. A. Lengyel, W. S. Horne, *J. Am. Chem. Soc.* **2013**, *135*, 12528–12531.
[14] E. J. Petersson, C. J. Craig, D. S. Daniels, J. X. Qiu, A. Schepartz, *J. Am. Chem. Soc.* **2007**, *129*, 5344–5345.
[15] E. J. Petersson, A. Schepartz, *J. Am. Chem. Soc.* **2008**, *130*, 821–823.
[16] G. W. Collie, K. Pulka-Ziach, C. M. Lombardo, J. Fremaux, F. Rosu, M. Decossas, L. Mauran, O. Lambert, V. Gabelica, C. D. Mackereth, *Nat. Chem.* **2015**, *7*, 871–878.
[17] G. W. Collie, R. Bailly, K. Pulka-Ziach, C. M. Lombardo, L. Mauran, N. Taib-Maamar, J. Dessolin, C. D. Mackereth, G. Guichard, *J. Am. Chem. Soc.* **2017**, *139*, 6128–6137.
[18] S. De, B. Chi, T. Granier, T. Qi, V. Maurizot, I. Huc, *Nat. Chem.* **2018**, *10*, 51–57.
[19] F. S. Menke, B. Wicher, V. Maurizot, I. Huc, *Angew. Chem. Int. Ed.* **2023**, *62*, e202217325. *Angew. Chem.* **2023**, *135*, e202217325.
[20] D. Mazzier, S. De, B. Wicher, V. Maurizot, I. Huc, *Chem. Sci.* **2019**, *10*, 6984–6991.

[21] D. Mazzier, S. De, B. Wicher, V. Maurizot, I. Huc, *Angew. Chem. Int. Ed.* **2020**, *59*, 1606–1610. *Angew. Chem.* **2020**, *132*, 1623–1627.
[22] F. S. Menke, D. Mazzier, B. Wicher, L. Allmendinger, B. Kauffmann, V. Maurizot, I. Huc, *Org. Biomol. Chem.* **2023**, *21*, 1275–1283.
[23] F. S. Menke, B. Wicher, L. Allmendinger, V. Maurizot, I. Huc, *Chem. Sci.* **2023**, *14*, 3742–3751.
[24] The term quaternary structure has been improperly used in the literature when describing e.g. helix bundles. Helix bundles lack a hierarchical component in that they do not contain tertiary-like substructures and may thus not be called quaternary.
[25] S. Wang, B. Wicher, C. Douat, V. Maurizot, I. Huc, *Angew. Chem. Int. Ed.* **2024**, *63*, e202405091; *Angew. Chem.* **2024**, *136*, e202405091.
[26] I. Huc, *Eur. J. Org. Chem.* **2004**, *2004*, 17–29.
[27] D.-W. Zhang, X. Zhao, J.-L. Hou, Z.-T. Li, *Chem. Rev.* **2012**, *112*, 5271–5316.
[28] D. Sanchez-Garcia, B. Kauffmann, T. Kawanami, H. Ihara, M. Takafuji, M.-H. Delville, I. Huc, *J. Am. Chem. Soc.* **2009**, *131*, 8642–8648.
[29] A. M. Kendhale, L. Poniman, Z. Dong, K. Laxmi-Reddy, B. Kauffmann, Y. Ferrand, I. Huc, *J. Org. Chem.* **2011**, *76*, 195–200.
[30] Maestro, Schrödinger, *LLC*, New York, NY, **2021**.
[31] L. R. MacGillivray, J. L. Atwood, *Nature* **1997**, *389*, 469–472.
[32] V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat, I. Huc, *Chem. Eur. J.* **2023**, *29*, e202300898.
[33] M. Vallade, P. Sai Reddy, L. Fischer, I. Huc, *Eur. J. Org. Chem.* **2018**, *2018*, 5489–5498.
[34] O. Al Musaimi, A. Basso, B. G. de la Torre, F. Albericio, *ACS Comb. Sci.* **2019**, *21*, 717–721.

Manuscript received: July 14, 2024
Accepted manuscript online: September 4, 2024
Version of record online: ■■, ■■

5.2. Supplementary Information

For:

Abiotic foldamer quaternary structures

Shuhe Wang, Lars Allmendinger and Ivan Huc

Table of Contents

1. List of Abbreviations

CD	circular dichroism
COSY	correlated spectroscopy
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DCM	dichloromethane
DIPEA	*N*,*N*-diisopropylethylamine
DMF	*N*,*N*-dimethylformamide
DMSO	dimethyl sulfoxide
DOSY	diffusion ordered spectroscopy
DSD	domain swapping dimer
eq.	equivalent
Fmoc	fluorenylmethoxycarbonyl
HBTU	2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HR-ESI	high resolution electrospray ionization
HSQC	heteronuclear single quantum correlation
Me	methyl
MeOH	methanol
min	minutes
MS	mass spectrometry
MW	microwave
NMP	*N*-methyl-2-pyrrolidone
NMR	nuclear magnetic resonance
NOESY	nuclear overhauser effect spectroscopy
r.t.	room temperature
SPS	solid phase synthesis
***t*Bu**	*tert*-butyl
TFA	trifluoroacetic acid
THF	tetrahydrofuran
UV/Vis	ultraviolet–visible

2. Supplementary figures

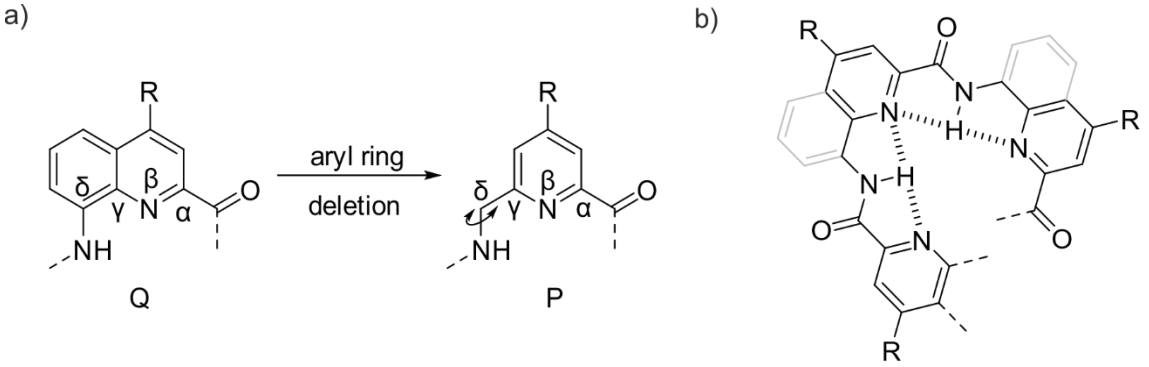


Figure S2. Chemical structure of Q and P units and the folding principle of their oligomers. a) Chemical structure of Q (or X with R = OH) and P (or Y with R = OH). b) Intramolecular H-bonding and helical folding principle of P/Q oligomers. Note that the amide carbonyl groups diverge from the folded structures and thus provide hydrogen bond acceptors.

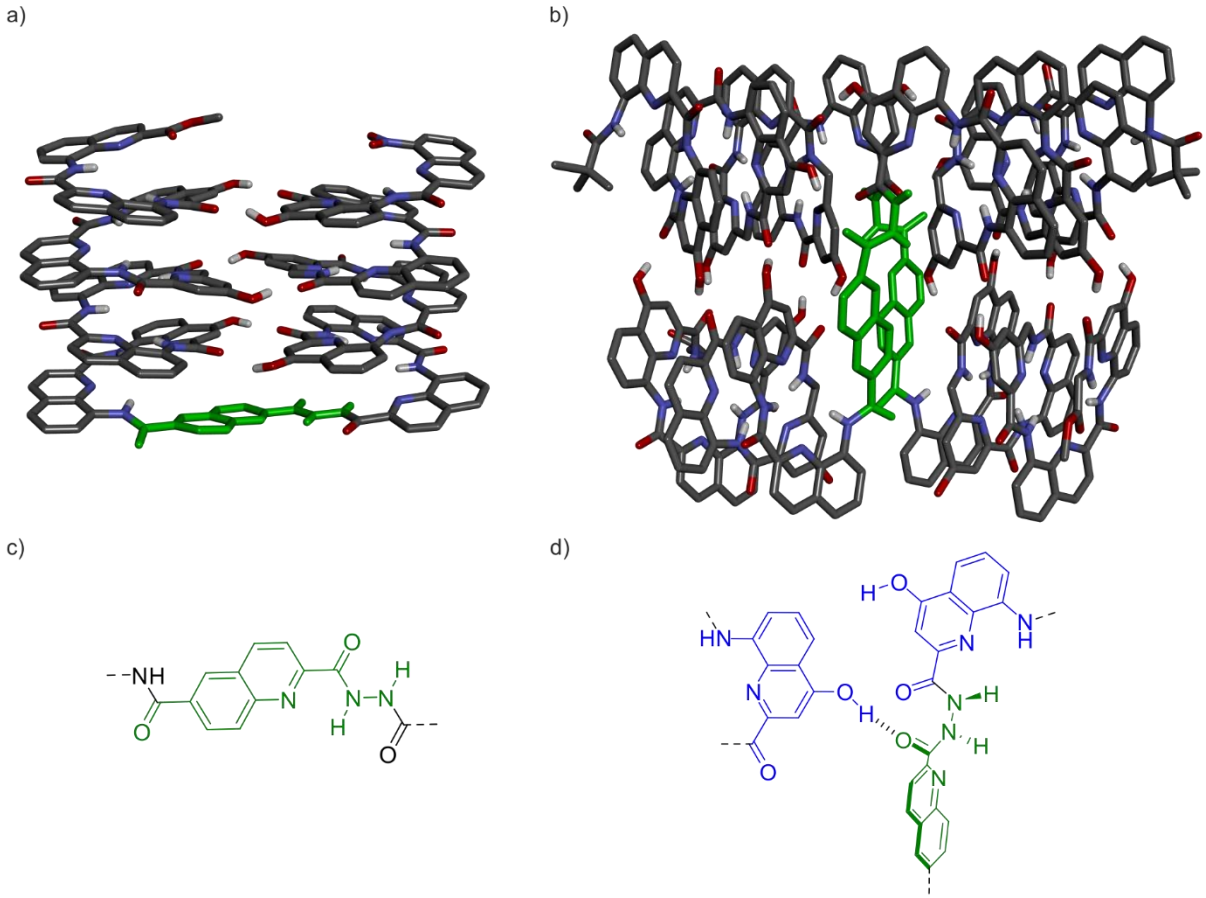


Figure S2. Crystal structures of helix-T2-helix and DSD folds showing two different conformations of T2. Crystal structures of: a) the helix-T2-helix fold of sequence 1[21]; and b) the DSD.[25] The side chains of Q^D and

Q^B and T2 are omitted for clarity, T2 is colored in green. c) The flat anti conformation of T2 observed in the structure of sequence 1. d) The gauche ~90° torsion about the N-N bond of T2 observed in the DSD.

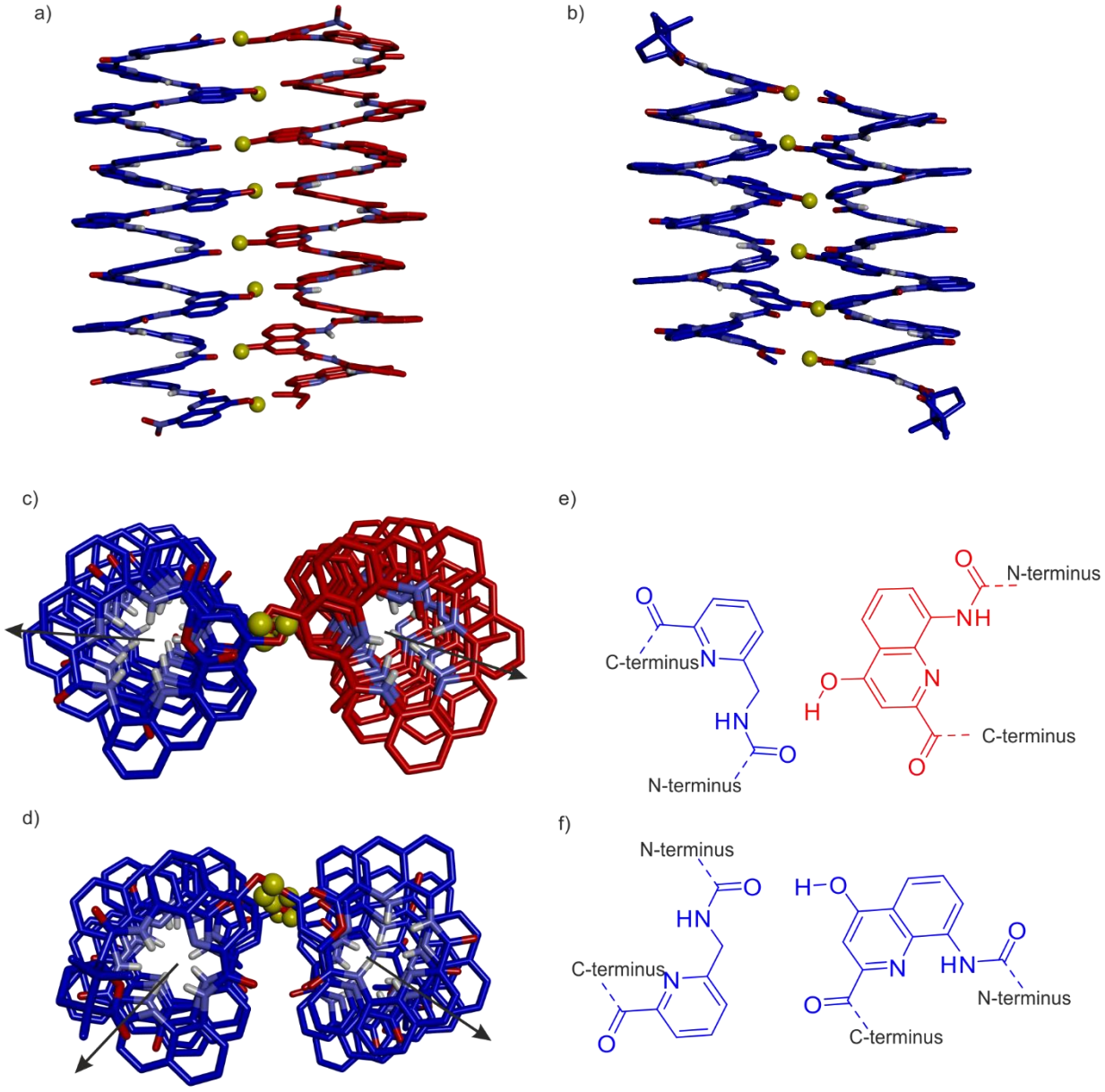


Figure S3. Crystal structures of homochiral and heterochiral shifted dimers.[19] Side (a) and top (b) views of a heterochiral shifted dimer structure. Side (b) and top (d) views of a homochiral shifted dimer structure. The protons of OH hydrogen bond donors are shown as yellow balls. P helices are colored in blue and M helix are colored in red. The arrows in (c) and (d) show that the relative angles of two interfaces are different. H-bond patterns of the heterochiral shifted dimer (e) and the homochiral shifted dimer (f). The units that belong to P helices are colored in blue and the unit belonging to an M helix is colored in red.

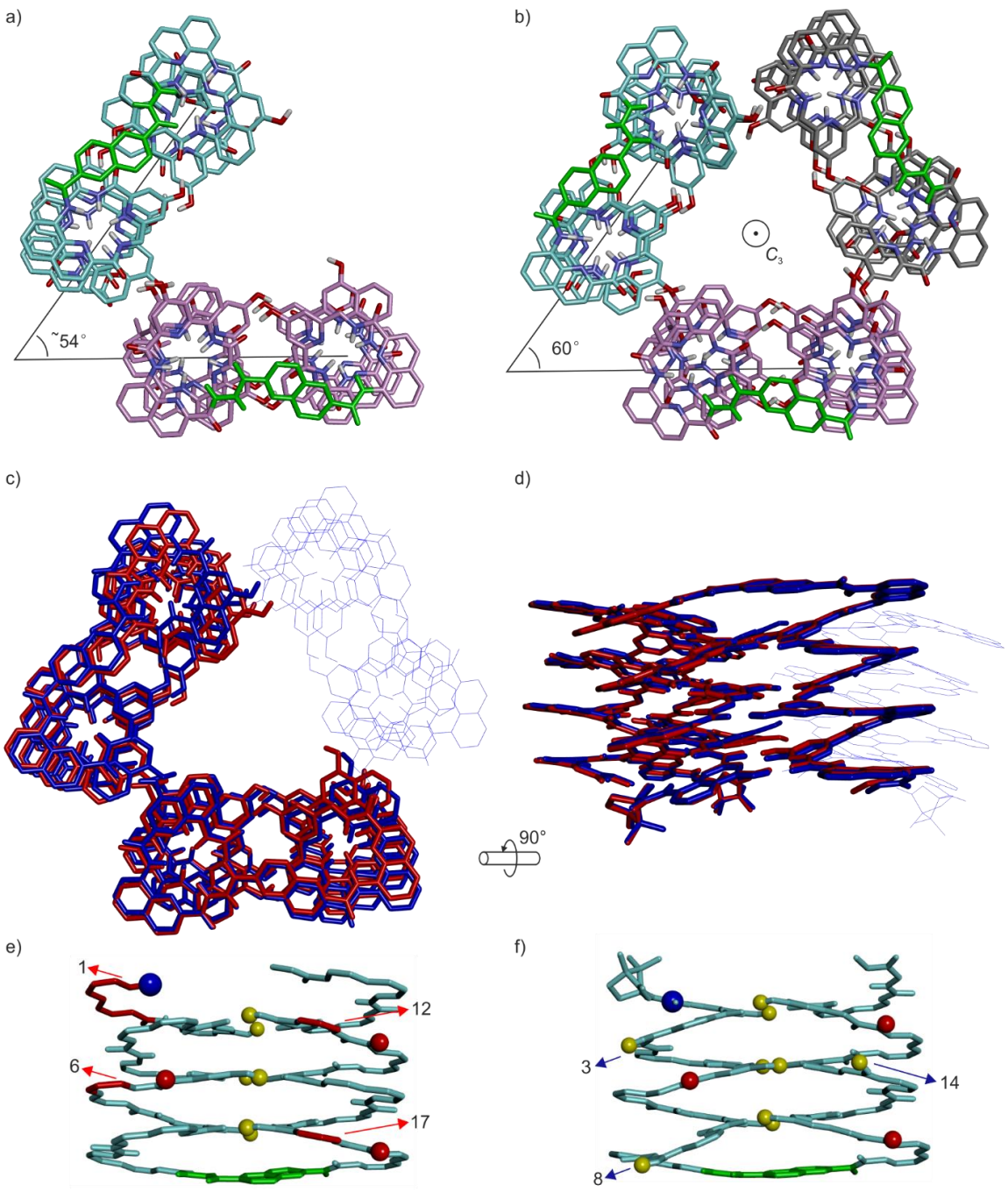


Figure S4. Quaternary structure design. a) Energy minimized model of $(2b)_2$ (MMFFs, TNCG). The intermolecular interface consists of a heterochiral shifted dimer pattern. The angle between two helix-turn-helix structures is around 54°. A slight conformational change of the intermolecular interface with presumably minimal energy penalty enables the formation of a C_3-symmetrical trimer with an exact 60° angle between two tertiary folds. b) Energy minimized model of $(2b)_3$ (MMFFs, TNCG). The C_3 axis is at the centre of the trimer. Identical sequences are colored in pink, cyan and grey, respectively. T2 is colored in green. Bottom view (c) and side view (d) of the overlay of molecular models in (a) and (b) colored in red and blue, respectively. The third tertiary

structure in the model of $(2b)_3$ is shown in a blue colored line representation for clarity. e) Crystal structure of sequence 1, highlighting the aromatic rings (colored in red) that were removed in sequences 2b and 3b by Q (or X) to P (or Y) mutation. f) Molecular model of sequence 2b highlighting the H-bond donors that were introduced for intermolecular interactions.

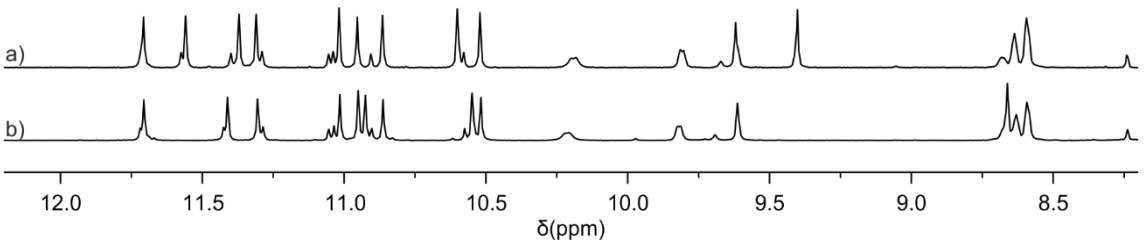


Figure S5. Extracts of ^{1}H NMR spectra (500 MHz, $CDCl_3$) of hydroxy group-protected precursors 2a (a) and 3a (b). In both cases, there is a major and a minor species corresponding to two different conformers.

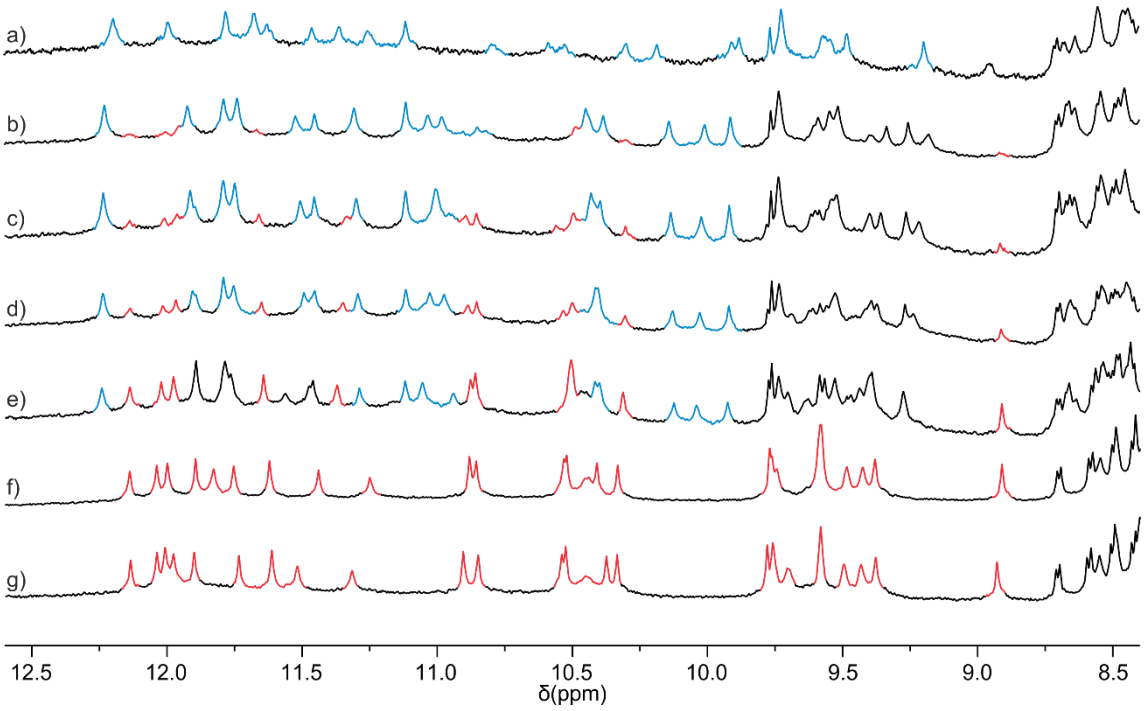


Figure S6. Extracts of ^{1}H NMR spectra (500 MHz 0.35 mM) of 2b in different wet $CDCl_3$ and dried $CDCl_3$ mixtures. The proportions (%, v/v) of wet $CDCl_3$ in dried $CDCl_3$ are 0 (a), 10 (b), 15 (c), 20 (d), 30 (e), 50 (f) and 100 (g). By increasing the amount of wet $CDCl_3$, a quantitative transition was observed, indicating that the red-colored species contains water molecules. For the preparation of wet/dried $CDCl_3$ see section 5.1.

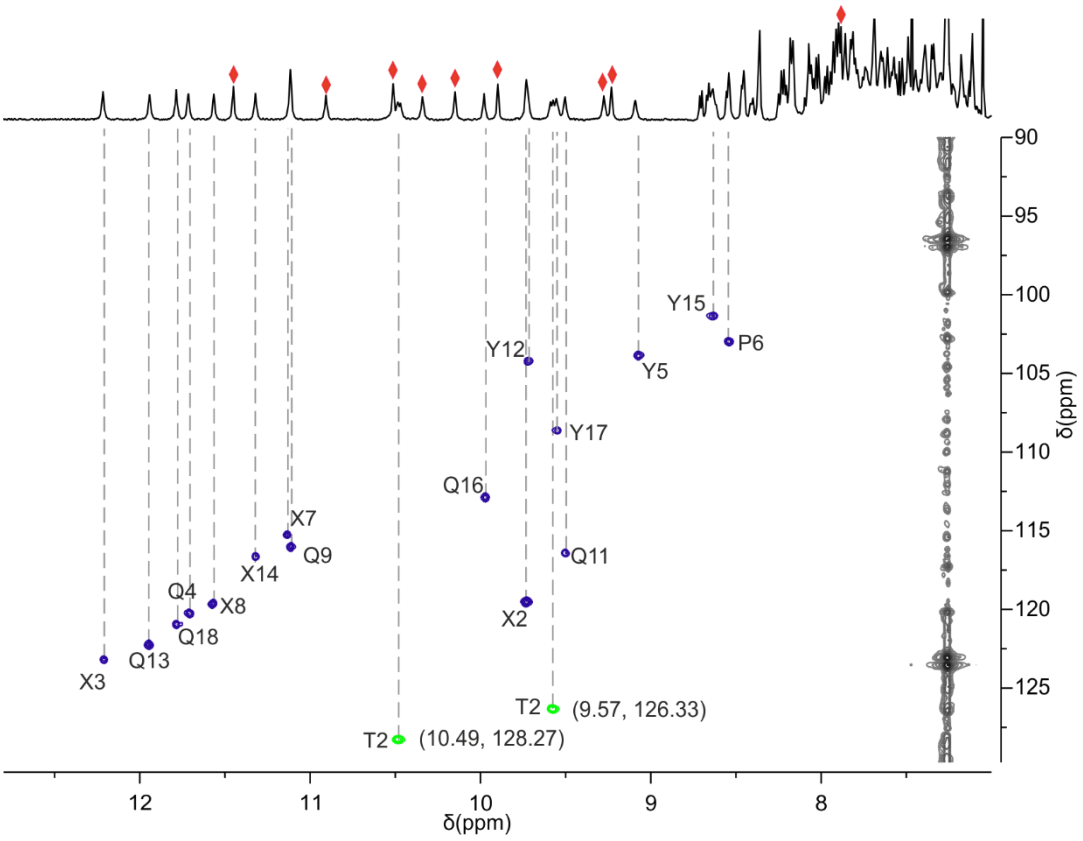


Figure S7. Extract of the $^{1}H,^{15}N$ HSQC spectrum (500 MHz, dried $CDCl_3$) of 2b. Only NH resonances are correlated. The green colored correlations were assigned to the NH-NH signals of T2 unit by $^{1}H,^{1}H$ COSY, which showed the corresponding vicinal NH-NH coupling (Figure S18). No NOE interaction of these two protons was observed suggesting the diacyl hydrazine group adopting a flat anti conformation (Figure S19). The correlation of NH signals of the glycine is presumably hidden because it overlaps with the signal of $CHCl_3$. Red diamonds indicate the signals that were assigned to hydrogen-bonded O<u>H</u> protons. The OH signal that has a chemical shift value of δ 7.88 ppm was assigned to Y5-O<u>H</u> by $^{1}H,^{1}H$ NOESY spectrum (Figure S20), indicating a weaker hydrogen bond. For the assignment of NH signal see section 3.5.

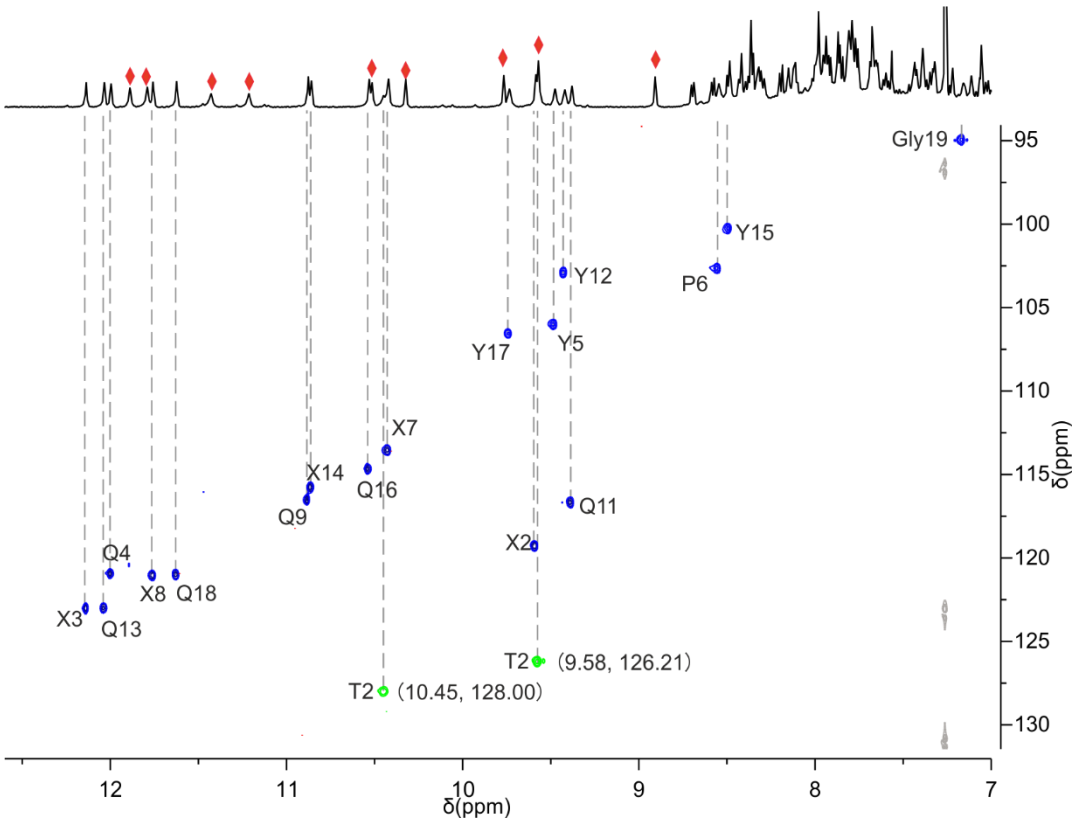


Figure S8. Extract of $^{1}H,^{15}N$ HSQC spectrum (500 MHz, wet $CDCl_3$) of 2b. Only NH resonances are correlated. The green correlations were assigned to the NH-NH signals of T2 by $^{1}H,^{1}H$ COSY which showed the corresponding NH-NH coupling (Figure S21). No NOE interaction of these two protons among each other was observed suggesting that the diacyl hydrazine group adopts a flat anti conformation (Figure S22). Red diamonds indicate the signals assigned to hydrogen-bonded OH protons. For the assignment of NH signal see 3.5.

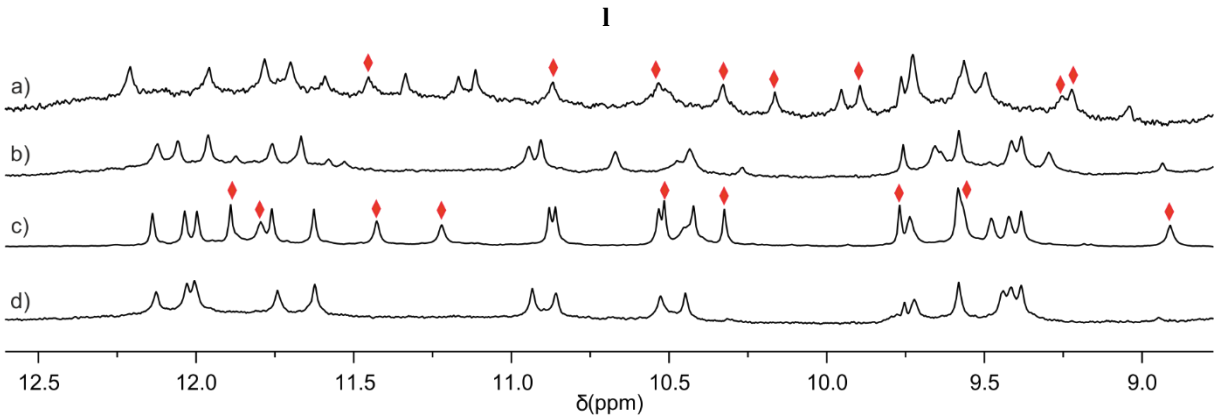


Figure S9. Deuterium exchange experiment. a) Extract of the ^{1}H NMR spectrum of 2b (500 MHz, dried $CDCl_3$). The signals assigned to OH protons are marked with a red dot. b) Extract of ^{1}H NMR spectrum of 2b (500 MHz, 0.5%CD_3OD/dried $CDCl_3$). Due to the slight shifts in the peaks caused by the change in solvent composition, the

exchange of O<u>H</u> assigned to Y5 (Figure S20) cannot be clearly assigned. c) Extract of the ^{1}H NMR spectrum of 2b (500 MHz, wet $CDCl_3$). The signals assigned to O<u>H</u> protons are marked with a red dot. d) Extract of ^{1}H NMR spectrum of 2b (500 MHz, D_2O saturated $CDCl_3$). For the preparation of D_2O saturated $CDCl_3$ see section 5.1.

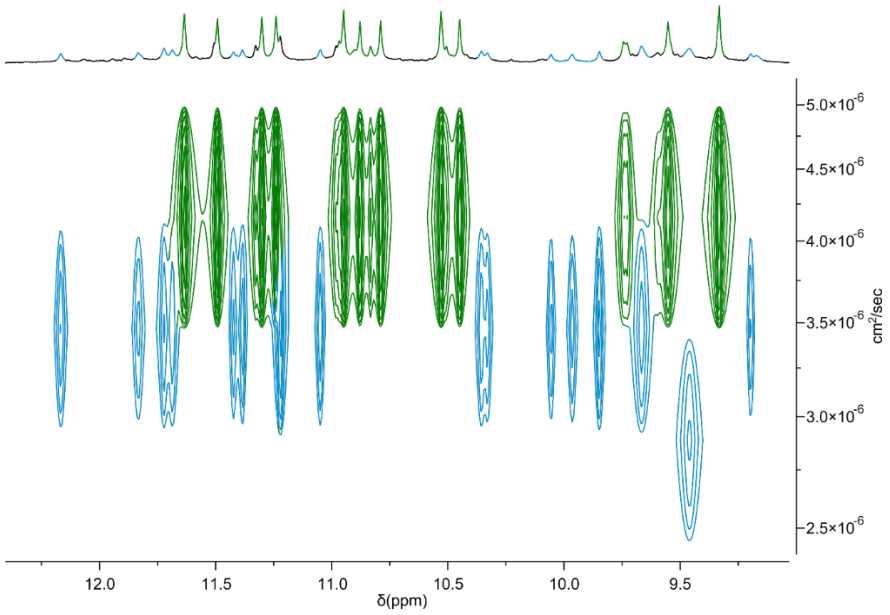


Figure S10. Extract of a ^{1}H DOSY spectrum (500 MHz, dried $CDCl_3$) of a 4 to 1 mixture of 2a (colored in green) and 2b (colored in blue). The blue-colored signals indicate a species that has a higher hydrodynamic radius although its molecular weight is smaller.

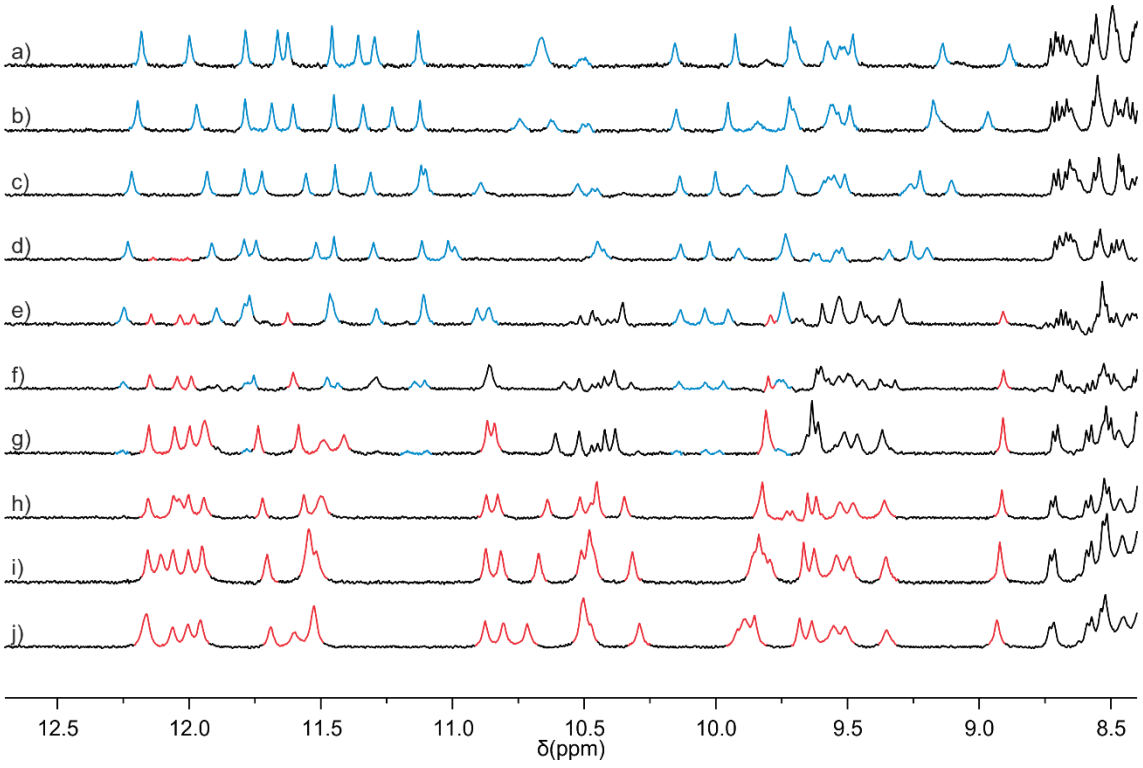


Figure S11. Extracts of ^{1}H NMR spectra (400 MHz, $CDCl_3$ from the bottle, 2.2 mM) of 2b at different temperatures. The temperatures were 328 K (a), 318 K (b), 308 K (c), 298 K (d), 283 K (e), 273 K (f), 263 K (g), 253 K (h), 243 K (i) and 233 K (j). At higher temperatures, 2b aggregates as a trimer and releases water molecules from the intermolecular interfaces. At lower temperatures, water molecules will bridge the H-bond interfaces, leading to the formation of the dimer.

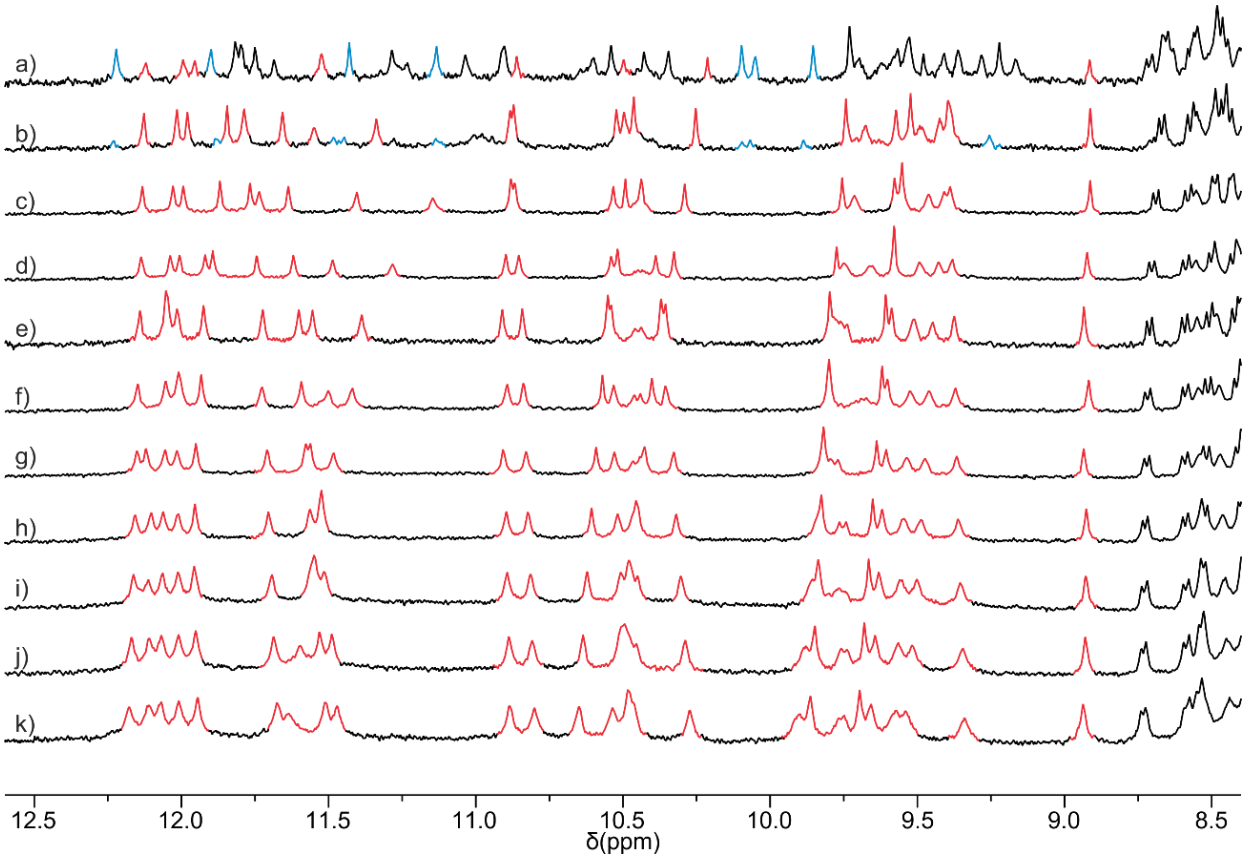


Figure S12. Extracts of ^{1}H NMR spectra (400 MHz, wet $CDCl_3$, 2.2 mM) of 2b in at different temperatures. The temperatures were 328 K (a), 318 K (b), 308 K (c), 298 K (d), 283 K (e), 273 K (f), 263 K (g), 253 K (h), 243 K (i), 233 K (j) and 223 K (k). The transition temperature varies depending on the water content. In wet $CDCl_3$, the transition temperature is above 328 K.

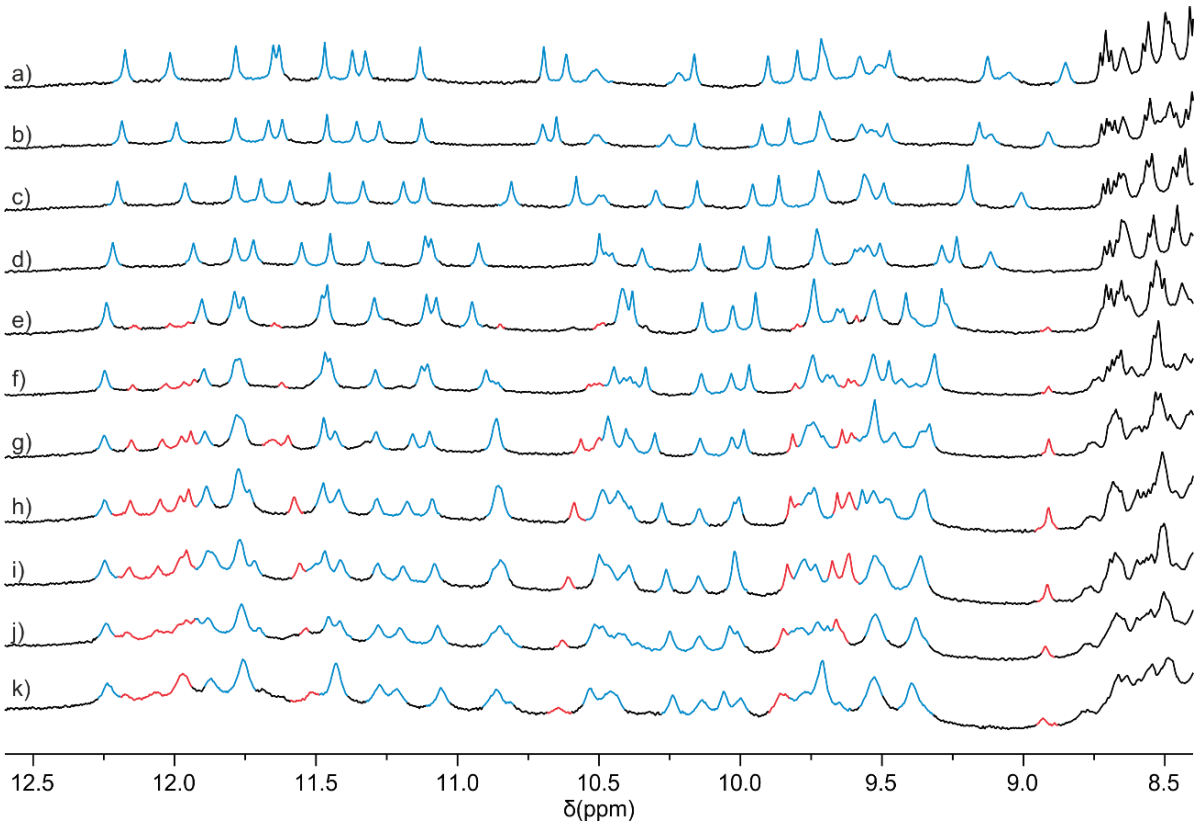


Figure S13. Extracts of ^{1}H NMR spectra (400 MHz, dried $CDCl_3$, 6.8 mM) of 2b in at different temperatures. The temperatures were 328 K (a), 318 K (b), 308 K (c), 298 K (d), 283 K (e), 273 K (f), 263 K (g), 253 K (h), 243 K (i), 233 K (j) and 223 K (k). The transition temperature varies depending on the water content. In dried $CDCl_3$, the transition temperature is below 223 K. The signals assigned to the dimer still appear at low temperatures since the $CDCl_3$ is not completely dry.

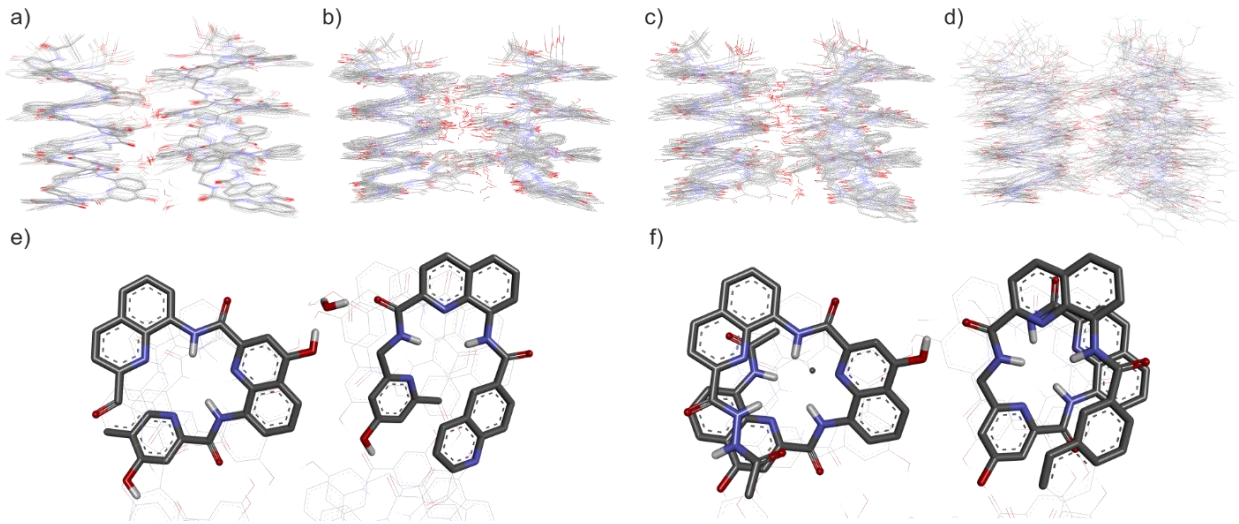


Figure S14. Stochastic dynamic simulation of $(3b)_2$•$(H_2O)_6$ at different temperatures. The overlay of ten snapshots of a 1 ns MD simulation of $(3b)_2$•$(H_2O)_6$ at 200 K (a), 300 K (b), 400 K (c) and 500 K (d). The tertiary folds are relatively stable at all temperatures. The water molecules have a tendency to leave the intermolecular interfaces when temperature increases. All water molecules remain in place at the intermolecular interface at 200 K. All water molecules leave the intermolecular interface at 500 K. The leaving of the water molecules leads the transition from water bridging shifted dimer interactions (e) to shifted dimer-like interactions (f) with torsions of the tertiary structures.

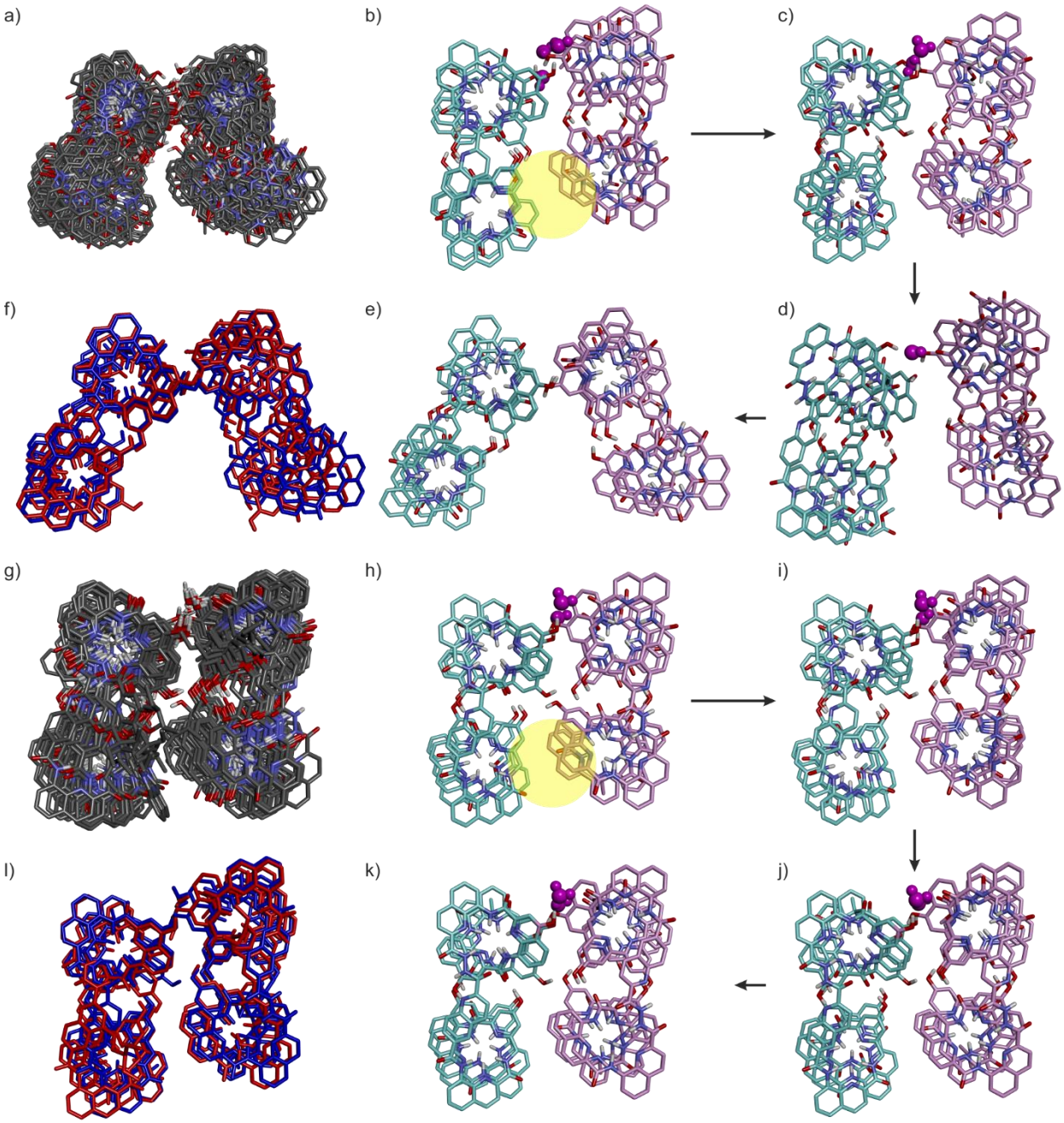


Figure S15. Stochastic dynamic simulation of a modified $(2b)_2$•$(H_2O)_3$ complex. The hydrogen bond donors and water molecules at one of the two intermolecular interfaces (highlighted by yellow shadow in (b)) have been removed before the MD simulation. a) Overlay of ten snapshots of the 1 ns MD simulation of modified $(2b)_2$•$(H_2O)_3$ at 400 K. b-e) Four snapshots showing the process of water leaving the interface which lead the transition from water-bridged shifted dimer interface (Figure 1e) to shifted dimer interface (Figure 1d). The angle between two tertiary folds adjusted back to near 54° at the end of the simulation. f) Overlay of (e) (colored in blue) and the molecular model from Figure S4a (colored in red). g) Overlay of ten snapshots of the 5 ns MD simulation of modified $(2b)_2$•$(H_2O)_3$ at 300 K. h-k) Four snapshots showing the process of the simulation. Only

one water molecules leave the interface and the relative angle of two tertiray structures remains. l) Overlap of molecular model before (colored in blue) and after (colored in red) the MD simulation.

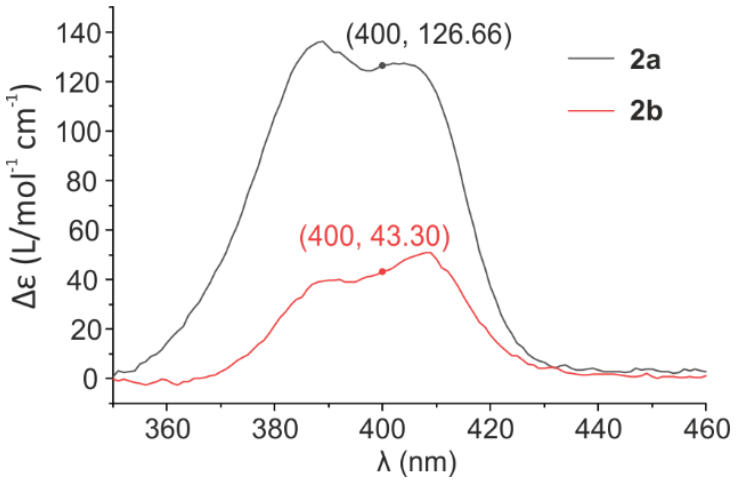


Figure S16. CD spectra of 2a and 2b in chloroform. The CD spectra of 2a (0.1 mM) and 2b (0.1 mM). The Δε value of 2a and 2b at 400 nm have been extracted. The ^{1}H NMR of sidechain protected sequence 2a showed two species, which are signed to PP (major species) and PM (minor species) diastereomeric conformers. A relatively strong positive CD band was observed as expected. The CD intensity of 2b is relatively weak suggesting that sequence 2b is P-turn-M and CD contributions of the two helices cancel each other. Because the P helix has more Q units than the M helix, the CD signal is not fully cancelled.

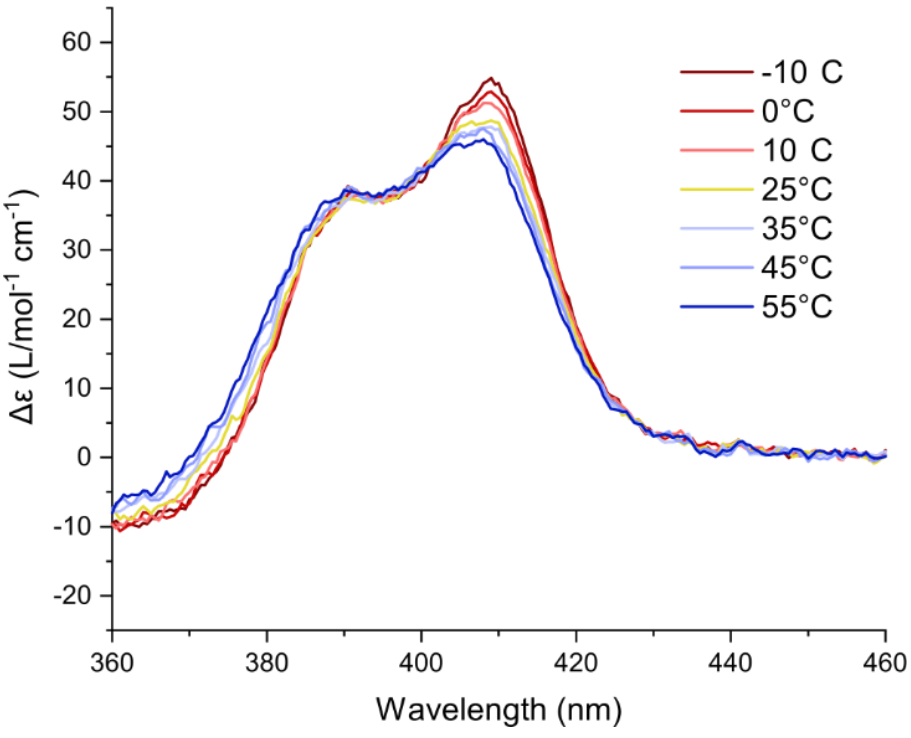


Figure S17. CD spectra of 2b in chloroform at different temperatures. The CD spectra of 2b (0.1 mM) measured at different temperatures showed no significantly change.

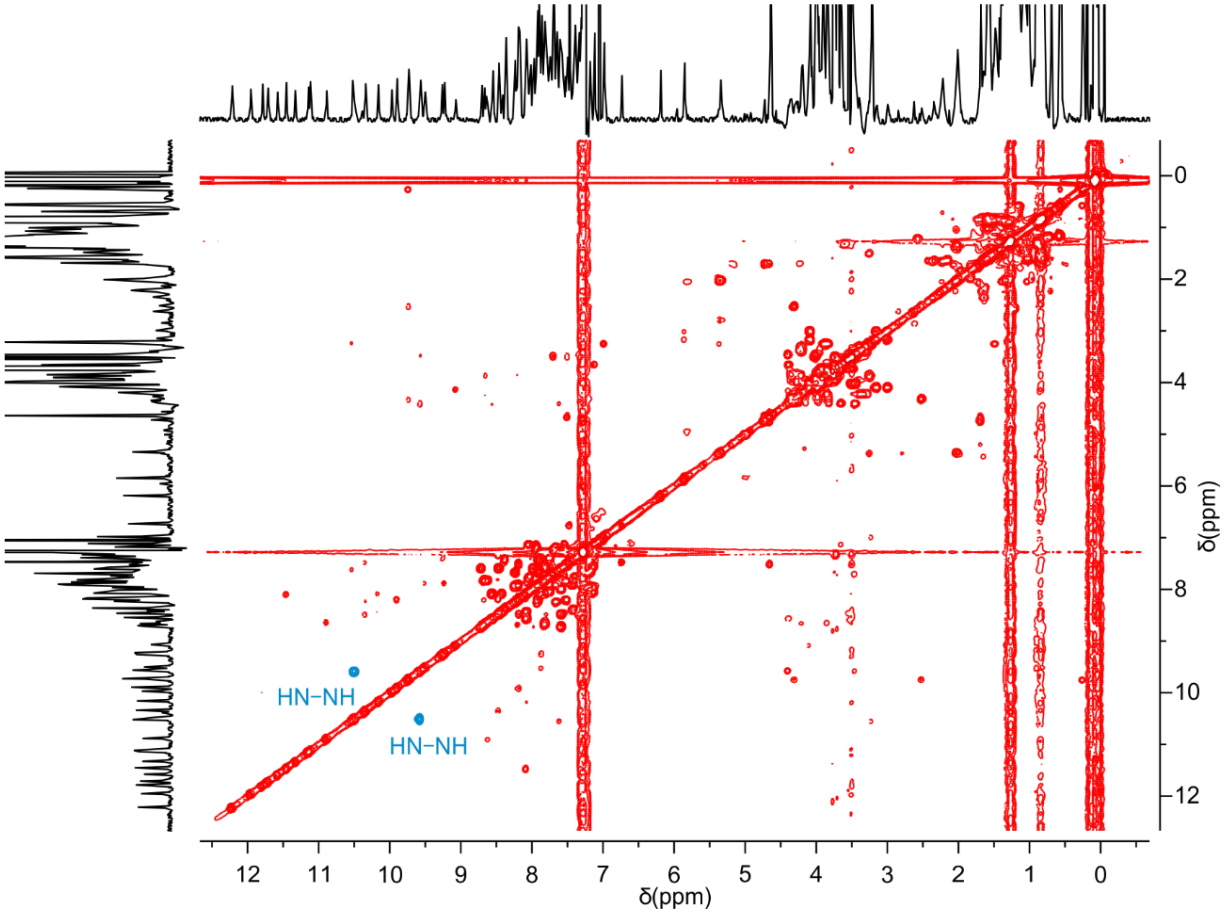


Figure S18 ^{1}H,^{1}H COSY spectrum of 2b (500 MHz, dried $CDCl_3$). The correlations belonging to the HN-NH moiety are colored in cyan. For the assignment of $C\underline{H}_2$ signals see section 3.5.

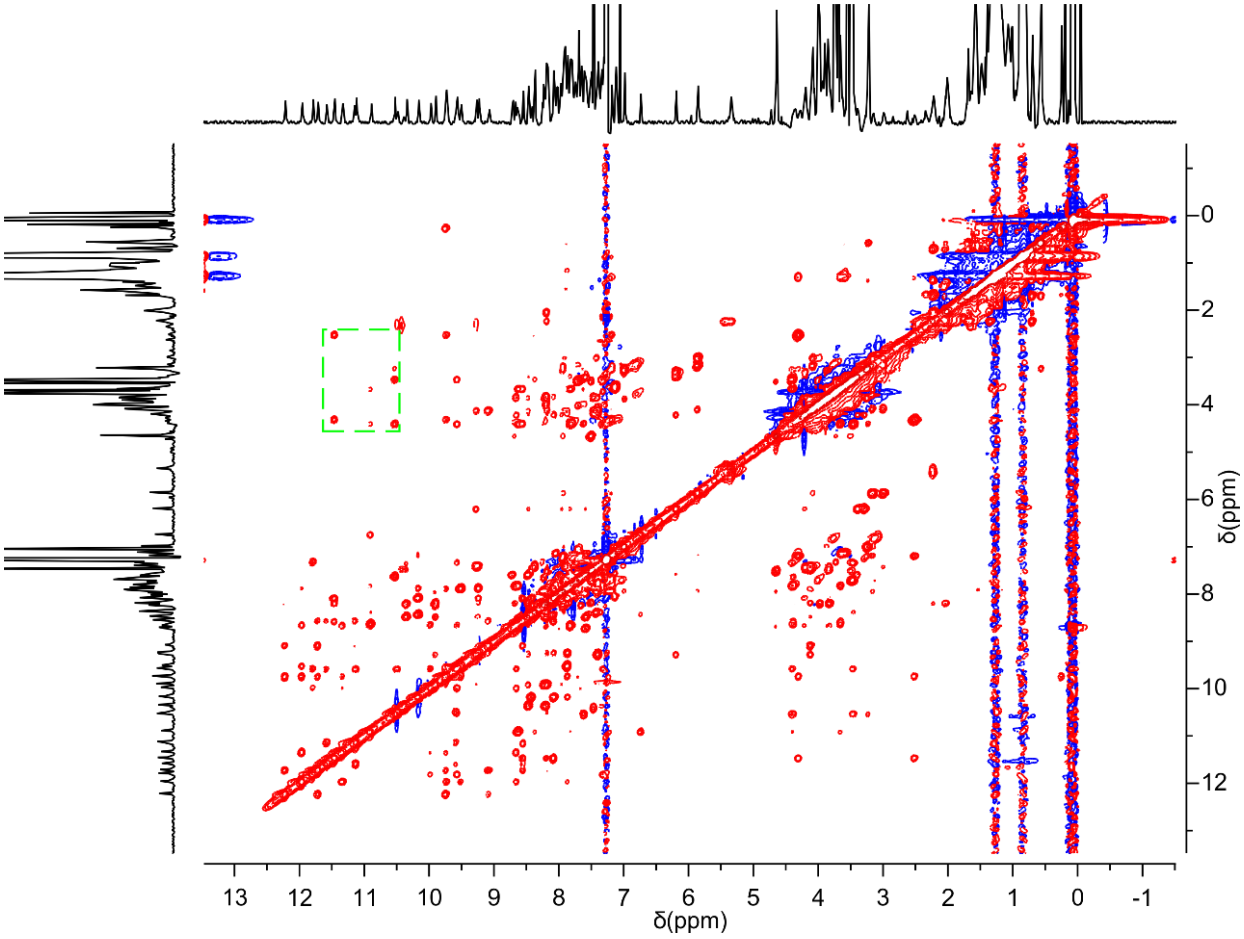


Figure S19 ^{1}H,^{1}H NOESY spectrum of 2b (500 MHz, dried $CDCl_3$). The region highlighted with a green dashed box is shown in Figure 3c. For the assignment of NH signals see sectio 3.5.

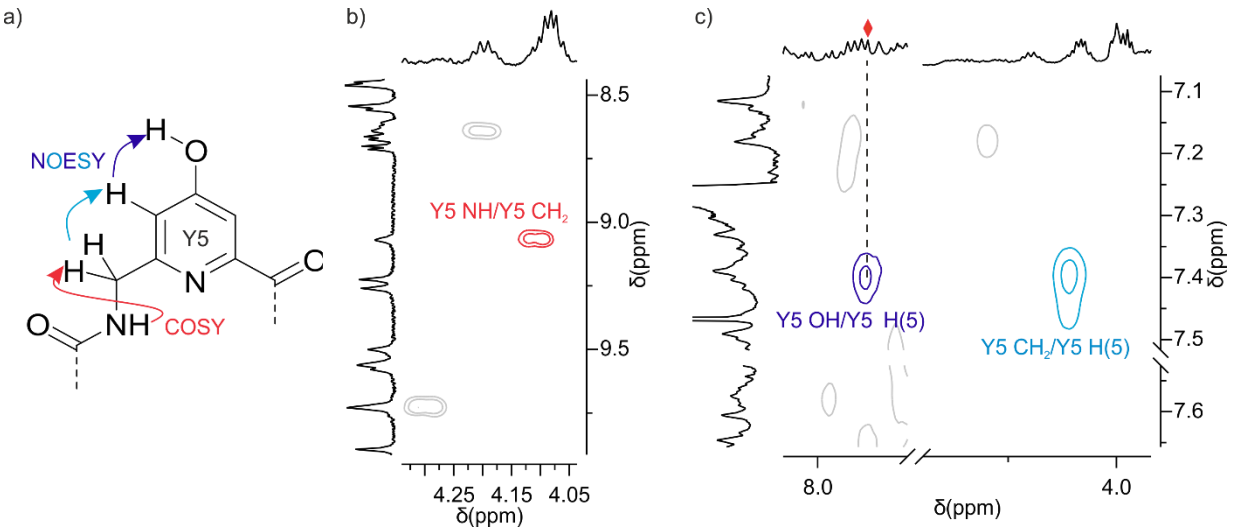


Figure S20 Tracing down the non-hydrogen-bonded OH signal in the region below δ 8.5 ppm. a) Assignment strategy for the "missing" OH signal (usually findable downfield shifted above 8.5 ppm) based on the structure of

the Y moiety of 2b. For the Y moiety (as well as the P moiety) correlations should be visible caused by the J-coupling of the NH protons with each of the unisochronous protons of the CH_2 group in the $^1H,^1H$ COSY spectrum. Furthermore, the CH_2 protons should yield NOE correlation with the aromatic proton of Y at position 5. This aromatic proton should then finally exhibit an NOE with the nearby OH proton. b) Excerpt of a $^1H,^1H$ COSY spectrum of 2b (500 MHz, dried $CDCl_3$) showing the visible correlations of protons of the CH_2 group with the NH proton, as explained in the assignment strategy under a). c) Two excerpts of a $^1H,^1H$ NOESY spectrum of 2b (500 MHz, dried $CDCl_3$) confirming the assignment strategy by showing the NOE correlation of OH proton/aromatic 5-H (colored in blue), and aromatic 5-H/CH_2 correlations (colored in cyan), respectively.

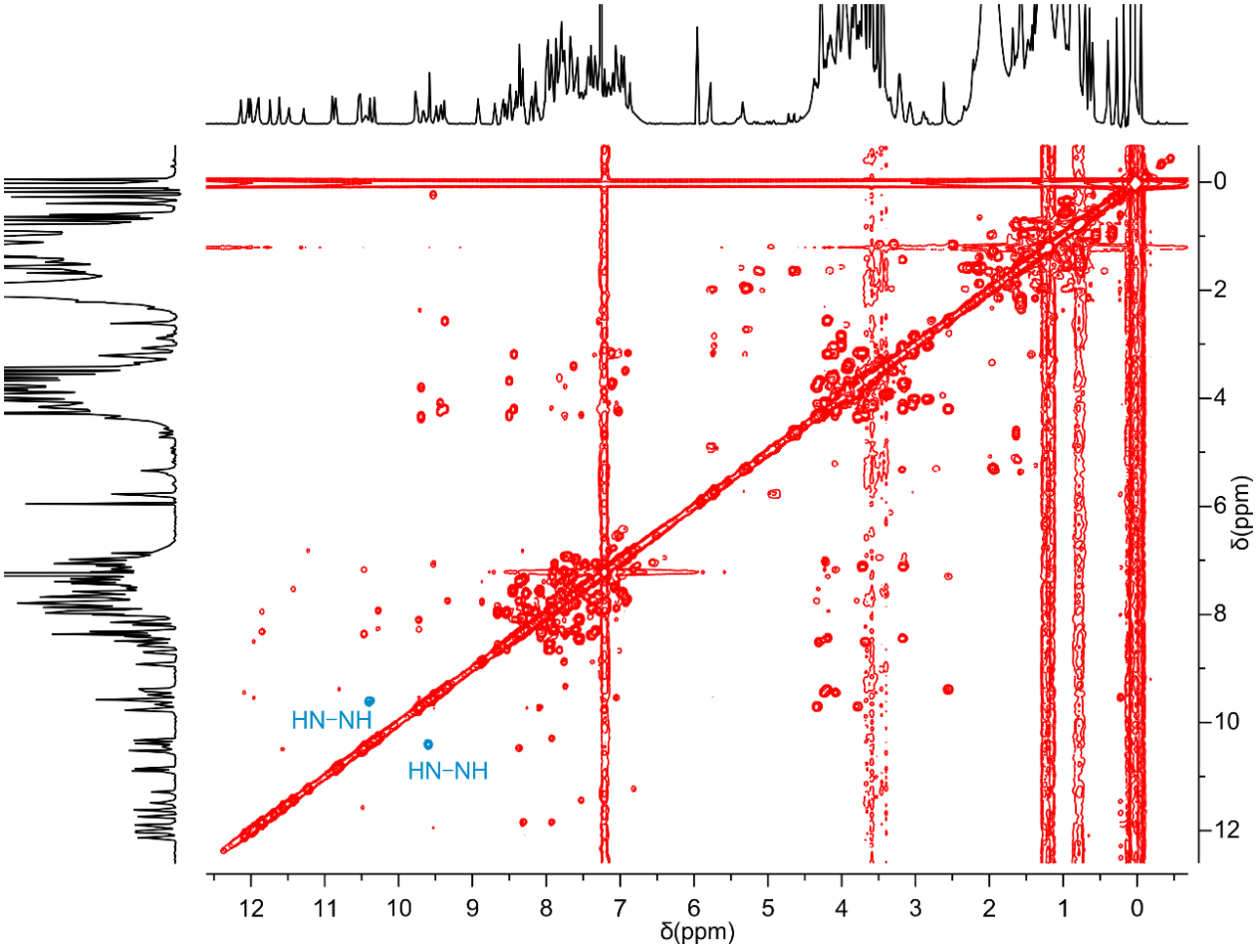


Figure S21 $^1H,^1H$ COSY spectrum of 2b (500 MHz, dried $CDCl_3$). The correlations of HN-NH are colored in cyan. For yhe assignment of $C\underline{H}_2$ signals see section 3.5.

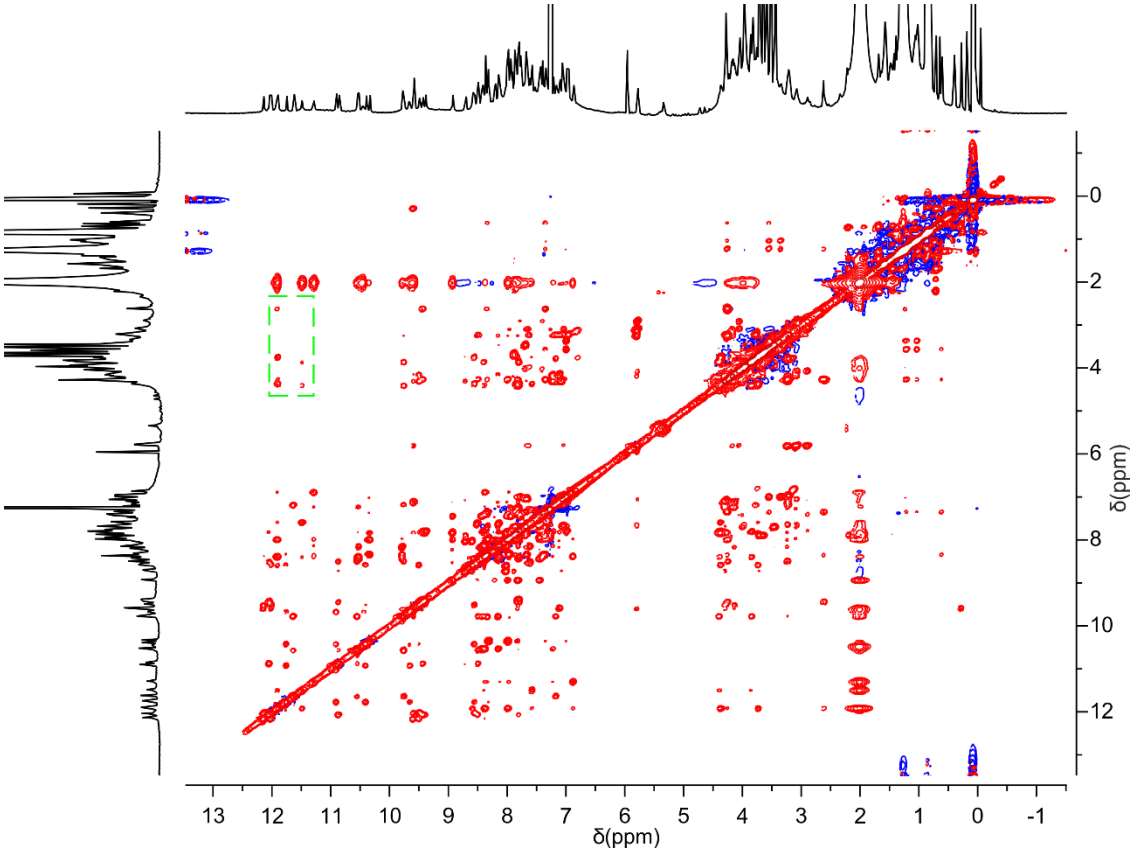


Figure S22 ^{1}H,^{1}H NOESY spectrum of 2b (500 MHz, wet $CDCl_3$). The part highlighted with a green dash line box is shown in Figure 3d. For the assignment of NH signals see sectio 3.5.

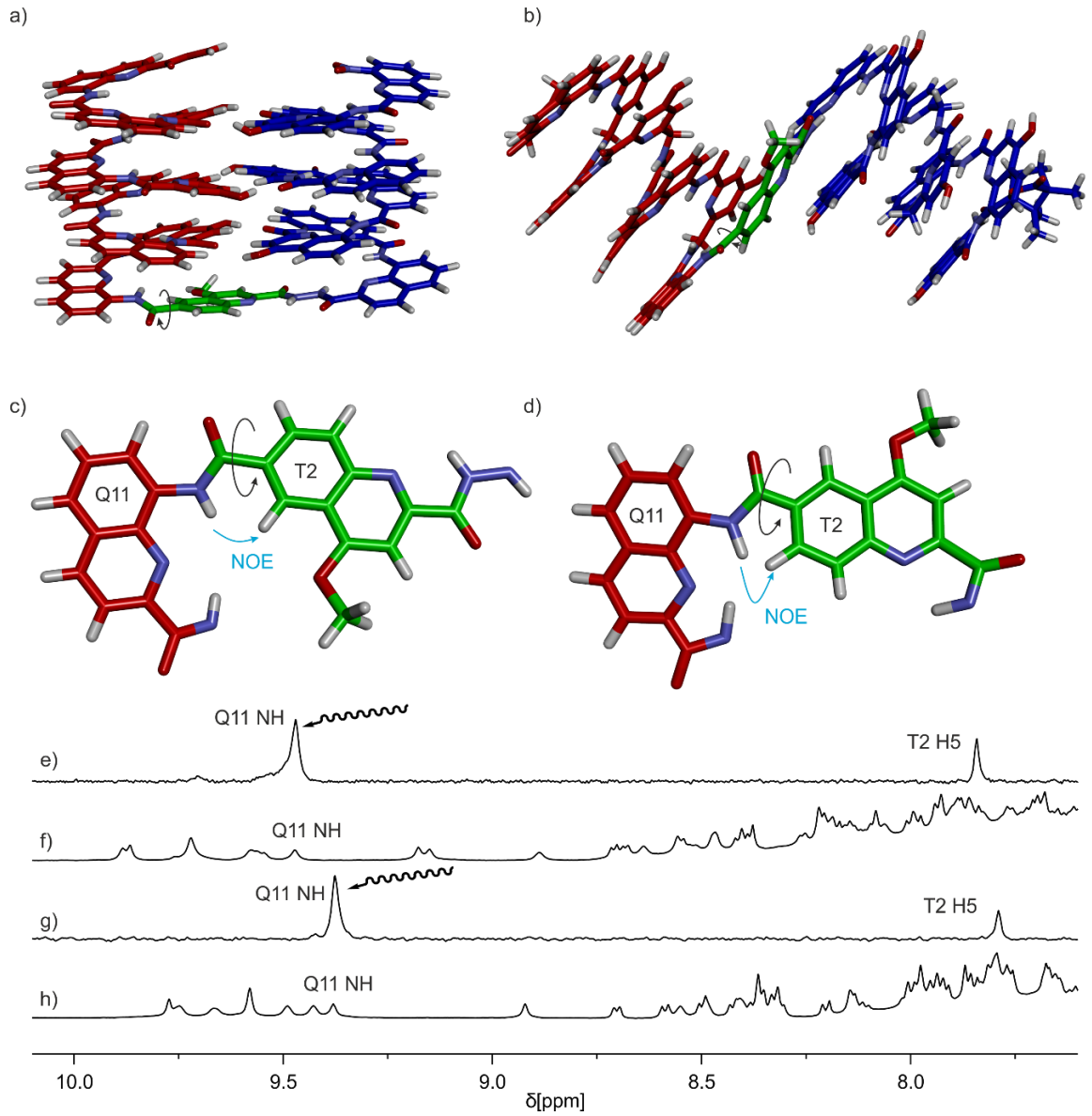


Figure S23 Determination of the conformation of T2 moiety. a, b) Two plausible conformations of P-turn-M structure. (c) and (d) Excerpts from the structure of (a) and (b), which illustrate the proton proximities of Q11 and T2 within the two conformers and the resulting expected NOEs. The rotatable bond that causes the conformational change is highlighted by a black arrow. The NH of Q11 shows an NOE with the nearest aromatic proton of T2 in both cases. In the U-shape conformation (a), the NH of Q11 will have an NOE with the proton in position 5 of T2, which should be a singlet, since there is no proton in ortho-position and the meta-coupling to H-7 is usually not detectable due to the small coupling constant. In contrast, in the stretched conformation (b), the NH will show an NOE with the proton in position 7 of T2, whose signal should be a doublet due to one coupling partner in ortho-position (here, of course, the probable meta-coupling which would ultimately result in doublets of doublets, is

also not expected). The side chain of T2 is replaced by –OMe and side chain of Q unit is omitted for clarity. 1D NOESY spectra (500 MHz) of the respective species in dried $CDCl_3$ (e) and wet $CDCl_3$ (g) showing the responses upon selective irradiating the NH resonances at δ 9.47 ppm (e) and δ 9.37 ppm (g), respectively. The signal marked with irradiation arrow in both 1D NOESYs illustrating which signal is irradiated. In both cases, the resonance of the proton in the aromatic region experiencing the NOE is a singlet, confirming the U-shape conformation of the P-turn-M structure. The ^{1}H NMR (500 MHz) of 2b in dried $CDCl_3$ (f) and wet $CDCl_3$ (h).

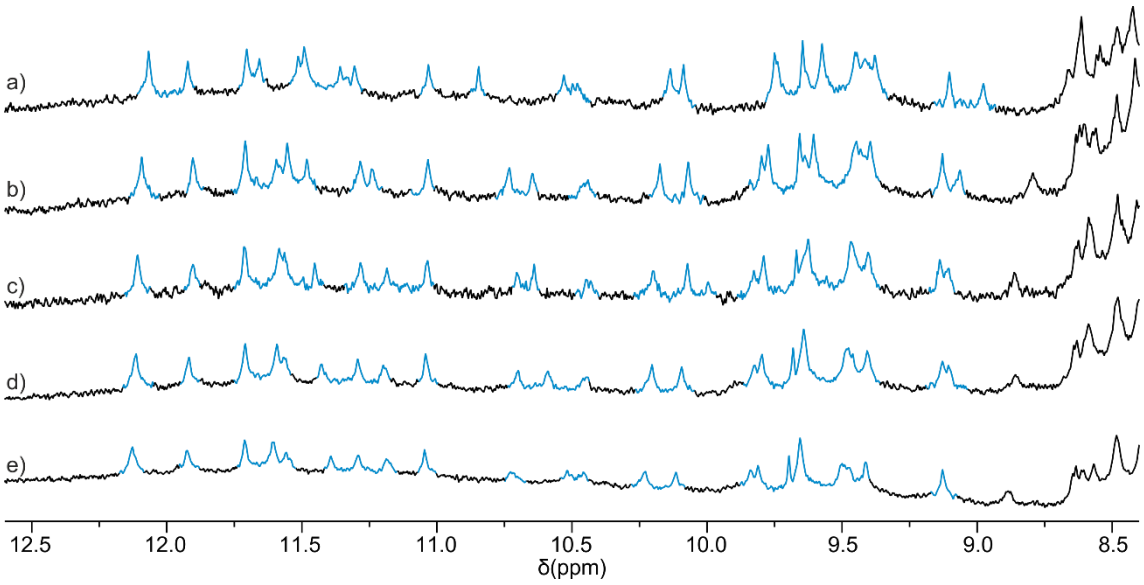


Figure S24. Extracts of ^{1}H NMR spectra (500 MHz, 0.35 mM) of 2b in different mixtures of dried CD_2Cl_2 in dried $CDCl_3$. The proportions (%, v/v) of dried CD_2Cl_2 in dried $CDCl_3$ are 100 (a), 75 (b), 50 (c), 25 (d), 0 (e). No obvious change was observed indicating that the trimer formation is independent of the type of chlorinated solvents. For the preparation of dried CD_2Cl_2 see section 5.1.

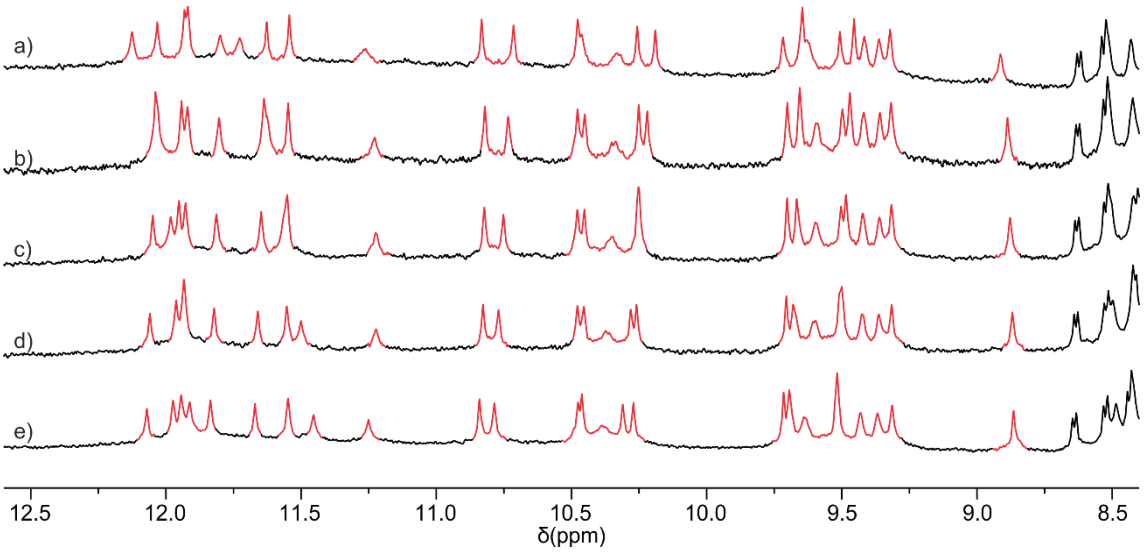


Figure S25. Extracts of ^{1}H NMR spectra (500 MHz, 0.35 mM) of 2b in different mixtures of wet CD_2Cl_2 in wet $CDCl_3$. The proportions (%, v/v) of wet CD_2Cl_2 in wet $CDCl_3$ are 100 (a), 75 (b), 50 (c), 25 (d), 0 (e). No obvious change was observed indicating that the dimer formation is independent of the type of chlorinated solvents. For the preparation of wet CD_2Cl_2 see section 5.1.

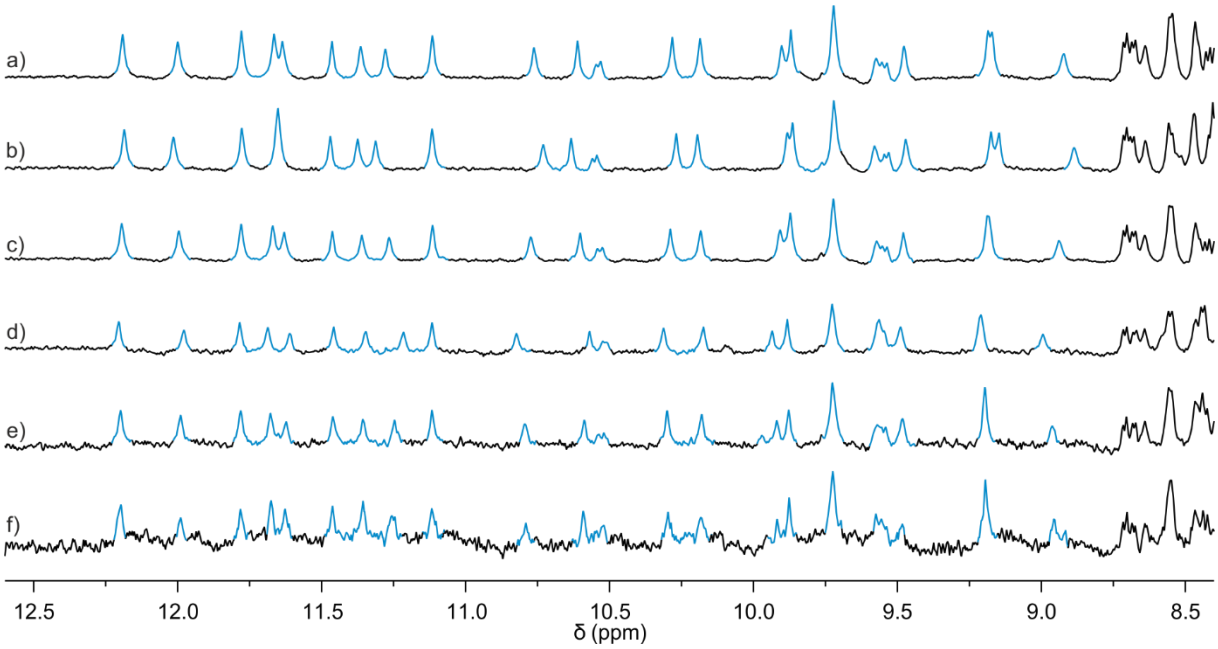


Figure SX. Extracts of ^{1}H NMR spectra (500 MHz, dried $CDCl_3$) of 2b at different concentration The concentrations of **2b** are 1.0 mM (a), 0.5 mM (b), 0.25 mM (c), 0.1 mM (d), 0.05 mM (e) and 0.02 mM (f), respectively.

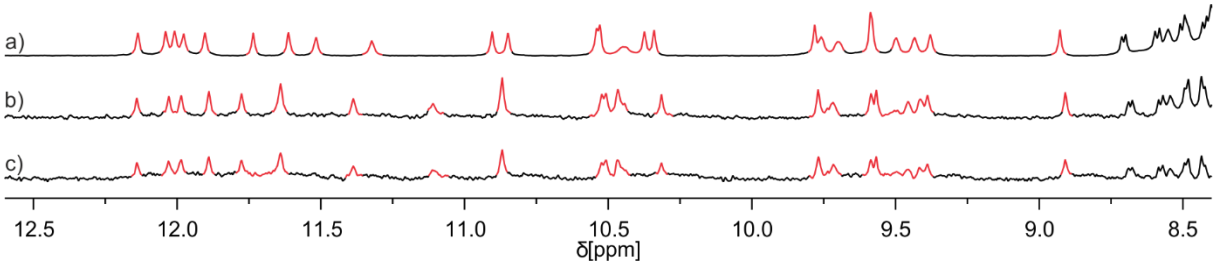


Figure SX. Extracts of ^{1}H NMR spectra (500 MHz, wet $CDCl_3$) of 2b at different concentration. The concentrations of **2b** are 1.0 mM (a), 0.5 mM (b), 0.05 mM (b) and 0.02 mM (c), respectively.

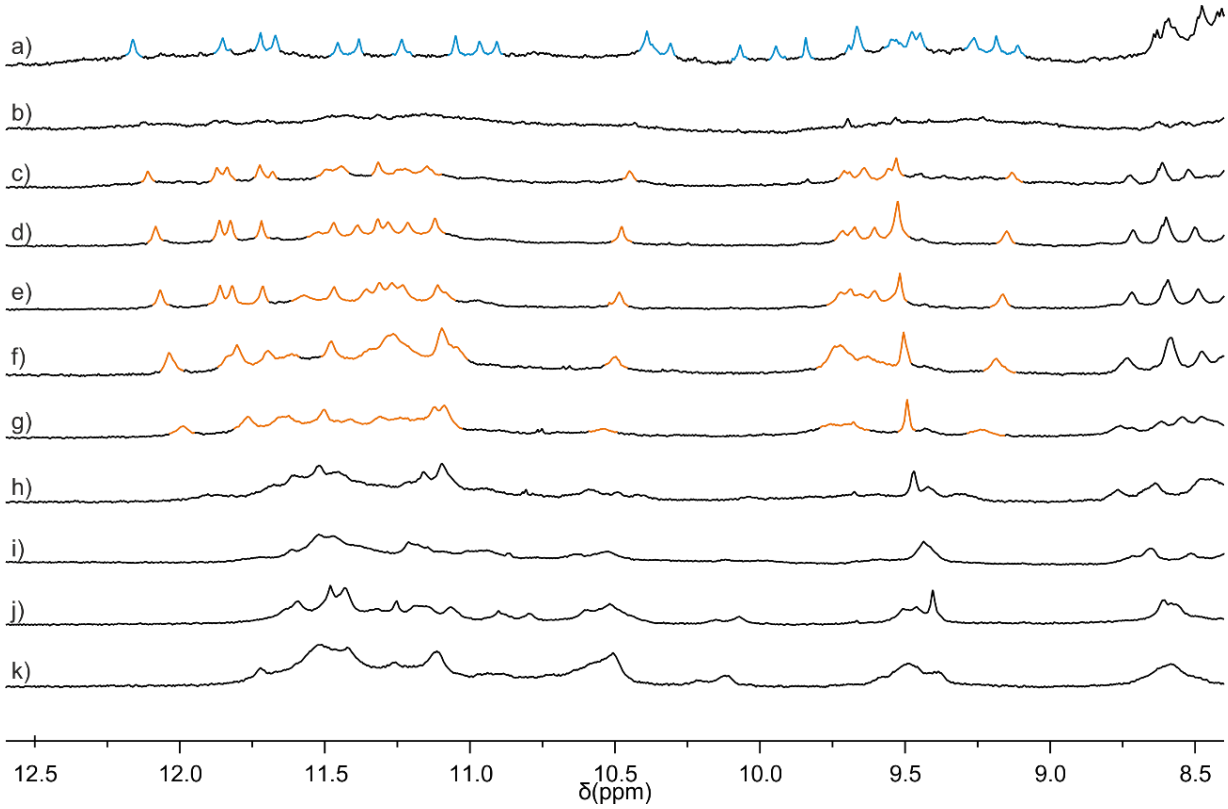


Figure S26. Extracts of ^{1}H NMR spectra (500 MHz) of 2b in mixtures of DMSO-d_6 in dried $CDCl_3$. The proportions (%, v/v) of DMSO-d_6 in dried $CDCl_3$ are 0 (a), 2 (b), 4 (c), 8 (d), 12 (e), 16 (f), 20 (g), 24 (h), 28 (i), 32 (j) and 36 (k). The signals belonging to species assigned to the trimer are colored in cyan, the signals for the monomeric tertiary structure in orange. With a DMSO content of more than 20%, the intramolecular H-bonds are disrupted. The broad NMR lines could be due to the equilibria between different conformers associated with flexible pyridine-based units. DMSO-induced dissociation and disruption are observed.

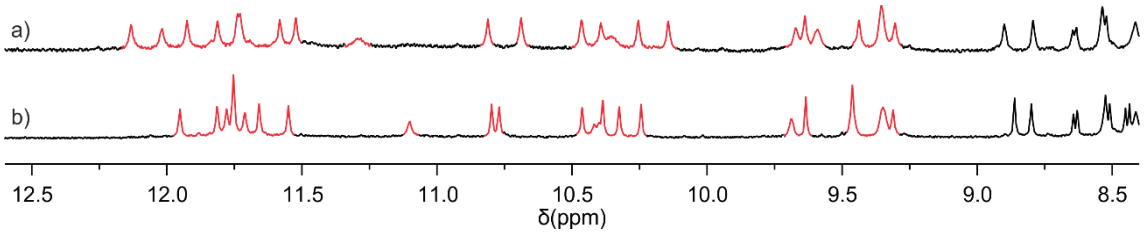


Figure S27. Extracts of ^{1}H NMR spectra (500 MHz) of 3b in and wet CD_2Cl_2 (a) wet $CDCl_3$ (b). Dimer formation appears to be independent of the type of chlorinated solvent.

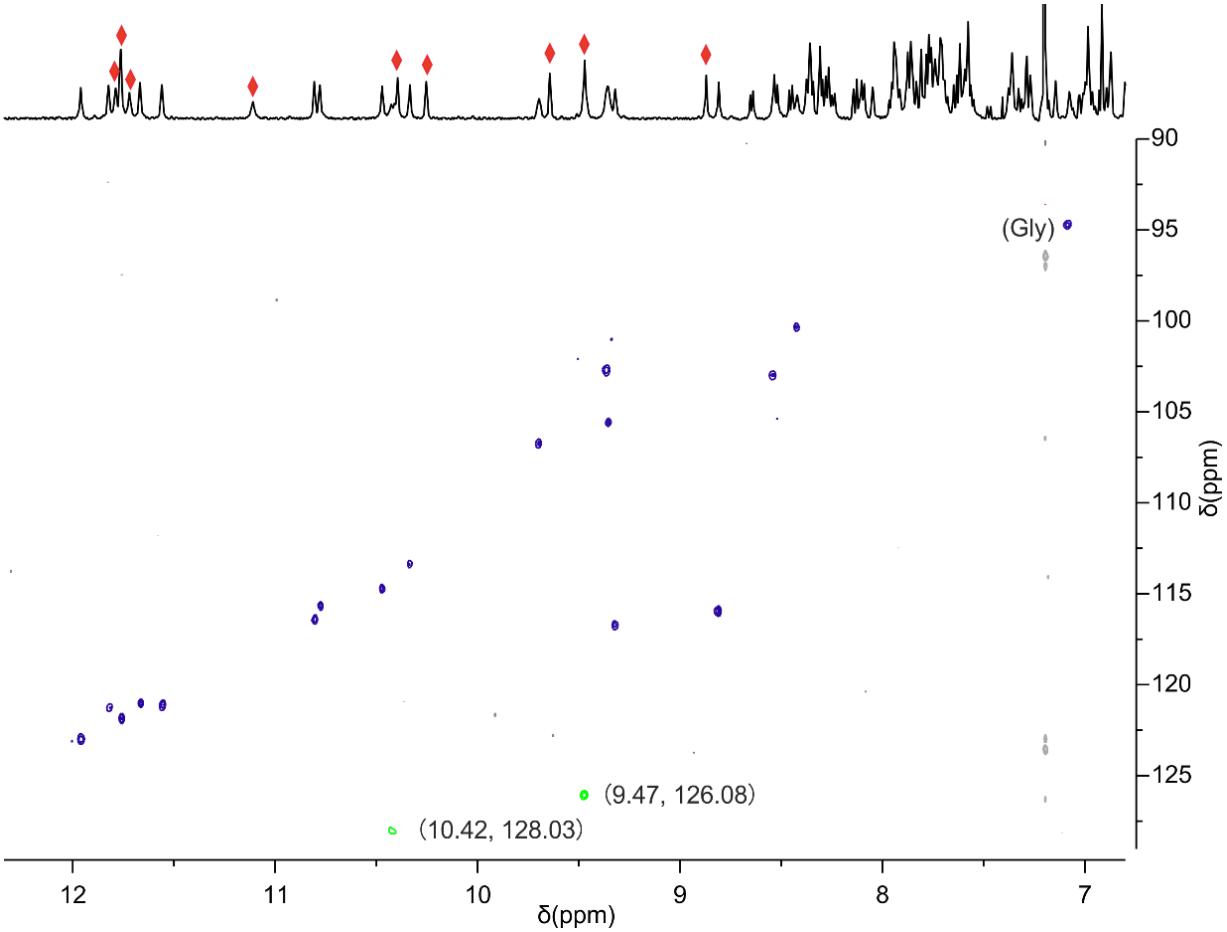


Figure S28. Extract of a $^{1}H,^{15}N$ HSQC spectrum (500 MHz, wet $CDCl_3$) of 3b. Only NH resonances are correlated. Red diamonds indicate the signals of H-bonded O<u>H</u> protons. The green correlations are assigned to the NH-NH signals of T2 unit in analogy to 2b (see Figure S8).

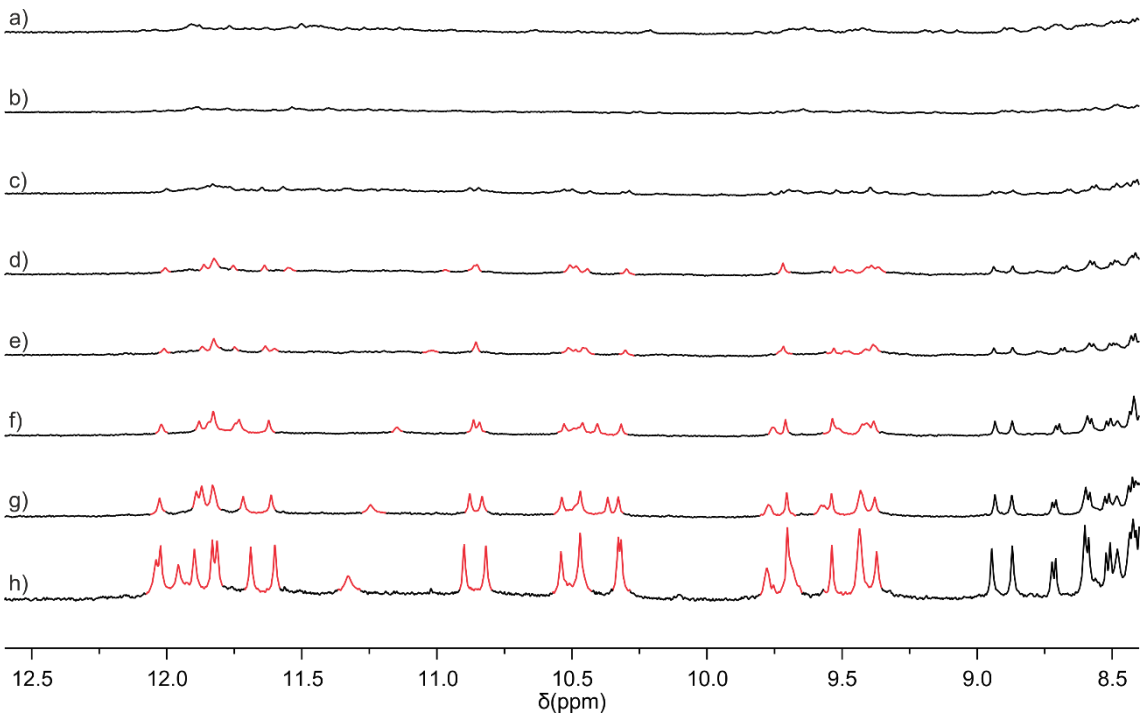


Figure S29. Extracts of ^{1}H NMR spectra (500 MHz) of 3b in different wet $CDCl_3$ and dried $CDCl_3$ mixtures. The proportions (%, v/v) of wet $CDCl_3$ in dry $CDCl_3$ are 0 (a), 5 (b), 10 (c), 15 (d), 20 (e), 30 (f), 50 (g) and 100 (h). A drier solvent might lead to some type of polymerization which is reflected in broad NMR signals. As the amount of water increases, the NMR signals are getting sharper and showing a spectrum that is more indicative of a dimer, suggesting that the water leads to the formation of a water-mediated dimer.

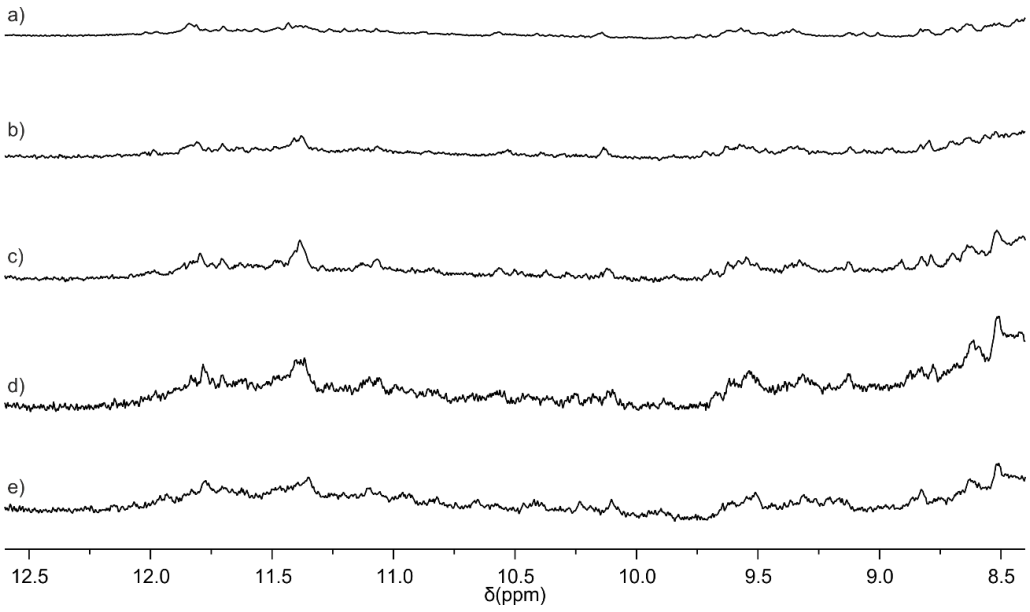


Figure S30. Extracts of ^{1}H NMR spectra (500 MHz, 0.35 mM) of 3b in different dried $CDCl_3$ and CD_2Cl_2 mixtures. The proportions (%, v/v) of dry CD_2Cl_2 in dry $CDCl_3$ are 100 (a), 75 (b), 50 (c), 25 (d), 0 (e). No

significant change was observed suggesting that the plausible polymerization is independent of the type of chlorinated solvent.

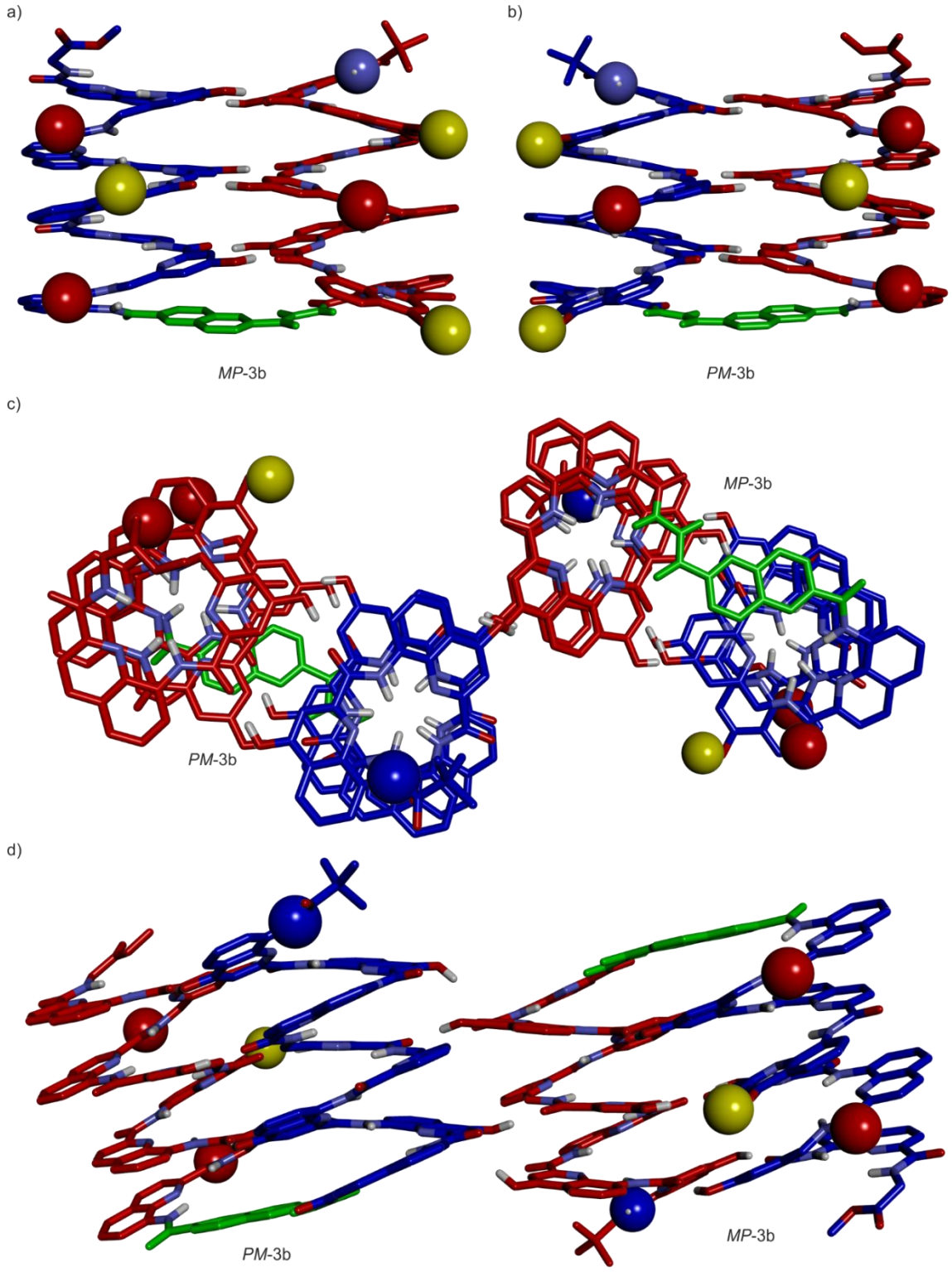


Figure S31. Molecular model of 3b enantiomers and possible aggregation mode. The energy-minimized models of MP-3b (a) and PM-3b (b). The top view (c) and side view (d) of a plausible MP-3b and PM-3b aggregation mediated by a head-to-tail heterochiral shifted dimer pattern, showing the possibility of

polymerization through two open ends. The P helices and M helices are shown in blue and red tube representation, respectively. The free H-bond donors and H-bond acceptors are shown as yellow balls and red balls, respectively. The N-terminus are shown as a blue ball. T2 units are colored in green.

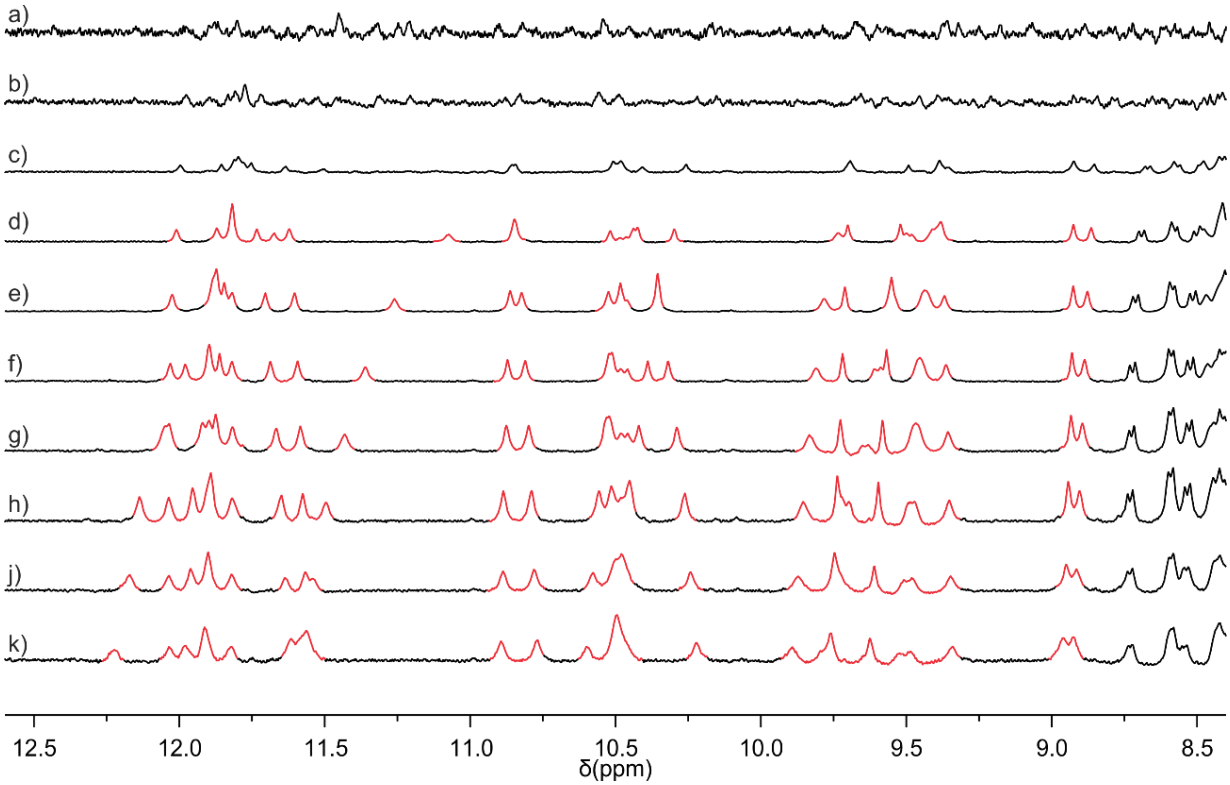


Figure S32. Extracts ^{1}H NMR (400 MHz, wet $CDCl_3$) spectra of 3b at different temperatures. The temperatures were 328 K (a), 318 K (b), 308 K (c), 298 K (d), 283 K (e), 273 K (f), 263 K (g), 253 K (h), 243 K (j) and 233 K (k). The signals get broad as temperature increases, suggesting a polymerization similar to that observed by reducing the water content in chloroform.

3. Supplementary methods

3.1. LC-MS analyses

LC-MS spectra were recorded on a Bruker microTOF II in positive ionization mode. The instrument was calibrated in positive mode by direct infusion of a calibration solution (Agilent Technologies ESI-L Low Concentration Tuning Mix). The HPLC line was an Ultimate 3000 RP-HPLC system (ThermoFisher Scientific) equipped with a Aeris^{TM} Widepore C4 column (2.1 x 150 mm, 3.6 µm) at a flow rate of 0.25 mL/min. 0.1 % formic acid and 0.025% TFA was added to the aqueous mobile phase (solvent A) and to acetonitrile (solvent B). The gradient is: 0–5 min, 50% to 100% solvent B; 5–14 min, 100% solvent B at 50 °C. The column eluent was

monitored by UV detection at 214, 254, and 300 nm with a diode array detector. The sample was prepared by adding 10 µL of a solution of the sample in DCM (0.1 mg/mL) to 1 mL acetonitrile containing 0.05–0.1% formic acid.

3.2. Molecular modelling

Models were simulated by MacroModel version 11.1 (Maestro, Schrödinger Inc.). Energy minimized structures were obtained using MacroModel energy minimization with the following parameters: force field: MMFFs; solvent: none or $CHCl_3$; electrostatic treatment: constant dielectric; dielectric constant: 1.0; charges from: force field; cutoff: normal; Van der Waals: 7.0; electrostatic: 12.0; H-bond: 4.0; mini method: TNCG; maximum iterations: 2500; converge on gradient; convergence threshold: 0.05; constraints: distances. As a starting point, the coordinates of the crystal structure of **1** (CCDC entry # 1955168) was used. Some hydroxy groups were inserted and some aromatic rings were removed. The modified sequence was first energy-minimized. In a second round, two identical modified helix-turn-helix structures were placed in a plausible arrangement, and distance constraints between plausible hydrogen-bonding partners were set to 1.8 on purpose. While setting the constraints, it was important to match the hydroxy group to their correct hydrogen-bonding carbonyl partner. Then all constraints were removed, and energy minimization was repeated offering the molecular model in Figure S4a. A third identical sequence was imported and placed in a plausible arrangement. The distance between all the plausible hydrogen-bonding partners (including all the intermolecular H-bonds and intramolecular H-bonds) were set to 1.8 on purpose. Then all constraints were removed and the energy minimization was repeated and obtaining the molecular model in Figure S4b. To obtain the molecular model of $(\mathbf{2b})_2\bullet(H_2O)_6$ (Figure 3b), six water molecules were placed in between the intermolecular interfaces. The distance constraints between plausible hydrogen-bonding partners (the hydroxy hydrogen to water oxygen atom, and the water hydrogen to carbonyl oxygen atom) were set to 1.8 on purpose. Then all constraints were removed, and energy minimization was repeated.

3.3. Molecular dynamic simulations

Molecular dynamic simulations were carried out using MacroModel version 11.1 (Maestro, Schrödinger Inc.). Energy minimized molecular model of $(\mathbf{2b})_2\bullet(H_2O)_6$ from 3.2 was used as the object for the MD simulation. At one intermolecular interface of the molecular model of $(\mathbf{2b})_2\bullet(H_2O)_6$, three water molecular were removed and

three hydroxy groups were replaced by hydrogen, obtaining the MD simulation object (**2b**)$_2$•(H_2O)$_3$. Stochastic dynamic simulations were obtained using MMFFs force field, $CHCl_3$ as solvent, extended cutoff and TNCG method. The simulations were performed for 1 ns or 5 ns at different temperatures (200 K, 300 K, 400 K and 500 K), time step of 1.5 fs and 1 ps as equilibration time. During the simulation, 10 structures were sampled and the sampled structures were minimized.

3.4. Nuclear magnetic resonance spectroscopy

NMR spectra of **2a**, **2b**, **3a** and **3b** were recorded on different NMR spectrometers: (I) an Avance III HD NMR spectrometer 500 MHz (Bruker BioSpin) with CryoProbe™ Prodigy for ^{1}H NMR, ^{1}H,^{15}N-HSQC, ^{1}H,^{1}H COSY, ^{1}H,^{1}H NOESY and DOSY spectra of foldamers. All NMR measurements were performed at 25 °C, unless otherwise stated. (II) a Varian NMR spectrometer 400 MHz (Varian, Inc.) for variable temperature measurements. Chemical shifts (δ) are described in part per million (ppm) relative to the ^{1}H residual signal of the deuterated solvent used, *i.e.*, DMSO-d_6 (2.50), CD_2Cl_2 (5.32) and $CDCl_3$ (7.26). For the DMSO-d_6 and $CDCl_3$ solvent mixture, the chemical shifts were calibrated according to DMSO-d_6 (2.50). For the CD_2Cl_2 and $CDCl_3$ solvent mixture, the chemical shifts were calibrated according to internal standard tetramethylsilane (0.00).^{1}H NMR splitting patterns with observed first-order coupling are entitled as singlet (s), doublet (d), triplet (t), or multiplet (m). Coupling constants (*J*) are ported in Hz.

Complete disruption of the hydrogen bonds was achieved by dissolving the sample in polar solvents (such as DMSO, pyridine or MeOH/chloroform mixture), followed by removal of the solvent. We usually measured the ^{1}H NMR spectra of the sequences at different times until no further change was observed. We generally consider that at this point the sequences reached equilibrium. Sometimes sequences need days to weeks to reach equilibrium. This is not the case here. Here, the time for complete equilibration is short and a stage prior to this can hardly be detected with NMR, as there are no significant changes in the spectra recorded after a few minutes compared to those obtained up to a week after resolving the sequence. This is independent of whether the solvent was wet or dry. A similar behavior was also observed for sequence **1**, which suggesting that the formation of the tertiary structure is relatively fast.

The equilibrium time (the measurement delay between the two different condition sets) for ^{1}H NMR spectra measurements for varying temperatures and different proportions of DMSO-d_6/$CDCl_3$ mixtures was usually several minutes. Due to similar properties of $CDCl_3$ and CD_2Cl_2, the individual samples in different

proportions of $CDCl_3/CD_2Cl_2$ mixtures were prepared separately, and the ^1H NMR spectra of all of the samples were measured after equilibrate the sample at r.t. overnight.

$^1H,^{15}N$ HSQC spectra were recorded with a phase-sensitive pulse sequence with sensitivity enhancement using trim pulses in inept transfer (hsqcetgpsi2) from the Bruker pulse program library. Data acquisition was performed utilizing non-uniform sampling (NUS; NUS amount: 50% with an automatically created NUSList) yielding 1024 (F2) x 128 (F1) data points in Echo/Antiecho gradient selection mode. The recycling delay was 2.0 s and 64 transients per increment were applied at a sweep width of 2.5 kHz in F2 and 7 kHz in F1 resulting in an acquisition time of 0.1462 s. NUS processing was performed using the fully automated NUS processing tool provided by MestReNova. Zero filling in F1 has been used to yield a final matrix of 1K x 1K real points.

$^1H,^1H$ COSY spectra were recorded with a phase-sensitive pulse sequence from the Bruker pulse program library (cosygpppqf). Data acquisition was performed with 1K (F2) x 512 (F1) data points and a mixing time of 0.2 s. The recycling delay was 1.5 s and 32 transients per increment were applied at a sweep width of 7.5 kHz in both dimensions resulting in an acquisition time of 0.14 s. A 90° shifted sine-square multiplication in both dimensions prior to FT and zero filing was applied to yield a final symmetrical 2D matrix of 4K x 4K data points.

$^1H,^1H$ NOESY spectra were recorded with a phase-sensitive pulse sequence from the Bruker pulse program library (noesygpph). Data acquisition was performed with 1K (F2) x 512 (F1) data points and a mixing time of 0.2 s. The recycling delay was 2 s and 16 transients per increment were applied at a sweep width of 7.5 kHz in both dimensions resulting in an acquisition time of 0.15 s. A 90° shifted sine-square multiplication in both dimensions prior to FT and zero filing was applied to yield a final symmetrical 2D matrix of 2K x 2K data points.

1D NOESY spectra were recorded with a pulse sequence with selective refocusing by a shaped pulse from the Bruker pulse program library (selnogpzs.2). Data acquisition was performed using a Gaus1_270.1000 shaped pulse on the selected resonance with a mixing time of 0.2 s. The recycling delay was 1 s and 128 transients were applied. An exponential window function of 1 Hz was applied to the raw fid.DOSY spectra were recorded applying a pulse sequence with stimulated echo using stimulated echo for diffusion from the Bruker pulse program library (stegp1s). The diffusion delay Δ (big delta) and the diffusion gradient pulse length δ (little delta) were set as follows: 250 ms and 1.0 ms. (Figure 4e); 50 ms and 1.0 ms (Figure 4f); 250 ms and 2.4 ms (Figure S10). The number of gradient steps were set to 32 with linear spacing starting from 2% reaching 95% of the full gradient strength in the final step. For each of the 32 gradient amplitudes, 256 transients of 65K complex data points were acquired. DOSY processing was performed with the DOSY processing tool from MestReNova (v.12.x64)

employing the "Peak Heights Fit" algorithm including the "overlapped peaks analysis" with 128 points in diffusion dimension and a window of $1.00 \cdot 10^{-10}$ to $1.00 \cdot 10^{+03}$ cm^2 s^{-1}.

3.5. Assignment of NH and CH_2 signals

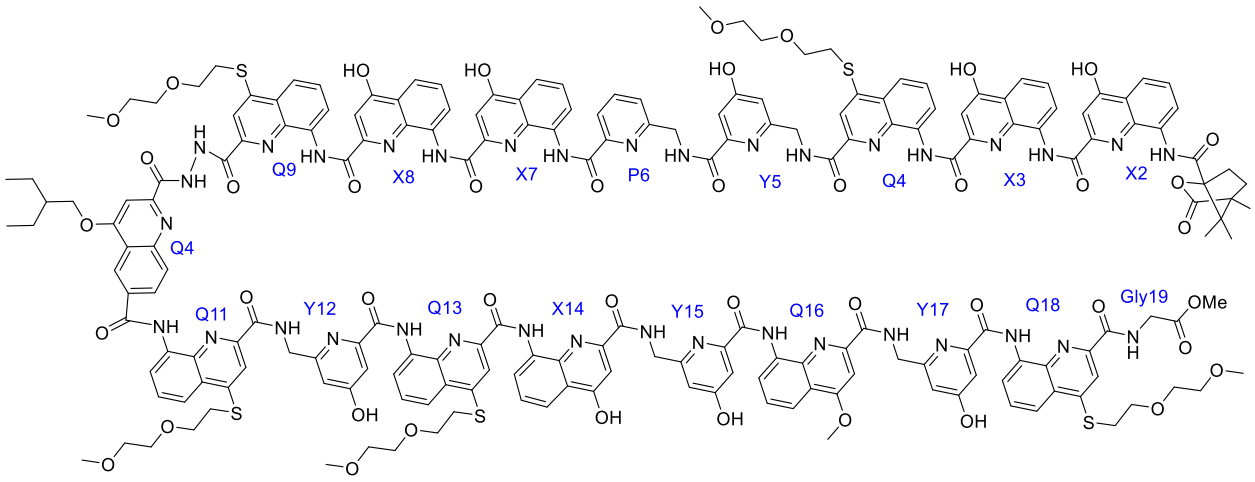


Figure S33. The structure of **2b** and the numbering of the units used in NMR assignment.

3.5.1. Assignment of NH signals

The assignment of the NH signals of the species of **2b** in dried $CDCl_3$ is based on the fact that the NH protons of two consecutive units are close to each other due to intramolecular H-bonding (Figure S34b) and will therefore have an NOE interactions. The signal assignments of two helices are achieved separately since no NOE can be observed across T2. First, the X2-NH was assigned due to the NOE correlation with one of the methyl groups of the N-terminal (1*S*)-camphanyl group (Figure S34c, d). The assignment of the N-terminus helix starts from X2-NH and follows the light blue dashed line to T2-HN-NH (Figure S34a). The Gly19-NH could be assigned based on the characteristic chemical shift at δ 7.30 ppm, which is not observed with sequences without Gly at C-terminus, such as sequence **1**. The assignment of the C-terminal helix starts from Gly19-NH following the green dashed line to Q11-NH (Figure S34a). The chemical shifts of each NH signals are reported in Table S1.

The assignment of the NH signals of the species of **2b** in wet $CDCl_3$ is based on the same strategy. The X2-NH and Gly19-NH (chemical shift value at δ 7.17 ppm) were assigned using the same method as mentioned above. The NOE network of the NHs of the N-terminal helix is highlighted with light blue dashed lines, that of the C-terminal helix with green dashed lines (Figure S35). The chemical shift values of each NH signals are reported in Table S2.

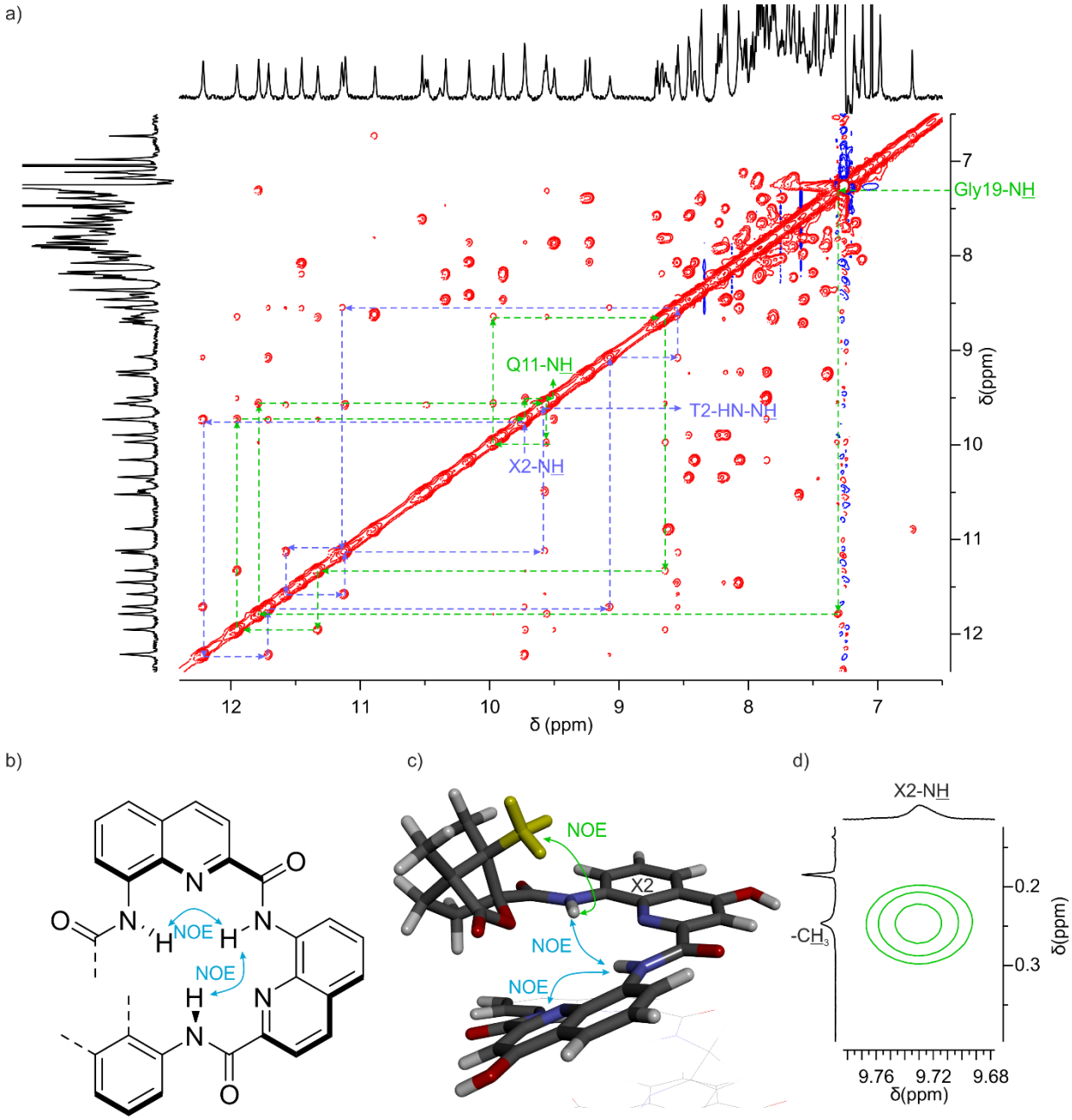


Figure S34. Assignment of N<u>H</u> signals by NOESY. a) Extract of an $^{1}H,^{1}H$ NOESY (500 MHz, dried $CDCl_3$) of **2b**. The NOE network of the NHs of the N-terminal helix is highlighted with light blue dashed lines, that of the C-terminal helix with green dashed lines. b) Arbitrary segment of the helical structure exemplarily illustrating intramolecular NOEs between one N<u>H</u> and NHs from two consecutive ones with cyan arrows. c) Structure illustrating the strategy for assigning the X2-N<u>H</u> signal. d) Corresponding cross peak from the NOESY spectrum confirming the strategy.

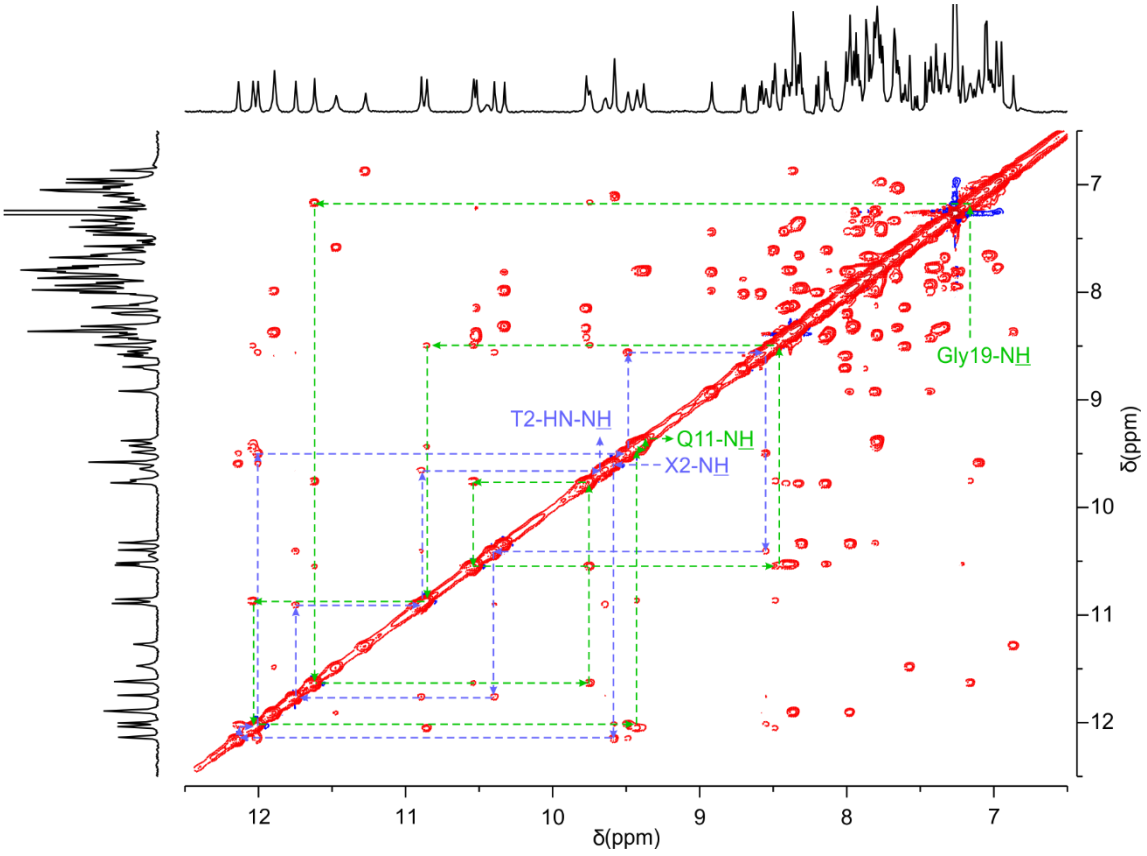


Figure S35. Assignment of N<u>H</u> signals by NOESY. Extract of an ^{1}H,^{1}H NOESY (500 MHz, wet $CDCl_3$) of **2b**. The NOE network of the NHs of the N-terminal helix is highlighted with light blue dashed lines, that of the C-terminal helix with green dashed lines.

3.5.2. Assignment of C<u>H</u>$_2$ signals

For the pyridine-based monomers (P and Y), correlations should be visible caused by the *J*-coupling of the NH protons with each of the unisochronous protons of the CH_2 group in the ^{1}H,^{1}H COSY spectrum (Figure S20a). The signal assignment of the species of **2b** in dried $CDCl_3$ is shown in Figure S36 and Figure S37 and the chemical shifts are reported in Table S1 and Table S2, respectively.

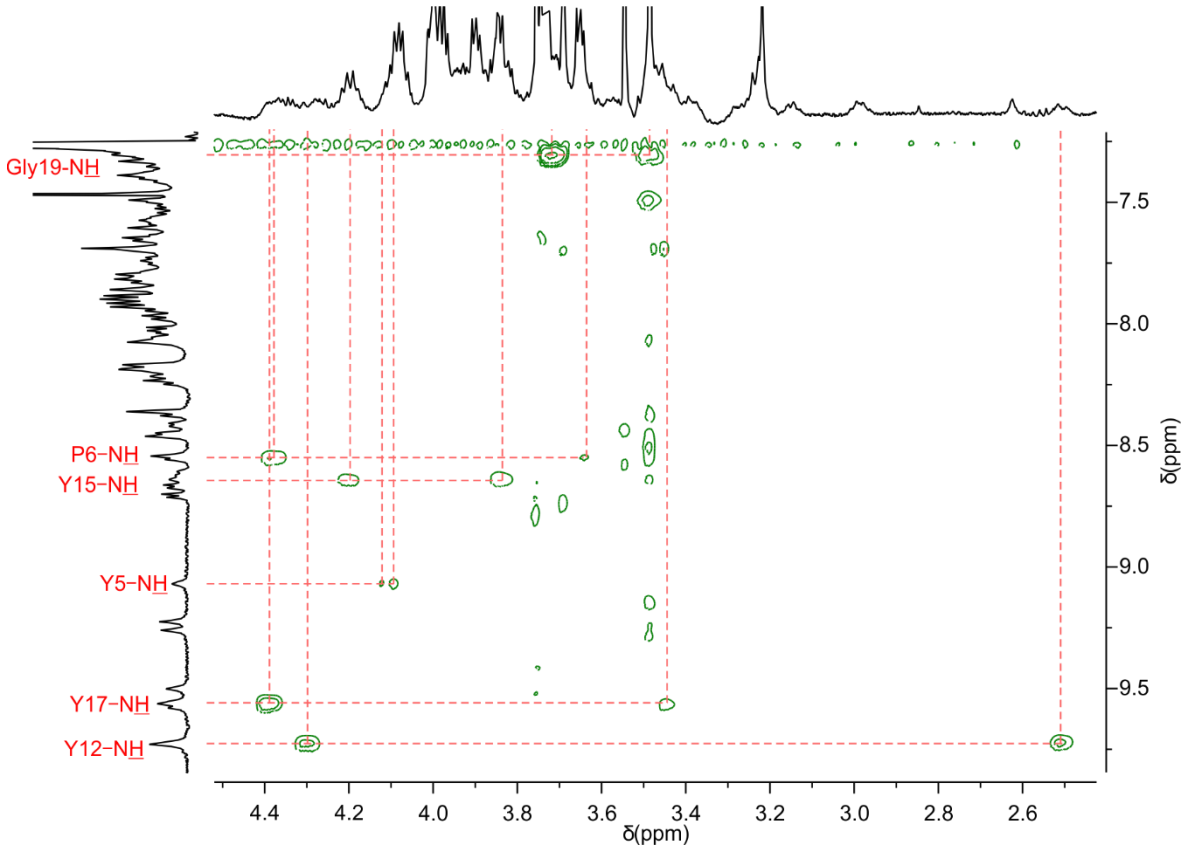


Figure S36. Assignment of C$\underline{H}_2$ signals of the P and Y units by COSY. Extract of an ^{1}H,^{1}H COSY (500 MHz, dried $CDCl_3$) of **2b**.

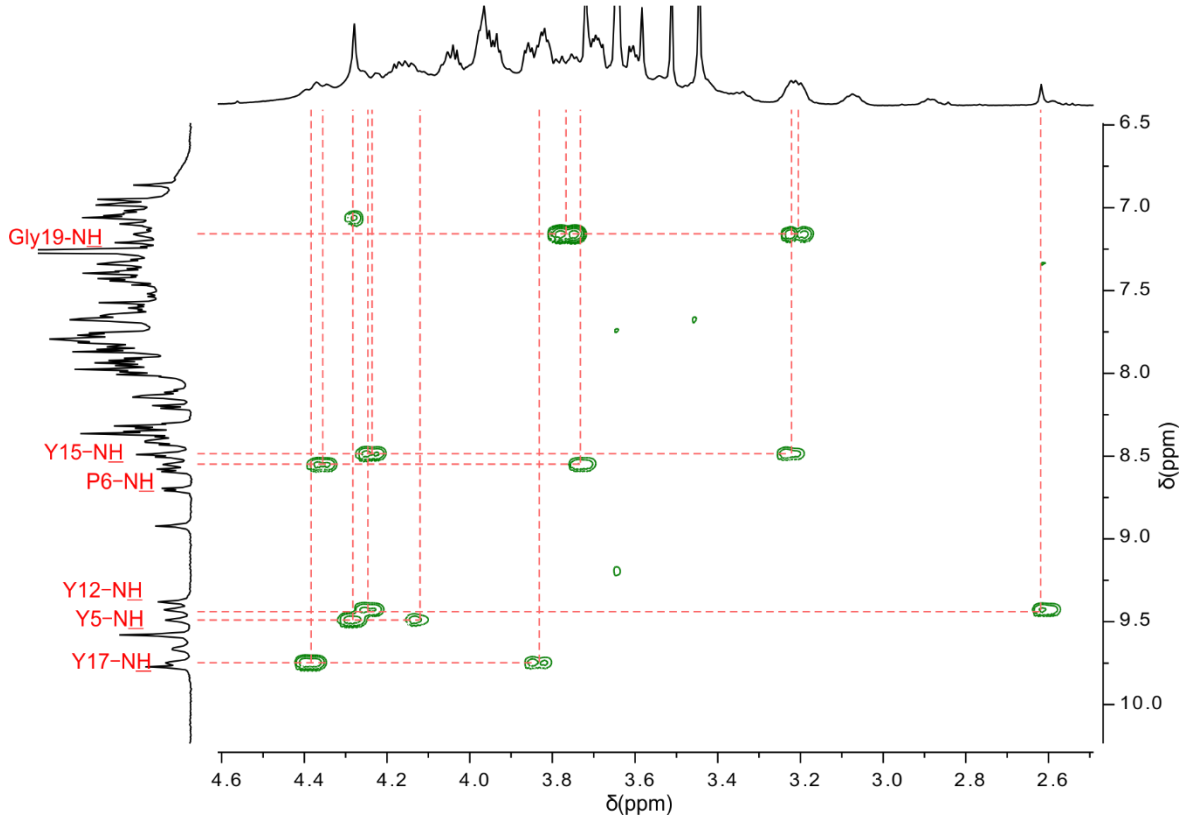


Figure S37. Assignment of C<u>H</u>$_2$ signals of P and Y units by COSY. Extract of an ^{1}H,^{1}H COSY (500 MHz, wet $CDCl_3$) of **2b**.

3.5.3. Supplementary tables

Table S1. Assignment of the chemical shifts of the species of **2b** in dried $CDCl_3$.

Monomer	Moiety	^{1}H (ppm)	^{15}N (ppm)
X2	NH	9.73	119.52
X3	NH	12.21	123.18
Q4	NH	11.71	120.97
Y5	NH	9.07	103.86
	C<u>H</u>$_2$	4.12, 4.09	
P6	NH	8.54	102.98
	C<u>H</u>$_2$	4.38, 3.64	
X7	NH	11.14	115.25
X8	NH	11.57	119.67
Q9	NH	11.11	116.01
T2	T2-HN-N<u>H</u>-	9.57	126.33

	T2-HN-NH-	10.49	128.27
Q11	NH	9.50	116.41
Y12	NH	9.72	104.23
	CH_2	4.31, 2.51	
Q13	NH	11.95	122.27
X14	NH	11.32	116.64
Y15	NH	8.63	101.36
	CH_2	4.20, 3.84	
Q16	NH	9.97	112.89
Y17	NH	9.55	108.62
	CH_2	4.40, 3.45	
Q18	NH	11.78	120.97
Gly19	NH	7.30	95.4
	CH_2	3.72, 3.50	

Table S2. Assignment of the chemical shifts of the species of **2b** in wet $CDCl_3$.

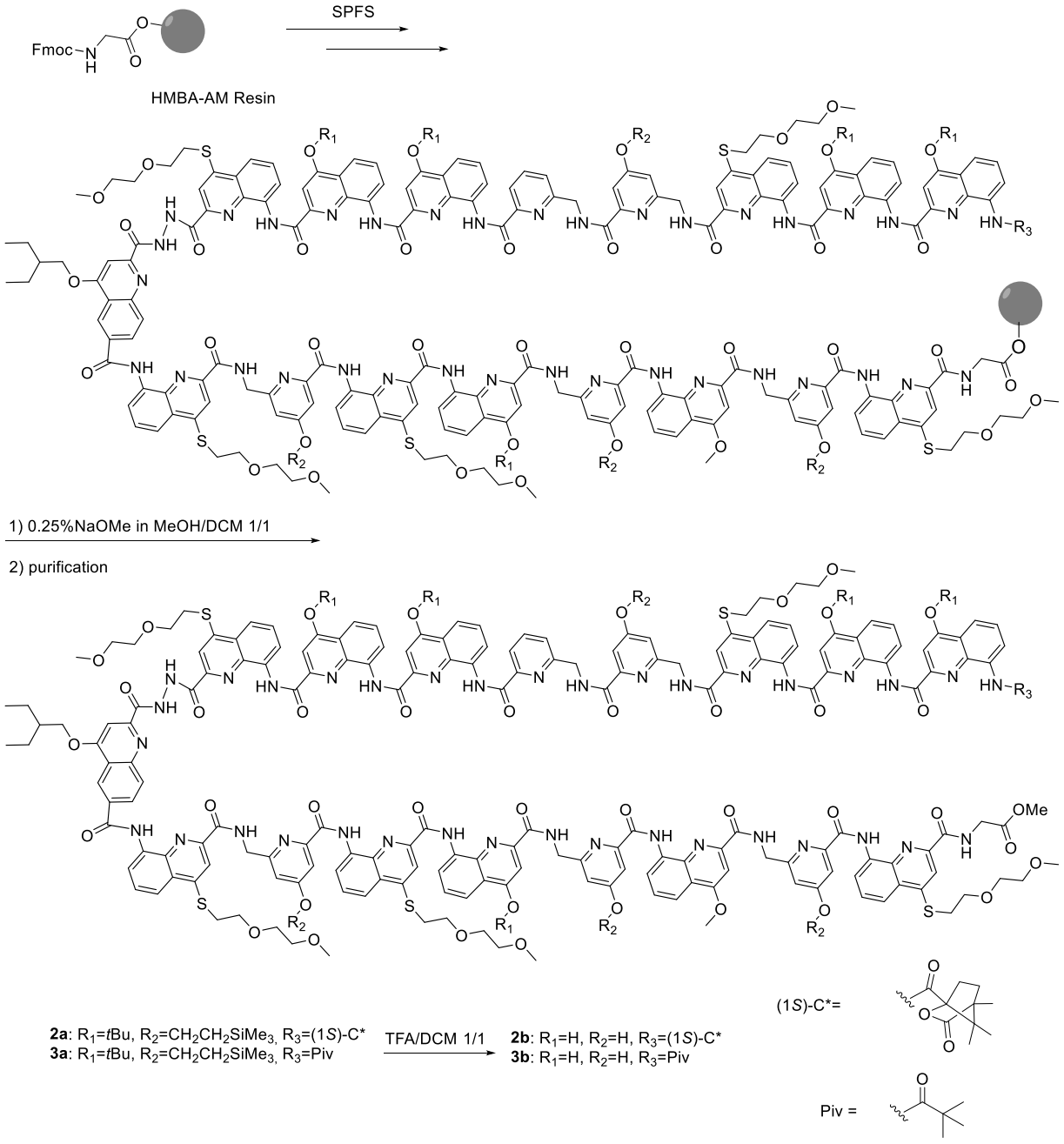
Monomer	Atom	^{1}H (ppm)	^{15}N (ppm)
X2	NH	9.59	119.27
X3	NH	12.14	123.00
Q4	NH	12.00	120.94
Y5	NH	9.49	106.01
	CH_2	4.23, 4.07	
P6	NH	8.56	102.65
	CH_2	4.28, 3.68	
X7	NH	10.43	113.54
X8	NH	11.76	121.03
Q9	NH	10.88	116.45
T2	T2-HN-NH-	9.58	126.21
	T2-HN-NH-	10.45	128.00
Q11	NH	9.39	116.67
Y12	NH	9.43	102.92
	CH_2	4.17, 2.55	
Q13	NH	12.04	123.00
X14	NH	10.87	115.76
Y15	NH	8.49	100.26
	CH_2	4.19, 3.17	
Q16	NH	10.54	114.66
Y17	NH	9.74	106.55
	CH_2	4.31, 3.76	

Q18	NH	11.63	120.99
Gly19	NH	7.17	94.98
	$C\underline{H}_2$	3.69, 3.13	

3.6. Circular dichroism spectroscopy

The CD spectra of **2a** and **2b** were recorded on a Jasco J-1500 spectrometer with 1 mm quartz cuvette. The following parameters were used: wavelength range from 460 to 280 nm. Scan speed: 100 nm/min; accumulation: 2; response time: 1.0 s; bandwidth: 2; temperature: 20 °C (unless specified); sensitivity: standard (100 mdeg); data pitch: 0.5 nm; nitrogen gas flow rate: 500 L/h. The concentration is 0.1 mM. $\Delta\varepsilon$ values (in $cm^2 \cdot mol^{-1}$) were obtained by using the formula: $\Delta\varepsilon = m°/(C.l.32980)$ where m = CD value in millidegrees; l = cuvette pathlength in cm; C = sample concentration in mol/L.

4. Synthetic scheme of foldamers synthesis



Scheme 1. Solid phase synthesis of aromatic oligomers.

5. Experimental procedures

5.1. General methods

Commercially available reagents were purchased from Sigma-Aldrich, Alfa-Aesar or TCI and were used without further purification unless specified. HMBA-AM resin (200-400 mesh, loading 0.8–1.2 mmol/g) was purchased from Iris-biotech. The anhydrous THF was dried over alumina columns (MBRAUN SPS-800 solvent purification system). DIPEA and chloroform were distilled over CaH_2 prior to use. Extra dry DMF was purchased from Sigma-Aldrich. Ultrapure water was obtained via a Stakpure OmniaPure-T UV-TOC ultrapure water system.

The $CDCl_3$ (99.80% D), CD_2Cl_2 (99.80% D) and DMSO-d_6 (99.80% D) for NMR measurement were purchased from Eurisotop®. The wet $CDCl_3$ and wet CD_2Cl_2 were prepared by shaking the chlorinated solvents with H_2O. It could be considered as water-saturated $CDCl_3$ or CD_2Cl_2 solution. The dried $CDCl_3$ and dried CD_2Cl_2 was prepared by stirring the chlorinated solvent with pre-dried basic Al_2O_3 overnight. The pre-dried basic Al_2O_3 was prepared by placing basic Al_2O_3 at 120°C overnight. Note that "dried" chlorinated solvent does not imply rigorously dry. The D_2O saturated $CDCl_3$ was prepared by shaking the dried $CDCl_3$ with D_2O.

Analytical reversed phase (RP) high performance liquid chromatography (HPLC) was performed on a Thermo Fisher Scientific Ultimate 3000 HPLC System using Macherey-Nagel Nucleodur C8 Gravity columns (4 × 50 mm, 5 μm). UV absorbance was monitored at 300 nm and 254 nm, if not stated otherwise. The semi-preparative HPLC was performed on a Waters system equipped with a 2545 Quaternary Gradient Module with automated fraction collector system on a XBridge® Prep C8 OBD™ column (19 × 150 mm, 5 μm) at a flow rate of 25 mL/min. 0.1% TFA was added to the aqueous mobile phase (referred to as mobile phase A) and to acetonitrile (referred to as mobile phase B). The gradient is: 0-5 min, 90% to 100% solvent B; 5-25 min, 100% solvent B at r.t.. The column eluent was monitored by UV detection at 254 and 300 nm with a diode array detector.

The ultraviolet–visible (UV/Vis) absorbance measurements were done with a Thermo Fisher Scientific Nanodrop One instrument using a 1 cm path length quartz cuvette. Circular dichroism (CD) spectra were measured on Jasco J-1500 spectrometers. Measurements were performed at 20 °C if not stated otherwise.

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 Chorus PurePep ® synthesizer followed by previously reported protocol.[32]

5.2. Synthesis of monomers

The Fmoc-Y-OH,[21] Fmoc-T2-OH,[21] Fmoc-Q^{Deg}-OH,[19] Fmoc-X-OH[19] and Fmoc-P-OH[33] were synthesized according to literature. All of the Fmoc-protected monomers was ≥ 98% pure before used in the solid phase synthesis.

5.3. Solid phase synthesis general methods

5.3.1. Loading of the resin via HBTU activation

HMBA-AM resin (500 mg, 0.4-0.6 mmol, 1 eq.) was swollen in 5 mL DCM for 1 h, transferred to the microwave vessel and washed 3 times with extra dry DMF. DIPEA (170 µL, 1.0 mmol, 2 eq.) was added to a mixture of Fmoc-Gly-OH (134 mg; 0.45 mmol, 0.9 eq.) and HBTU (228 mg, 1.2 eq.) in extra dry DMF (5 mL) and the resulting solution was shaken for 30 s before to be poured to the resin-containing reaction vessel. The reaction mixture was subjected to treatment in a microwave oven (50 °C, 20 min, 25 W). The resin was filtered and washed with DMF (5 x 2 mL) and DCM (10 x 2 mL). Capping was performed by adding a mixture of DCM/pyridine/benzoyl chloride (3:1:1 (v/v/v), 5 mL) to the resin followed by shaking for 30 min at r.t., and subsequent washing with DCM (20 x 2 mL). For monitoring the efficiency of the 1st loading, small amount of resin (around 2 mg) was taken and dried under vacuum. The loading was estimated at this scale.

5.3.2. Esimation of the Loading

To a small amount of Fmoc-Gly-HMBA AM resin or Fmoc-Q^{B}-SASRIN resin (1–2 mg), a freshly prepared solution of DMF/piperidine (8:2 (v/v), 3.0 mL) was added. The mixture was shaken and incubated for 5 min. Then the absorption was measured at 301 nm using a NanoDrop One Microvolume UV-Vis Spectrophotometer and a Hellma quartz glass cuvette 104 (path length 10 mm). Three replicates were measured, then the loading was calculated with the following equation:

$$\text{loading}\left(\text{in } \frac{\text{mmol}}{\text{g}}\right) = \frac{\text{Abs}_{301\,\text{nm}} \times \text{V}}{\varepsilon_{301\text{nm}} \times l \times \text{m}}$$

$$\varepsilon_{301\text{nm}} = 7800 \text{ L/mol/cm}^{[34]}$$

5.3.3. Solid phase synthesis via in-situ activation

Fmoc-Gly-HMBA-AM resin (100 mg, loading 0.3 mmol/g, 30 µmol) was first swollen in DCM (3 mL) for 1 h, the resin was transferred into the microwave vessel and washed 3 times with DMF and 3 times with NMP.

The deprotection of the Fmoc group was performed by adding 3 mL DBU/NMP (2:98, v/v) to the resin and incubation for 3 min. The resin was next filtered off and the deprotection step was repeated once. After filtration, the resin was washed with DCM (3 x 2 mL) and then with anhydrous THF (5 x 2 mL). This deprotection step was performed after each aromatic monomer coupling.

The resin was next suspended in anhydrous THF (1 mL) and 2,3,5-collidine (5 eq. with respect to the resin-loading) was added to the resin supernatant. The Fmoc-protected monomer (2 eq. with respect to the resin-loading) and PPh_3 (4 eq. with respects to the resin-loading) were successively added in a vial to be solubilized in freshly distilled $CHCl_3$ (1 mL).

Trichloroacetonitrile (4.5 eq. with respect to the resin loading) was next added to the vial and the resulting acyl chloride solution was shaken for 30 s before to be poured to the resin-containing reaction vessel. The reaction vessel was then placed in the microwave oven and subjected to MW irradiation for 15 min (50 °C, 50 W). The resin was then washed 3 times with anhydrous THF. This entire coupling step was then repeated once more. For the coupling of Fmoc-T2-OH, the same acid chloride activation process was followed but the coupling was, this time, performed at r.t. by shaking the resin for 2 h. Then the resin was washed with anhydrous THF (3 x 2 mL).

For the final coupling of pivaloyl- (Piv-) or (1*S*)-camphanic ((1*S*)-C*-) amides, the resin was suspended in anhydrous THF (1 mL) and 2,3,5-collidine (5 eq. with respect to the resin-loading) was added to the resin suspensions. A solution of pivaloyl chloride (2 eq. with respect to the resin-loading) or (1*S*)-camphanic chloride (2 eq. with respect to the resin-loading, purchased from Sigma-Aldrich, 98%, *ee*: 99%) in freshly distilled $CHCl_3$ (1 mL) was added to the supernatant and the resin was shaken at r.t. for 2 h. The resin was filtered off, washed 3 times with dry THF, and the same process was repeated once. After coupling, the resin was vigorously washed 3 times with DMF and 3 times with DCM.

5.3.4. Mini-cleavage

To perform a mini cleavage, the resin (1–2 mg) was swollen in 1 mL MeOH/DCM (1:1, v/v) solution followed by addition of 10 µL NaOMe (25% (m/m)) and incubated at r.t. for 10 min. The cleavage solution was

diluted with DCM, washed with aqueous citric acid solution (5%), dried over $MgSO_4$, filtered and the solvent was finally removed under reduced pressure.

5.3.5. Full-cleavage

The resin (around 100 mg) was dried under vacuum and slowly added to 400 mL cleavage solution (preparation see below) under N_2 atmosphere. The mixture was stirred under N_2 atmosphere for 2 h before it was added to 100 mL aqueous citric acid solution (5%). The aqueous layer was extracted with DCM (3 x 50 mL). The combined organic phases were washed with brine, dried over $MgSO_4$, filtered and the solvent was evaporated under reduced pressure. The crude was recovered as solid.

Preparation of cleavage solution: 200 mL dry MeOH was added to 200 mL dry DCM under N_2 atmosphere. 2 mL NaOMe (25% (m/m)) in methanol was added and the mixture was well-mixed by magnetic stirring. Sufficient amount of cleavage solution (at least 400 mL cleavage solution for 100 mg resin) was important to avoid the formation of oligomer acid as the side-product.

5.4. Synthesis of oligomers

(1*S*)-Camph-$\underline{XX}Q^D\underline{Y}P\underline{XX}Q^D$-T2-$Q^D\underline{Y}Q^D\underline{XY}Q^M\underline{Y}Q^D$-Gly-OMe (**2a**): Compound **2a** was synthesized using the SPS procedures reported in paragraph 5.3 on Fmoc-Gly-HMBA AM resin. The crude product was obtained after full cleavage and purification by RP-HPLC. (35 mg, 15%).**^{1}H NMR** (500 MHz, $CDCl_3$) δ 11.71 (s, 1H), 11.56 (s, 1H), 11.37 (s, 1H), 11.31 (s, 1H), 11.02 (s, 1H), 10.95 (s, 1H), 10.86 (s, 1H), 10.60 (s, 1H), 10.52 (s, 1H), 10.19 (d, *J* = 8.3 Hz, 1H), 9.81 (d, *J* = 6.3 Hz, 1H), 9.62 (s, 1H), 9.40 (s, 1H), 8.64 (t, *J* = 3.3 Hz, 1H), 8.61–8.58 (m, 2H), 8.09–8.02 (m, 3H), 7.96–7.93 (m, 2H), 7.91–7.86 (m, 2H), 7.80–7.75 (m, 4H), 7.74–7.67 (m, 4H), 7.63 (d, *J* = 8.1 Hz, 1H), 7.61–7.57 (m, 3H), 7.57–7.51 (m, 5H), 7.49–7.46 (m, 5H), 7.44 (d, *J* = 7.9 Hz, 1H), 7.40 (d, *J* = 7.0 Hz, 1H), 7.35 (s, 1H), 7.33–7.28 (m, 3H), 7.23 (dd, *J* = 7.6, 2.4 Hz, 2H), 7.19 (d, *J* = 1.9 Hz, 1H), 7.17 (s, 2H), 7.13 (d, *J* = 6.5 Hz, 1H), 7.04–6.96 (m, 5H), 6.94–6.88 (m, 4H), 6.80 (s, 1H), 6.67 (s, 1H), 6.57 (s, 1H), 6.46 (s, 1H), 6.44 (s, 1H), 6.42 (s, 1H), 6.36 (s, 1H), 6.29 (s, 1H), 6.25 (s, 1H), 6.16 (s, 2H), 5.63 (s, 1H), 4.25 (dd, *J* = 16.2, 5.1 Hz, 1H), 4.10–4.02 (m, 5H), 4.00–3.86 (m, 25H), 3.81–3.74 (m, 12H), 3.72–3.65 (m, 9H), 3.63–3.60 (m, 4H), 3.58 (s, 3H), 3.55 (s, 3H), 3.54 (s, 2H), 3.50 (s, 3H), 3.50 (s, 6H), 3.45 (s, 3H), 3.39 (s, 3H), 3.35–3.06 (m, 11H), 3.02–2.88 (m, 3H), 2.45–2.39 (m, 1H), 2.16 (d, *J* = 15.0 Hz, 1H), 2.02 (d, *J* = 14.1 Hz, 1H), 1.97–1.93 (m, 1H), 1.89–1.85 (m, 3H), 1.81 (s, 6H), 1.77 (s, 9H), 1.74 (s, 9H), 1.67 (s, 9H), 1.64 (s, 9H), 1.61 (d, *J* = 9.8

Hz, 6H), 1.49 (s, 2H), 1.13–1.00 (m, 2H), 0.70 (t, *J* = 7.4 Hz, 2H), 0.59 (s, 3H), 0.54 (s, 1H), 0.50 (t, *J* = 7.4 Hz, 2H), 0.22 (s, 9H), 0.07 (s, 9H), -0.02 (s, 9H), -0.46 (s, 9H). **HRMS** (ESI+) calcd. for $C_{241}H_{274}N_{36}O_{44}S_5Si_4$ $[M+2H]^{2+}$ 2324.8841, found 2324.9068.

(1*S*)-Camph-XXQ^{D}YPXXQ^{D}-T2-Q^{D}YQ^{D}XYQ^{M}YQ^{D}-Gly-OMe (**2b**): Compound **2a** was treated with a solution of TFA/DCM (1:1 (v/v), 2 mL) at r.t. overnight. The solvent was removed under vacuum. The solid was precipitated from diethyl ether, was subsequently filtered and washed 2 times with diethyl ether to yield the desired product as a yellow solid. (30 mg, quant.). **^{1}H NMR** (500 MHz, dry $CDCl_3$) δ 12.21 (s, 1H), 11.94 (s, 1H), 11.79 (s, 1H), 11.71 (s, 1H), 11.56 (s, 1H), 11.45 (s, 1H), 11.32 (s, 1H), 11.12 (s, 1H), 10.91 (s, 1H), 10.51 (s, 1H), 10.47 (d, *J* = 9.4 Hz, 1H), 10.34 (s, 1H), 10.15 (s, 1H), 9.98 (s, 1H), 9.90 (s, 1H), 9.73 (s, 2H), 9.58 (d, *J* = 8.3 Hz, 1H), 9.55 (s, 1H), 9.50 (s, 1H), 9.28 (s, 1H), 9.23 (s, 1H), 9.09 (s, 1H), 8.71 (d, *J* = 6.7 Hz, 1H), 8.68–8.61 (m, 2H), 8.56–8.53 (m, 2H), 8.46 (d, *J* = 6.3 Hz, 2H), 8.41 (d, *J* = 6.0 Hz, 1H), 8.37 (d, *J* = 6.7 Hz, 2H), 8.26–8.20 (m, 3H), 8.17 (d, *J* = 8.2 Hz, 3H), 8.08–8.04 (m, 3H), 8.02 (d, *J* = 7.0 Hz, 2H), 7.97 (d, *J* = 7.1 Hz, 2H), 7.95 (s, 1H), 7.93 (s, 1H), 7.91 (s, 1H), 7.90 (s, 1H), 7.88 (s, 1H), 7.86 (s, 1H), 7.83 (s, 1H), 7.81 (d, *J* = 2.3 Hz, 2H), 7.80 (s, 1H), 7.79 (s, 1H), 7.76–7.72 (m, 2H), 7.69 (d, *J* = 4.3 Hz, 2H), 7.66–7.63 (m, 3H), 7.61 (s, 1H), 7.59–7.55 (m, 2H), 7.53 (d, *J* = 8.6 Hz, 2H), 7.49 (t, *J* = 6.7 Hz, 2H), 7.45 (s, 2H), 7.40 (d, *J* = 7.9 Hz, 3H), 7.35 (dd, *J* = 5.2, 3.2 Hz, 2H), 7.19–7.16 (m, 2H), 7.11 (s, 2H), 6.98 (d, *J* = 2.3 Hz, 2H), 6.71 (s, 1H), 6.17 (s, 1H), 5.86 (s, 1H), 5.37–5.30 (m, 2H), 4.64 (s, 3H), 4.36–4.16 (m, 4H), 4.12–4.05 (m, 4H), 4.02–3.96 (m, 9H), 3.95–3.79 (m, 13H), 3.75 (s, 3H), 3.75–3.72 (m, 4H), 3.69 (s, 3H), 3.66–3.63 (m, 4H), 3.54 (s, 3H), 3.49 (s, 3H), 3.48 (s, 3H), 3.47–3.35 (m, 4H), 3.22 (s, 3H), 3.18–3.12 (m, 1H), 3.02–2.95 (m, 1H), 2.26–2.19 (m, 4H), 2.03–1.97 (m, 4H), 1.61–1.07 (m, 20H), 0.83–0.53 (m, 4H). **^{1}H NMR** (500 MHz, wet $CDCl_3$) δ 12.14 (s, 1H), 12.04 (s, 1H), 12.00 (s, 1H), 11.89 (s, 1H), 11.80 (s, 1H), 11.76 (s, 1H), 11.63 (s, 1H), 11.43 (s, 1H), 11.22 (s, 1H), 10.88 (s, 1H), 10.86 (s, 1H), 10.53 (s, 1H), 10.51 (s, 1H), 10.45 (d, *J* = 7.6 Hz, 1H), 10.42 (s, 1H), 10.32 (s, 1H), 9.77 (s, 1H), 9.73 (s, 1H), 9.58 (s, 1H), 9.57 (d, *J* = 7.6 Hz, 1H), 9.48 (s, 1H), 9.42 (s, 1H), 9.38 (s, 1H), 8.91 (s, 1H), 8.70 (d, *J* = 7.0 Hz, 1H), 8.58 (d, *J* = 7.5 Hz, 1H), 8.55 (d, *J* = 3.3 Hz, 1H), 8.51–8.47 (m, 2H), 8.44–8.41 (m, 2H), 8.40–8.35 (m, 3H), 8.34–8.29 (m, 3H), 8.20 (d, *J* = 8.1 Hz, 1H), 8.16–8.10 (m, 3H), 8.02–7.97 (m, 5H), 7.96–7.91 (m, 5H), 7.88–7.84 (m, 3H), 7.83–7.76 (m, 7H), 7.70–7.62 (m, 5H), 7.60 (t, *J* = 7.5 Hz, 1H), 7.57 (s, 1H), 7.45–7.42 (m, 2H), 7.40–7.32 (m, 4H), 7.23 (d, *J* = 9.0 Hz, 2H), 7.15 (s, 1H), 7.12 (s, 1H), 7.07–7.04 (m, 2H), 7.02 (d, *J* = 7.5 Hz, 1H), 6.98 (d, *J* = 6.1 Hz, 2H), 6.95 (s, 1H), 6.88 (s, 1H), 5.79 (d, *J* = 6.9 Hz, 2H), 4.42–4.32 (m, 1H), 4.28 (s, 3H), 4.26–4.15 (m, 4H), 4.15–4.10 (m, 5H), 4.07–4.02 (m, 4H), 3.99–3.93 (m, 8H), 3.87–3.73 (m, 10H), 3.72 (s, 3H), 3.70 (s,

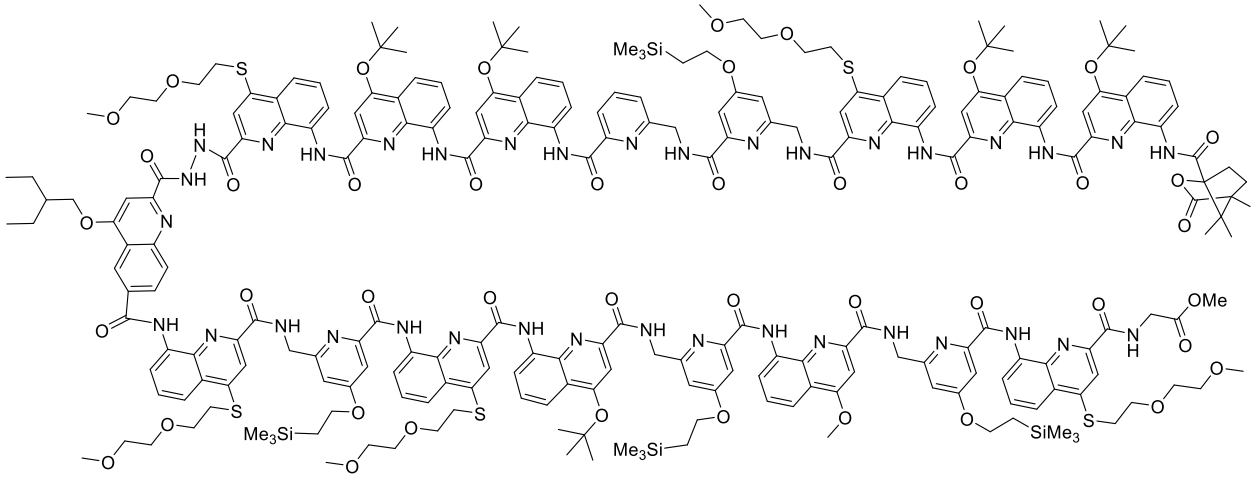
3H), 3.70–3.68 (m, 4H), 3.65 (s, 3H), 3.64 (s, 3H), 3.62–3.60 (m, 3H), 3.59 (s, 2H), 3.55–3.53 (m, 1H), 3.51 (s, 3H), 3.50–3.46 (m, 1H), 3.45 (s, 3H), 3.34 (t, *J* = 7.6 Hz, 1H), 3.26–3.18 (m, 5H), 3.11–3.04 (m, 2H), 2.93–2.84 (m, 0H), 2.62 (s, 3H), 1.29–1.22 (m, 6H), 0.88 (t, *J* = 6.8 Hz, 3H), 0.71 (s, 3H), 0.64 (s, 3H), 0.60 (t, *J* = 7.5 Hz, 3H), 0.39 (t, *J* = 7.4 Hz, 3H), 0.28 (s, 3H). **HRMS** (ESI+) calcd. for $C_{201}H_{186}N_{36}O_{44}S_5$ $[M+2H]^{2+}$ 1984.5326, found 1984.6086.

Piv-<u>XX</u>Q^{D}<u>Y</u>P<u>XX</u>Q^{D}-T2-Q^{D}<u>Y</u>Q^{D}<u>XY</u>Q^{M}<u>Y</u>Q^{D}-Gly-OMe (**3a**): Compound **3a** was synthesized using the SPS procedures reported in paragraph 5.3 on Fmoc-Gly-HMBA AM resin. The crude product was obtained after full cleavage and purification by RP-HPLC. (15 mg, 13%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 11.71 (s, 1H), 11.41 (s, 1H), 11.31 (s, 1H), 11.01 (s, 1H), 10.95 (s, 1H), 10.92 (s, 1H), 10.86 (s, 1H), 10.55 (s, 1H), 10.52 (s, 1H), 10.21 (d, *J* = 7.1 Hz, 1H), 9.82 (d, *J* = 5.4 Hz, 1H), 9.61 (s, 1H), 8.66 (s, 1H), 8.63 (t, *J* = 3.3 Hz, 1H), 8.59 (s, 2H), 8.11 (d, *J* = 7.3 Hz, 1H), 8.08–8.03 (m, 4H), 7.97 (t, *J* = 3.3 Hz, 1H), 7.91–7.85 (m, 4H), 7.80–7.77 (m, 2H), 7.76 (s, 2H), 7.73 (d, *J* = 8.4 Hz, 1H), 7.70 (s, 1H), 7.69 (s, 1H), 7.65 (d, *J* = 8.0 Hz, 1H), 7.63–7.59 (m, 3H), 7.58–7.51 (m, 4H), 7.47–7.38 (m, 5H), 7.31 (dd, *J* = 17.6, 7.9 Hz, 3H), 7.23 (d, *J* = 7.1 Hz, 2H), 7.19 (d, *J* = 1.7 Hz, 1H), 7.16 (s, 2H), 7.13 (d, *J* = 7.1 Hz, 1H), 7.09–7.02 (m, 4H), 6.99 (d, *J* = 7.8 Hz, 1H), 6.96–6.89 (m, 5H), 6.80 (s, 1H), 6.67 (s, 1H), 6.55–6.51 (m, 1H), 6.49 (d, *J* = 5.9 Hz, 0H), 6.46 (s, 1H), 6.43 (s, 1H), 6.42 (s, 1H), 6.36 (s, 1H), 6.31–6.27 (m, 1H), 6.24 (s, 1H), 6.23–6.21 (m, 1H), 6.16 (s, 1H), 6.10 (d, *J* = 1.7 Hz, 1H), 5.60 (s, 1H), 4.25 (dd, *J* = 16.1, 5.1 Hz, 1H), 4.10–4.03 (m, 4H), 4.02–3.93 (m, 9H), 3.92 (s, 3H), 3.91–3.85 (m, 8H), 3.82–3.74 (m, 9H), 3.71–3.65 (m, 7H), 3.63–3.60 (m, 4H), 3.58 (s, 3H), 3.55 (s, 3H), 3.50 (s, 3H), 3.49 (s, 3H), 3.45 (s, 3H), 3.44–3.39 (m, 2H), 3.38 (s, 3H), 3.35–3.20 (m, 5H), 3.19–3.15 (m, 2H), 3.15–3.06 (m, 5H), 3.01–2.88 (m, 2H), 2.45–2.39 (m, 1H), 2.11 (d, *J* = 16.0 Hz, 1H), 2.02 (d, *J* = 13.5 Hz, 1H), 1.87 (s, 2H), 1.80 (s, 9H), 1.78 (s, 9H), 1.77 (s, 6H), 1.74 (s, 9H), 1.74 (s, 9H), 1.64 (s, 9H), 1.61 (s, 2H), 1.50 (s, 2H), 1.14–0.99 (m, 4H), 0.69 (t, *J* = 7.4 Hz, 2H), 0.58 (s, 9H), 0.50 (t, *J* = 7.4 Hz, 2H), 0.22 (s, 9H), 0.07 (s, 9H), -0.02 (s, 9H), -0.46 (s, 9H). **HRMS** (ESI+) calcd. for $C_{236}H_{270}N_{36}O_{42}S_5Si_4$ $[M+2H]^{2+}$ 2276.8602, found 2276.8962.

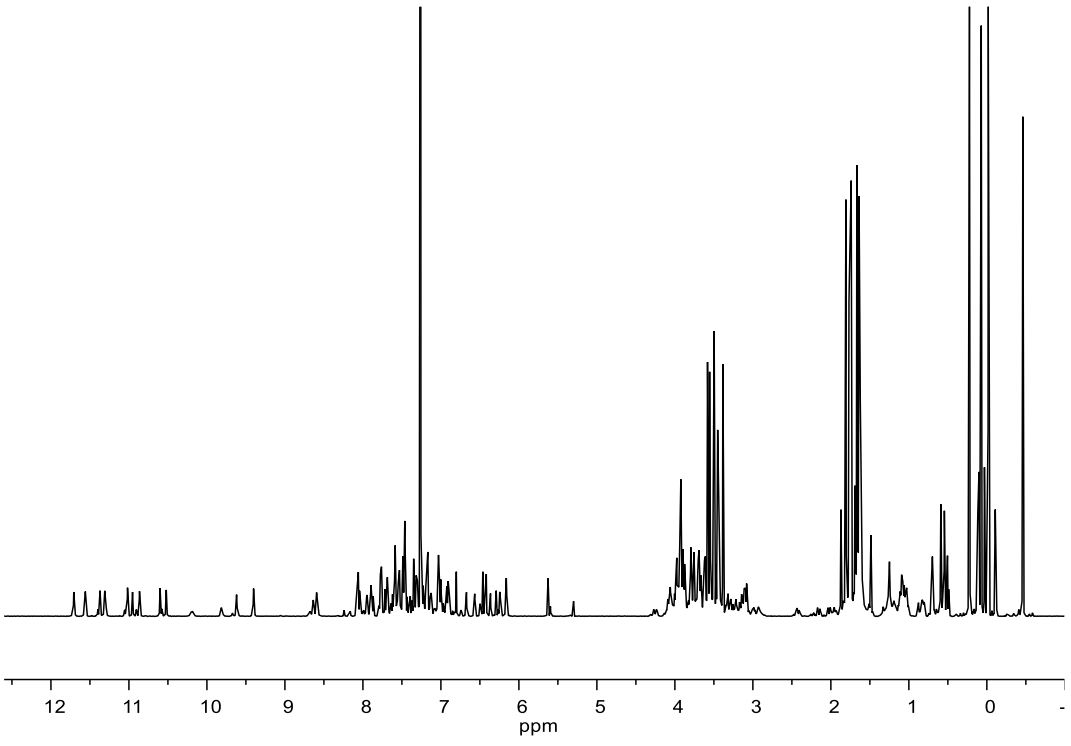
Piv-XXQ^{D}YPXXQ^{D}-T2-Q^{D}YQ^{D}XYQ^{M}YQ^{D}-Gly-OMe (**3b**): Compound **3a** was treated with a solution of TFA/DCM (1:1 (v/v), 2 mL) at r.t. overnight. The solvent was removed under vacuum. The solid was precipitated from diethyl ether, was subsequently filtered and washed 2 times with diethyl ether to yield the desired product as a yellow solid. (13 mg, quant.). **^{1}H NMR** (500 MHz, wet $CDCl_3$) δ 11.96 (s, 1H), 11.82 (s, 1H), 11.79 (s, 1H), 11.76 (s, 2H), 11.72 (s, 1H), 11.67 (s, 1H), 11.56 (s, 1H), 11.11 (s, 1H), 10.81 (s, 1H), 10.78 (s, 1H), 10.47 (s, 1H), 10.42 (d, *J* = 9.5 Hz, 1H), 10.39 (s, 1H), 10.33 (s, 1H), 10.25 (s, 1H), 9.70 (s, 1H), 9.64 (s, 1H), 9.47 (s, 2H),

9.36 (s, 2H), 9.32 (s, 1H), 8.87 (s, 1H), 8.81 (s, 1H), 8.65 (d, *J* = 7.2 Hz, 1H), 8.53 (d, *J* = 8.0 Hz, 2H), 8.45 (d, *J* = 7.5 Hz, 1H), 8.42 (s, 1H), 8.37 (s, 1H), 8.36 (s, 2H), 8.34 (s, 1H), 8.30 (d, *J* = 5.8 Hz, 2H), 8.28 (s, 1H), 8.27 (s, 1H), 8.24 (d, *J* = 7.3 Hz, 1H), 8.13 (d, *J* = 8.3 Hz, 1H), 8.10 (d, *J* = 6.3 Hz, 2H), 8.05 (s, 1H), 7.97–7.91 (m, 3H), 7.86 (p, *J* = 7.6, 6.6 Hz, 4H), 7.82–7.79 (m, 1H), 7.78 (s, 1H), 7.78–7.73 (m, 4H), 7.72–7.67 (m, 4H), 7.65–7.54 (m, 6H), 7.38–7.35 (m, 3H), 7.35–7.31 (m, 1H), 7.30–7.28 (m, 2H), 7.27 (s, 1H), 7.14 (s, 1H), 7.08–7.05 (m, 1H), 7.02 (dd, *J* = 8.5, 2.5 Hz, 1H), 7.00 (s, 1H), 6.99–6.96 (m, 3H), 6.92 (d, *J* = 6.9 Hz, 2H), 6.89 (d, *J* = 7.2 Hz, 1H), 6.87 (s, 1H), 6.80 (s, 1H), 5.78 (s, 1H), 5.73 (s, 1H), 4.34–4.24 (m, 2H), 4.22 (s, 2H), 4.21–4.08 (m, 6H), 4.03–3.95 (m, 6H), 3.93–3.86 (m, 9H), 3.82–3.68 (m, 10H), 3.65 (s, 3H), 3.65–3.61 (m, 3H), 3.59 (d, *J* = 2.0 Hz, 6H), 3.56–3.53 (m, 2H), 3.50 (s, 3H), 3.46 (s, 3H), 3.44–3.40 (m, 1H), 3.39 (s, 3H), 3.37–3.34 (m, 1H), 3.31–3.21 (m, 2H), 3.20–3.10 (m, 5H), 3.08–2.98 (m, 2H), 2.86–2.68 (m, 2H), 2.59–2.51 (m, 2H), 2.18–2.14 (m, 2H), 1.39–1.18 (m, 20H). **HRMS** (ESI+) calcd. for $C_{196}H_{182}N_{36}O_{42}S_5$ $[M+2H]^{2+}$ 1936.5660, found 1936.5981.

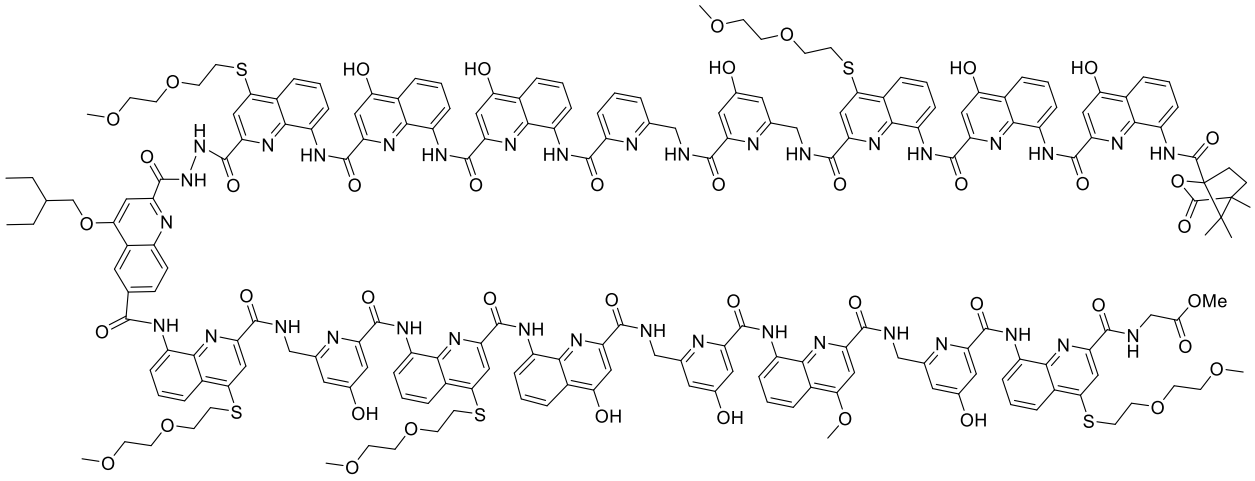
6. NMR spectra of new compounds



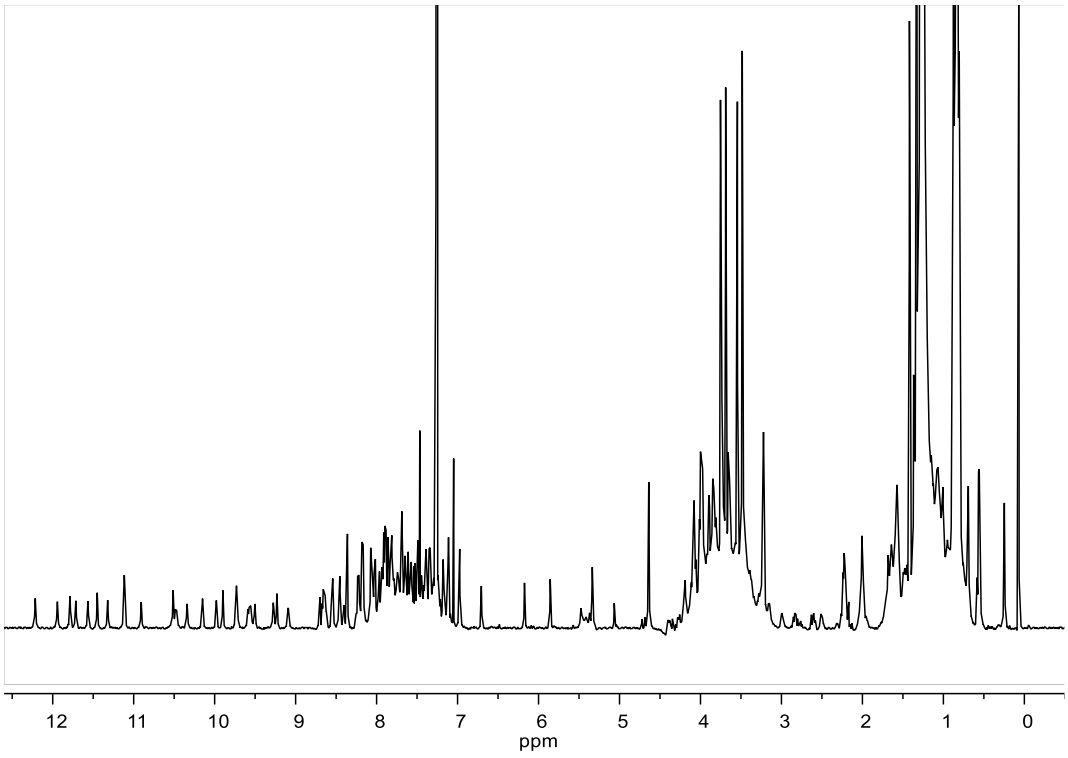
Chemical structure of **2a**.



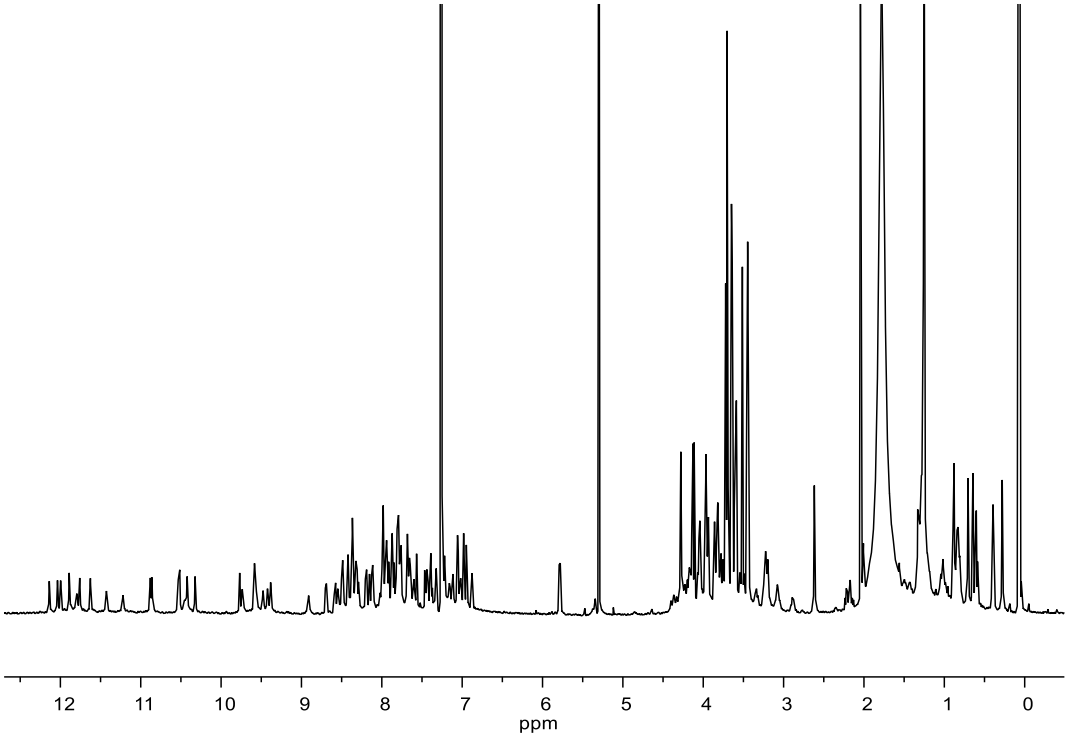
^{1}H NMR spectrum (500 MHz, $CDCl_3$) of **2a**.



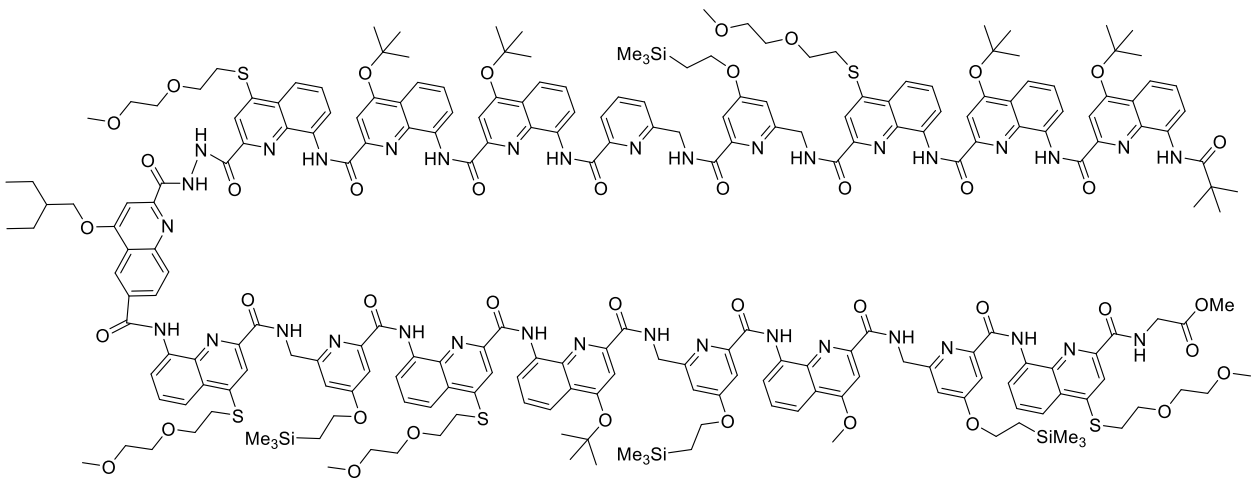
Chemical structure of **2b**.



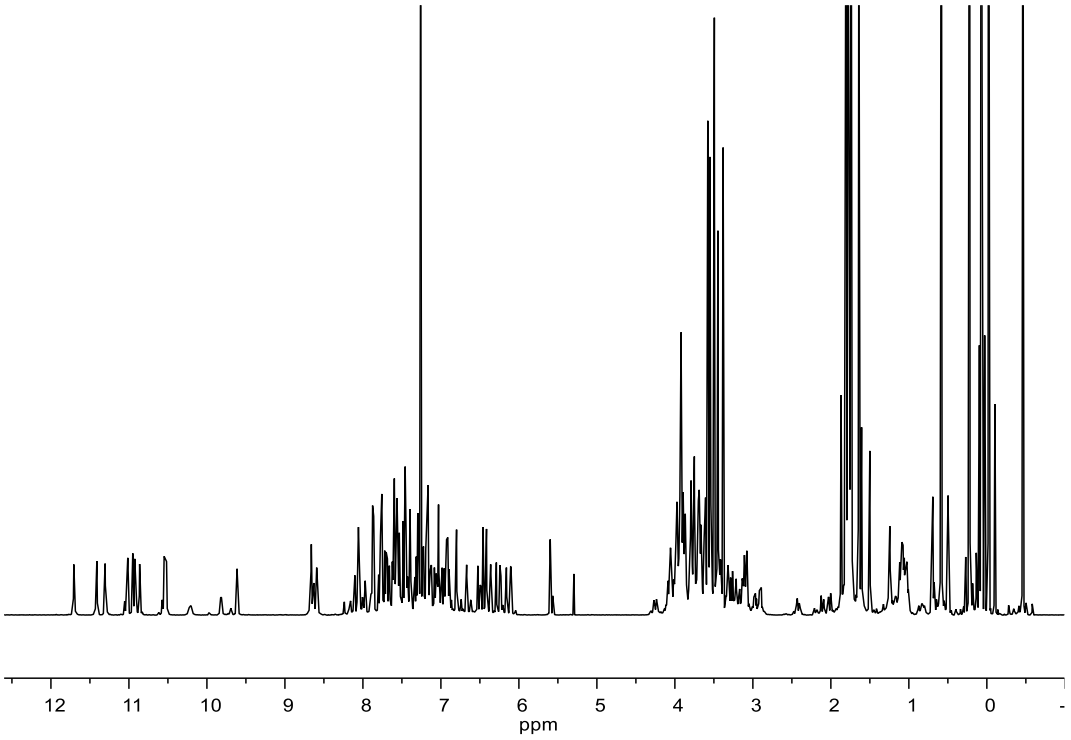
^{1}H NMR spectrum (500 MHz, dried $CDCl_3$) of **2b**.



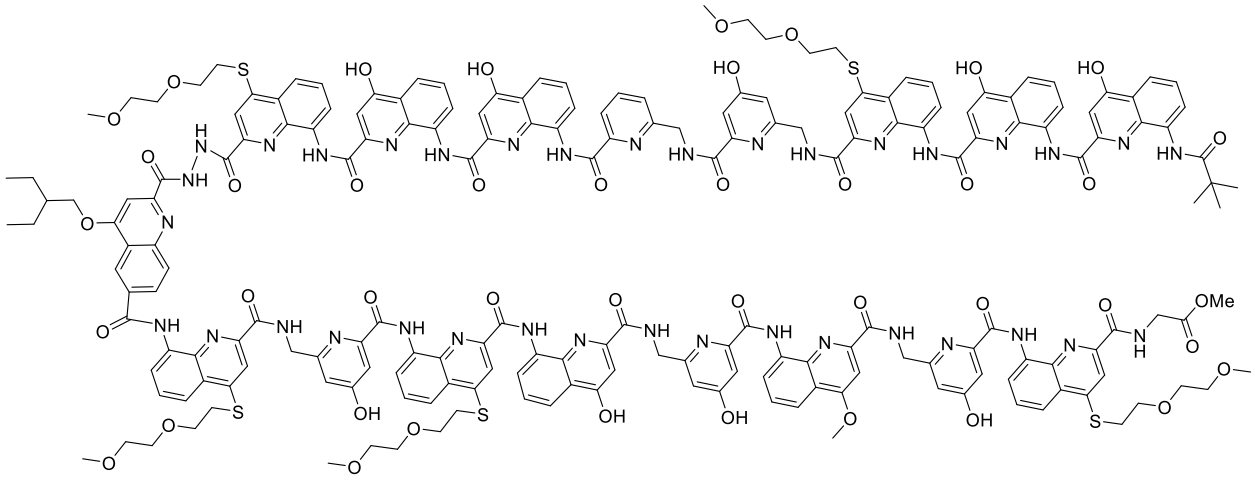
^{1}H NMR spectrum (500 MHz, wet $CDCl_3$) of **2b**.



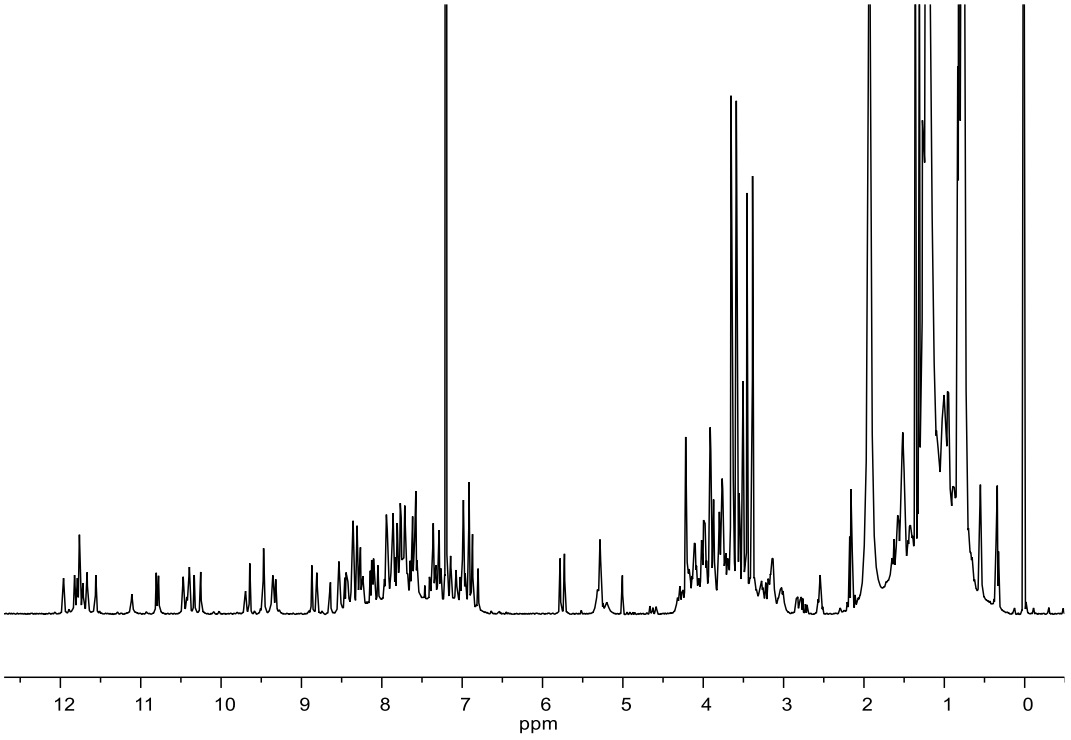
Chemical structure of **3a**.



^{1}H NMR spectrum (500 MHz, $CDCl_3$) of **3a**.



Chemical structure of **3b**.



^{1}H NMR spectrum (500 MHz, wet $CDCl_3$) of **3b**.

6. Design of an abiotic unimolecular three-helix bundle

As mentioned in Chapter 3, the hydroxy group-decorated single helices will aggregate into homochiral tilted dimers and homochiral parallel trimers in the absence of the turn unit T1. Our previous study took the tilted dimer structure as the starting point to design and synthesize a series of helix-turn-helix structures mediated by a tilted dimer interface. The design of a unimolecular three-helix bundle tertiary structure described in this chapter follows a similar strategy, using the parallel trimer structure as another favored aggregate. However, the three helices in the parallel trimer have the same handedness and orientation, resulting in the N-termini being distant from the C-termini of the two adjacent helices. Covalently connecting the N-terminus of a helix to the C-terminus of an adjacent helix would require a long and difficult-to-design linker. Nevertheless, inverting the orientation and handedness of one of the helices can solve this problem by reducing the distances between the N-terminus and C-termini of adjacent helices while maintaining a very similar side chain arrangement (Figure 15). With the help of simple computations, a molecular model of the heterochiral trimeric bundle can be built. The intermolecular hydrogen bond pattern of the heterochiral trimer model is very similar to that of homochiral trimer. Thus, two linkers with different flexibility were designed based on the heterochiral trimer model (Figure 15).

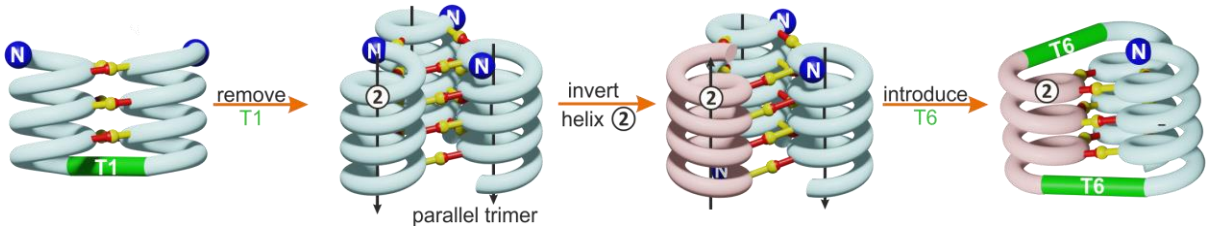


Figure 15. The helices will aggregate in to parallel trimer in the absence of turn unit T1. Inverting the handedness and orientation of one helix will produce very similar intermolecular hydrogen bond patterns to those of parallel trimer. Two relatively simple liners could be designed based on the heterochiral trimer model, which could potentially promote the formation of three-helix tertiary structure.

Our findings are summarized in a manuscript which is accepted by *Chemical Science*. Linker units with varying flexibilities were synthesized as Fmoc-protected amino acids, facilitating oligomer synthesis based on solid-phase synthesis. These linkers were validated with two series of model sequences, Q_3-turn-Q_3 and Q_8-turn-Q_8. The results from solution and solid-state studies indicated that

both linkers could promote the formation of the heterochiral helix-turn-helix structure as expected, although there were slight variations in the intramolecular hydrogen bond patterns. The final sequences were prepared using a fragment condensation strategy. Due to differences in the reactivity of the terminal amine groups on the linkers, the sequences with different linker units were split into different fragments. This approach simplified the purification process, yielding pure oligoamide sequences of 6.9 kDa with 10% yield. Solution studies on the final sequences suggested that flexible linkers may hinder the formation of such complex tertiary folds. In contrast, the findings supported that sequences with rigid linkers folded into the designed structure (Figure 15).

Contributions: The project was planned by I. Huc. Synthetic monomer precursors have been provided by D. Gill. Monomer synthesis have been performed by me, F. Menke and D. Gill. Foldamer synthesis was performed by me. C. Glas, L. Allmendinger and me carried out NMR measurements. Crystallographic studies and structure refinement were performed by J. Sigl. The measurement of CD spectra was performed by me. The LC-MS analysis was performed by C. Douat. Me, I. Huc and V. Mariout contributed to experiment design and interpretation. The research was supervised by I. Huc. The manuscript was written by me and I. Huc. All the authors proofread and improved the manuscript. This work was supported the China Scholarship Council (CSC, predoctoral fellowship to S. W.).

6.1. Publication (accepted)

Design of an abiotic unimolecular three-helix bundle

Author: Shuhe Wang, Johaness Sigl, Lars Allmendinger, Victor Mariout and Ivan Huc

Accepted: *Chemical Science*, **2025**. Advance Article. (doi.org/10.1039/d4sc07336c).

Chemical Science

ROYAL SOCIETY OF CHEMISTRY

EDGE ARTICLE

View Article Online
View Journal

Cite this: DOI: 10.1039/d4sc07336c

All publication charges for this article have been paid for by the Royal Society of Chemistry

Design of an abiotic unimolecular three-helix bundle†

Shuhe Wang,[a] Johannes Sigl,[a] Lars Allmendinger,[a] Victor Maurizot [b] and Ivan Huc *[a]

Starting from the solid state structure of C_3-symmetrical homochiral parallel trimolecular bundle of three aromatic helices held together by intermolecular hydrogen bonds, we have used simple rational principles and molecular modelling to design a similar heterochiral structure where one helix had an opposite orientation and handedness. A rigid and a flexible linker to connect these helices and transform the bundle into a unimolecular object were designed and synthesized. Model sequences with two helices and one linker were then prepared. Their conformations were investigated in solution by nuclear magnetic resonance and circular dichroism, in the solid state by X-ray crystallography, and by molecular dynamics simulations, overall supporting the initial design. A final 6.9 kDa unimolecular three-helix bundle was then prepared using a fragment condensation approach. Solution studies support the formation of the targetted tertiary fold in the case of the rigid linker, thereby validating the overall approach.

Received 29th October 2024
Accepted 13th November 2024

DOI: 10.1039/d4sc07336c

rsc.li/chemical-science

Introduction

Abiotic foldamers are defined as artificial folded architectures chemically remote from proteins and nucleic acids, the biopolymers from which they are inspired. Foldamers with aryl rings in their main chain constitute the most developed class of abiotic foldamers.[1–7] Interest for such compounds stems from the expectation that using distinct chemical backbones might give access to distinct shapes and functions, possibly beyond the reach of biopolymers. For example, many aromatic foldamers fold in organic solvents. The relative rigidity associated with the introduction of aryl rings in the main chain facilitates the prediction of the preferred conformations of such foldamers. Over the years, various types of aromatic monomers have been produced and sequences have been synthesized that can fold into helices[8–11] or sheets.[12,13] Helices possessing a sizable cavity can be used for endomolecular recognition,[14–21] while the decoration of aromatic helices with proteinogenic and water solubilizing side chains can be used to recognize large surface areas of proteins[22] and interfere with *e.g.* amyloid proteins or DNA binding proteins.[23–26] Such helices have also been shown to aggregate in solution to form multistranded helices[9,27–29] or stacked structures.[30,31]

In peptides, sophisticated functions are rarely associated with an isolated α-helix or β-strand. Instead, most protein functions emerge at the level of tertiary structures that are complex objects consisting of several helices or sheets. While the design of artificial proteins has been thriving,[32–35] tertiary structures based on abiotic backbones are still in their infancy. Our own efforts have consisted in assembling secondary motifs based on oligoamides of 8-amino-2-quinoline carboxylic acid bearing different solubilising side chains in position 4 (Q^D, Q^B, Q^M in Fig. 1a). Oligomers of these δ-amino acids fold into extremely stable 2.5 aromatic helices in all types of solvents (Fig. S1†).[36,37] Such helices are easily produced by automated solid phase synthesis,[38] and constitute convenient building blocks to be assembled into larger structures.[39]

In α-peptides, the bundling of several α-helices by multimolecular self-assembly or by folding of a single sequence containing multiple α-helical segments separated by loops is one of the best understood motifs. Such structures can be designed reliably,[40–42] and have inspired the development of β-peptidic[43–45] and urea-based[46–48] foldamer helix bundles as well as bundles of 3_{10} helices.[49] We have followed the same path and used helix–helix interactions to produce the first abiotic tertiary structures.[50] Monomer X, an analogue of Q, was developed for this purpose (Fig. 1a). The 4-hydroxy group of X protrudes from the aromatic helix and may hydrogen bond to the amide carbonyl group of another helix in chlorinated solvents. Similarly, P and Y constitute analogues of Q and X, respectively, in which the quinoline benzene ring has been trimmed to avoid possible steric clashes within helix bundles.

[a]Department Pharmazie, Ludwig-Maximilians-Universität München, Butenandtstraße 5–13, Munich D-81377, Germany. E-mail: ivan.huc@cup.lmu.de

[b]Université de Bordeaux, CNRS, Bordeaux Institut National Polytechnique, CBMN (UMR 5248), Institut Européen de Chimie Biologie, 2 Rue Escarpit, Pessac 33600, France

† Electronic supplementary information (ESI) available: Supplementary figures, detailed experimental protocols and characterization of new compounds. CCDC 2391393 and 2391429. For ESI and crystallographic data in CIF or other electronic format see DOI: https://doi.org/10.1039/d4sc07336c

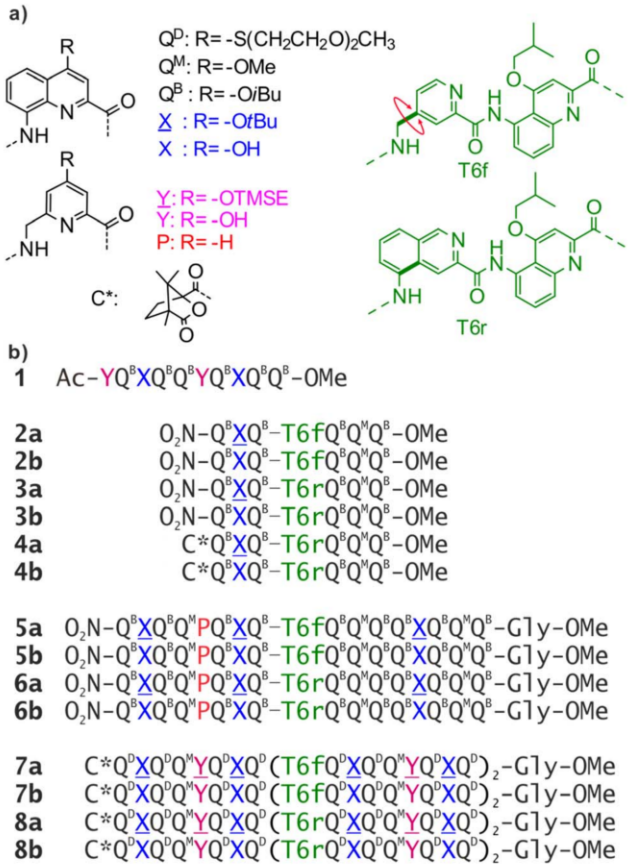


Fig. 1 (a) Structure of Q^D, Q^B, Q^M, X, Y, P, T6f, and T6r monomers as well as N-terminal chiral C* group. X and Y are the hydroxy protected precursors of X and Y, respectively. TMSE = 2-trimethylsilylethyl. (b) Oligoamide foldamer sequences. The Gly at the C terminus stands for glycine. In sequences ending with an 8-nitro group, this group replaces the terminal amine.

By carefully arranging the hydroxy groups of X and Y units at the helix surface and choosing a proper turn unit to covalently link two helices, a first helix-turn-helix motif was designed in which two helices are held with their axes parallel to each other by inter-helix hydrogen bonds (Fig. 2a and S2†).[50] In this first generation, the helices of the bundle were by design set to both have the same right-handed (*P*) or left-handed (*M*) handedness while in a second generation of helix-turn-helix motif, a *P* helix was linked to an *M* helix (Fig. S2†).[51] These tertiary structures are robust enough to undergo further assembly into quaternary motifs upon introducing additional hydroxy groups to form intermolecular hydrogen bonds.[52] Nevertheless, we discovered that the parallel arrangement of two helices is not the most stable. When the connecting turn element (T1 or T2 in Fig. S2†) is absent as, for example, in sequence **1** (Fig. 1b), single helices associate in parallel trimers (Fig. 2c) or in titled dimers (Fig. 2b), not in parallel dimers.[50] These different aggregates coexist in solution making the selective formation of one or the other challenging. The reason for their prevalence is that they both allow the release of some unfavorable helix torsion present in the helix-turn-helix structures without hampering hydrogen bonding.[53,54]

The parallel trimer was a serendipitous discovery and the first motif in which three abiotic helices were assembled in such a way that each helix interacts with the two others (Fig. 2i). Here we present how we successfully transformed this trimolecular object into a unimolecular helix-turn-helix-turn-helix tertiary structure analogous to a pattern common in proteins, for example in the B domain of protein A (Fig. S3†).[55,56] We first explored the behavior of shorter models **2–6** before implementing the same design principle in **7** and **8** which contain three helical segments. This achievement not only generated the most complex abiotic tertiary structure known to date. It also allowed us to streamline the design principles, computational approach, as well as synthetic and purification strategies. Routinely accessing such 6.9 kDa abiotic folds may become a realistic prospect in the near future.

Results and discussion

Redesign of the relative helix orientation

The design of a unimolecular three-helix bundle derived from the crystal structure of the parallel trimer of **1** required to meet two distinct challenges. The design and synthesis of proper linkers, *i.e.*, turn units, is addressed in the next section. This section deals with the first challenge which is to organize the helices in space so that they may conveniently be connected. The crystal structure of **1** is a C_3-symmetrical trimeric assembly (Fig. 2c and S4†). In this structure, the three identical helices named ①, ②, and ③ in Fig. 2 have the same handedness (all right-handed, *P*, or all left-handed, *M*) and are arranged in a head-to-head manner. Thus, the N terminus of helix ② is close to the N termini of helices ① and ③ but distant from their C termini. Covalently connecting two N termini and two C termini using diacid and diamine linkers, respectively, would be a way to transform $(\mathbf{1})_3$ into a unimolecular three-helix bundle. However, this approach was not considered because it would require a complicated synthetic approach. Another approach would be to covalently connect the N terminus of a helix to the C terminus of an adjacent helix, but this would require a long and difficult-to-design linker. Inverting the orientation of one helix would solve that problem by reducing the distance between N and C termini. However, the hydrogen bonding motif would then no longer promote helix–helix associations. Instead, donors would face other donors, and acceptors would face acceptors. The solution comes from the inversion of both the orientation and the handedness of one helix as this also preserves the position of the hydrogen bond donors and acceptors (Fig. 2e, f, i–l). This operation is well-known in the so-called α-helical retro-inverso peptides,[57–59] and has been used by us to design the *PM*/*MP* helix-turn-helix tertiary fold.[51] Associations between *P* and *M* peptidic helices have also been reported.[60–63] Thus, a molecular model of a trimeric bundle where helix ② has an orientation and a handedness both opposite to those of the two other helices was built and energy-minimized in Maestro (Fig. 3a and S4†).[64] The intermolecular hydrogen bond patterns extracted from the heterochiral trimer molecular model were very similar to those of the homochiral C_3-symmetrical trimer. Nevertheless, one should recall that **1**

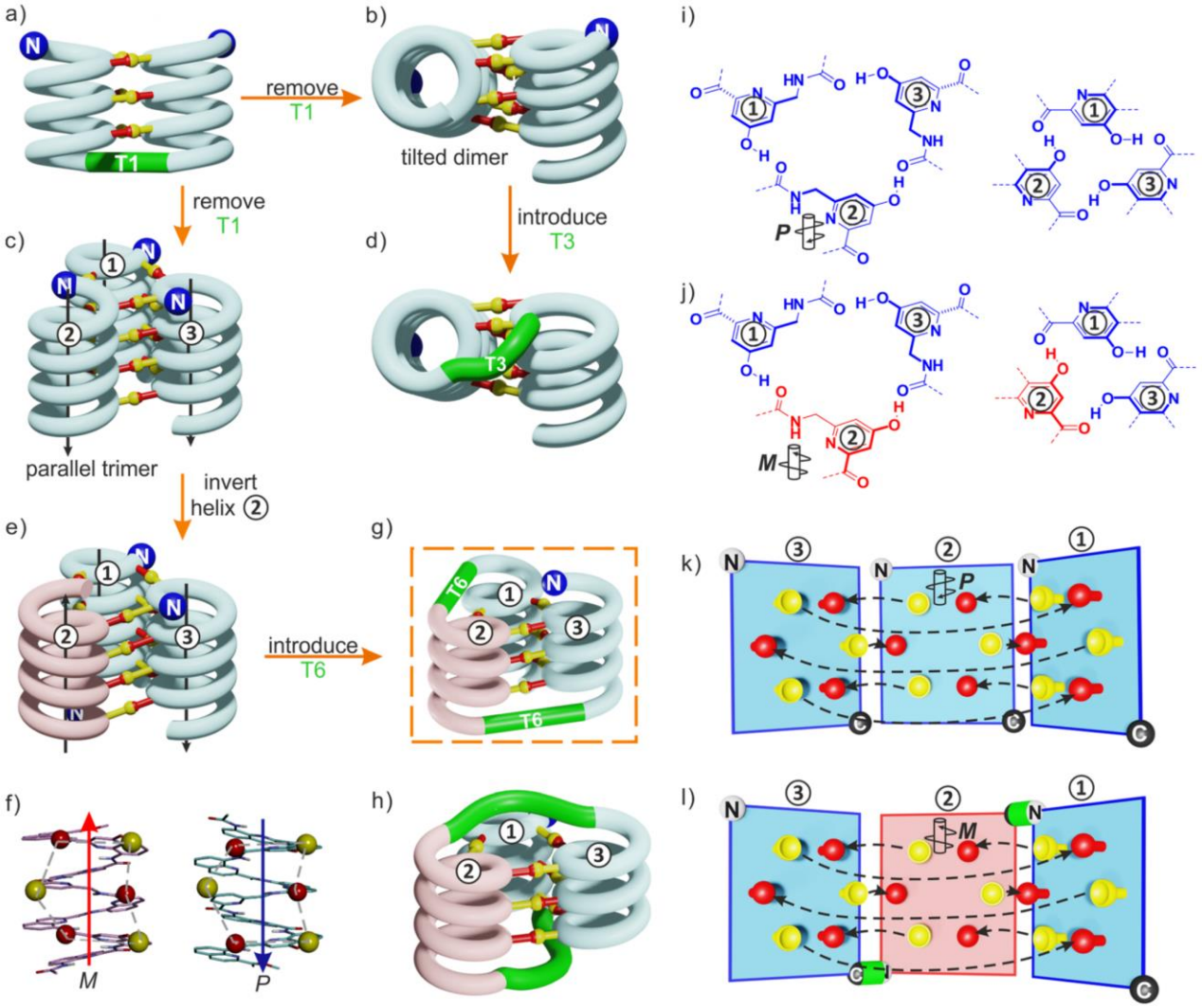


Fig. 2 (a) Schematic representation of a homochiral helix-T1-helix structure. After removing the T1 turn unit, single helices may assemble in tilted dimers (b) or in a parallel trimer (c). As shown in (c), the three helices are parallel in the trimer (black arrows) and the N termini (blue balls) are on the same side of the structure. (d) An oligoethylene glycol-based T3 turn unit was introduced to stabilize an intramolecular tilted arrangement.[65] (e) Model of a trimer after inverting the orientation and the handedness of one helix. *P* helices and *M* helices are colored in blue and red, respectively. The C terminus of helix ② is then close to the N termini of both helices ① and ③. (f) Illustration of the similar arrangement of hydrogen bond donors (yellow balls) and acceptors (red balls) after inverting the handedness and orientation of a helix. (g) Connection of the C terminus of helix ② to the N terminus of helix ③ and of the N terminus of helix ② to the C terminus of helix ①. (h) Connection of the C terminus of helix ② to the N terminus of helix ① and of the N terminus of helix ② to the C terminus of helix ③. (i) Hydrogen bonding motif between X units and between Y units in the parallel trimer shown in (c). (j) Hydrogen bonding motif between X units and between Y units in the trimer shown in (e). The color code is the same as in (e). (k) "Open-book" view of the hydrogen bond patterns of the homochiral trimer. (l) "Open-book" view of the hydrogen bond patterns of the representation shown in (g). In (k) and (l), the helix face is represented as a plane, with the N terminus shown as a white ball, the C terminus shown as a black ball and the linker T6 shown as a green stick. Planes with a blue and red frame correspond to *P* and *M* helices, respectively. Dashed-arrows indicate hydrogen bonds.

trimerizes into a homochiral *PPP*/*MMM* $(\mathbf{1})_3$ helix bundle. The *PPM*/*MMP* heterochiral trimer was not observed in solution or in the solid state and must therefore be inherently less stable.

Linker design and synthesis

Having reduced the distance between the C and N termini to be connected, the next step was the design of the linker. In an earlier study, we succeeded to design linkers for a tilted helix dimer motif (Fig. 2d),[65] and could experience how delicate this exercise can be. For the earlier parallel helix-turn-helix designs, the turn was selected first and the helix–helix interactions were designed next.[50,51] In the present case, this approach was not considered as we preferred to take advantage of an existing pattern of hydrogen bonds. First, sequence length was adjusted so as to minimize the distance between the termini to be connected – plus or minus one unit may result in significant distance variations. This generated a design with a total of nine hydrogen bonds, compared to twelve intermolecular hydrogen bonds in the structure of $(\mathbf{1})_3$. As depicted in Fig. 2g and h, two distinct helix–helix connection strategies may be envisaged

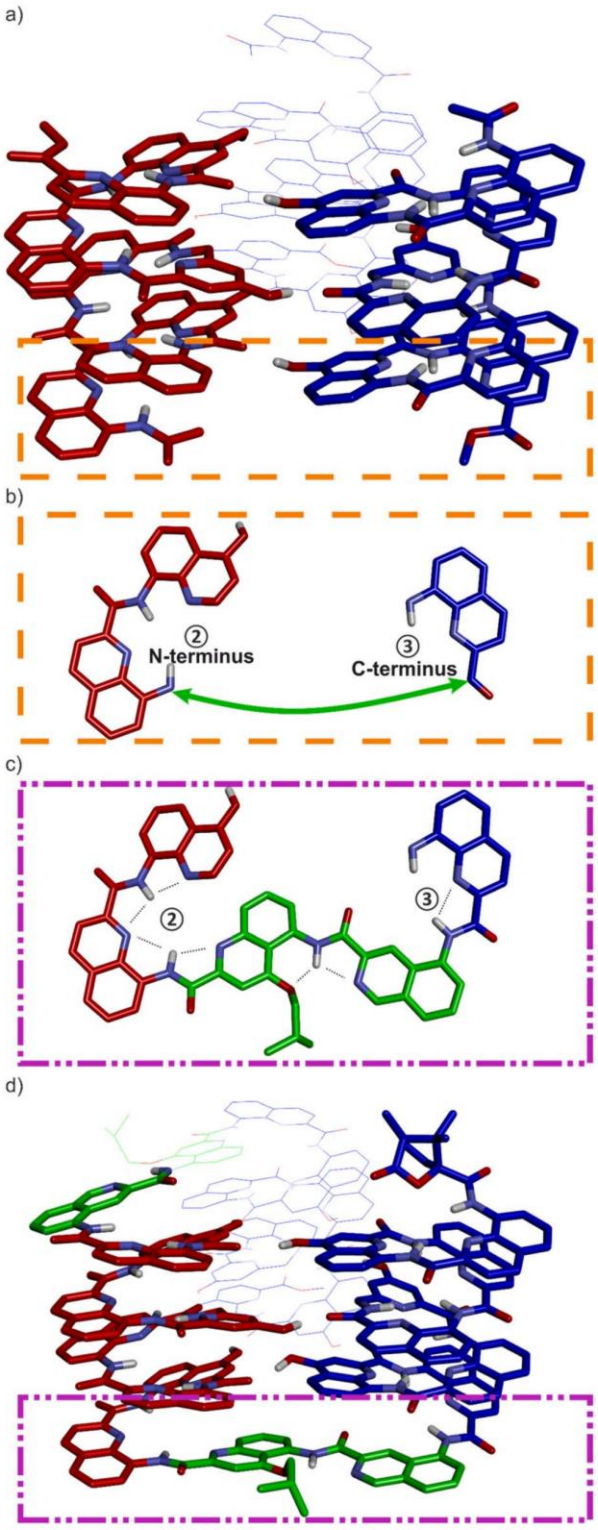


Fig. 3 (a) Energy-minimized model of a heterochiral (blue = *P*, red = *M*) trimer of Ac-QXQQYQXQ-OMe. Helices in the front are in stick representation. The helix in the back is in line representation. (b) Top view of the terminal units within the orange dashed line box of (a). A green arrow shows where the turn unit should be placed. (c) Top view of the terminal units within the purple dashed line box of (d). (d) Energy-minimized model of **8b**. Color coding is as in (a). The T6r turn is shown in green. In all structures, side chains on Q units have been omitted for clarity.

based on an amino acid turn unit: (N terminus)-①-turn-②-turn-③-(C terminus) or (N terminus)-③-turn-②-turn-①-(C terminus) (Fig. 2g). In other words, the N terminus of the central helix ② is close to the C termini of both ① and ③ but its spatial relationships to the two differ. In both strategies, the same linker would be used twice but the linkers for one strategy or the other differ. We discarded the arrangement shown in Fig. 2h because the suitable linkers between two helices that we found were also overlapping with the terminal cross section of the other helix, de facto preventing elongation of the C and N termini of the final object. The other arrangement shown in Fig. 2g does not have this impediment. Note that, at this stage, no consideration of absolute configuration is needed and the system can be treated as racemic. If ① is *P*-helical, ② is *M* and ③ is *P*. If ① is *M*-helical, ② is *P* and ③ is *M*.

The linker should preferably be relatively rigid to reduce the possibility of arrangements other than those intended. The linker should also not create strain that would risk destabilizing the structure. Keeping this in mind, models with different connections between the helix termini (green arrow in Fig. 3b) were built and energy-minimized. When strain was not apparent in the minimized structure under the form of *e.g.* helix distortion, a further test was to cut a bond in the linker and energy-minimize again. Any important conformational change at this stage would be interpreted as strain and the corresponding linker would be discarded or subjected to improvement. Finally, the synthetic accessibility was also considered to select the linker. A number of trials led to the design of two analogous turn units, T6f and T6r (Fig. 1a, 3c and d). T6f is a diamide of 4-aminomethyl pyridine 2-carboxylic acid and 5-amino quinoline 2-carboxylic acid, while T6r is a diamide of 5-amino isoquinoline 3-carboxylic acid and 5-amino quinoline 2-carboxylic acid. T6f is more flexible due to the presence of an additional rotatable bond at the CH_2 group. The linker units were produced with a free acid function and an Fmoc-protected amine, that is, ready for solid-phase synthesis (Fig. 4, see ESI† for details).

The synthesis of 5-nitroquinoline **10** is similar to that of the 8-nitro precursor of Q^B,[66] starting from 3-nitroaniline instead of 2-nitroaniline. Of note, the cyclization that produces **9** yields an equal amount of the 7-nitro isomer which can be used in other foldamer architectures.[67] The two isomers can be separated by selective crystallization. Similarly, **14** represents a regioisomer of Fmoc-P-OH, albeit obtained by a different route. Hydrogenation of the commercially available ethyl 4-cyano-2-pyridinecarboxylate in MeOH and in presence of Boc_2O produced **12** without observable transesterification. The synthesis of precursor **20** starts from commercially available tetrahydroisoquinoline 3-carboxylic acid. The nitration step is selective and the 5-nitro regioisomer was isolated in 80% yield by precipitation. The acylation of **11** by **14** using PyBOP activation or by **20** produced Fmoc-T6f-OH and Fmoc-T6r-OH, respectively. Overall, these syntheses worked quite well and needed not be repeated for this study. About half of the steps required chromatographic purification. We surmise that a full optimization would allow to improve some yields further, to increase the scales, and to avoid some chromatographic steps.

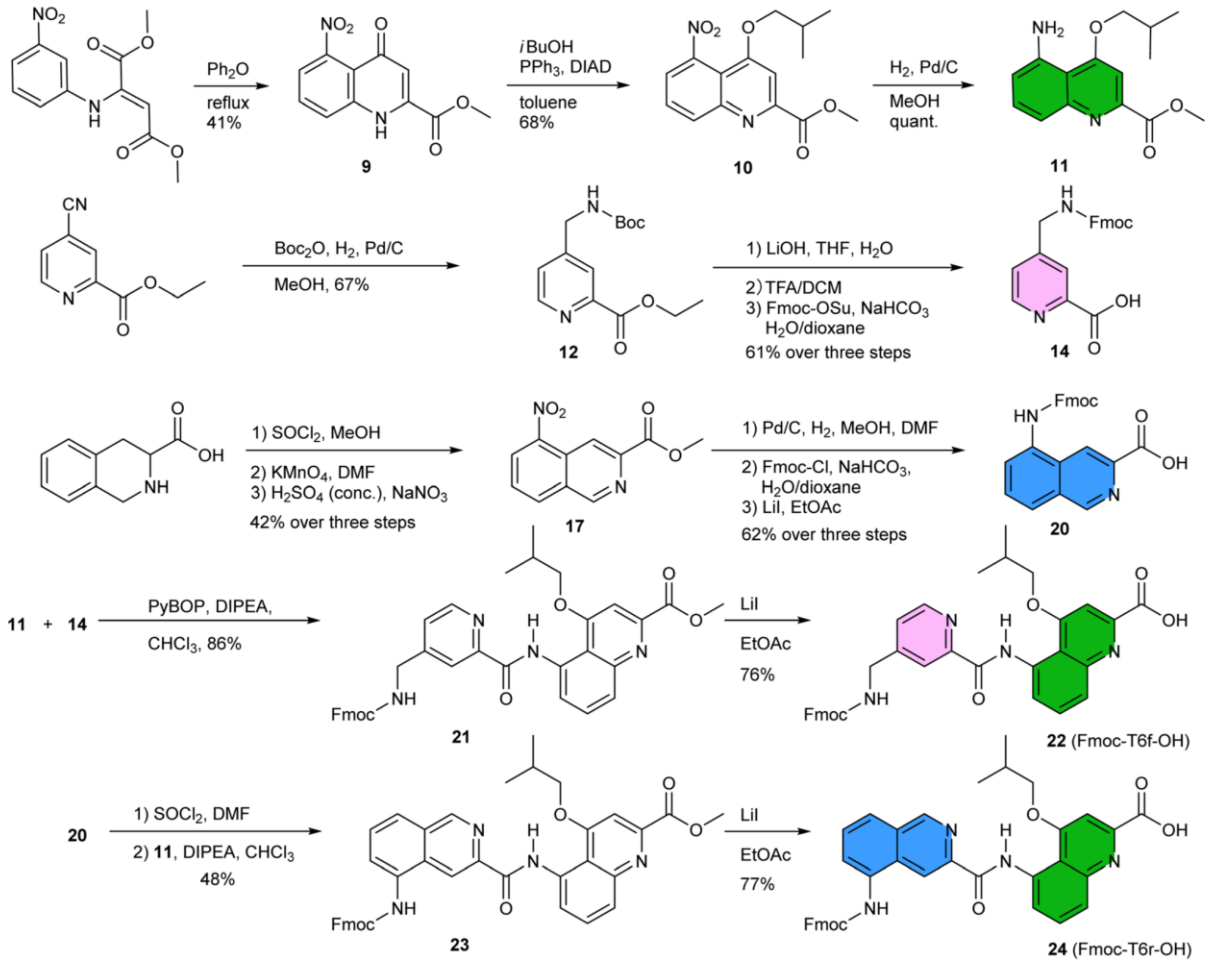


Fig. 4 Synthesis of Fmoc-T6f-OH and Fmoc-T6r-OH. Compound **13** (not shown) is the saponification product of **12**. Compounds **18** and **19** (not shown) are the products of the hydrogenation of **17** and of the subsequent Fmoc installation, respectively.

Validation of the linker design in helix-turn-helix models

Sequences **7b** and **8b** (Fig. 1b) were designed to fold into the desired ~6.9 kDa three-helix tertiary structure. They both have three identical QXQQYQXQ subdomains linked by T6f in **7b** or T6r in **8b** and a (1*S*)-camphanyl group at the N terminus that quantitatively biases the handedness of the N-terminal helix to *P*.[68] As for **4b**, chirality can be useful to investigate conformations by circular dichroism (CD). Before undertaking the synthesis and investigation of such large compounds, we sought a validation of the turn units T6r and T6f in smaller helix-turn-helix motifs.

We thus prepared three sequences equivalent in length to Q_3-turn-Q_3, **2b**, **3b** and **4b**. Sequences **2b** and **3b** are achiral, with T6f and T6r as turn units, respectively. Sequence **4b** is a chiral analogue of **3b**, bearing a (1*S*)-camphanyl group so that the N-terminal helix has *P* handedness.[68] All three sequences have a single X monomer in position two of the first helical domain. The hydroxy group of this X unit can potentially form an intramolecular hydrogen bond with a carbonyl group of the second Q_3 helical domain. In these short oligomers, the second residue after the turn is a Q^M monomer in replacement of X in **7b** and **8b**, where this residue should interact with the third helix.

Sequences **2a**, **3a** and **4a** (Fig. 1b), the protected precursors of **2b**, **3b** and **4b**, were synthesized on solid phase using previously reported methods.[38] The synthesis was performed on a super acid sensitive resin (SASRIN™) so that mildly acidic resin cleavage preserved *t*Bu-ether protection of the X monomer. The C-terminal carboxy group was then methylated and the sequences were purified in their protected form. After TFA-mediated cleavage of the *t*Bu group, they were purified in their deprotected form. The ^{1}H NMR spectra of **2a**, **3a** and **4a** in $CDCl_3$ showed one set of sharp signals (Fig. S5†), and so did the spectra of **2b**, **3b** and **4b** (Fig. 5a). However, the signal patterns of the CH_2 protons belonging to *i*Bu side chains indicate that *P*/*M* handedness interconversion of the helical segments of **2a** and **3a** is fast on the NMR time scale at 298 K whereas it is slow for **2b** and **3b**, indicating higher conformational stability of the latter (Fig. S6†). For **2b**, **3b** and **4b**, ^{1}H,^{15}N Heteronuclear Single Quantum Coherence (HSQC) spectra allowed for the indirect assignment of the OH resonance as an exchangeable proton that does not correlate with ^{15}N (Fig. 5a and S7–9†). In all cases, the OH proton signal is found above 10 ppm indicating that this proton is involved in a hydrogen bond.

The solid-state structures of **2b** and **3b** were elucidated using single crystal X-ray diffraction analysis. Their conformations are

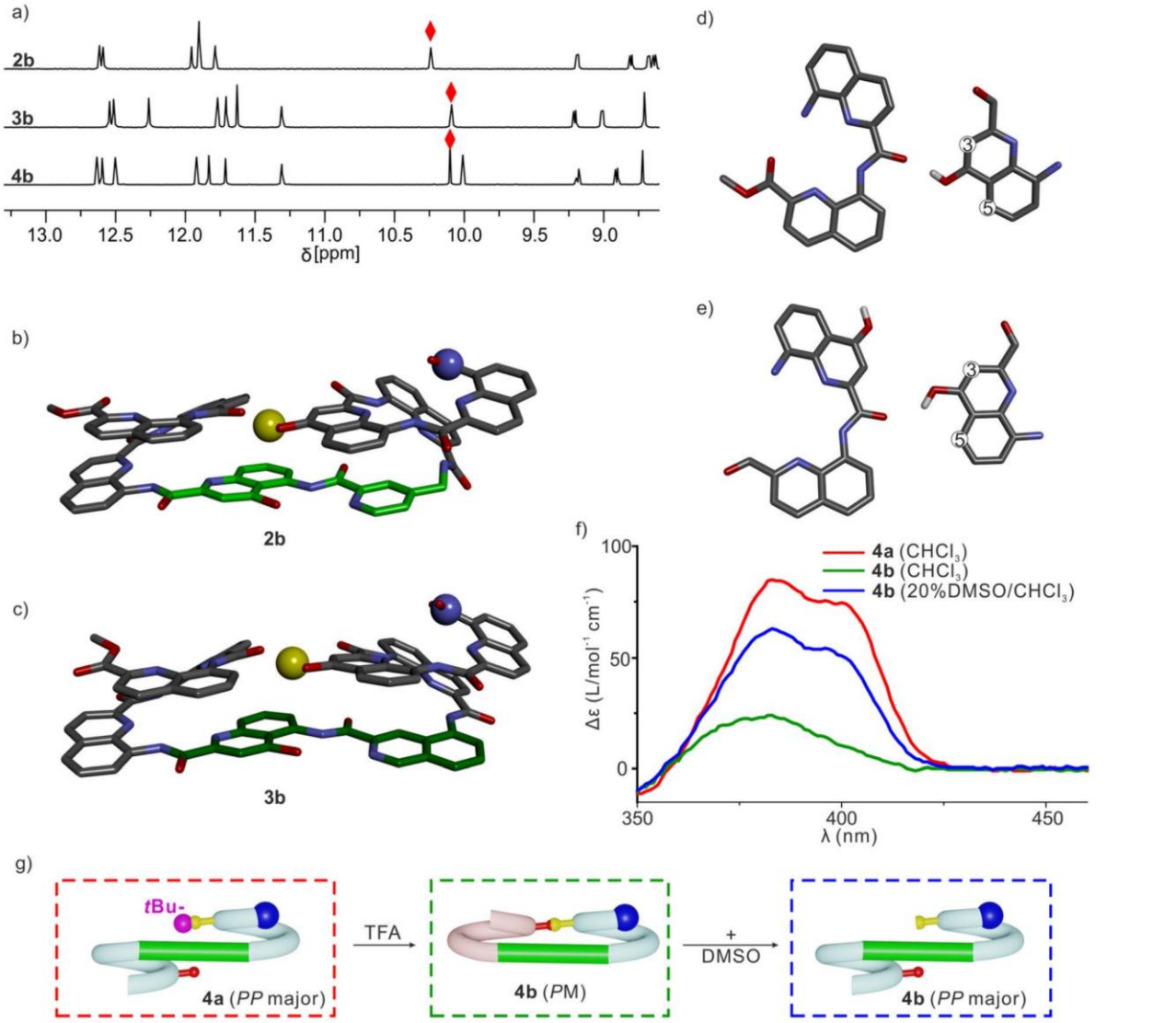


Fig. 5 (a) ^{1}H NMR spectra (500 MHz, $CDCl_3$) of **2b**, **3b** and **4b**. Hydrogen-bonded *OH* signals are marked with red diamonds. Solid state structures of **2b** (b) and **3b** (c). Both structures are shown in stick representation. Hydrogen atoms (except the OH proton), side chains and included solvent molecules are omitted for clarity. T6f and T6r units are colored in light green and dark green, respectively. The hydrogen bond donor is shown as a yellow ball. The N terminus is shown as a blue ball. Only the *PM* enantiomers are shown. (d) Intramolecular hydrogen bonding extracted from (c), which resemble that observed in previously described shifted dimers. (e) Intramolecular hydrogen bonding in molecular models of **7b** and **8b**. (f) CD spectra of **4a** and **4b** in chloroform or in 20% DMSO/chloroform. (g) Cartoon representation explaining the different CD intensities. The *t*Bu protecting group is shown as a purple ball. Its bulkiness favors the *PP* conformer of **4a**, whereas **4b** adopts a *PM* conformation in $CDCl_3$. DMSO disrupts the *PM* conformation of **4b** allowing for the prevalence of the *PP* conformer as in **4a**.

similar and both are very close to the initial molecular models. The helix-turn-helix motif is heterochiral, that is, *PM* or *MP*, meaning that T6r and T6f promote a reversal of helix handedness as desired. The hydroxy group of X is involved in the expected intramolecular hydrogen bond (Fig. 5b and c). A slight deviation from the initial design concerns the relative orientation of the hydrogen bond donor and acceptor (Fig. 5d and e). The observed orientation, with the hydroxy group pointing to the position 3 of the quinoline ring to which it belongs (Fig. 5d), is similar to that of a previously described arrangement that we called "shifted" interface (Fig. S10†).[69] In contrast, in the predicted arrangement (Fig. 5e) and in the structure of $(\mathbf{1})_3$ (Fig. 2i),[50] the hydroxy group points toward the position 5 of the quinoline ring. Molecular Dynamics (MD) simulations starting from the solid-state structure of **3b** showed that the two hydrogen bond patterns alternate within a simulation time as short as 1 ns, suggesting fast dynamics in these smaller model compounds (Fig. S11†).

The CD spectra of chiral sequences **4a** and **4b** in solution were consistent with the solid-state structure of **3b** (Fig. 5f). The relatively weak CD band of **4b** is in agreement with its two helical segments having opposite handedness and largely cancelling each other's CD contribution. In contrast, the bulky *t*Bu ether of **4a** may disfavor a *P*-turn-*M* arrangement with both helices on the same side (Fig. 5g). Having both Q_3 helices on opposite sides of the turn unit would then favor a *P*-turn-*P*

arrangement,[70,71] as reflected in its intense positive CD band. Adding DMSO to a $CHCl_3$ solution of **4b** disrupts the intramolecular hydrogen bond, resulting in an increase of the proportion of the *P*-turn-*P* conformation and thus in an enhancement of CD intensity (Fig. 5f and g). Altogether, these results hint at T6r and T6f performing well at promoting the desired conformations. We thus proceeded with the examination of longer model systems.

Sequences **5b** (with T6f) and **6b** (with T6r) are equivalent in length to Q_8-turn-Q_8 and represent achiral analogues of the two N-terminal helix-turn-helix segments of **7b** and **8b**, respectively. In **5b** and **6b**, the OH groups intended to promote interactions between helices ③ and ② of **7b** and **8b** have been preserved, and the OH groups intended to promote interactions with helix ① have been removed. Thus, **5b** and **6b** contain three X units in total, compared to three X units and two Y units for the first two helices of **7b** and **8b**. The protected precursors **5a** and **6a** (Fig. 1b) were synthesized on solid phase using the same protocol as for **2a**, **3a** and **4a**, but on a different resin. With a Gly-HMBA AM resin (4-(hydroxymethyl)benzoyl-aminomethyl polystyrene with a preloaded glycine), resin cleavage in presence of methanol under basic conditions directly yields the C-terminal methyl ester while preserving *t*Bu-ether protection on the X monomers. Sequences **5** and **6** thus have a glycine methyl ester at their C terminus.

Handedness interconversion in each of the longer helices of **5a** and **6a** is slow on the NMR timescale, unlike in the shorter helices of, *e.g.*, **2a**. This allows for the observation of diastereomeric conformers as distinct sets of signals at 298 K (Fig. S5†). Thus, the ^{1}H NMR spectrum of **5a** showed two species in a 1 : 0.8 ratio, which correspond to *PP*/*MM* and *PM*/*MP* diastereomers. In contrast, the spectrum of **6a** shows one set of signals assigned to the *PP*/*MM* conformers by extension of the preference of **4a** for the *PP* conformation. The different spectra of **5a** and **6a** reflect different behaviors of T6f and T6r. Due to its rigidity, the latter conveys handedness from one helix to the next, whereas the former disrupts helix handedness communication.[70,71] Nevertheless, after side chain deprotection, the ^{1}H NMR spectra of **5b** and **6b** both show a single set of signals in $CDCl_3$ (Fig. 6b), indicating that a single diastereomeric conformer prevails and thus that quantitative helix handedness communication through the turn units takes place, as observed above for **2b** and **3b**. O*H* resonances were assigned as for **2a** and **3a** (Fig. 6b, S12 and 13†) and appeared above 10 ppm, indicating their involvement in hydrogen bonds. Diffusion-Ordered Spectroscopy (DOSY) shows that **6b** has a larger diffusion coefficient than **6a** (Fig. S14†). Since **6a** is a monomer, this suggests that **6b** is monomeric as well and thus that all hydrogen bonds are intramolecular. The prevalent conformers of **5b** and **6b** in $CDCl_3$ were also prevalent in CD_2Cl_2 (Fig. S15 and 16†). This was examined because the relative stability of $(\mathbf{1})_3$ had been shown to differ in these two solvents.[50]

We then evaluated the stability of these conformers by ^{1}H NMR upon adding DMSO-d_6 to $CDCl_3$ solutions (Fig. 6c, S17 and 18†). DMSO competes for hydrogen bonding and is known to disrupt tertiary folds mediated by hydrogen bonds, resulting in variations of chemical shift values.[50,51,53,72] The conformational change was observed in both species, yet their transition

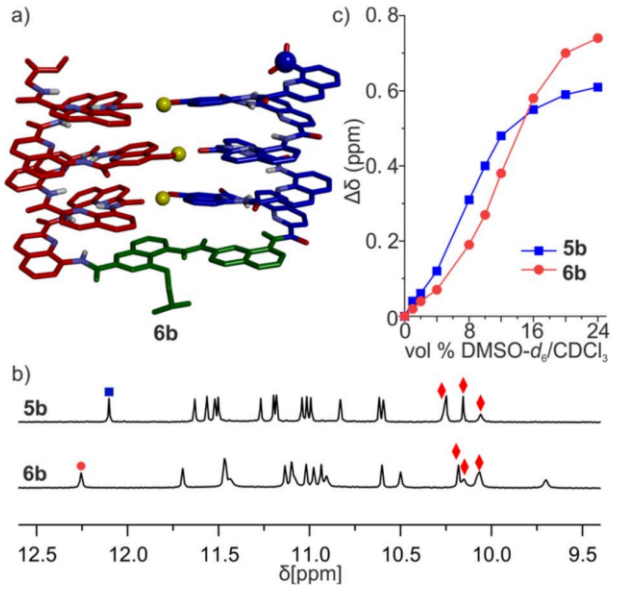


Fig. 6 (a) Energy-minimized model of the **6b**-T conformer of **6b**. The molecule is shown in stick representation. Non-polar hydrogen atoms and side chains (except that of the T6r turn) have been removed for clarity. Hydroxy protons are shown as yellow balls. The intramolecular hydrogen bond pattern is similar to what we observed in the parallel trimer. (b) Extracts of ^{1}H NMR spectra (500 MHz, $CDCl_3$) of **5b** and **6b**. The signals marked with red diamonds were assigned to hydroxy protons involved in hydrogen bonding. (c) Variations of the chemical shift of ^{1}H NMR signals marked with blue boxes and red dots in (b) upon the addition of DMSO-d_6 to $CDCl_3$ solution of **5b** and **6b**, respectively.

points differed slightly: around 8% (vol/vol) DMSO for **5b**, and 12% for **6b**. The sequence having the more rigid linker thus appears to be more stable.

The results above are consistent with **5b** and **6b** folding as a U-shaped helix-turn-helix structure stabilized by the turn unit geometry and three intramolecular inter-helix hydrogen bonds. Crystals of **5b** and **6b** were obtained but did not diffract at atomic resolution and their solid-state structure could not be solved. Their conformations were thus further investigated by MD simulations. Building on what was observed for **3b**, two different models of **6b** were built. In one model, the OH groups were made to point towards the position 5 of the quinoline ring they belong to, as in the models of **7b** and **8b** and in the structure of $(\mathbf{1})_3$ (Fig. 6a). This conformation was termed **6b**-T, T standing for "Three helix pattern". In the other model, the OH groups were made to point towards the position 3 of the quinoline ring they belong to, as in the crystal structures of **2b** and **3b**. This conformation was termed **6b**-S, S standing for "Shifted pattern". Both models could be energy-minimized while preserving the orientation of the OH groups and the three inter-helix hydrogen bonds. In **6b**-T, a view of the helix down its axis showed a 15-crown-5-like shape of the inner rim, suggesting that the helix held its preferred curvature of 2.5 units per turn (Fig. S19†). This was less obvious in **6b**-S, hinting at a possible distortion of the main chain. Furthermore, the two helices had their axes parallel in **6b**-T and at an angle in **6b**-S. During MD

simulation, the **6b**-S was found to convert to **6b**-T within 2 ns while the reverse process was not observed (Fig. S20†), suggesting a higher stability of **6b**-T.

In summary, the evidence gathered so far suggests that T6f and T6r indeed mediate the helix-turn-helix arrangement desired to form three-helix bundles. This entails preventing the formation of aggregates and of other hydrogen-bonded arrangements that may occur in the absence of linker.[50] The shorter helix-T6-helix models may also adopt a conformation with a different orientation of the hydrogen bonds without changing the donors and acceptors involved, but we expect these alternate conformations to be disfavored with the longer helices. Some differences were observed between T6f and T6r, with the more rigid T6r emerging as a better option to control conformation.

Synthesis and folding of an abiotic unimolecular three-helix bundle

The synthetic approach to prepare **7a** and **8a** was a matter of strategic choices. While relatively long, these sequences should in principle still be well within the range of automated solid phase foldamer synthesis and may be prepared in one go, like **5a** and **6a**.[38] However, our experience has shown that Reverse-Phase (RP) HPLC purification – the method that gave the best results so far – can be extremely challenging for long hydrophobic sequences. In turn, incomplete purification makes it difficult to assign multiple species seen in NMR spectra to impurities or to alternate conformers or aggregates. With their diethylene glycol side chains, the Q^D monomers help enhance hydrophilicity, and thus amenability to RP-HPLC, while not compromising solubility in organic solvent or crystal growth ability. Yet, the hydrophobic protecting groups of X and Y are unavoidable. We therefore opted for a fragment condensation approach to synthesize **7a** and **8a**. This entailed the preparation and careful RP-HPLC purification of relatively short fragments, and their subsequent coupling on solid phase. In the final mixture, deleted products where an entire fragment is missing should differ in size from the product sufficiently to allow for purification by recycling gel permeation chromatography (GPC) in chloroform.

Sequences **7a** and **8a** both have three identical octaamide segments linked by T6f or T6r turns. However, fragment condensation with aromatic helices works much better on aliphatic amines and this led to adopting different strategies for the two targets (Fig. 7). T6f is terminated by an aliphatic amine where segments can be coupled. Accordingly, we prepared the Fmoc-QXQQYQXQ-OH repeat motif (fragment A in Fig. 7a) on SASRIN™ resin and purified it in its fully protected form. Sequence **7a** was then assembled on a Gly-HMBA AM resin using alternatively fragment A activated with BOP and T6f activated as an acid chloride. The terminal Fmoc group was removed with 2% DBU/NMP and (1*S*)-camphanic chloride was used to finally cap the N terminus and induce a *P* handedness in the N terminal helix. After cleavage from the resin as a methyl ester using the same conditions as for **5a** and **6a**, **7a** was purified by GPC and isolated in high purity in an overall yield of 10%.

Because T6r is terminated by an aromatic amine, a different strategy was required to prepare **8a**. The sequence was divided

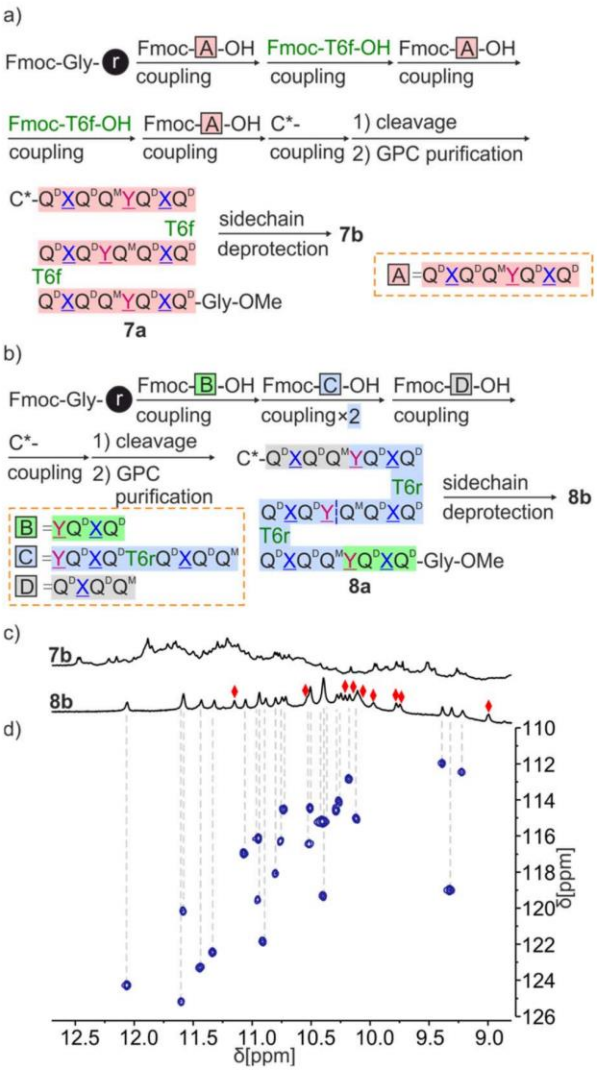


Fig. 7 Scheme representation of fragment condensation strategies for the synthesis of **7a**, **7b** (a) and **8a**, **8b** (b). (c) Extracts of ^{1}H NMR spectra (500 MHz, CD_2Cl_2) of **7b** and **8b**. (d) Extract of ^{1}H,^{15}N HSQC (500 MHz, CD_2Cl_2) of **8b**. The region above 8.8 ppm is shown, see Fig. S25† for a full spectrum. Only NH resonances correlate, red diamonds indicate the signals of hydrogen-bonded OH protons.

into four fragments of three different kinds (B, C and D in Fig. 7b) with the plan to condense the fragments at the benzylic amine of Y units. The synthesis proceeded as for **7a** with the prior synthesis of the fragments on SASRIN™ resin and their purification as a free carboxylic acid, followed by their assembly on Gly-HMBA AM resin. The final isolated yield after GPC purification was 9%.

Consistent with the behavior of **5a**, the ^{1}H NMR spectrum of **7a** showed four sets of signals with similar intensities (ratio of 1/0.9/0.85/0.8) that can be explained by the coexistence of *PPP*, *PMP*, *PPM* and *PMM* diastereomeric conformers (Fig. S5†). In contrast, and in agreement with the behavior of **6a**, the ^{1}H NMR spectrum of **8a** showed one set of signals assigned to the *PPP* conformer (Fig. S5†). After side chain deprotection, the ^{1}H NMR spectra of **7b** in both $CDCl_3$ and CD_2Cl_2 showed numerous

overlapping signals indicating the presence of several species (Fig. 7c and S21†). No significant changes were observed after allowing the solution to stand for 2 weeks (Fig. S22†). It is not clear whether the species correspond to aggregates, severely kinetically trapped states, alternate folds or a combination of these. Adding DMSO-d_6 to a solution of **7b** to disrupt the hydrogen bonds led to a simplification of the spectrum (Fig. S23†).

Sequence **8b** behaved differently from **7b**. Its spectrum in CD_2Cl_2 showed one set of signals indicating a well-defined unimolecular species or a symmetrical aggregate (Fig. 7c and S24†), while the spectrum in $CDCl_3$ was complex, reflecting the coexistence of several conformers or aggregates in solution. The $^1H,^{15}N$ HSQC spectrum in CD_2Cl_2 allowed us to assign all twenty-nine $^{15}\underline{N}H$ resonances and, indirectly, the nine hydrogen bonded O*H* resonances (Fig. 7d and S25†). The DOSY spectrum of an **8a** + **8b** mixture in CD_2Cl_2 confirms the smaller size of **8b** and thus its probable monomeric nature (Fig. S26†). Moreover, the ^1H NMR spectrum did not significantly change upon adding 4% DMSO-d_6 even at concentrations as low as 50 μmol (Fig. 8b and S27†). This proportion of DMSO is usually insufficient to disrupt tertiary folds but potentially destabilizing for some aggregates.[50,53,54] The effect of further increasing the proportion of DMSO was monitored by ^1H NMR and CD. The ^1H NMR spectra showed that a transition occurred between 12 and 16% of DMSO (Fig. S28†), that is, more than required for the disruption of the folded conformer of **6b**, hinting at a higher stability of **8b**. The CD spectra corroborated our interpretation of NMR data. The spectrum of protected precursor **8a** showed a strong positive band consistent with a *PPP* conformation (Fig. 8c). In comparison, the CD spectrum of **8b** showed a relatively weak band in CH_2Cl_2 which we assigned to the desired *PMP* conformer in which the contributions of the helical segments partly cancel each other. As described above for **4b**, the intensity of this band increased upon adding DMSO as the *PMP* conformer is at least in part disrupted, leading to an increase of the proportion of *PPP* conformer (Fig. 8c–e).

Taken altogether, these data point to **8b** being a unimolecular tertiary fold with two helix handedness reversals held together by nine intramolecular hydrogen bonds consistent with the initially proposed molecular model (Fig. 8a). Well-defined folding of **8b**, however, appears to be restricted to CH_2Cl_2. In $CDCl_3$, its ^1H NMR spectrum is broad and shows multiple species. These species already emerge upon adding $CDCl_3$ to a solution of **8b** in CD_2Cl_2 (Fig. S24†). Whether these species correspond to alternate folds or aggregates is unknown at this stage. One may point to other examples that we previously reported of changes in aggregation behavior of aromatic helices between $CDCl_3$ and CD_2Cl_2.[69] It is probably because of other possibilities these molecules have to fold and aggregate that the rigid T6r turn brought a critical advantage to control the conformation of **8b** in CD_2Cl_2. Indeed, the *PMP* three-helix bundle of **8b** does not form at all in the absence of linker – sequence **1** trimerizes in a *PPP*/*MMM* three-helix bundle. Apparently, the more flexible T6f turn of **7b** failed to bring a sufficient advantage. One important lesson to draw from this study is thus that the design of a well-defined tertiary fold amounts as much to stabilizing the desired fold as to preventing the formation of alternate structures.

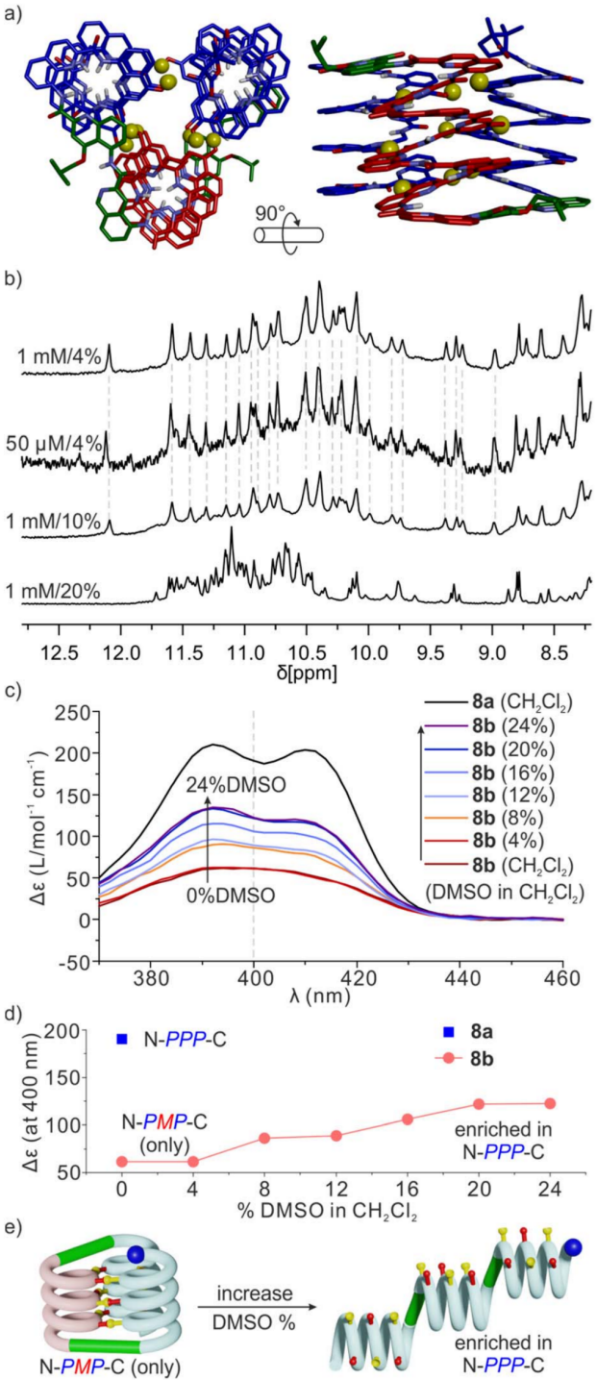


Fig. 8 (a) Molecular model of **8b**. The molecule is shown in stick representation. Non-polar hydrogen atoms and side chains (except that of the T6r turn) have been removed for clarity. Hydroxy protons are shown as yellow balls. (b) Excerpts of ^1H NMR spectra (500 MHz, CD_2Cl_2) of **8b** in DMSO-d_6/CD_2Cl_2 at different concentrations and different vol% of DMSO-d_6. (c) CD spectra of **8a** and **8b**, 50 μM in CH_2Cl_2 (DCM) with different proportions of DMSO. (d) $\Delta\varepsilon$ values at 400 nm extracted from (c) as a function of the vol% of DMSO in CH_2Cl_2. (e) Schematic representation of the process of hydrogen bond and tertiary fold disruption.

Conclusions

Starting from a well-characterized C_3-symmetrical *PPP*/*MMM* parallel trimolecular bundle of three aromatic helices held together by intermolecular hydrogen bonds, we have used simple rational principles and molecular modelling to design a similar *PMP* structure where the central helix had an opposite orientation and handedness. We next identified possible linkers to connect the termini of the helices to transform a trimolecular bundle into a unimolecular tertiary fold in which each helix hydrogen bonds to the two others. The synthesis and characterization first of reduced models and then of the full size 6.9 kDa abiotic fold validated the overall approach. The final unimolecular three-helix bundle was shown to be stable in CH_2Cl_2 even though other solvents or the use of flexible linkers impacted its behavior. These results represent an important milestone in the exploration of abiotic tertiary foldamers. The whole process that led to the unimolecular three-helix bundle highlights well what can be designed and how simple molecular modelling tools work quite effectively. The solvent dependence suggests that the design of **8b** can be further improved to enhance its robustness. Progress in this direction is being made and will be reported in due course.

Data availability

Crystallographic data for **2b** and **3b** has been deposited at the CCDC under accession numbers 2391393 and 2391429, respectively, and can be obtained from **https://www.ccdc.cam.ac.uk**.

Author contributions

SW performed the syntheses. JS collected diffraction data and solved the solid-state structures. SW and LA performed solution studies. SW and VM performed modelling studies. IH supervised the research. SW and IH wrote the manuscript. All authors proofread and improved the manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We thank L. Bodero and C. Douat for assitance with automated solid phase synthesis and D. Gill, F. Menke and M. Denis for providing some precursors. We acknowledge financial support from the China Scholarship Council (CSC, predoctoral fellowship to S. W.) and from the France-Germany International Research Project “Foldamers Structures and Functions” (IRP FoldSFun).

Notes and references

1 M. Ohkita, J. M. Lehn, G. Baum and D. Fenske, *Chem.–Eur. J.*, 1999, **5**, 3471–3481.
2 D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes and J. S. Moore, *Chem. Rev.*, 2001, **101**, 3893–4012.
3 I. Huc, *Eur. J. Org Chem.*, 2004, **2004**, 17–29.
4 I. Huc and L. Cuccia, in *Foldamers Structure, Properties, and Applications*, ed. S. Hecht and I. Huc, Wiley-VCH, Weinheim, 2007, chapter 1, pp. 3–33.
5 B. Gong, *Acc. Chem. Res.*, 2008, **41**, 1376–1386.
6 I. Saraogi and A. D. Hamilton, *Chem. Soc. Rev.*, 2009, **38**, 1726–1743.
7 D. W. Zhang, X. Zhao, J. L. Hou and Z. T. Li, *Chem. Rev.*, 2012, **112**, 5271–5316.
8 J. C. Nelson, J. G. Saven, J. S. Moore and P. G. Wolynes, *Science*, 1997, **277**, 1793–1796.
9 V. Berl, I. Huc, R. G. Khoury, M. J. Krische and J.-M. Lehn, *Nature*, 2000, **407**, 720–723.
10 J. Zhu, R. D. Parra, H. Zeng, E. Skrzypczak-Jankun, X. C. Zeng and B. Gong, *J. Am. Chem. Soc.*, 2000, **122**, 4219–4220.
11 Y. Hamuro, S. J. Geib and A. D. Hamilton, *J. Am. Chem. Soc.*, 1996, **118**, 7529–7541.
12 L. Sebaoun, V. Maurizot, T. Granier, B. Kauffmann and I. Huc, *J. Am. Chem. Soc.*, 2014, **136**, 2168–2174.
13 J. Atcher, P. Mateus, B. Kauffmann, F. Rosu, V. Maurizot and I. Huc, *Angew. Chem., Int. Ed.*, 2021, **60**, 2574–2577.
14 R. B. Prince, S. A. Barnes and J. S. Moore, *J. Am. Chem. Soc.*, 2000, **122**, 2758–2762.
15 J.-L. Hou, X.-B. Shao, G.-J. Chen, Y.-X. Zhou, X.-K. Jiang and Z.-T. Li, *J. Am. Chem. Soc.*, 2004, **126**, 12386–12394.
16 N. Chandramouli, Y. Ferrand, G. Lautrette, B. Kauffmann, C. D. Mackereth, M. Laguerre, D. Dubreuil and I. Huc, *Nat. Chem.*, 2015, **7**, 334–341.
17 Q. Gan, X. Wang, B. Kauffmann, F. Rosu, Y. Ferrand and I. Huc, *Nat. Nanotechnol.*, 2017, **12**, 447–452.
18 W. Wang, C. Zhang, S. Qi, X. Deng, B. Yang, J. Liu and Z. Dong, *J. Org. Chem.*, 2018, **83**, 1898–1902.
19 K. M. Kim, G. Song, S. Lee, H. G. Jeon, W. Chae and K. S. Jeong, *Angew. Chem., Int. Ed.*, 2020, **59**, 22475–22479.
20 Y. Zhong, B. Kauffmann, W. Xu, Z.-L. Lu, Y. Ferrand, I. Huc, X. C. Zeng, R. Liu and B. Gong, *Org. Lett.*, 2020, **22**, 6938–6942.
21 V. Koehler, A. Roy, I. Huc and Y. Ferrand, *Acc. Chem. Res.*, 2022, **55**, 1074–1085.
22 P. S. Reddy, B. Langlois d'Estaintot, T. Granier, C. D. Mackereth, L. Fischer and I. Huc, *Chem.–Eur. J.*, 2019, **25**, 11042–11047.
23 S. Kumar, A. Henning-Knechtel, I. Chehade, M. Magzoub and A. D. Hamilton, *J. Am. Chem. Soc.*, 2017, **139**, 17098–17108.
24 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.
25 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.
26 J. Ahmed, T. C. Fitch, C. M. Donnelly, J. A. Joseph, T. D. Ball, M. M. Bassil, A. Son, C. Zhang, A. Ledreux, S. Horowitz, Y. Qin, D. Paredes and S. Kumar, *Nat. Commun.*, 2022, **13**, 2273.

27 Y. Zhong, T. A. Sobiech, B. Kauffmann, B. Song, X. Li, Y. Ferrand, I. Huc and B. Gong, *Chem. Sci.*, 2023, **14**, 4759–4768.
28 Q. Gan, C. Bao, B. Kauffmann, A. Grélard, J. Xiang, S. Liu, I. Huc and H. Jiang, *Angew. Chem., Int. Ed.*, 2008, **47**, 1715–1718.
29 K. Zhang, C. Ma, N. Li, C. Lu, D. Li, S. Fu and Q. Gan, *Chem. Commun.*, 2019, **55**, 10968–10971.
30 D. Bindl, P. K. Mandal, L. Allmendinger and I. Huc, *Angew. Chem., Int. Ed.*, 2022, **61**, e202116509.
31 A. Galan, K. Lulic, J. Wang, B. Wicher, I. Huc, J. Duhamel and V. Maurizot, *Chem. Commun.*, 2023, **59**, 5253–5256.
32 P.-S. Huang, S. E. Boyken and D. Baker, *Nature*, 2016, **537**, 320–327.
33 A. T. Hilditch, A. Romanyuk, S. J. Cross, R. Obexer, J. J. McManus and D. N. Woolfson, *Nat. Chem.*, 2024, **16**, 89–97.
34 A. Pillai, A. Idris, A. Philomin, C. Weidle, R. Skotheim, P. J. Leung, A. Broerman, C. Demakis, A. J. Borst, F. Praetorius and D. Baker, *Nature*, 2024, **632**, 911–920.
35 L. Jiang, E. A. Althoff, F. R. Clemente, L. Doyle, D. Rothlisberger, A. Zanghellini, J. L. Gallaher, J. L. Betker, F. Tanaka and C. F. Barbas III, *Science*, 2008, **319**, 1387–1391.
36 H. Jiang, J.-M. Léger and I. Huc, *J. Am. Chem. Soc.*, 2003, **125**, 3448–3449.
37 T. Qi, V. Maurizot, H. Noguchi, T. Charoenraks, B. Kauffmann, M. Takafuji, H. Ihara and I. Huc, *Chem. Commun.*, 2012, **48**, 6337–6339.
38 V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat and I. Huc, *Chem.–Eur. J.*, 2023, **29**, e202300898.
39 N. Delsuc, J.-M. Léger, S. Massip and I. Huc, *Angew. Chem., Int. Ed.*, 2007, **46**, 214.
40 E. A. Naudin, K. I. Albanese, A. J. Smith, B. Mylemans, E. G. Baker, O. D. Weiner, D. M. Andrews, N. Tigue, N. J. Savery and D. N. Woolfson, *Chem. Sci.*, 2022, **13**, 11330–11340.
41 R. B. Hill, D. P. Raleigh, A. Lombardi and W. F. DeGrado, *Acc. Chem. Res.*, 2000, **33**, 745–754.
42 K. I. Albanese, R. Petrenas, F. Pirro, E. A. Naudin, U. Borucu, W. M. Dawson, D. A. Scott, G. J. Leggett, O. D. Weiner, T. A. Oliver and D. N. Woolfson, *Nat. Chem. Biol.*, 2024, **20**, 991–999.
43 D. S. Daniels, E. J. Petersson, J. X. Qiu and A. Schepartz, *J. Am. Chem. Soc.*, 2007, **129**, 1532–1533.
44 P. S. Wang and A. Schepartz, *Chem. Commun.*, 2016, **52**, 7420–7432.
45 J. P. Miller, M. S. Melicher and A. Schepartz, *J. Am. Chem. Soc.*, 2014, **136**, 14726–14729.
46 G. W. Collie, K. Pulka-Ziach, C. M. Lombardo, J. Fremaux, F. Rosu, M. Decossas, L. Mauran, O. Lambert, V. Gabelica, C. D. Mackereth and G. Guichard, *Nat. Chem.*, 2015, **7**, 871–878.
47 L. Cussol, L. Maura-Ambrosino, J. Buratto, A. Y. Belorusova, M. Neuville, J. Osz, S. Fribourg, J. Fremaux, C. Dolain, S. R. Goudreau, N. Rochel and G. Guichard, *Angew. Chem., Int. Ed.*, 2021, **60**, 2296–2303.
48 Y. Toledo-González, J.-M. Sotiropoulos, D. Bécart, G. Guichard and P. Carbonnière, *J. Org. Chem.*, 2022, **87**, 10726–10735.
49 P. Kumar, N. G. Paterson, J. Clayden and D. N. Woolfson, *Nature*, 2022, **607**, 387–392.
50 S. De, B. Chi, T. Granier, T. Qi, V. Maurizot and I. Huc, *Nat. Chem.*, 2018, **10**, 51–57.
51 D. Mazzier, S. De, B. Wicher, V. Maurizot and I. Huc, *Angew. Chem., Int. Ed.*, 2020, **59**, 1606–1610.
52 S. Wang, L. Allmendinger and I. Huc, *Angew. Chem., Int. Ed.*, 2024, e202413252.
53 F. S. Menke, D. Mazzier, B. Wicher, L. Allmendinger, B. Kauffmann, V. Maurizot and I. Huc, *Org. Biomol. Chem.*, 2023, **21**, 1275–1283.
54 S. Wang, B. Wicher, C. Douat, V. Maurizot and I. Huc, *Angew. Chem., Int. Ed.*, 2024, e202405091.
55 M. Tashiro, R. Tejero, D. E. Zimmerman, B. Celda, B. Nilsson and G. T. Montelione, *J. Mol. Biol.*, 1997, **272**, 573–590.
56 D. Zheng, J. M. Aramini and G. T. Montelione, *Protein Sci.*, 2004, **13**, 549–554.
57 N. Doti, M. Mardirossian, A. Sandomenico, M. Ruvo and A. Caporale, *Int. J. Mol. Sci.*, 2021, **22**, 8677.
58 J. Rai, *Chem. Biol. Drug Des.*, 2019, **93**, 724–736.
59 M. Chorev and M. Goodman, *Acc. Chem. Res.*, 1993, **26**, 266–273.
60 W. S. Horne, J. L. Price, J. L. Keck and S. H. Gellman, *J. Am. Chem. Soc.*, 2007, **129**, 4178–4180.
61 J. L. Price, E. B. Hadley, J. D. Steinkruger and S. H. Gellman, *Angew. Chem., Int. Ed.*, 2010, **49**, 368.
62 D. F. Kreitler, Z. Yao, J. D. Steinkruger, D. E. Mortenson, L. Huang, R. Mittal, B. R. Travis, K. T. Forest and S. H. Gellman, *J. Am. Chem. Soc.*, 2019, **141**, 1583–1592.
63 D. E. Mortenson, J. D. Steinkruger, D. F. Kreitler, D. V. Perroni, G. P. Sorenson, L. Huang, R. Mittal, H. G. Yun, B. R. Travis, M. K. Mahanthappa, K. T. Forest and S. H. Gellman, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 13144–13149.
64 *Maestro, Schrödinger*, LLC, New York, NY, 2021.
65 F. S. Menke, B. Wicher, L. Allmendinger, V. Maurizot and I. Huc, *Chem. Sci.*, 2023, **14**, 3742–3751.
66 T. Qi, T. Deschrijver and I. Huc, *Nat. Protoc.*, 2013, **8**, 693–708.
67 B. Teng, J. Atcher, L. Allmendinger, C. Douat, Y. Ferrand and I. Huc, *Org. Biomol. Chem.*, 2023, **21**, 3525–3530.
68 A. M. Kendhale, L. Poniman, Z. Dong, K. Laxmi-Reddy, B. Kauffmann, Y. Ferrand and I. Huc, *J. Org. Chem.*, 2011, **76**, 195–200.
69 F. S. Menke, B. Wicher, V. Maurizot and I. Huc, *Angew. Chem., Int. Ed.*, 2023, **62**, e202217325.
70 V. Maurizot, C. Dolain, Y. Leydet, J.-M. Léger, P. Guionneau and I. Huc, *J. Am. Chem. Soc.*, 2004, **126**, 10049–10052.
71 C. Dolain, J. M. Leger, N. Delsuc, H. Gornitzka and I. Huc, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 16146–16151.
72 D. Mazzier, S. De, B. Wicher, V. Maurizot and I. Huc, *Chem. Sci.*, 2019, **10**, 6984–6991.

6.2. Supplementary Information

for:

Design of an abiotic unimolecular three-helix bundle

Shuhe Wang, Johannes Sigl, Lars Allmendinger, Victor Maurizot and Ivan Huc

Contents

1 List of Abbreviations

CD	circular dichroism
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DCM	dichloromethane
DIPEA	*N*,*N*-diisopropylethylamine
DMF	*N*,*N*-dimethylformamide
DMSO	dimethyl sulfoxide
DOSY	diffusion ordered spectroscopy
HR-ESI	high resolution electrospray ionization
eq.	equivalent
Fmoc	fluorenylmethoxycarbonyl
HBTU	2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HSQC	heteronuclear single quantum correlation
Me	methyl
MeOH	methanol
min	minutes
MS	mass spectrometry
MW	microwave
NMP	*N*-methyl-2-pyrrolidone
NMR	nuclear magnetic resonance
r. t.	room temperature
SPS	solid phase synthesis
***t*Bu**	*tert*-butyl
TFA	trifluoroacetic acid
THF	tetrahydrofuran
UV/Vis	ultraviolet–visible

2 Supplementary figures

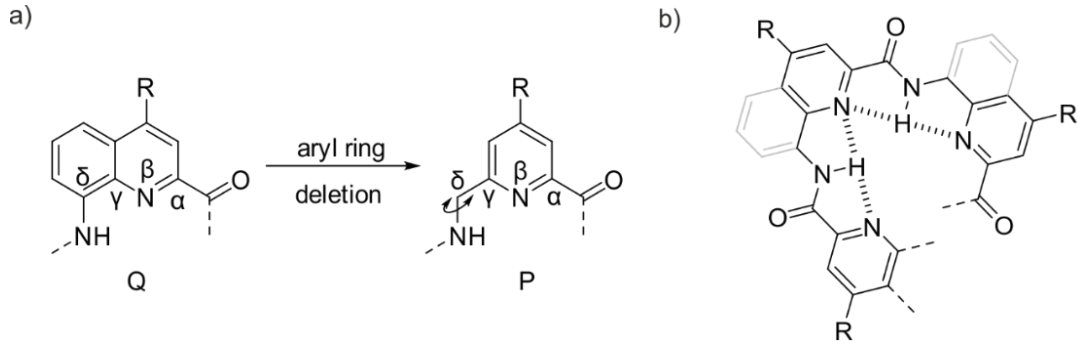


Figure S3. Chemical structure of Q and P units and the folding principle of their oligomers. a) Chemical structure of Q (or X with R = OH) and P (or Y with R = OH). b) Intramolecular H-bonding and helical folding principle of P/Q oligomers. Note that the amide carbonyl groups diverge from the folded structures and thus provide hydrogen bond acceptors.

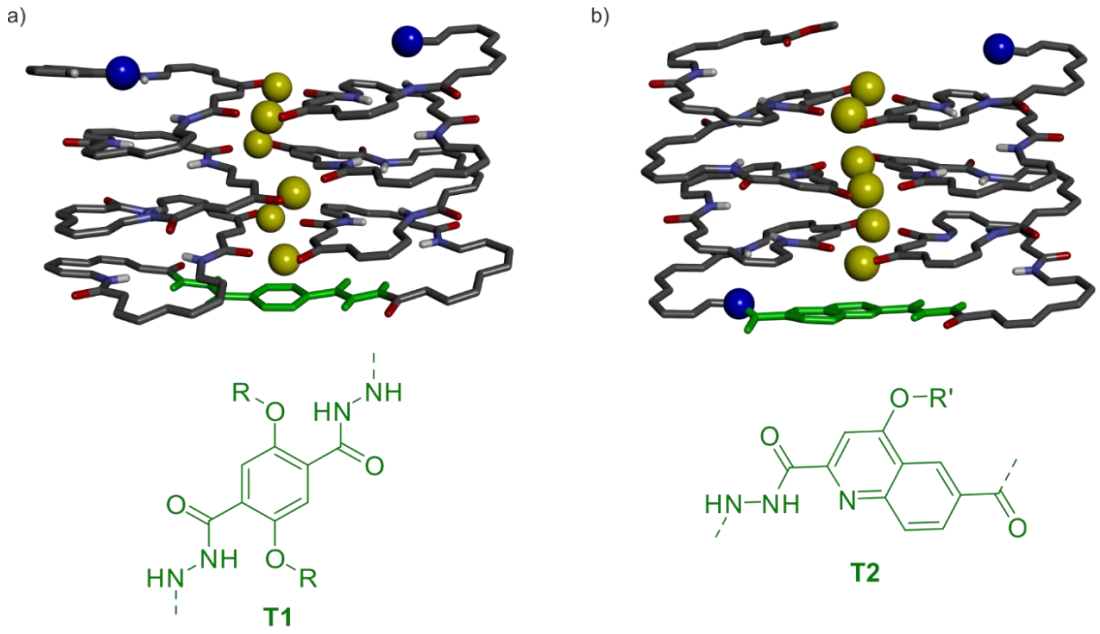


Figure S2. a) The side view of a helix-T1-helix tertiary structure and the chemical structure of T1.[1] **b) The side view of a helix-T2-helix tertiary structure and the chemical structure of T2.**[2] The N-terminus of each helix is highlighted with a blue ball. The hydroxy proton hydrogen bond donors are shown as yellow balls. Turn units are colored in green. Only the outer rim of the helix is shown. The side chains on the Q unit and turn unit are omitted for clarity.

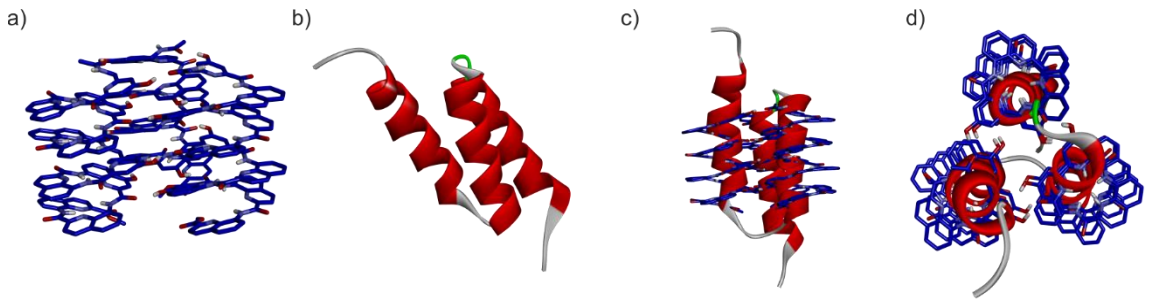


Figure S3. Comparison of an abiotic three helix bundle and the Z domain derived from protein A. a) The structure of $(\mathbf{1})_3$[1]. b) A modified subdomain of protein A (b, PDB:2B89). The side view (c) and top view (d) of the overlay of the structures shown in (a) and (b).

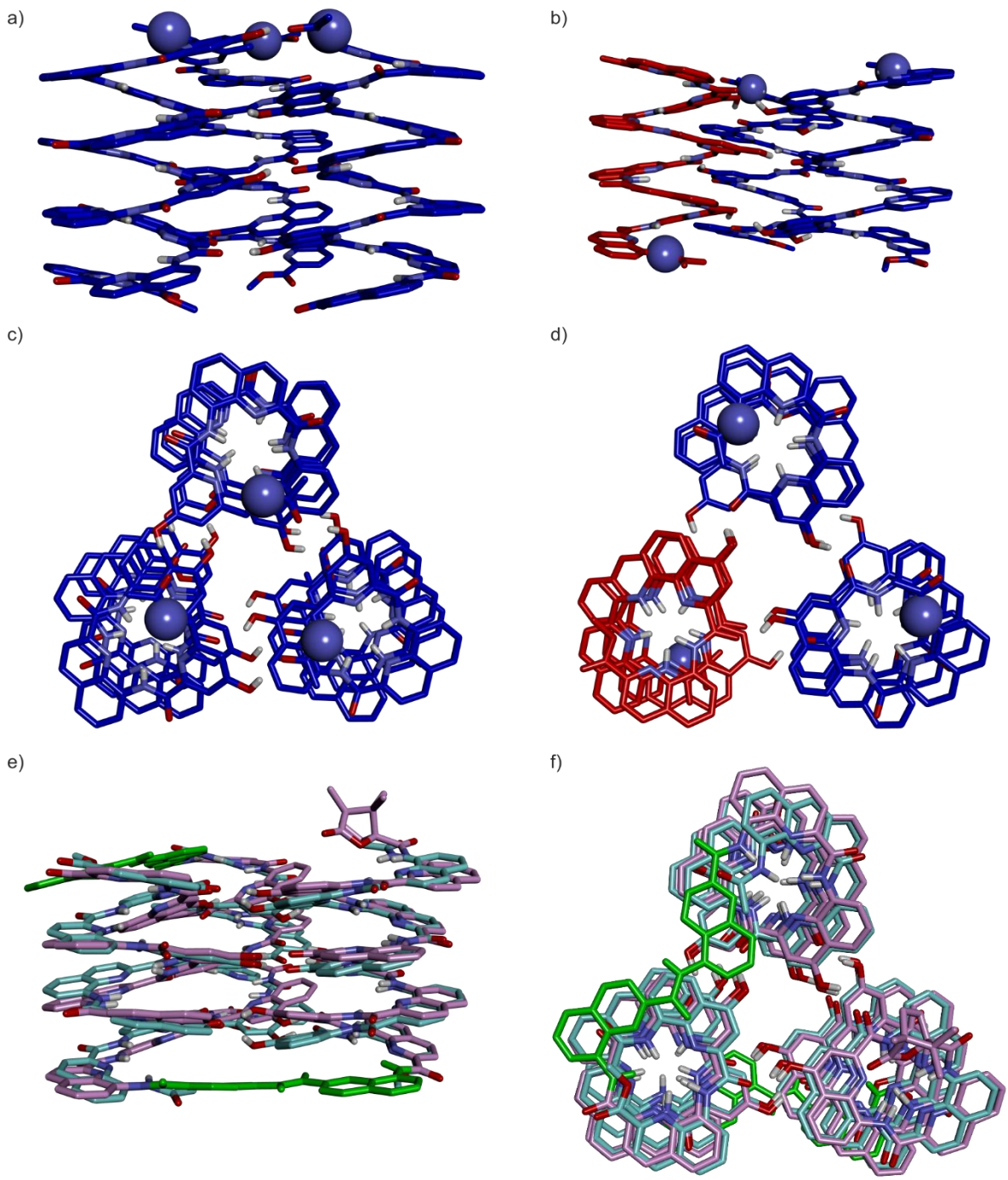


Figure S4. Design of the unimolecular three helix bundle and of the turn units. The side view (a) and top view (c) of the crystal structure of (**1**)$_3$.[1] The side view (b) and top view (d) of an energy-minimised heterochiral trimer molecular model of sequence (Ac-QXQQYQXQ-OMe)$_3$. The *P* helices are colored in blue and *M* helices are colored in red. The N-termini are shown as blue balls. Side chains are omitted for clarity. Side view (e) and top view (f) of the overlay of the structure in (b) and of the energy-minimised model of sequence **8b**. The (Ac-QXQQYQXQ-OMe)$_3$ is colored in cyan and sequence **8b** is colored in pink (helical segments) and green (T6r turn unit).

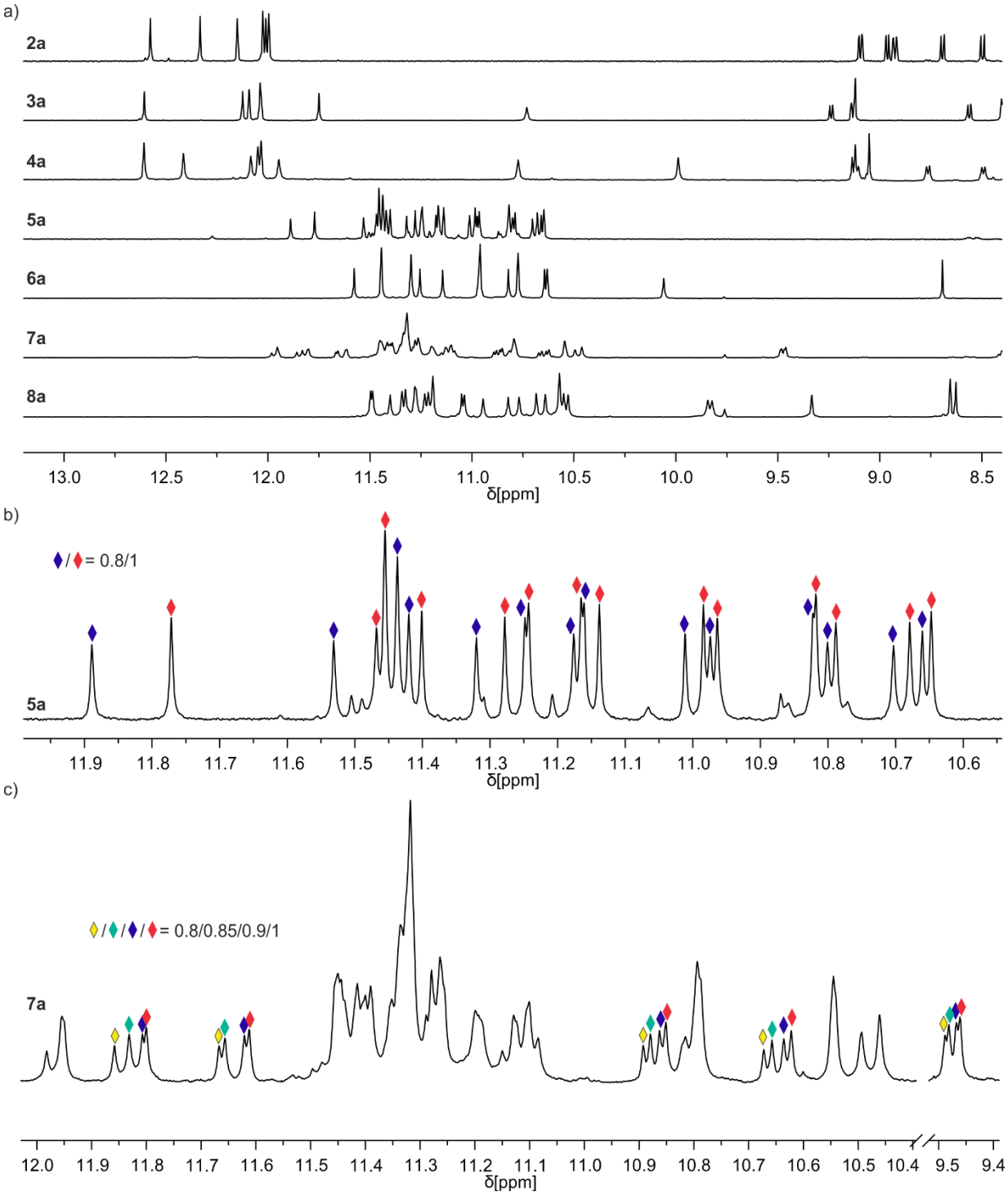


Figure S5. ^{1}H NMR spectra of side chain protected sequences. a) Extracts of the ^{1}H NMR spectra (500 MHz, $CDCl_3$) of **2a**, **3a**, **4a**, **5a**, **6a**, **7a** and **8a**. The sequences **2a**, **3a** and **4a** show one set of signals because the equilibrium between their diastereomeric conformers is fast on the NMR timescale, not necessarily because only one species is present. Sequence **5a** shows two sets of signals from two major species which were assigned to *PP*/*MM* and *PM*/*MP* conformers. Sequence **6a** shows one major species assigned to the *PP*/*MM* conformer. Sequence **7a** shows four sets of signals assigned to *PPP*, *PMP*, *PPM* and *PMM* conformers. Sequence **8a** show

one major species assigned to the *PPP* conformer. The enlarged spectra of **5a** (b) and **7a** (c) show the presence of different isomers and their proportions.

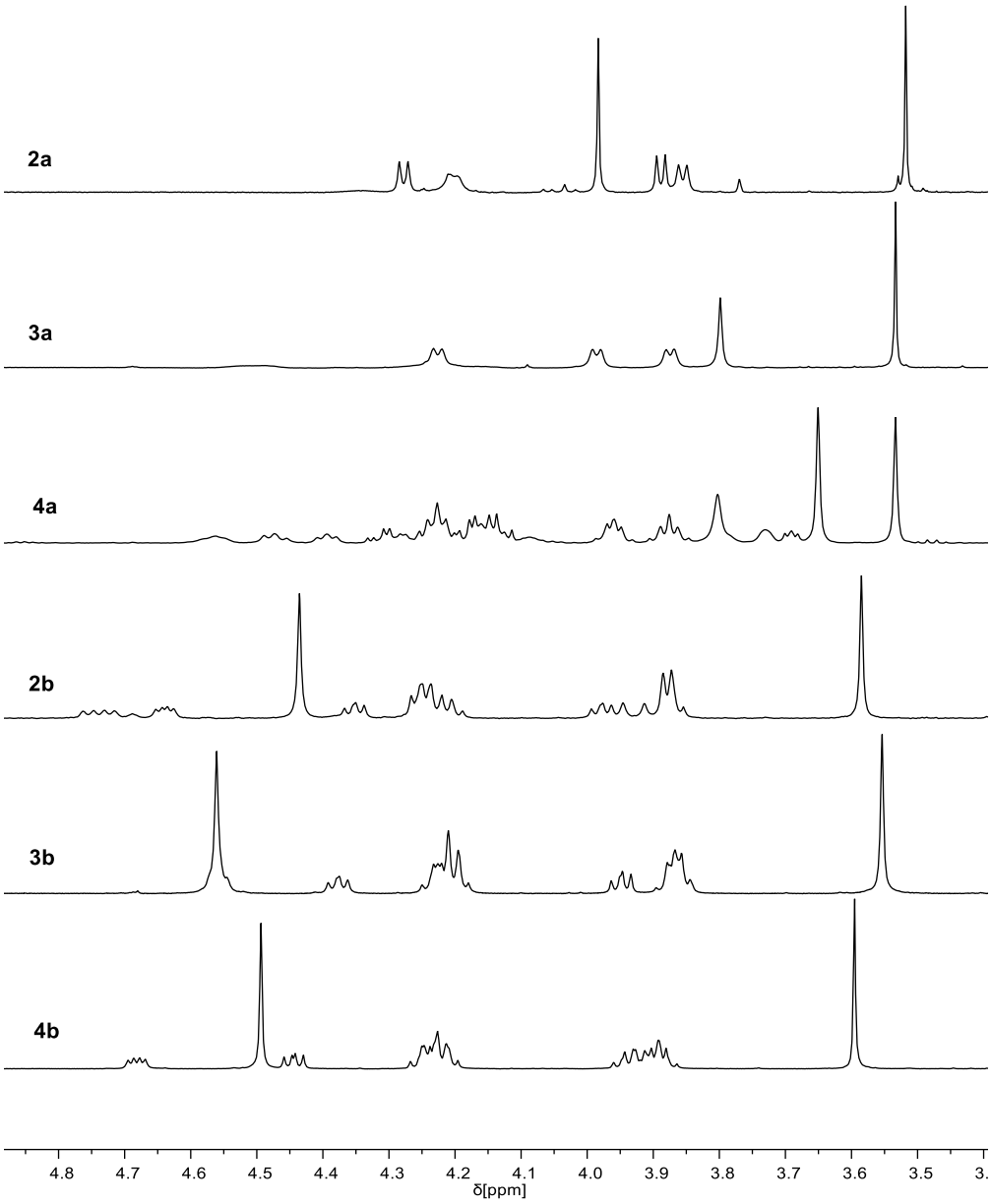


Figure S6. Excerpts of the ^{1}H NMR spectra (500 MHz, $CDCl_3$) of 2a, **3a**, **4a**, **2b**, **3b and 4b.** The singlet near 3.6 ppm is the methyl ester resonance. The other singlet belongs to the Q^M methoxy side chain. The deshielding of the Q^M methyl group in **2b**, **3b**, and **4b** is consistent with its proximity to the carbonyl group of the X unit on the other helix, as seen in the solid state structures of **2b** and **3b**. All other resonances belong to the diastereotopic CH_2 groups of *i*Bu side chains and T6f units. For **2a** and **3a**, CH_2 groups appear as doublets indicating fast exchange between *PP*/*MM* and *PM*/MP conformers at 298 K. Note that in both cases, one CH_2 group is broadened (coalesced) to such an extent that it is not visible. In **4b**, the CH_2 groups appear as doublets of doublets of

diastereotopic protons. This is a consequence of this molecule being chiral while exchange is still fast on the NMR timescale. In contrast, the doublet of doublets in the spectra of achiral **2b** and **3b** indicate slow exchange on the NMR time scale. In the slow exchange regime, a single set of signals indicates that a single, racemic, diastereomeric conformer is present.

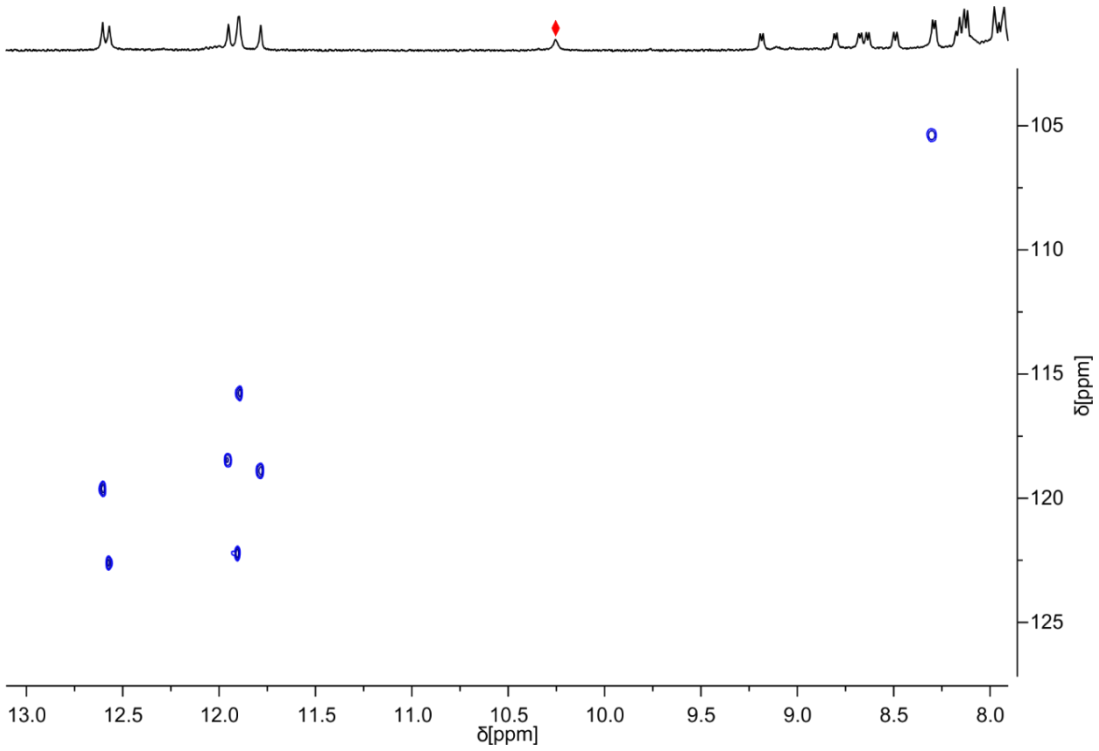


Figure S7. Indirect identification of the signals of the hydrogen-bonded OH proton of 2b. Extract of the $^{1}H,^{15}N$ HSQC NMR spectrum (500 MHz, $CDCl_3$) of **2b**. Only NH resonances correlate, the red diamond indicates the signal of the hydrogen bonded OH proton.

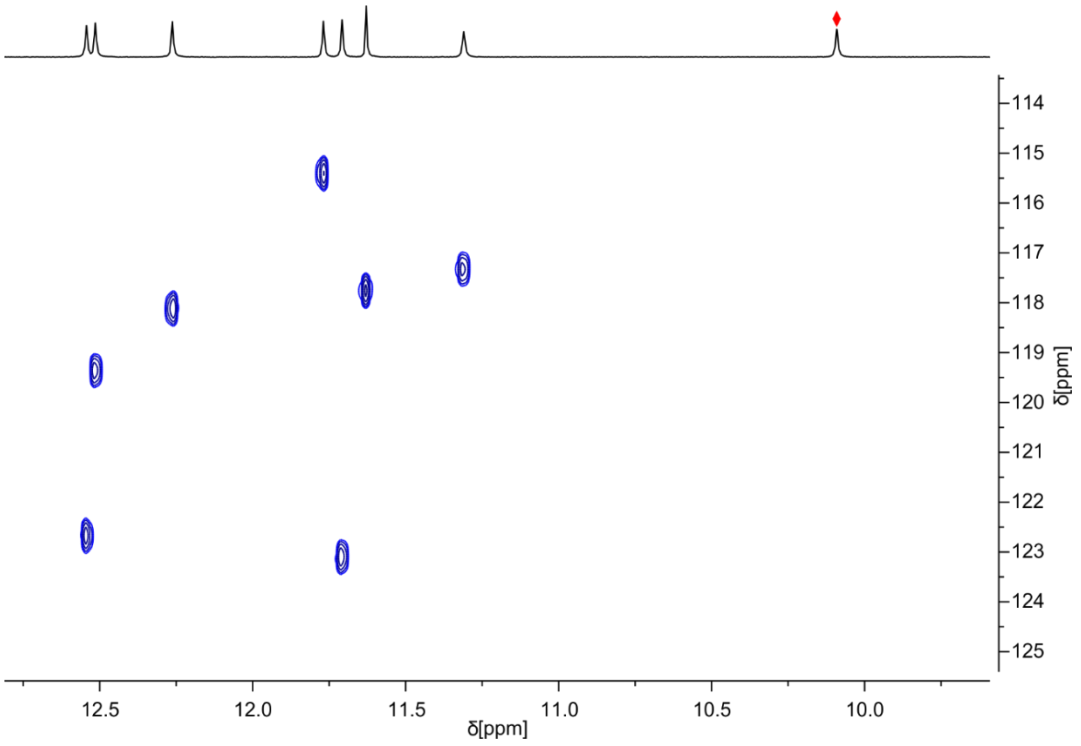


Figure S8. Indirect identification of the signals of hydrogen-bonded OH proton of 3b. Extract of the $^{1}H,^{15}N$ HSQC NMR spectrum (500 MHz, $CDCl_3$) of **3b**. Only NH resonances correlate, the red diamond indicates the signal of the hydrogen bonded OH proton.

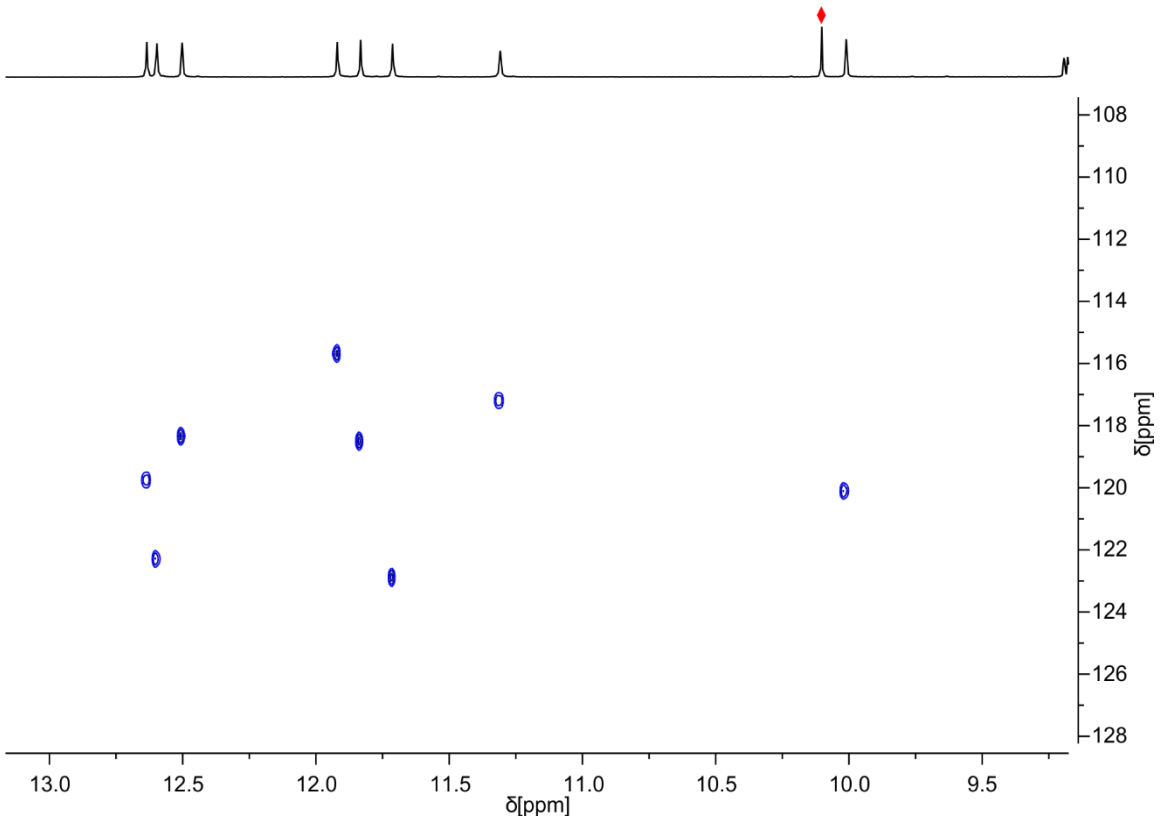


Figure S9. Indirect identification of the signals of hydrogen-bonded OH proton of 4b. Extract of the $^{1}H,^{15}N$ HSQC NMR spectrum (500 MHz, $CDCl_3$) of **4b**. Only NH resonances correlate, the red diamond indicates the signal of the hydrogen bonded OH proton.

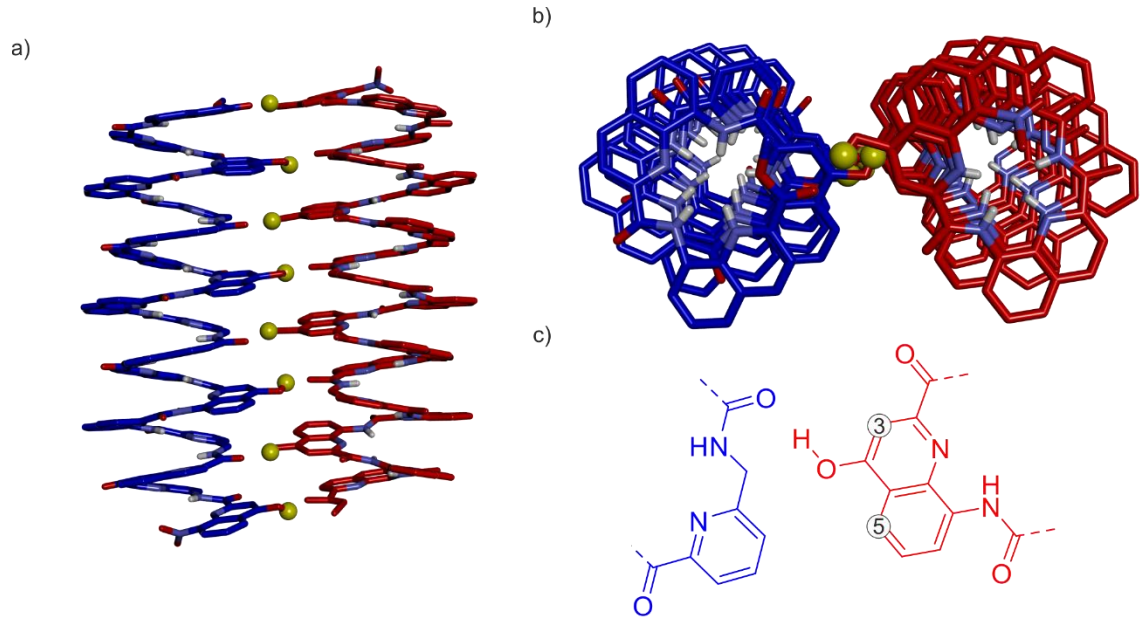


Figure S10. The crystal structure of a shifted dimer.[3] The side view (a) and top view (b) of the shifted dimer structure. The backbones of *P* and *M* helices are colored blue and red, respectively. c) The intermolecular hydrogen bond pattern.

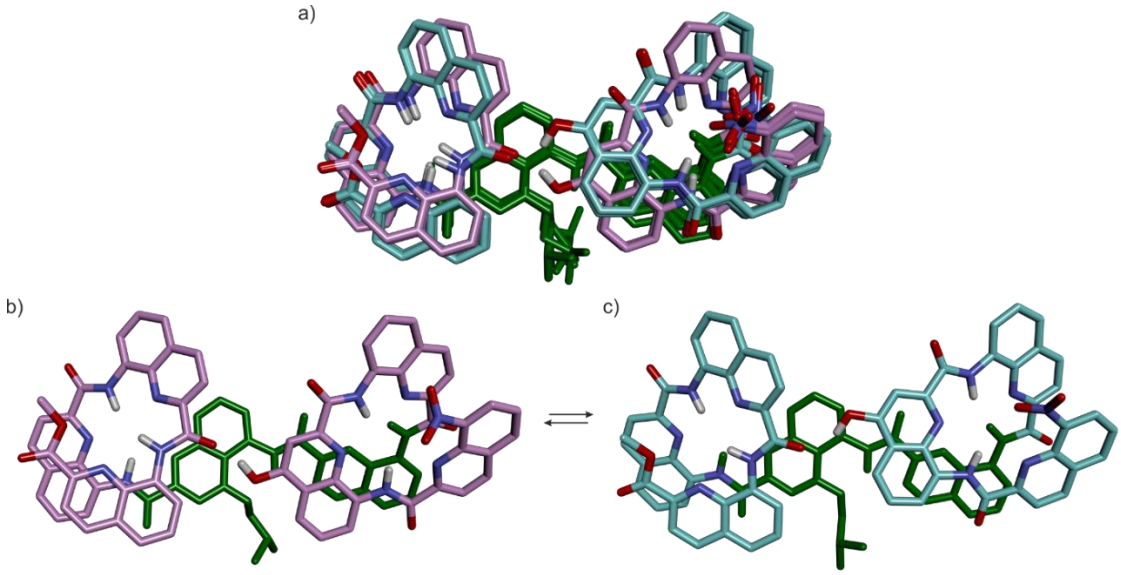


Figure S11. Molecular dynamic (MD) simulation of 3b. The MD simulation was performed based on the crystal structure of **3b** in Maestro (MMFFs, $CHCl_3$, TNCG, 300 K, 1 ns). a) The overlay of ten structures sampled during

the MD simulation. b) The crystal structure of **3b**. c) One helix-turn-helix structure mediated by the trimer pattern extracted from the MD simulation. The two patterns interconvert rapidly.

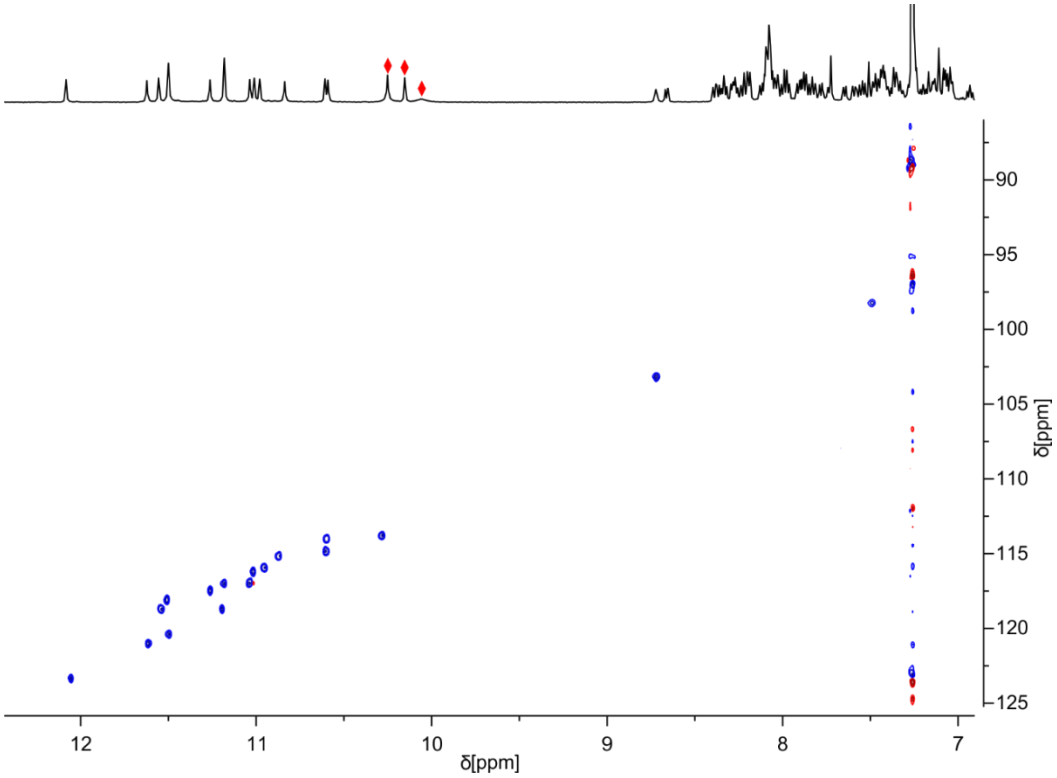


Figure S12. **Indirect identification of the signals of hydrogen-bonded OH protons of 5b in $CDCl_3$.** Extract of the $^{1}H,^{15}N$ HSQC NMR spectrum (500 MHz, $CDCl_3$) of **5b**. Only NH resonances correlate, red diamonds indicate the signals of hydrogen bonded OH protons. The correlation of the glycine NH signal is probably hidden under the streak of the t1 noise caused by the intense resonance of $CHCl_3$.

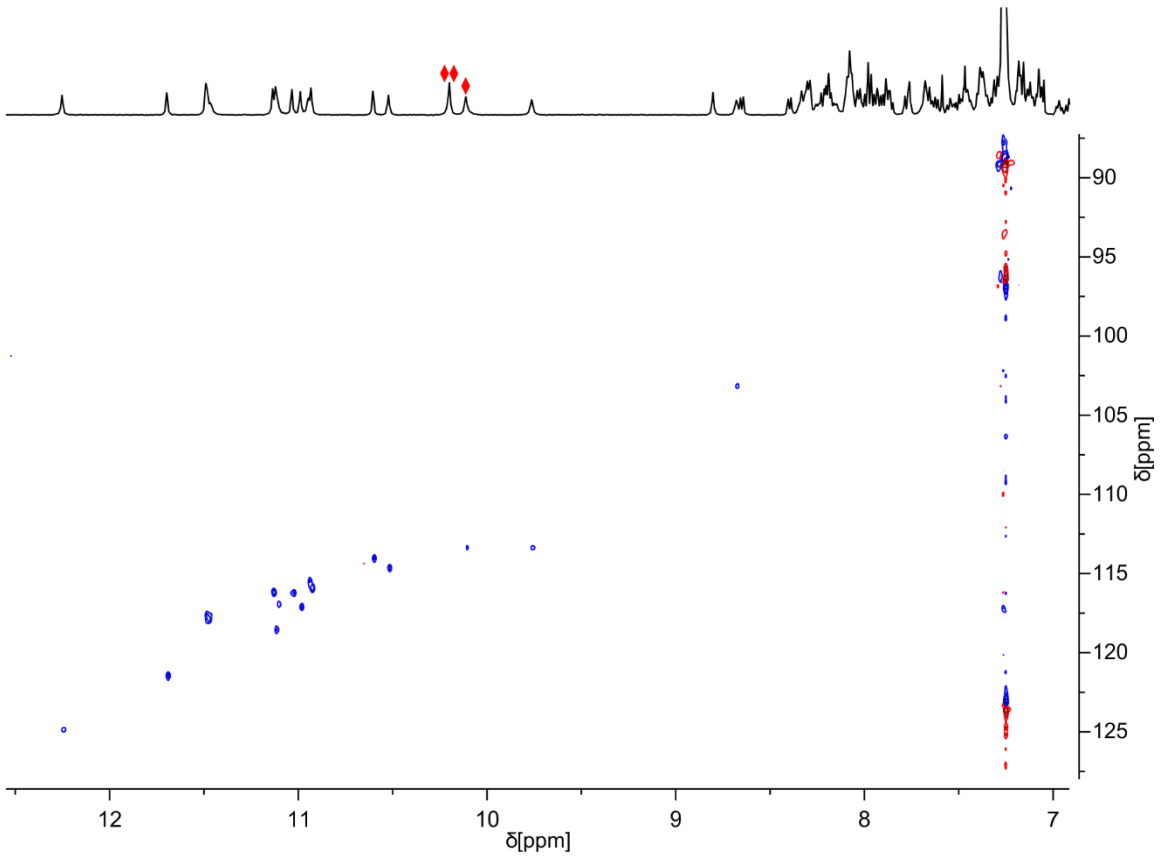


Figure S13. Indirect identification of the signals of hydrogen-bonded OH protons of 6b in $CDCl_3$. Extract of the $^{1}H,^{15}N$ HSQC NMR spectrum (500 MHz, $CDCl_3$) of **6b**. Only NH resonances correlate, red diamonds indicate the signals of hydrogen-bonded OH protons. The correlation of the glycine NH signal is probably hidden under the streak of the t1 noise caused by the intense resonance of $CHCl_3$.

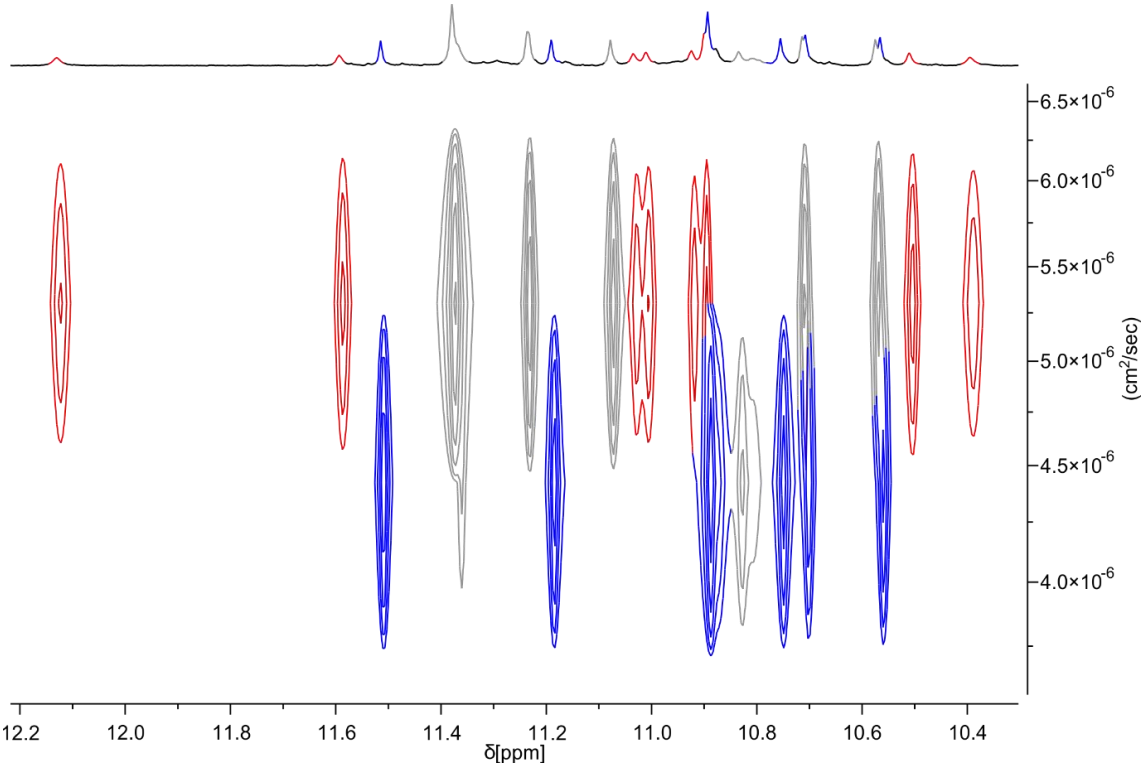


Figure S14. ^{1}H DOSY spectrum (500 MHz, $CDCl_3$) of a mixture of 6a and 6b with a ratio of 1/0.8. Some peaks assigned to **6a** are highlighted in blue and those assigned to **6b** are highlighted in red. Protected sequence **6a** is monomeric and its diffusion coefficient is substantially lower than that of **6b**, indicating that the latter is also monomeric and supporting a compact helix-turn-helix conformation of **6b**.

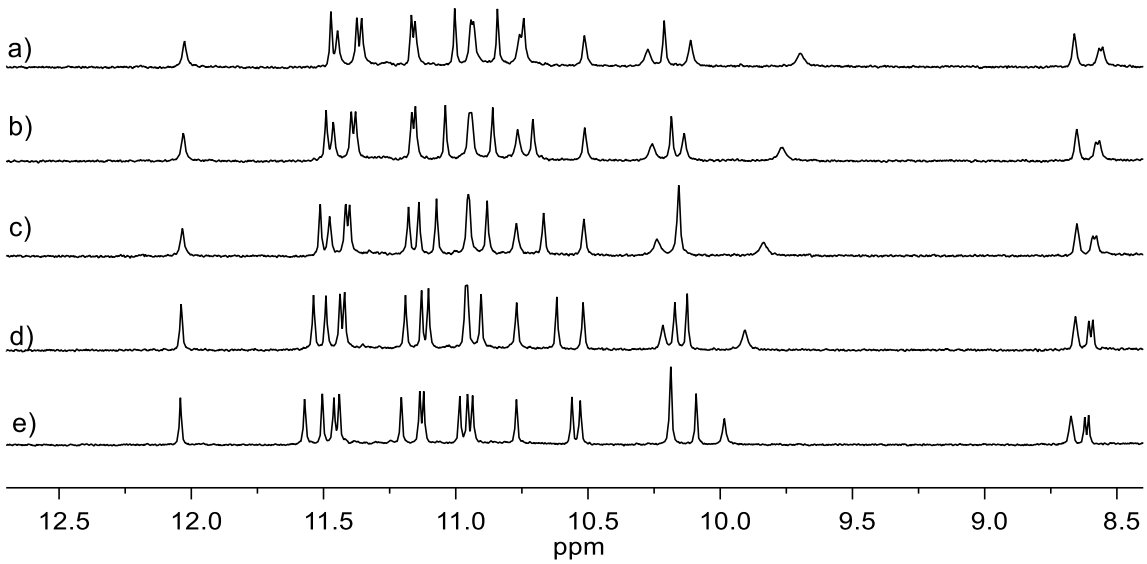


Figure S15. ^{1}H NMR spectra of 5b in $CDCl_3/CD_2Cl_2$ mixtures. Extracts of the ^{1}H NMR spectra (500 MHz) of **5b** in $CDCl_3/CD_2Cl_2$ mixtures showing the amide and hydroxy proton resonances. The vol% of CD_2Cl_2 are 100

(a), 75 (b), 50 (c), 25 (d), and 0 (e). In other sequences, it was observed that the aggregation behavior or the conformation change radically with $CDCl_3$ or CD_2Cl_2,[1,3] but it is not the case for **5b**.

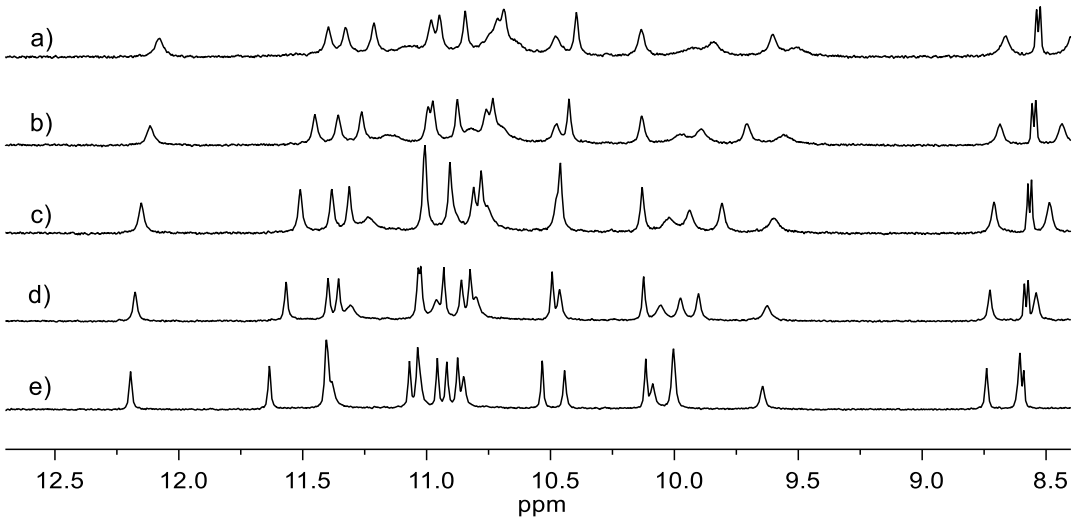


Figure S16. ^{1}H NMR spectra of 6b in $CDCl_3$/CD_2Cl_2 mixtures. Extracts of the ^{1}H NMR spectra (500 MHz) of **6b** in $CDCl_3$/CD_2Cl_2 mixtures showing the amide and hydroxy proton resonances. The vol% of CD_2Cl_2 are 100 (a), 75 (b), 50 (c), 25 (d), and 0 (e). In other sequences, it was observed that the aggregation behavior or the conformation change radically with $CDCl_3$ or CD_2Cl_2,[1,3] but it is not the case for **6b**.

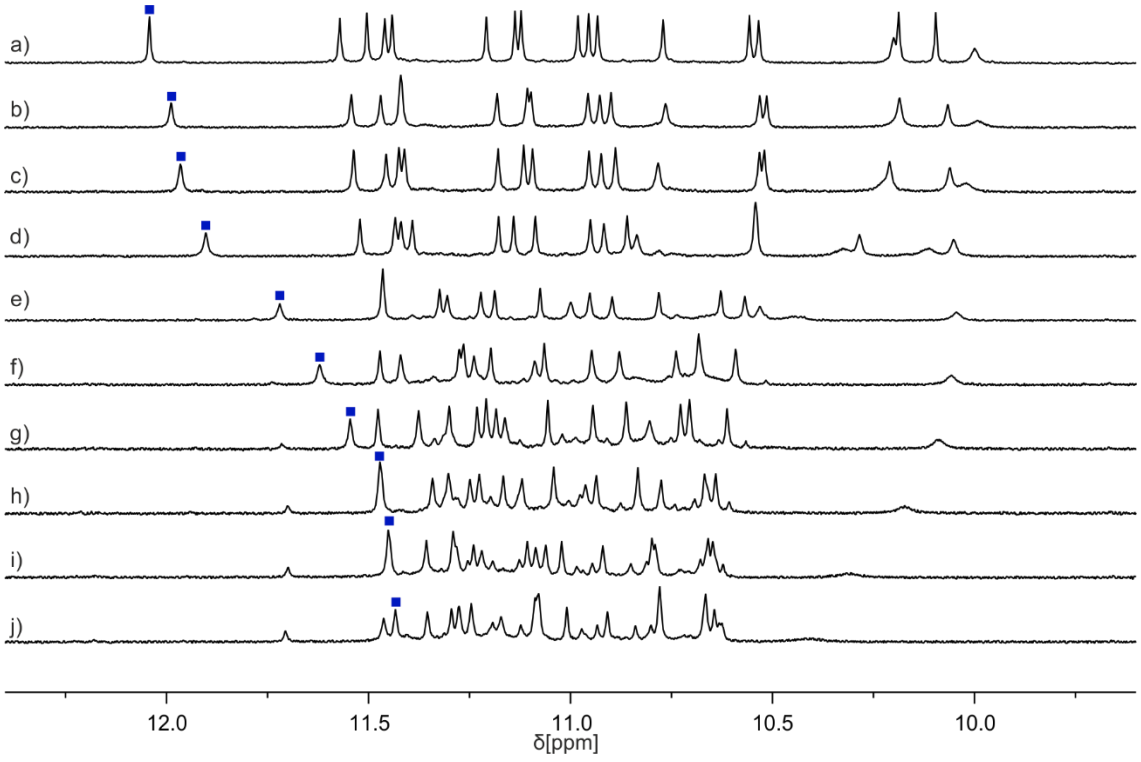


Figure S17. ^{1}H NMR spectra of 5b in DMSO-d_6/$CDCl_3$ mixtures. Extracts of the ^{1}H NMR spectra (500 MHz, 298K) of **5b** in DMSO-d_6/$CDCl_3$ mixtures. The vol% of DMSO-d_6 are 0 (a), 1 (b), 2 (c), 4 (d), 8 (e), 10 (f), 12 (g), 16 (h), 20 (i), and 24 (j). The variation of the chemical shift of the signal marked with a blue box is shown in Figure 6c.

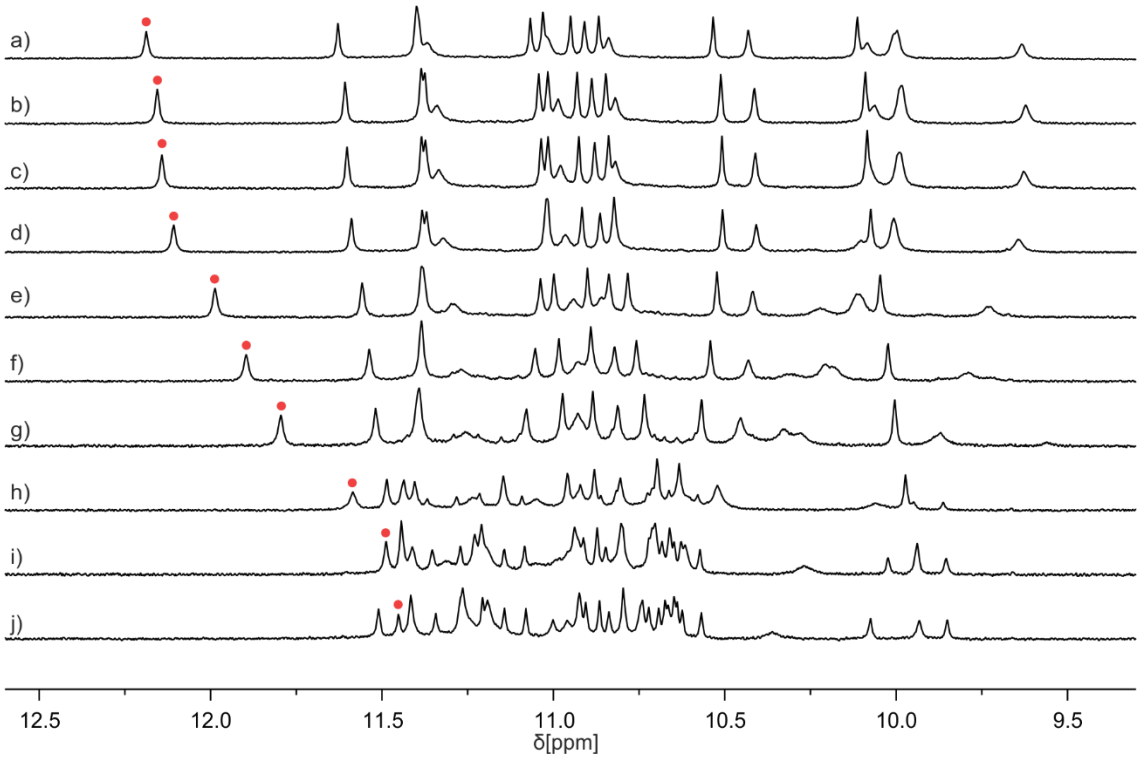


Figure S18. ^{1}H NMR spectra of 6b in DMSO-d_6/$CDCl_3$ mixtures. Extracts of the ^{1}H NMR spectra (500 MHz, 298K) of **6b** in DMSO-d_6/$CDCl_3$ mixtures. The vol% of DMSO-d_6 are 0 (a), 1 (b), 2 (c), 4 (d), 8 (e), 10 (f), 12 (g), 16 (h), 20 (i), and 24 (j). The variation of the chemical shift of the signal marked with a red dot is shown in Figure 6c.

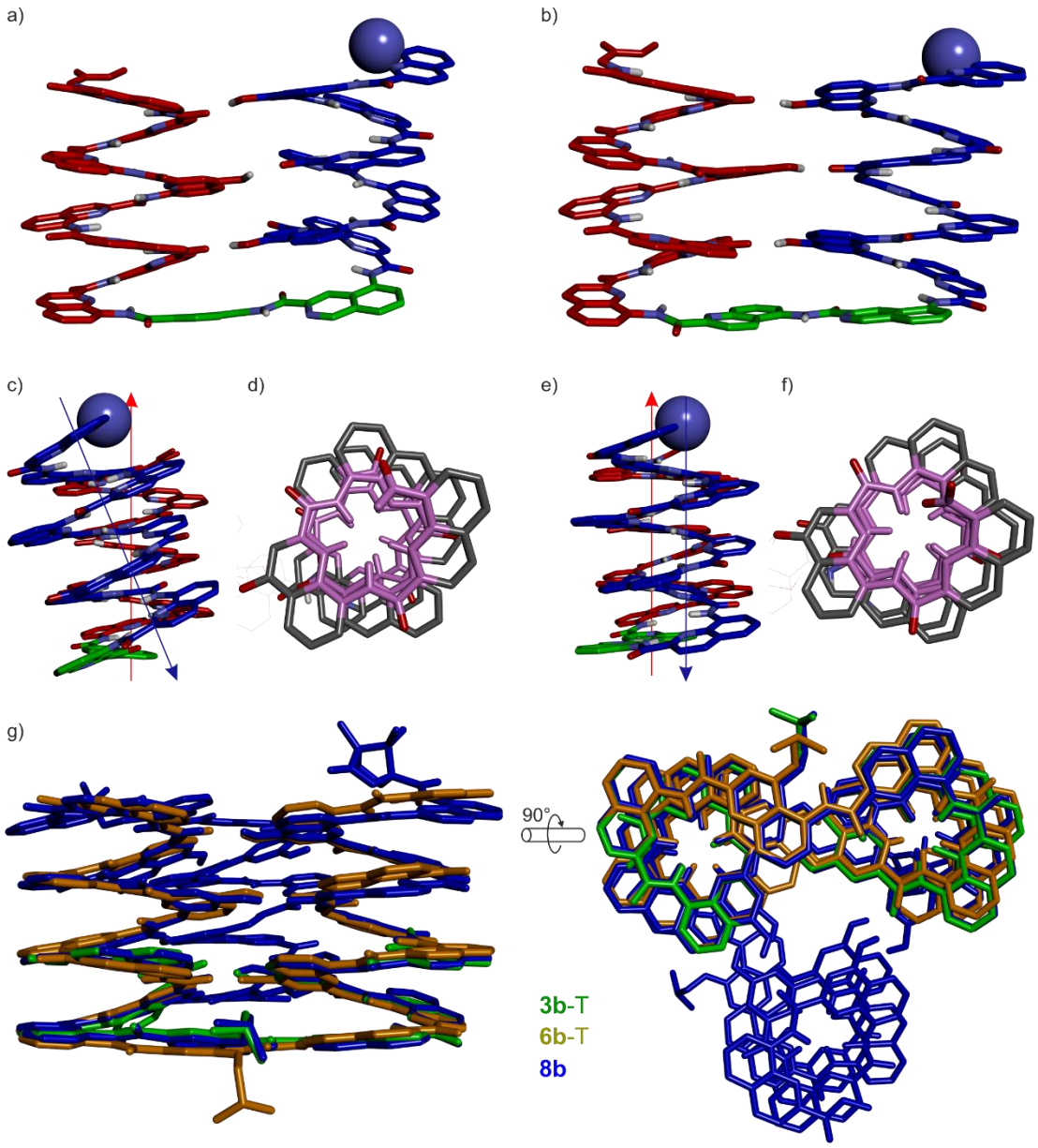


Figure S19. Molecular models of two conformations of 6b mediated different hydrogen bond patterns. Energy minimized model of conformer **6b**-S mediated by shifted a hydrogen bonding pattern (a) and of conformer **6b**-T mediated by a three helix bundle pattern (b). The main chain in the *P* helix, the *M* helix and in T6r is colored in blue, red and green, respectively. The N-terminus is highlighted with a blue ball. The side views of **6b**-S (c) and **6b**-T (e) highlight the relative orientation of the helixes (axes shown as a blue arrow for the *P* helix and a red arrow for the *M* helix). The two helices are parallel in **6b**-T, not in **6b**-S. Views down the axis of the N-terminal helix of **6b**-S (d) and **6b**-T (f). The inner rim is highlighted in pink. It shows a 15-crown-5 shape in **6b**-T, not in **6b**-S, hinting at potential strain in the latter. g) Overlay of molecular models of **3b**-T (**3b** mediated by a three helix bundle pattern, by analogy with **6b**-T, in green), **6b**-T (in brown) and **8b** (in blue).

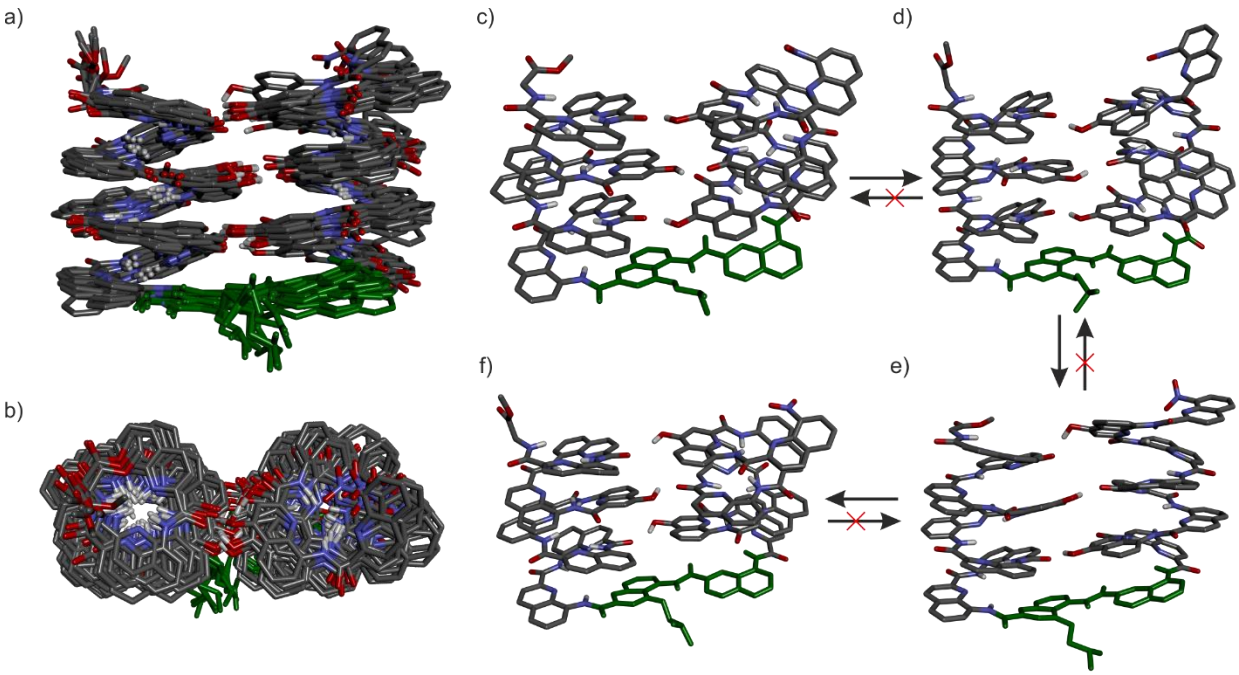


Figure S20. Molecular dynamic (MD) simulation of 6b. The MD simulation in Maestro (MMFFs, $CHCl_3$, TNCG, 300 K, 2 ns) was performed starting from the molecular model of **6b**-S shown in Figure S19. The side view (a) and top view (b) of the overlay of ten structures sampled during the MD simulation. c-f) Four structures extracted from the MD simulation. The structure gradually converted to structures that were similar to **6b**-T during the simulation.

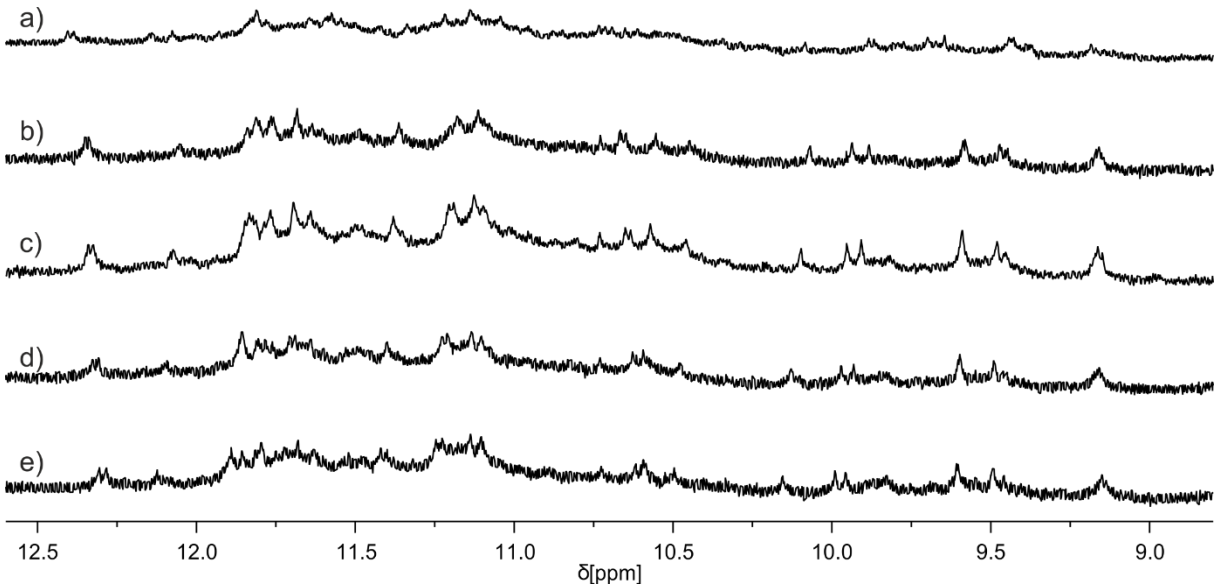


Figure S21. ^{1}H NMR spectra of 7b in $CDCl_3/CD_2Cl_2$ mixtures. Extracts of the ^{1}H NMR spectra (500 MHz) of **7b** in $CDCl_3/CD_2Cl_2$ mixtures showing the amide and hydroxy proton resonances. The vol% of CD_2Cl_2 are 100 (a), 75 (b), 50 (c), 25 (d), and 0 (e).

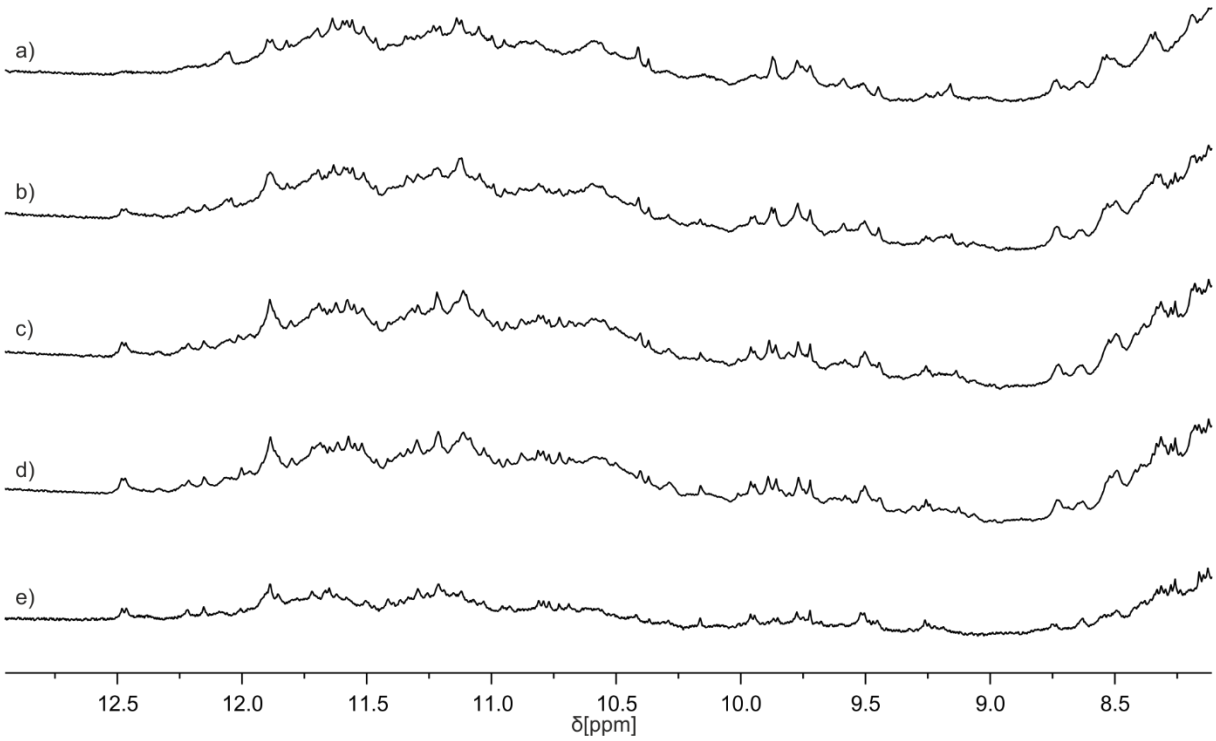


Figure S22. The spectrum of 7b changes marginally with time. Extracts of the ^{1}H NMR spectra (500 MHz, CD_2Cl_2) of **7b**. The equilibration time was 1 h (a), 1 day (b), 5 days (c), 1 week (d), and 2 weeks (e).

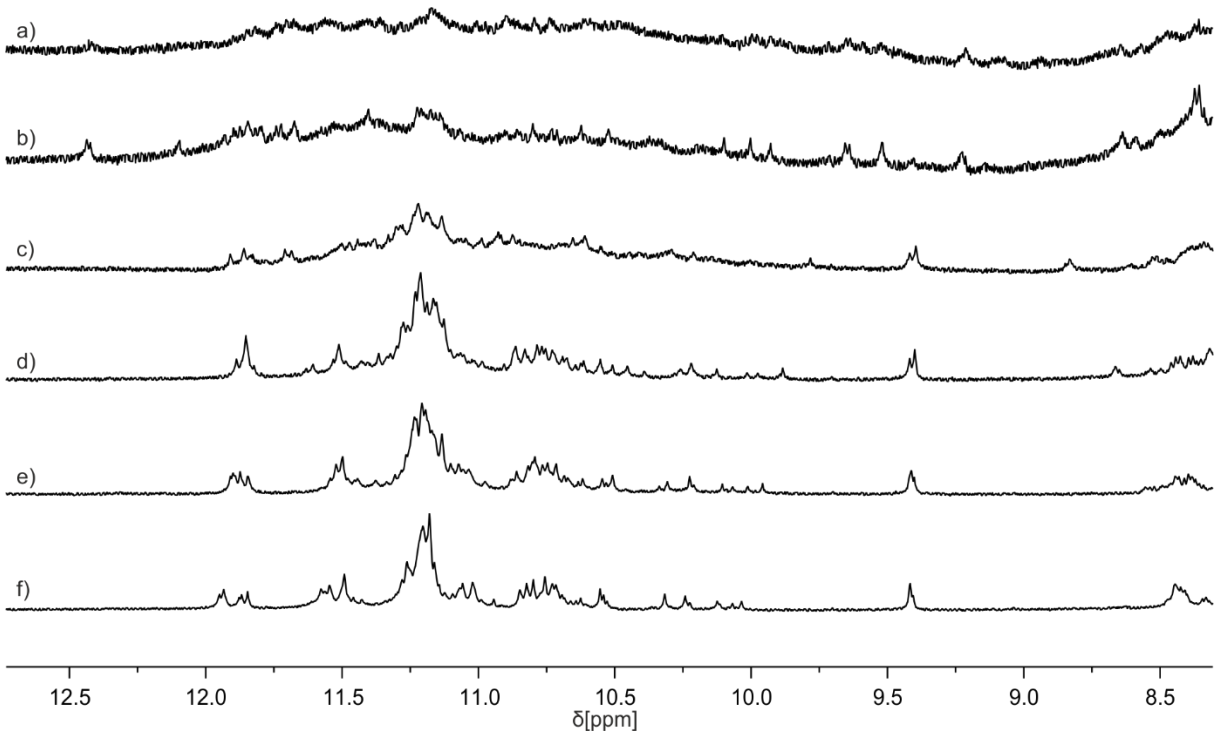


Figure S23. ^{1}H NMR observation of the DMSO-induced dissociation or disruption of 7b. Extracts of the ^{1}H NMR spectra (500 MHz) showing amide resonances of **7b** in CD_2Cl_2/DMSO-d_6 mixture. The vol% of DMSO-d_6 are 2 (a), 4 (b), 8 (c), 12 (d), 16 (e), and 20 (f).

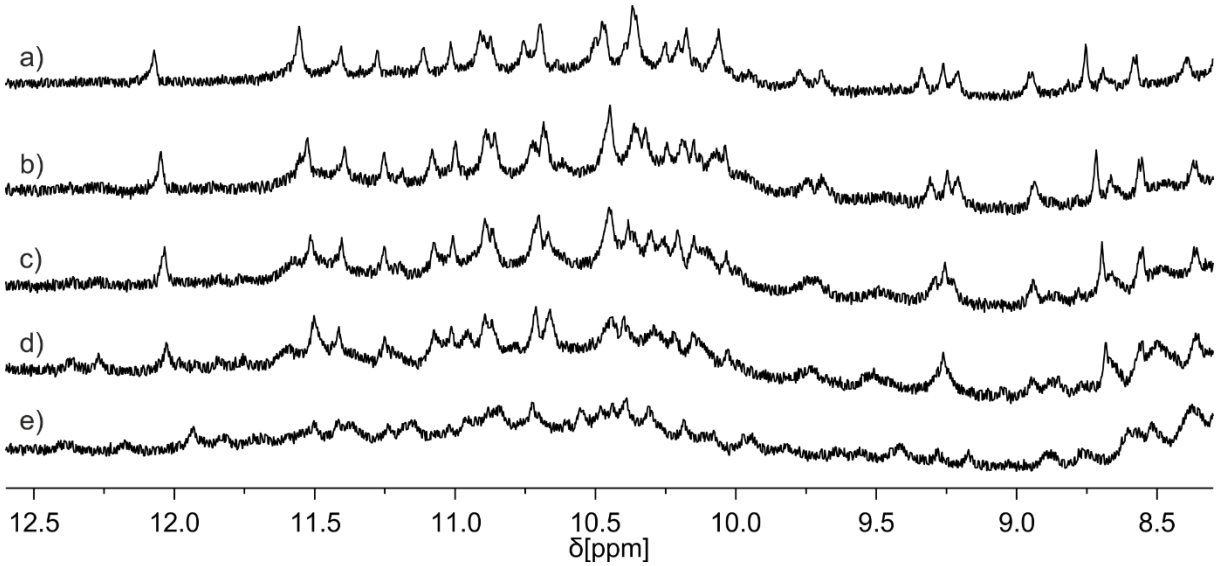


Figure S24. ^{1}H NMR spectra of 8b in $CDCl_3$/CD_2Cl_2 mixtures. Extracts of the ^{1}H NMR spectra (500 MHz) of **8b** in $CDCl_3$/CD_2Cl_2 mixtures showing the amide and hydroxy proton resonances. The vol% of CD_2Cl_2 are 100 (a), 75 (b), 50 (c), 25 (d), and 0 (e).

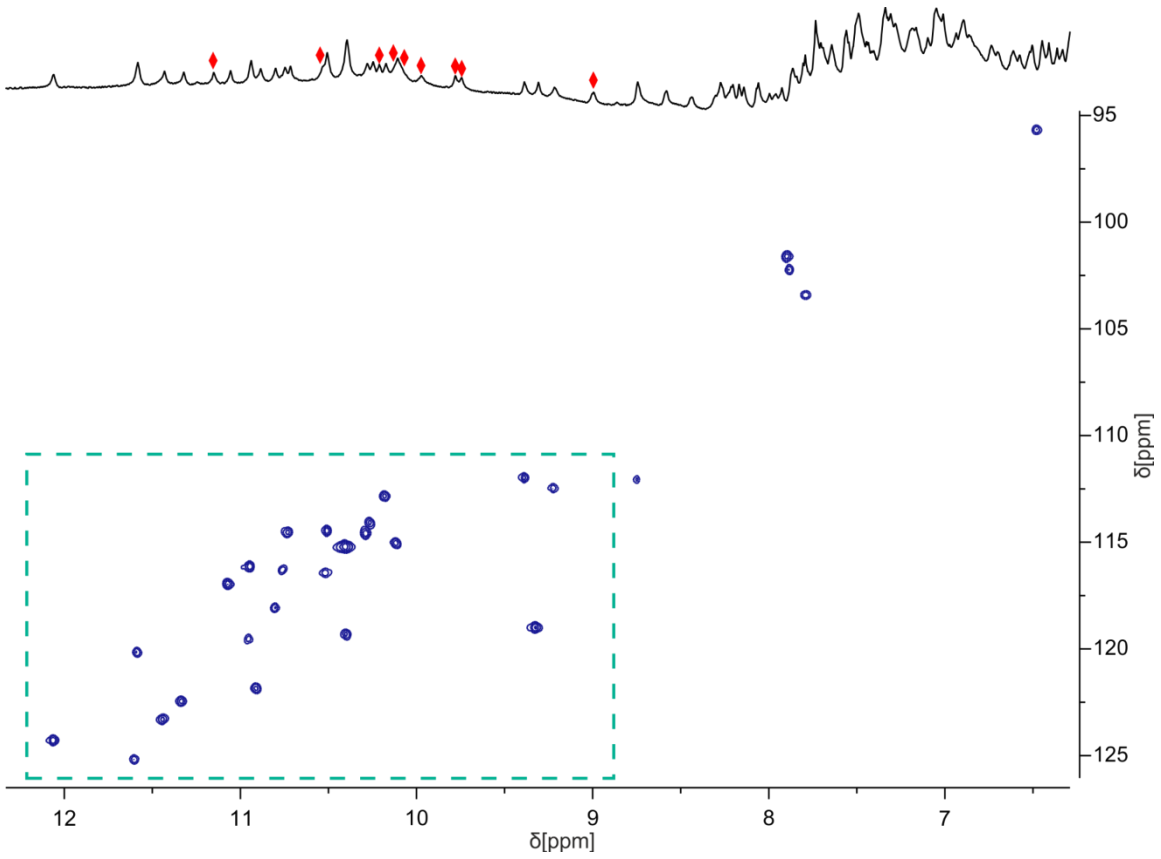


Figure S25. Indirect identification of the signals of hydrogen-bonded OH protons of 8b in $CDCl_3$. Extract of the $^{1}H,^{15}N$ HSQC NMR spectrum (500 MHz, CD_2Cl_2) of **8b**. Only NH resonances correlate, red diamonds indicate the signals of hydrogen-bonded OH protons. The region highlighted with green dashed box is shown in Figure 7d.

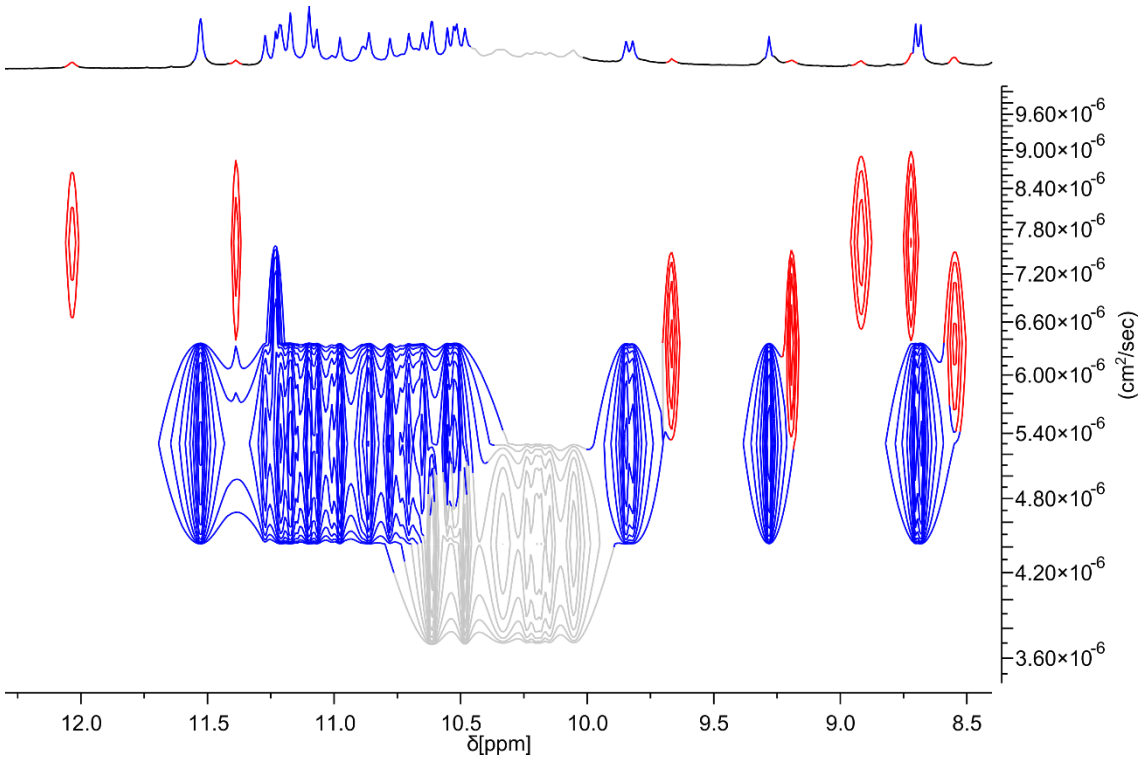


Figure S26. ^{1}H DOSY spectrum (500 MHz, CD_2Cl_2) of a mixture of 8a and 8b in a ratio of 1/0.5. The peaks corresponding to **8a** are highlighted in blue and those to **8b** are highlighted in red. Protected sequence **8a** is monomeric and its diffusion coefficient is substantially lower than that of **8b**, indicating that the latter is also monomeric and supporting a compact helix-turn-helix-turn-helix conformation of **8b**.

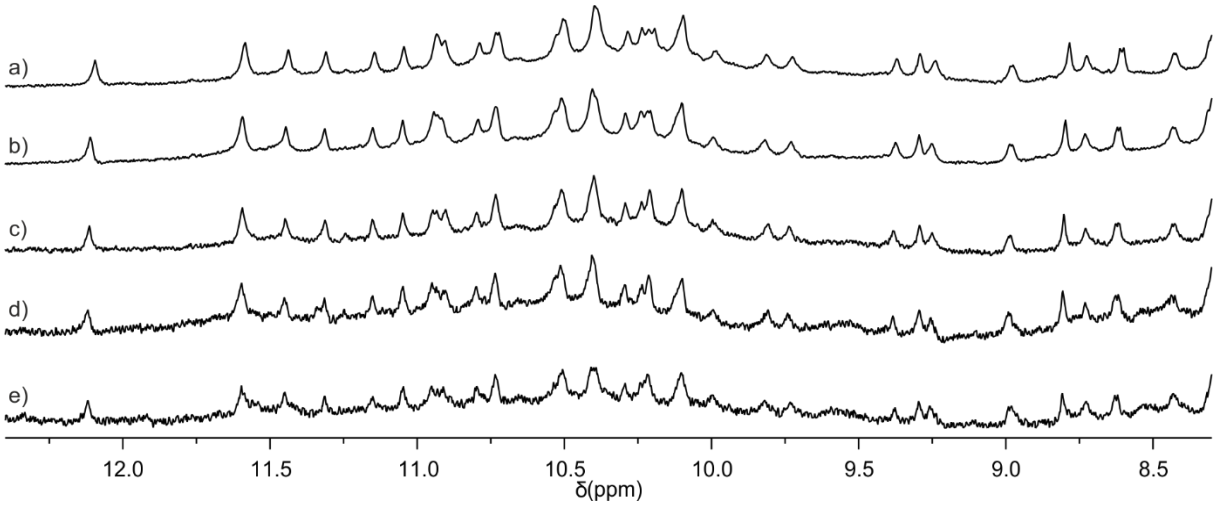


Figure S27. ^{1}H NMR spectra of 8b at different concentrations. Extracts of the ^{1}H NMR spectra of **8b** (500 MHz, 4% DMSO-d_6/CD_2Cl_2) at 1 mM (a), 0.5 mM (b), 0.25 mM (c), 0.1 mM (d), and 0.05 mM (e). No significant changes were observed.

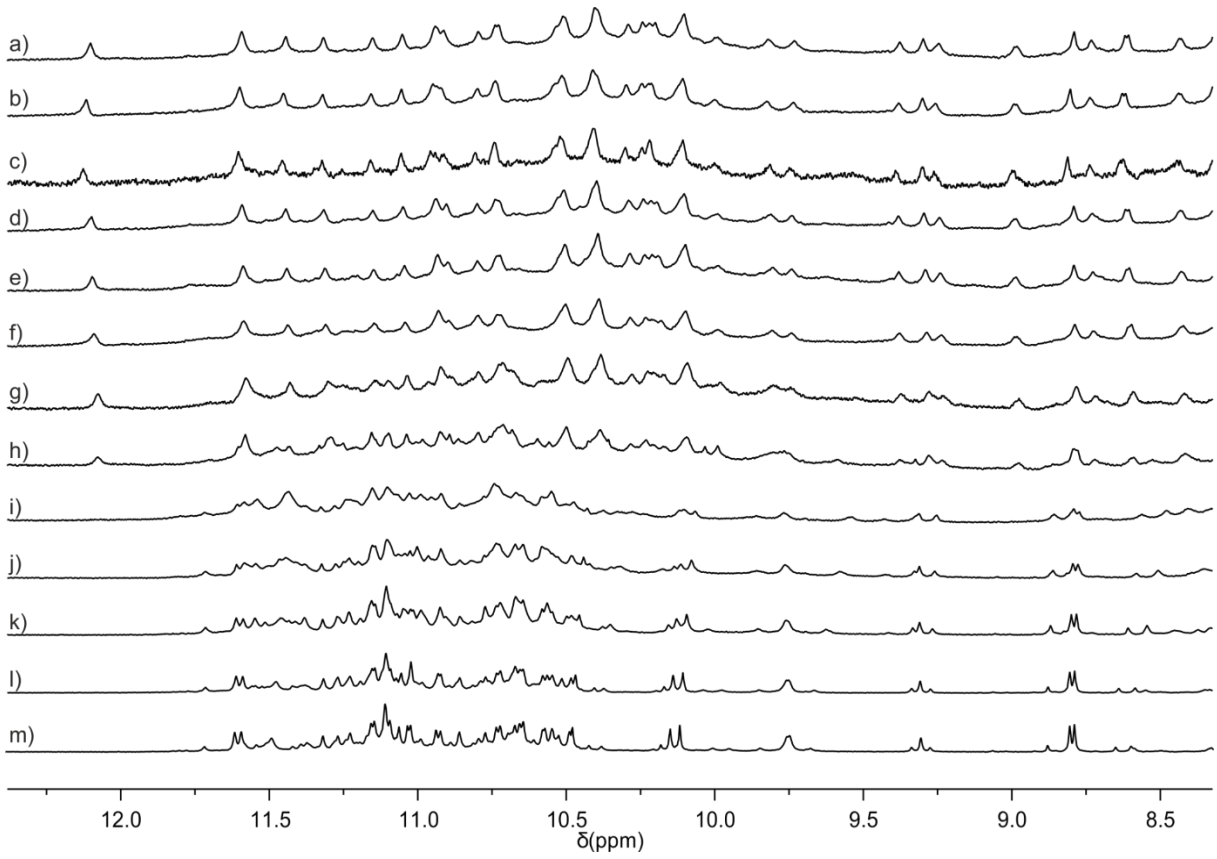


Figure S28. ^{1}H NMR observation of the DMSO-induced disruption of 8b. Extracts of the ^{1}H NMR spectra (500 MHz) showing amide resonances of **8b** in CD_2Cl_2/DMSO-d_6 mixture. The vol% of DMSO-d_6 are 1 (a), 2 (b), 4 (c), 6 (d), 8 (e), 10 (f), 12 (g), 14 (h), 16 (i), 18 (j), 20 (k), 22 (l), and 24 (m).

3 Supplementary methods

3.1 LC-MS analyses

LC-MS spectra were recorded on a Bruker microTOF II in positive ionization mode. The instrument was calibrated in positive mode by direct infusion of a calibration solution (Agilent Technologies ESI-L Low Concentration Tuning Mix). The HPLC line was an Ultimate 3000 RP-HPLC system (ThermoFisher Scientific) equipped with a Aeris^{TM} Widepore C4 column (2.1 x 150 mm, 3.6 µm) at a flow rate of 0.25 mL/min. 0.1 % formic acid and 0.025% TFA were added to the aqueous mobile phase (solvent A) and to acetonitrile (solvent B). The gradient is: 0-5 min, 70% to 100% solvent B; 5–14 min, 100% solvent B at 50 °C. The column eluent was monitored by UV detection at 214, 254, and 300 nm with a diode array detector. The sample was prepared by adding 10 µL of a solution of the sample in DCM (0.1 mg/mL) to 1 mL acetonitrile containing 0.05–0.1% formic acid.

3.2 Molecular modelling

Models were simulated by using Maestro version 11.5 (Schrödinger Inc.). Energy minimized structures were obtained using MacroModel energy minimization with the following parameters: force field: MMFFs; solvent: none; electrostatic treatment: constant dielectric; dielectric constant: 1.0; charges from: force field; cutoff: normal; Van der Waals: 7.0; electrostatic: 12.0; H-bond: 4.0; mini method: TNCG; maximum iterations: 2500; converge on gradient; convergence threshold: 0.05; constraints: distances. As a starting point, the crystal structure of **1** (CCDC entry # 1955168) was used. The Y unit at the N-terminus was replaced by a nitro group and the handedness and orientation of one helix was converted. The molecular model shown in Figure 3a, S4b,d was obtained after energy-minimization. After inserting two T6r linker units to the heterochiral trimer molecular model, the structure was energy-minimized to obtain the molecular model of **8b** shown in Figure 3d, 8a, S4e,f. As a starting point, the crystal structure of **3b** (CCDC entry # 2391429) was used and the C- and N-terminus were prolonged. The molecular model of **6b**-S shown in Figure S19a was obtained after energy-minimization. After removing the C-terminus helix and T6r unit of molecular model of **8b**, the molecular model of **6b**-T shown in Figure 6a, S19b was obtained after energy-minimization. The molecular model of **3b**-T shown in Figure S19g was obtained using a similar approach.

3.3 Molecular dynamic simulations

Molecular dynamic simulations were carried out using MacroModel version 11.1 (Maestro, Schrödinger Inc.). The crystal structure of **3b** (CCDC entry # 2391429) and the energy-minimized molecular model of **6b-S** and **6b-T** were used as the object for the MD simulation. Stochastic dynamic simulations were obtained using MMFFs force field, $CHCl_3$ as the solvent, extended cutoff and TNCG method. The simulations were performed for 1 ns or 2 ns at 300 K, the time step of 1.5 fs and 1 ps as equilibration time. During the simulation, 10 structures were sampled and the sampled structures were minimized.

3.4 Nuclear magnetic resonance spectroscopy

NMR spectra were recorded on different NMR spectrometers: (1) an Avance III HD NMR spectrometer 400 MHz (Bruker BioSpin) for ^{1}H NMR and ^{13}C NMR spectra of some small molecules. (2) an Avance III HD NMR spectrometer 500 MHz (Bruker BioSpin) with CryoProbe™ Prodigy for ^{1}H NMR, ^{13}C NMR, ^{1}H,^{15}N HSQC, and DOSY spectra of some small molecules and foldamers. All NMR measurements were performed at 25 °C unless specified. Chemical shifts are described in part per million (ppm, δ) relative to the ^{1}H residual signal of the deuterated solvent used – meaning DMSO-d_6 (δ 2.50 ppm), pyridine-d_5 (δ 8.74 ppm), CD_2Cl_2 (δ 5.32 ppm) and $CDCl_3$ (δ 7.26 ppm). For the DMSO-d_6/$CDCl_3$ and DMSO-d_6/CD_2Cl_2 solvent mixtures, the chemical shifts were calibrated according to DMSO-d_6 (δ 2.50 ppm). For the CD_2Cl_2 and $CDCl_3$ solvent mixture, the chemical shifts were calibrated according to internal standard tetramethylsilane (δ 0.00 ppm).^{1}H NMR splitting patterns with observed first-order coupling are entitled as singlet (s), broad singlet (bs), doublet (d), triplet (t), doublet of doublets (dd) or multiplet (m). Coupling constants (*J*) are ported in Hz.

The ^{1}H NMR spectra of each sample were measured at different times respectively until no further change was observed within a week. We generally consider that at this point the compound reached equilibrium. When the sample reached equilibrium, re-dissolving the compound solid results in the equilibrated spectrum immediately without going through the equilibration process again. Complete disruption of the hydrogen bonds was achieved by dissolving the sample in polar solvents (such as DMSO, pyridine or MeOH/chloroform mixture), followed by removal of the solvent. When all of the hydrogen bonds were completely disrupted, it has to go through the equilibrium process again to reach the equilibrium. The equilibration time (the measurement time gap between two different conditions) of ^{1}H NMR spectra at different temperatures and in different proportions of DMSO-d_6/$CDCl_3$ and DMSO-d_6/CD_2Cl_2 solvents were usually several minutes. Due to similar properties of

$CDCl_3$ and CD_2Cl_2, the individual samples in different proportions of $CDCl_3/CD_2Cl_2$ mixture were prepared and the ^1H NMR spectra of all of the samples were measured over time whereas no change was observed.

$^1H,^{15}N$ HSQC spectra were recorded with a phase-sensitive pulse sequence with sensitivity enhancement using trim pulses in inept transfer (hsqcetgpsi2) from the Bruker pulse program library. Data acquisition was performed utilizing non-uniform sampling (NUS; NUS amount: 50% with an automatically created NUSList) yielding 1024 (F2) x 128 (F1) data points in Echo/Antiecho gradient selection mode. The recycling delay was 2.0 s and 64 transients per increment were applied at a sweep width of 2.5 kHz in F2 and 7 kHz in F1 resulting in an acquisition time of 0.1462 s. NUS processing was performed using the fully automated NUS processing tool provided by MestReNova. Zero filling in F1 has been used to yield a final matrix of 1K x 1K real points.

The DOSY spectrum was recorded by applying a pulse sequence with stimulated echo using stimulated echo for diffusion from the Bruker pulse program library (stegp1s). The diffusion delay Δ (big delta) and the diffusion gradient pulse length δ (little delta) were set as follows: 100 ms and 2.0 ms for the DOSY of **6a**/**6b** mixture, 150 ms and 2.0 ms for the DOSY of **8a**/**8b** mixture. The number of gradient steps was set to 32 with linear spacing starting from 2% reaching 95% of the full gradient strength in the final step. For each of the 32 gradient amplitudes, 256 transients of 65K complex data points were acquired. DOSY processing was performed with the DOSY processing tool from MestReNova (v.12.x64) employing the "Peak Heights Fit" algorithm including the "use existing peaks" and "autocorrect peak positions" with 128 points in diffusion dimension and a window of $1.00 \cdot 10^{-10}$ to $1.00 \cdot 10^{+00}$ $cm^2\ s^{-1}$.

3.5 CD studies

The CD spectra of **4a** and **4b** were recorded on a Jasco J-1500 spectrometer with 1 mm quartz cuvette. The following parameters were used: wavelength range from 460 to 280 nm. Scan speed: 100 nm/min; accumulation: 2; response time: 1.0 s; bandwidth: 2; temperature: 20 °C; sensitivity: standard (100 mdeg); data pitch: 0.5 nm; nitrogen gas flow rate: 500 L/h. The CD spectra of **8a** and **8b** were recorded on a Jasco J-810 spectrometer with 1 mm quartz cuvette. The following parameters were used: wavelength range from 460 to 280 nm. Scan speed: 200 nm/min; accumulation: 3; response time: 1.0 s; bandwidth: 2; temperature: 20 °C; sensitivity: standard (100 mdeg); data pitch: 0.1 nm; nitrogen gas flow rate: 500 L/h. The sample solution of **4b** and **8b** in different proportions of $DMSO/CHCl_3$ or $DMSO/CH_2Cl_2$ solvents was prepared separately. The concentration is 0.1 mM for **4a**/**4b** and 0.05 mM for **8a**/**8b**. Δε values (in $cm^2 \cdot mol^{-1}$) were obtained by using the formula: Δε =

m°/(C.l.32980) where m = CD value in millidegrees; l = cuvette pathlength in cm; C = sample concentration in mol/L.

3.6 X-ray crystallography

Crystals of **2b** and **3b** were grown via liquid-liquid diffusion in NMR tubes from $CHCl_3$-MeCN and CH_2Cl_2-MeCN respectively. X-ray diffraction data was collected on a Rigaku XtaLAB Synergy R, HyPix-Arc 150 with Cu-Kα radiation from a PhotonJet Rotating-anode X-ray Source at 100 K. Data processing was performed in CrysAlisPro (V. 1.171.43.130a).[4] The structures were solved by direct methods and preliminarily refined in the AutoChem-6[5] pipeline in CrysAlisPro. Further refinements were carried out with the SHELXL[6] package through Olex2 1.5-ac6-020[7]. All non-hydrogen atoms were refined anisotropically. Hydrogen atom positions were calculated geometrically and refined using the riding model. DFIX restraints were used to optimize model geometry and RIGU instructions were applied to some sidechain atom groups. Due to the high content of disordered chlorinated solvents in both structures, solvent masking was implied using the BYPASS implementation in Olex2. In structure 3b, atoms beyond the ring adjacent oxygen (O00A) of the -O*i*Bu sidechain on the iso-Quinoline linker were deleted, since no complete chain could be modelled due to disorder. Final cif files were evaluated with the checkCIF algorithm through the IUCR online tool. Atomic coordinates and structure factors were deposited to the Cambridge Crystallographic Data Centre (CCDC) and can be accessed under the CCDC codes 2391393 (**2b**) and 2391429 (**3b**).

Table S1. Crystal data and refinement details

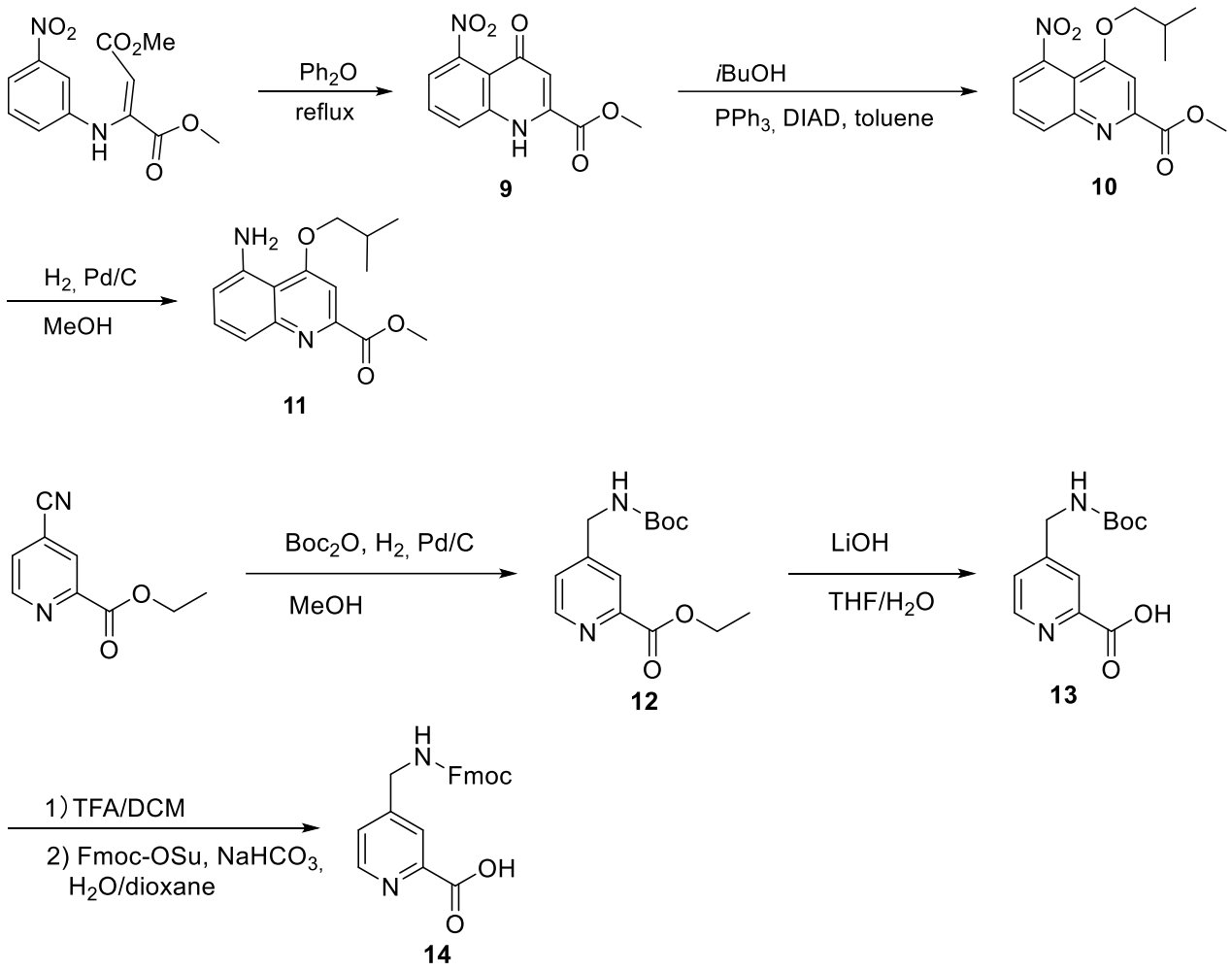
Identification code	**2b**	**3b**
Chemical formula	$C_{99}H_{92}N_{16}O_{18} \cdot CHCl_3$ solvent	$C_{102}H_{92}N_{16}O_{18} \cdot (CH_2Cl_2\ 3\ MeCN)$ solvent
Formula weight	1793.92	1829.95
Crystal growth conditions	liquid-liquid diffusion $CHCl_3$-MeCN	liquid-liquid diffusion CH_2Cl_2-MeCN
Crystal size (mm^3)	0.6 × 0.15 × 0.1	0.42 × 0.37 × 0.21
Crystal system	Triclinic	
Spacegroup	*P*-1	
Unit cell dimensions (Å, °)	a = 14.3247(2)	a = 13.16060(10)
	b = 17.1233(2)	b = 17.3713(2)
	c = 23.0646(2)	c = 24.3400(2)
	α = 96.8220(10)	α = 95.6610(10)

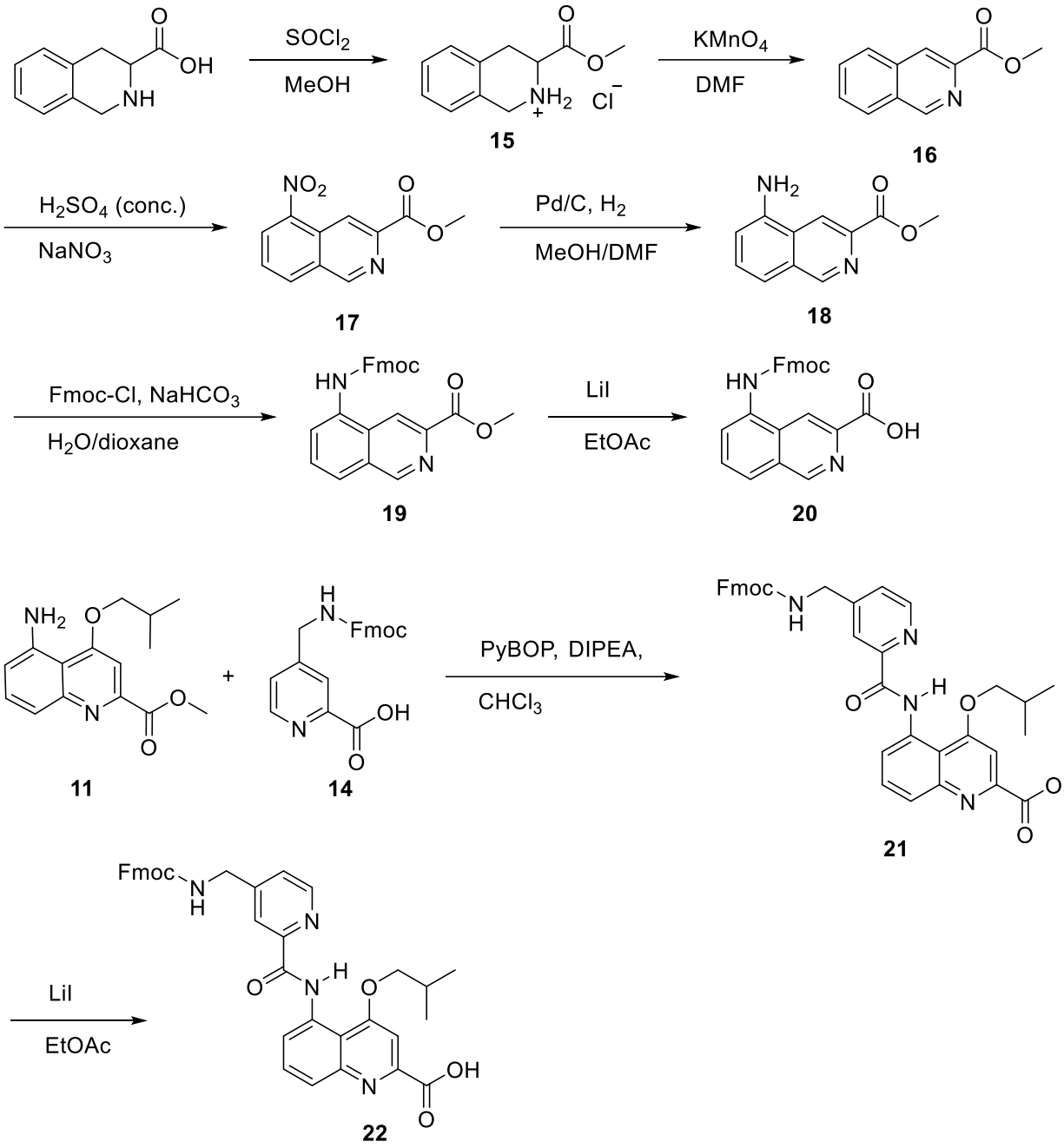
	β = 98.8460(10)	β = 93.9890(10)
	γ = 108.1960(10)	γ = 105.3540(10)
Volume (Å^3)	5226.07(11)	5307.66(7)
Z	2	2
ρ(calc) (g/cm^3)	1.215	1.234
Absorption coefficient (mm^{-1})	1.378	1.155
F (000)	1998.0	2063.0
Radiation	Cu Kα (λ = 1.54184)	
T (K) collection	100.0	
2Θ range for data collection/°	6.188 to 151.082	6.82 to 151.304
Reflections collected	106475	113760
Independent reflections	20937 [R_{int} = 0.0359, R_{sigma} = 0.0235]	21386 [R_{int} = 0.0243, R_{sigma} = 0.0189]
Data/restraints/parameters	20937/4/1305	21386/2/1335
Goodness-of-fit on F^2	1.052	1.073
Final R indexes [I>=2σ (I)]	R_1 = 0.0761, wR_2 = 0.2384	R_1 = 0.0806, wR_2 = 0.2312
Final R indexes [all data]	R_1 = 0.0854, wR_2 = 0.2489	R_1 = 0.0864, wR_2 = 0.2372
Largest diff. peak/hole	0.87/-1.28	2.92/-1.53
CCDC accession #	2391393	2391429
Identification code	2b	3b
Chemical formula	$C_{99}H_{92}N_{16}O_{18} \cdot CHCl_3$ solvent	$C_{102}H_{92}N_{16}O_{18} \cdot (CH_2Cl_2$ 3 MeCN) solvent
Formula weight	1793.92	1829.95
Crystal growth conditions	liquid-liquid diffusion $CHCl_3$-MeCN	liquid-liquid diffusion CH_2Cl_2-MeCN
Crystal size (mm^3)	0.6 × 0.15 × 0.1	0.42 × 0.37 × 0.21
Crystal system	Triclinic	
Spacegroup	*P*-1	
Unit cell dimensions (Å, °)	a = 14.3247(2)	a = 13.15900(8)
	b = 17.1233(2)	b = 17.36265(14)
	c = 23.0646(2)	c = 24.33049(17)
	α = 96.8220(10)	α = 95.6603(6)
	β = 98.8460(10)	β = 94.0009(5)
	γ = 108.1960(10)	γ = 105.3544(6)
Volume (Å^3)	5226.07(11)	5307.66(7)
Z	2	2
ρ(calc) (g/cm^3)	1.215	1.280
Absorption coefficient (mm^{-1})	1.378	1.176
F (000)	1998.0	2139.0

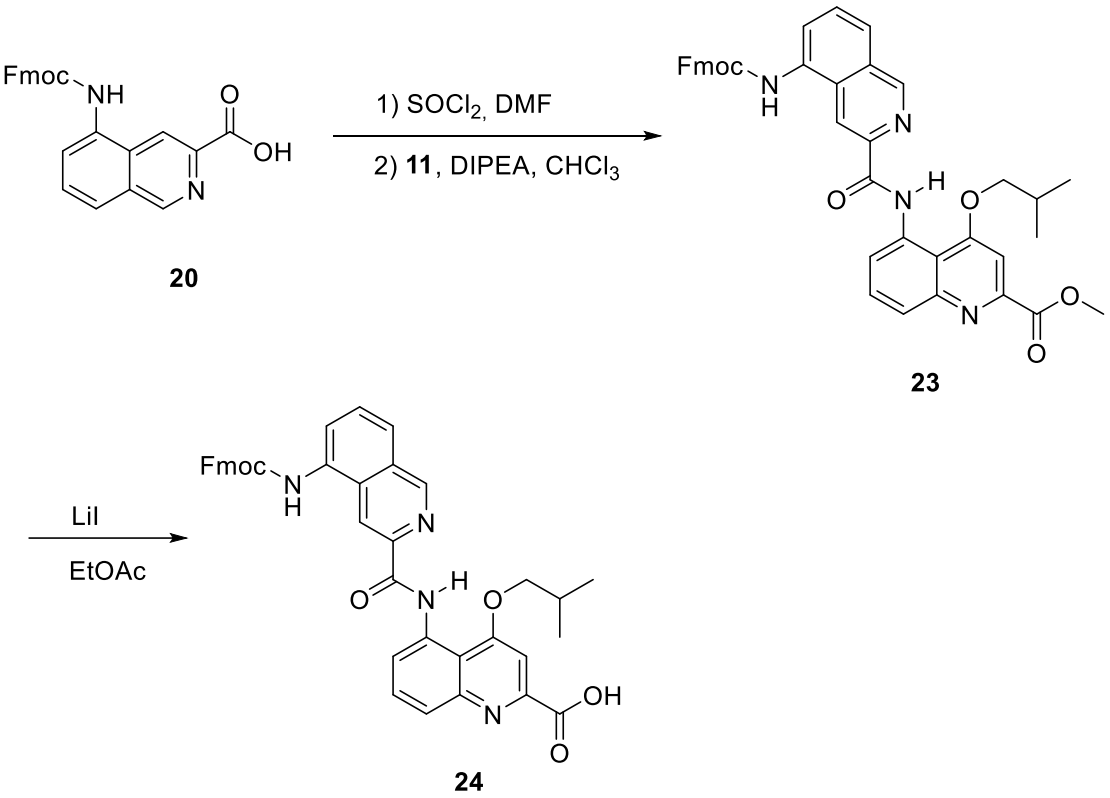
Radiation	Cu Kα (λ = 1.54184)	
T (K) collection	100.0	
2Θ range for data collection/°	6.188 to 151.082	6.82 to 151.304
Reflections collected	106475	113668
Independent reflections	20937 [R_{int} = 0.0359, R_{sigma} = 0.0235]	21377 [R_{int} = 0.0281, R_{sigma} = 0.0203]
Data/restraints/parameters	20937/4/1305	21377/67/1406
Goodness-of-fit on F^2	1.052	1.039
Final R indexes [I>=2σ (I)]	R_1 = 0.0761, wR_2 = 0.2384	R_1 = 0.0691, wR_2 = 0.2012
Final R indexes [all data]	R_1 = 0.0854, wR_2 = 0.2489	R_1 = 0.0746, wR_2 = 0.2070
Largest diff. peak/hole	0.87/-1.28	1.36/-1.05
CCDC accession #	2391393	2391429

4 Synthetic Schemes

4.1 Synthesis of turn units

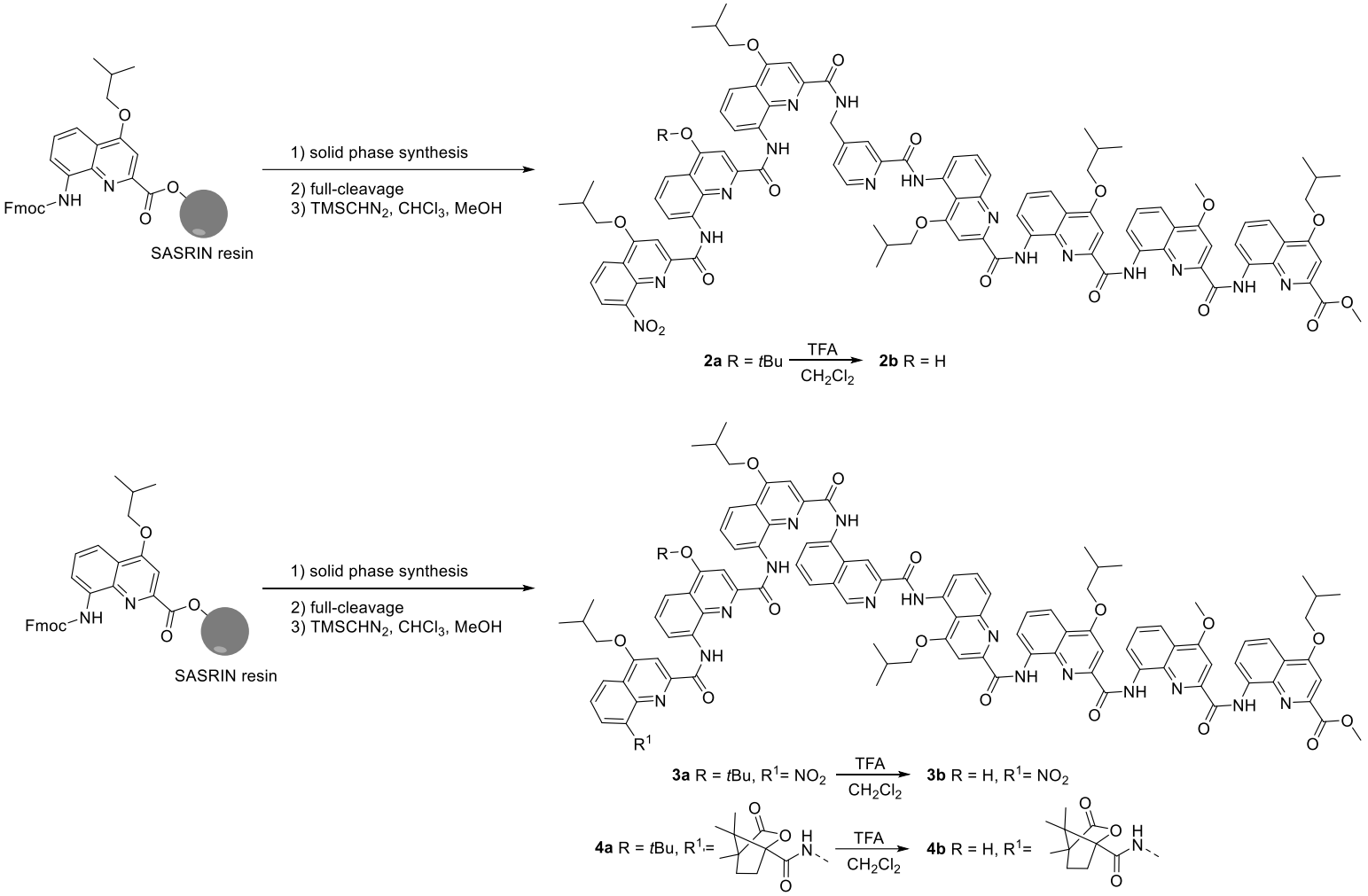



SOCl2
MeOH
KMnO4
DMF
15
16
H2SO4 (conc.)
NaNO3
Pd/C, H2
MeOH/DMF
17
18
Fmoc-Cl, NaHCO3
H2O/dioxane
LiI
EtOAc
19
20
11
14
PyBOP, DIPEA,
CHCl3
21
LiI
EtOAc
22

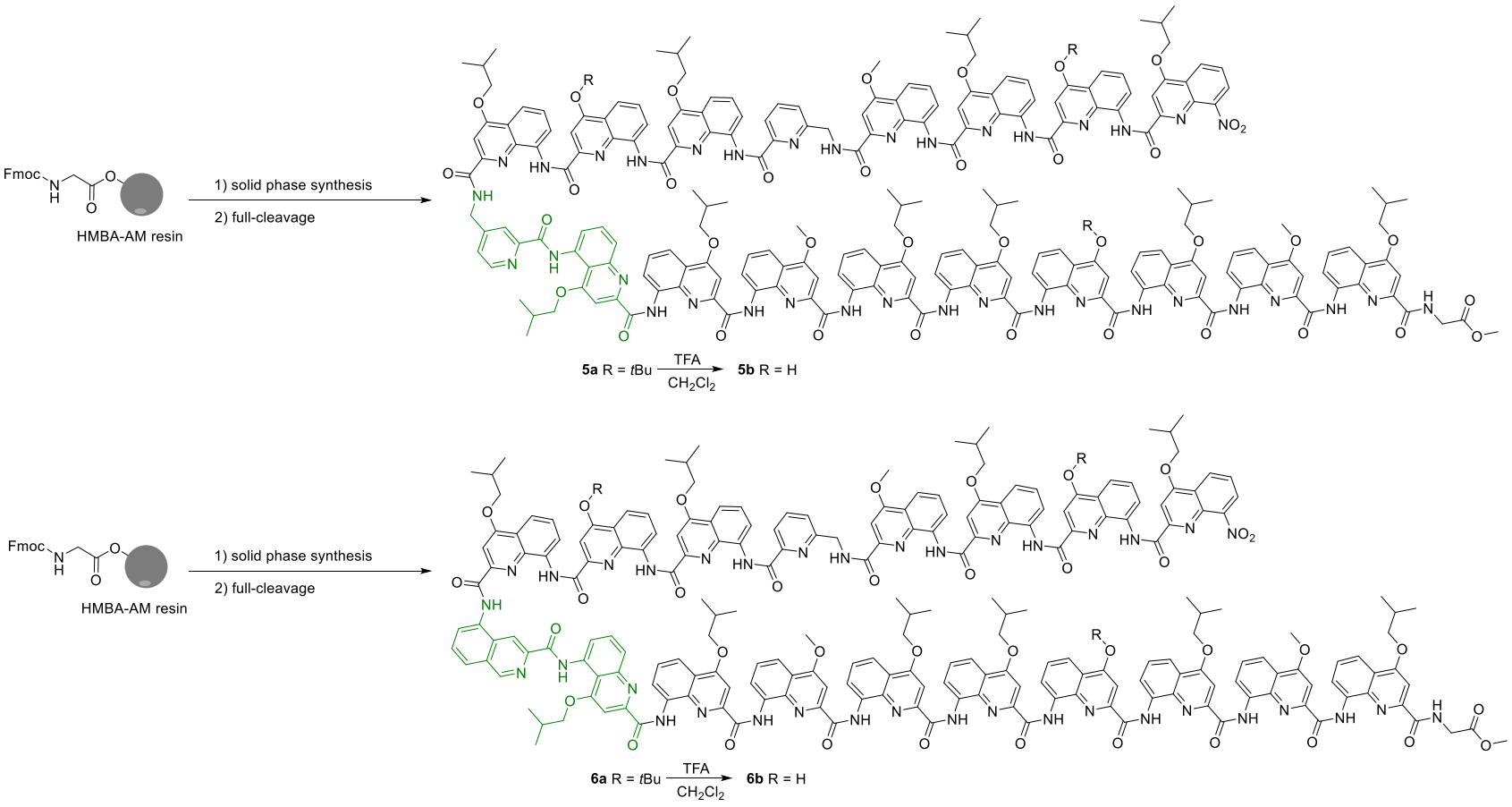


Scheme 1. Synthesis of linker units Fmoc-T6f-OH (**22**) and Fmoc-T6r-OH (**24**).

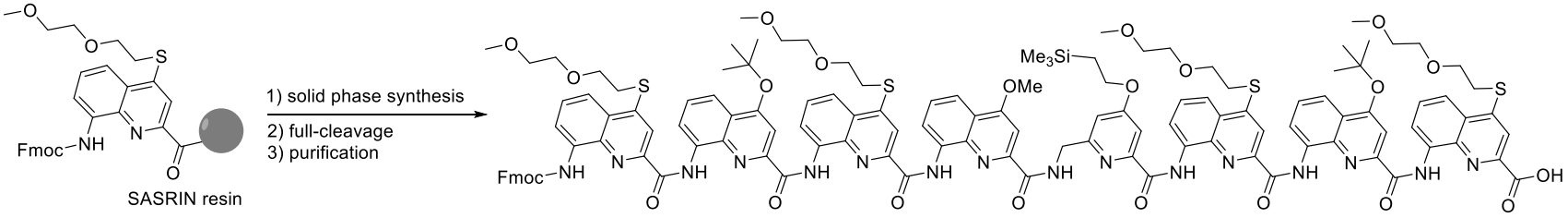
4.2 synthesis of foldamers

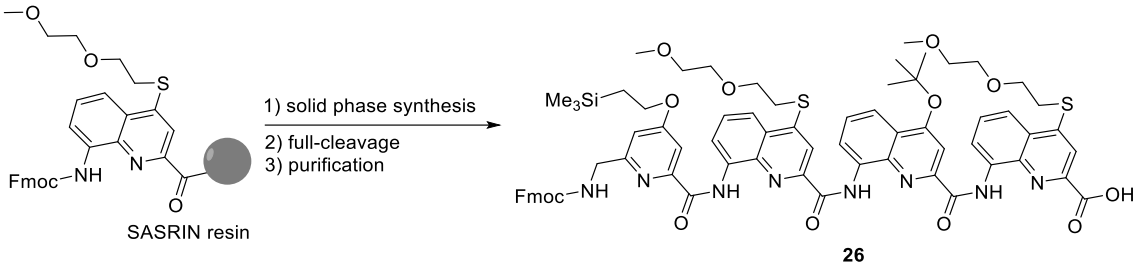


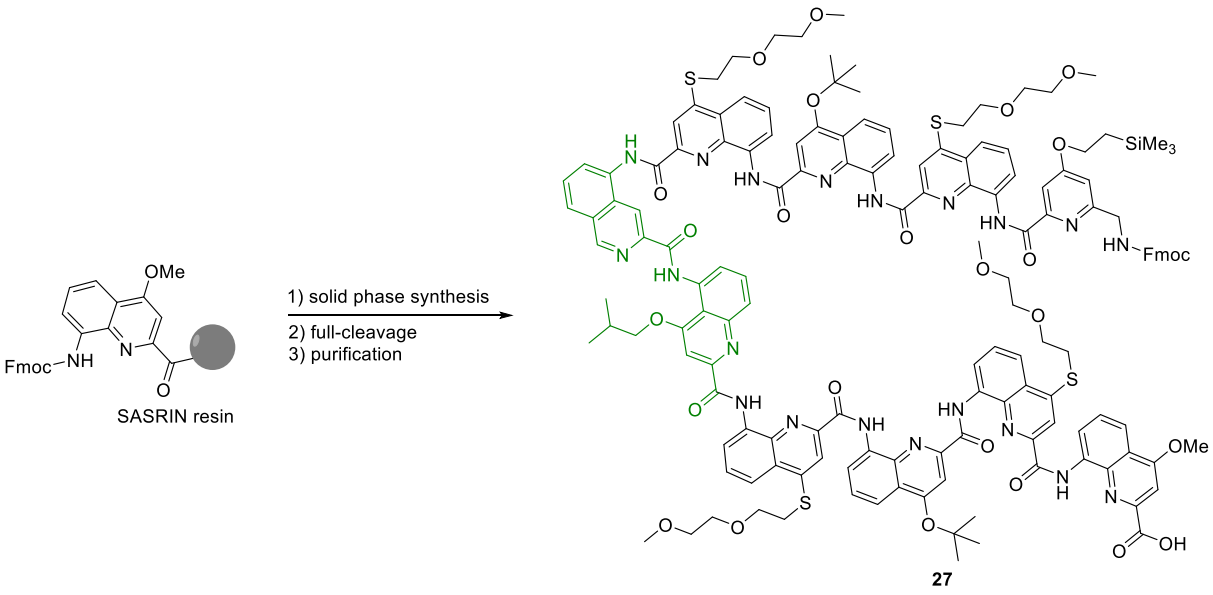
Scheme 2. Synthesis of **2a**, **2b**, **3a**, **3b**, **4a**, and **4b**.

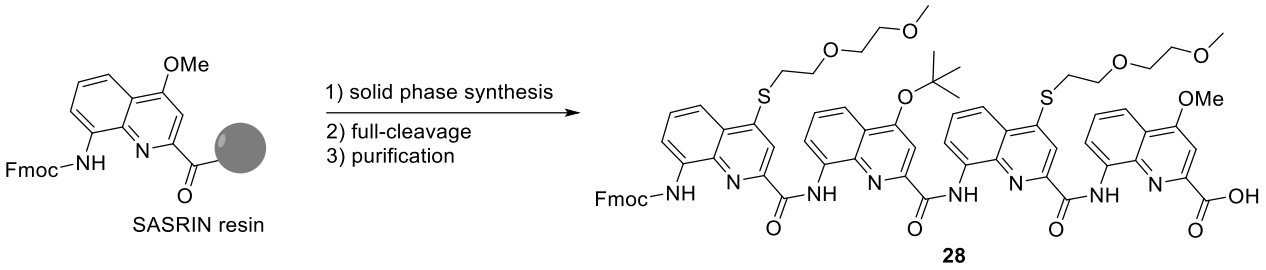


Scheme 3. Synthesis of **5a**, **5b**, **6a**, and **6b**.

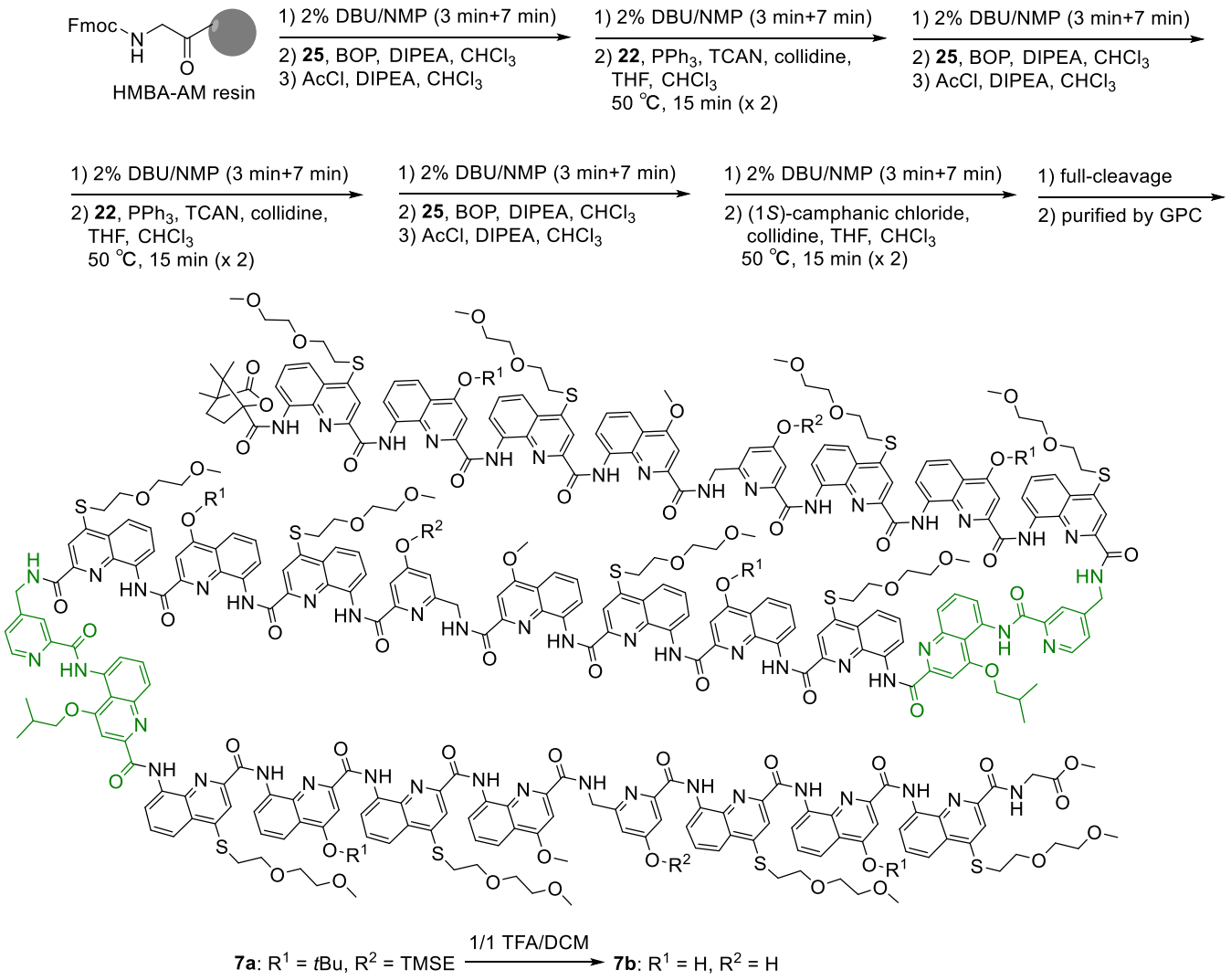

1) solid phase synthesis
2) full-cleavage
3) purification
SASRIN resin
25


1) solid phase synthesis
2) full-cleavage
3) purification
SASRIN resin
26

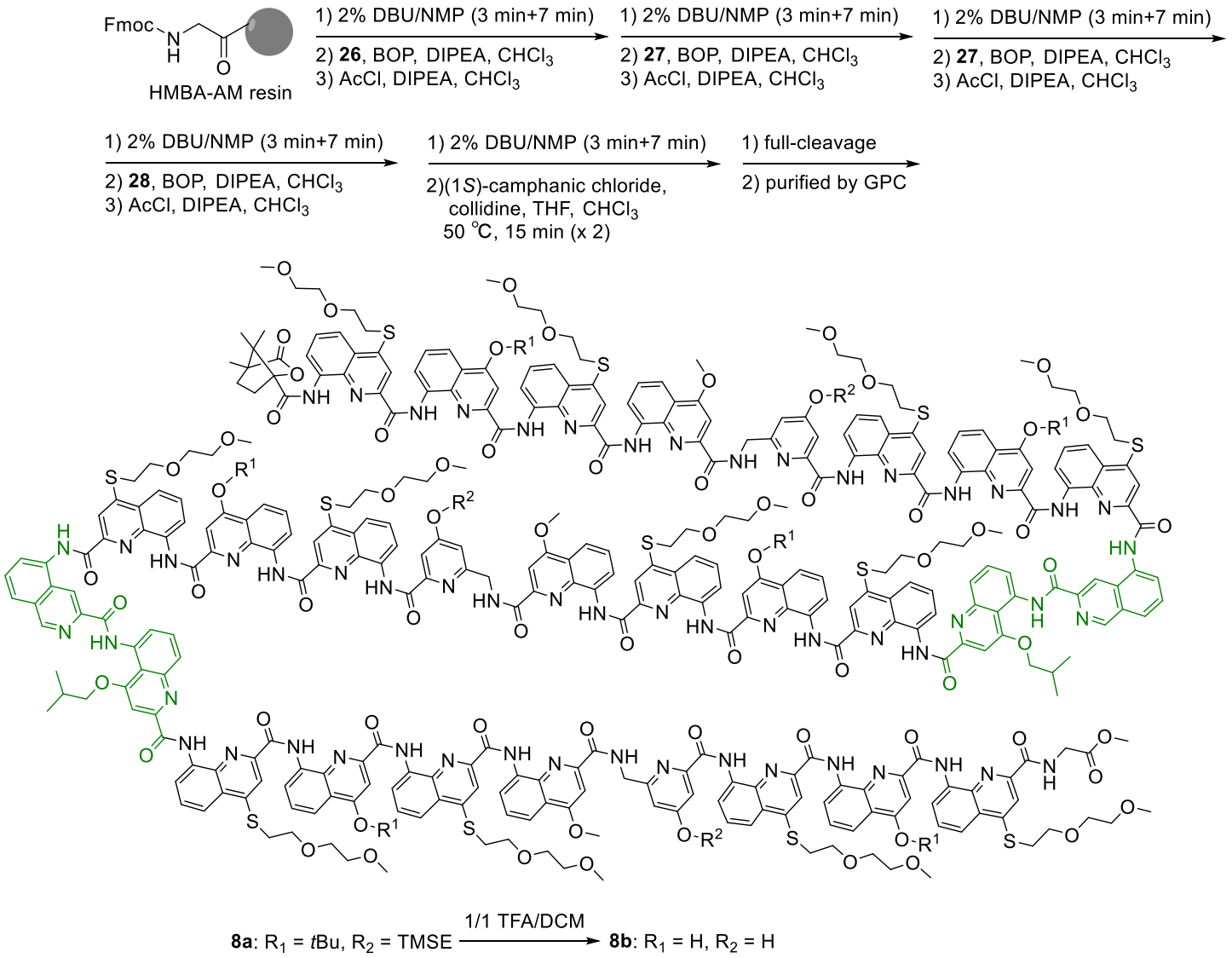

1) solid phase synthesis
2) full-cleavage
3) purification
SASRIN resin
27



Scheme 4. Synthesis of the fragments.



Scheme 5. Synthesis of **7a** and **7b** by fragment condensation strategy.



Scheme 6. Synthesis of **8a** and **8b** by fragment condensation strategy.

5 Synthetic Procedures

5.1 General methods

Commercially available reagents were purchased from Sigma-Aldrich, Alfa-Aesar or TCI and were used without further purification unless specified. HMBA-AM resin (200–400 mesh, loading 0.8–1.2 mmol/g) was purchased from Iris-biotech. SASRIN resin (200–400 mesh, 0.80-1.20 mmol/g) was purchased from Bachem. THF, DCM and toluene were dried over alumina columns (MBRAUN SPS-800 solvent purification system). *N*,*N*-diisopropylethylamine and chloroform were distilled over CaH_2 prior to use. Extra dry DMF was purchased from Sigma-Aldrich. Ultrapure water was obtained via a Stakpure OmniaPure-T UV-TOC ultrapure water system. Reactions were monitored by thin layer chromatography (TLC) on Merck silica gel 60-F254 plates and observed under UV light. Column chromatography purifications were carried out on Merck GEDURAN Si60 (40–63 μm).

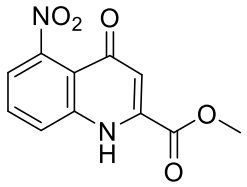
Analytical reversed-phase (RP) high-performance liquid chromatography (HPLC) was performed on a Thermo Fisher Scientific Ultimate 3000 HPLC System using Macherey-Nagel Nucleodur C8 Gravity columns (4 × 50 mm, 5 μm). UV absorbance was monitored at 300 nm and 254 nm, if not stated otherwise. The semi-preparative HPLC was performed on a Waters system equipped with a 2545 Quaternary Gradient Module with an automated fraction collector system on a XBridge® Prep C8 OBD™ column (19 × 150 mm, 5 μm) at a flow rate of 25 mL/min. 0.1 % TFA was added to the aqueous mobile phase (referred to as mobile phase A) and to acetonitrile (referred to as mobile phase B). The gradient is: 0–5 min, 90% to 100% solvent B; 5–25 min, 100% solvent B at r.t.. The column eluent was monitored by UV detection at 254 and 300 nm with a diode array detector. Preparative recycling gel permeation chromatography (GPC) was carried out on JAIGEL 20*600 mm columns (Japan Analytical Industry) in chloroform containing 1% ethanol and 0.25% trimethylamine as mobile phase, with a flow rate of 10 mL/min. Monitoring by UV detection was carried out at 214 nm, 254 nm, 300 nm and 400 nm.

The ultraviolet-visible (UV/Vis) absorbance measurements were done with a Thermo Fisher Scientific Nanodrop One instrument using a 1 cm path-length quartz cuvette. Circular dichroism (CD) spectra were measured on Jasco J-810 or Jasco J-1500 spectrometers. Measurements were performed at 20 °C if not stated otherwise.

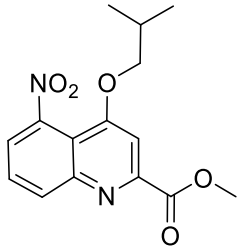
Solid phase synthesis (SPS) was performed manually under MW irradiation on a CEM Discover (Liberty Bio) microwave oven using a reaction vessel and an internal fiber optic probe for temperature control, or with a fully automated synthesizer followed by previously reported protocol.[8]

5.2 synthesis of monomers and the turn units.

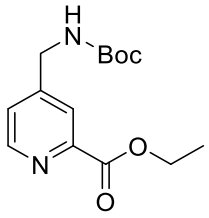
The Fmoc-<u>Y</u>-OH,[2] Fmoc-Q^M-OH,[9] Fmoc-Q^D-OH,[3] Fmoc-Q^B-OH,[10] Fmoc-<u>X</u>-OH[3] and Fmoc-P-OH[11] were synthesized according to literature. All of the Fmoc-protected monomers were≥ 98% pure before being used in the solid phase synthesis.



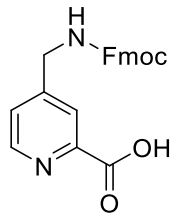
Compound 9. Diphenyl ether (200 mL) was heated to its boiling point, then dimethyl 2-((3-nitrophenyl)amino)maleate[12] (10 g, 35.7 mmol) was added. The reaction mixture was maintained at boiling for 20 min, after which it was allowed to cool to room temperature. Cyclohexane (200 mL) was added and the product was filtered off. The product was washed thoroughly with cyclohexane (100 mL × 3). DMF (100 mL) was added to the residue and the side product methyl 4-isobutoxy-7-nitroquinoline-2-carboxylate was precipitated and removed by filtration. The DMF solution was collected and concentrated in vacuo, yielding the desired compound as a brownish solid (3.62 g, 41% yield). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 12.53 (s, 1H), 8.14 (d, *J* = 8.6, 1H), 7.83 (dd, *J* = 8.6, 7.5 Hz, 1H), 7.59 (d, *J* = 7.5 Hz, 1H), 6.65 (s, 1H), 3.97 (s, 3H). **^{13}C NMR** (101 MHz, DMSO-d_6) δ 174.6, 162.1, 147.8, 140.9, 138.3, 132.7, 122.4, 118.1, 115.9, 111.5, 53.7. **HRMS** (ESI+) calcd. for $C_{11}H_8N_2O_5$ $[M+H]^+$ 249.0506, found 249.0505.



Compound 10. Compound **9** (1 g, 4.0 mmol, 1 equiv.) and PPh_3 (1.59 g, 6.0 mmol, 1.5 equiv.) were dissolved in dry toluene (15 mL) under N_2. Diisopropyl azodicarboxylate (1.23 mL, 6.0 mmol, 1.5 equiv.) was added dropwise to the solution. The mixture was stirred at 70 °C for 1 h and the formation of a white precipitate was observed. As isobutanol (0.6 mL, 6.0 mmol, 1.5 equiv.) was added the precipitation disappeared and the reaction mixture was stirred overnight at 70 °C. The solvent was removed and the crude product was purified by silica gel chromatography (EtOAc/cyclohexane 2:3). The product was crystallized from acetonitrile and obtained as a light-yellow solid (833 mg, 68% yield). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.37 (dd, *J* = 8.6, 1.0 Hz, 1H), 7.77 (dd, *J* = 8.6, 7.5 Hz, 1H), 7.66 (s, 1H), 7.64 (dd, *J* = 7.5, 1.0 Hz, 1H), 4.09 (s, 3H), 4.02 (d, *J* = 6.2 Hz, 2H), 2.19 (nonet, *J* = 6.6 Hz, 1H), 1.07 (d, *J* = 6.8 Hz, 6H). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 165.6, 161.5, 150.6, 149.1, 146.5, 133.5, 129.2, 122.1, 113.7, 103.1, 76.7, 53.7, 28.2, 19.3. **HRMS** (ESI+) calcd. for $C_{15}H_{16}N_2O_5$ $[M+H]^+$ 305.1132, found 305.1132.

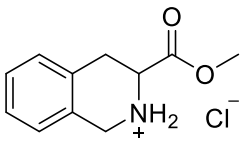


Compound 12. Ethyl 4-cyano-2-pyridinecarboxylate (300 mg, 1.7 mmol) and Boc_2O (930 mg, 4.3 mmol) were dissolved in dry MeOH (4 mL) under N_2 atmosphere. Then the Pd/C (30 mg) was added and the N_2 was replaced by H_2. The reaction mixture was stirred at r.t. overnight under H_2 atmosphere. The solution was then filtered and concentrated. The crude product was purified by silica gel chromatography (cyclohexane/EtOAc 2:3). The product was obtained as a white solid (320 mg, 67%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 8.70 (d, *J* = 4.9 Hz, 1H), 8.06–8.02 (m, 1H), 7.42–7.37 (m, 1H), 5.03 (s, 1H), 4.48 (q, *J* = 7.2 Hz, 2H), 4.43–4.38 (m, 2H), 1.47 (s, 9H), 1.45 (t, *J* = 7.2 Hz, 3H). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 165.2, 155.9, 150.0 149.9, 148.5, 125.1, 123.3, 80.3, 62.1, 43.4, 28.4, 14.4. **HRMS** (ESI+) calcd. for $C_{14}H_{20}N_2O_4$ $[M+H]^+$ 281.1496, found 281.1467.

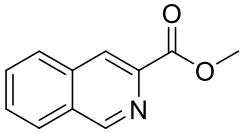


Compound 14. Compound **12** (320 mg, 1.1 mmol) was dissolved in THF/H_2O 5:1 (30 mL), then LiOH (55 mg, 2.2 mmol) was added. The reaction mixture was stirred at r.t. for 2 h, with completion indicated by TLC. The reaction mixture was diluted with water. The pH was adjusted to 4 by adding citric acid solution (5%, w/w). The product was extracted with DCM (3 x 50 mL). The organic layers were combined and dried over $MgSO_4$, filtered and concentrated. The product **13** was obtained as a solid in almost quantitative yield and used

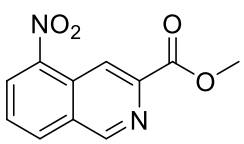
directly in the next step without further purification. Compound **13** (7.5 g, 29.7 mmol) was dissolved in DCM (37 mL). TFA (12 mL) was slowly added to the solution and stirred at r.t. until TLC indicated complete Boc deprotection. The solvent was removed under reduced pressure to provide the amino acid as a TFA salt. Then the amino acid was dissolved in a mixture of 1,4-dioxane (300 mL) and saturated $NaHCO_3$ solution (200 mL). The solution was cooled to 0 °C and a solution of Fmoc-OSu (11.0 g, 32.6 mmol, in 57 mL 1,4-dioxane) was added dropwise over 1 h. The reaction mixture was stirred at r.t. overnight. The resulting mixture was diluted with water and the pH was adjusted to 3 by dropwise addition of citric acid solution (5%, w/w). The precipitate was collected by filtration and washed with water. The product was obtained as a white solid. (6.8 g, 61%). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 8.63 (d, J = 5.0 Hz, 1H), 8.03 (t, J = 6.1 Hz, 1H), 7.97 (s, 1H), 7.89 (d, J = 7.5 Hz, 2H), 7.70 (d, J = 7.5 Hz, 2H), 7.46–7.39 (m, 3H), 7.33 (t, J = 7.5, 2H), 4.38 (d, J = 6.9 Hz, 2H), 4.30 (d, J = 6.3 Hz, 2H), 4.25 (t, J = 6.9 Hz, 1H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 166.2, 156.5, 150.4, 149.4, 148.5, 143.8, 140.8, 127.6, 127.1, 125.1, 125.0, 124.7, 122.7, 120.2, 65.5, 46.8, 42.7. **HRMS** (ESI+) calcd. for $C_{22}H_{18}N_2O_4$ $[M+H]^+$ 375.1339, found 375.1338.



Compound 15. (*RS*)-1,2,3,4-tetrahydroisoquinoline-3-carboxylic acid (5 g, 28.2 mmol) was dissolved in MeOH (200 mL) and the reaction solution was cooled to 0 °C. Thionyl chloride (10 mL, 51.3 mmol) was added dropwise. The reaction was stirred at r.t. overnight. The solvent was evaporated and excess thionyl chloride was removed in vacuo. The residue was washed with diethyl ether to afford the product as a white solid (5.4 g, 99%). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 10.22 (s, 2H), 7.26 (s, 4H), 4.57 (dd, J = 11.1, 5.1 Hz, 1H), 4.37–4.28 (m, 2H), 3.81 (s, 3H), 3.30 (dd, J = 16.9, 5.1 Hz, 1H), 3.17 (dd, J = 16.9, 11.1 Hz, 1H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 168.9, 130.5, 128.8, 128.4, 127.5, 126.9, 126.6, 53.1, 53.0, 43.8, 28.0. **HRMS** (ESI+) calcd. for $C_{11}H_{13}NO_2$ $[M+H]^+$ 192.1019, found 192.1018.

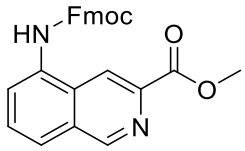


Compound 16. Compound **15** (12.7 g, 66.5 mmol) was dissolved in DMF (250 mL) and the solution was cooled to 0 °C. KMnO4 (7.35 g, 46.5 mmol) was added slowly and the reaction mixture was heated to r.t., then stirred for 48 h. The solvent was removed in vacuo and the resulting solid was dissolved in a mixture of 10% MeOH in DCM. The solution was filtered through a celite pad to remove insoluble material. The filtrate was then concentrated and purified by silica gel chromatography (acetone/DCM 1:30) to offer the product (5.5 g, 53%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 9.33 (s, 1H), 8.60 (s, 1H), 8.06 (m, 1H), 7.98 (m, 1H), 7.77 (m, 2H), 4.06 (s, 3H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 165.7, 152.7, 141.1, 134.9, 131.5, 130.0, 129.5, 128.1, 127.7, 123.6, 52.4. **HRMS** (ESI+) calcd. for $C_{11}H_9NO_2$ $[M+Na]^+$ 210.0525, found 210.0525.

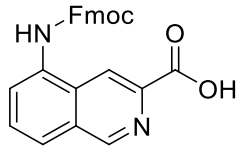


Compound 17. Compound **16** (4.72 g, 25.2 mmol) was dissolved in concentrated H_2SO_4 (47 mL) and the solution was cooled to 0 °C. $NaNO_3$ (2.36 g, 27.8 mmol) was added slowly while maintaining the temperature below 5 °C. The reaction mixture was stirred at r.t. for 2 h and then poured slowly into ice water. The pH was adjusted to 7 by adding saturated $NaHCO_3$ solution and the aqueous phase was extracted by a mixture of 10% MeOH in DCM (3 x 200 mL). The combined organic layers were collected, dried over $MgSO_4$, filtered and concentrated. The residue was washed with diethyl ether to afford the product as a yellow solid (4.67 g, 80%). **^{1}H NMR** (500 MHz, DMSO-d_6)

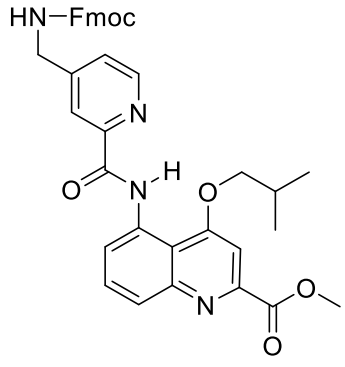
δ 9.65 (s, 1H), 9.11 (s, 1H), 8.78 (d, *J* = 7.6 Hz, 1H), 8.72 (d, *J* = 8.1 Hz, 1H), 8.06 (t, *J* = 7.9 Hz, 1H), 3.97 (s, 3H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 165.6, 154.2, 145.2, 144.3, 135.9, 130.3, 130.2, 129.6, 127.7, 118.7, 53.3. **HRMS** (ESI+) calcd. for $C_{11}H_8N_2O_4$ $[M+Na]^+$ 255.0376, found 255.0377.



Compound 19. Compound **17** (2.27 g, 9.8 mmol) was dissolved in dry MeOH (30 mL) and dry DMF (30 mL) under N_2. Then Pd/C (227 mg) was added and the N_2 was replaced by H_2. The reaction mixture was stirred overnight under H_2 atmosphere, then filtered and concentrated. The product **18** was obtained with a quantitative yield and used directly in the next step. Compound **18** (1.5 g, 7.4 mmol) was dissolved in 1,4-dioxane (200 mL) and $NaHCO_3$ solution (5%, w/w, 30 mL). The reaction mixture was cooled to 0 °C. A solution of Fmoc-Cl (2.11 g, 8.2 mmol) in 1,4-dioxane (50 mL) was added dropwise at 0 °C and the reaction mixture was stirred at r.t. overnight. The pH was then adjusted to 4 by adding 5% citric acid solution. The aqueous phase was extracted with DCM (3 x 100 mL). The combined organic phases were collected, dried over $MgSO_4$, filtered and concentrated. The residue was washed with diethyl ether to afford the product as a white solid. (1.45 g, 78%).**^{1}H NMR** (500 MHz, DMSO-d_6) δ 10.14 (s, 1H), 9.40 (s, 1H), 8.79 (s, 1H), 8.04 (d, *J* = 8.1 Hz, 1H), 7.92 (d, *J* = 7.5 Hz, 3H), 7.79 (m, 3H), 7.43 (t, *J* = 7.5 Hz, 2H), 7.35 (t, *J* = 7.5 Hz, 2H), 4.53 (d, *J* = 6.9 Hz, 2H), 4.37 (t, *J* = 6.9 Hz, 1H), 3.95 (s, 3H). **^{13}C NMR** z(126 MHz, DMSO-d_6) δ 165.7, 154.6, 152.9, 143.8, 140.9, 140.8, 134.4, 130.0, 129.9, 129.4, 127.8, 127.2, 125.4, 125.3, 124.3, 120.2, 118.9, 66.2, 52.5, 46.6. **HRMS** (ESI+) calcd. for $C_{26}H_{20}N_2O_4$ $[M+H]^+$ 425.1496, found 425.1492.

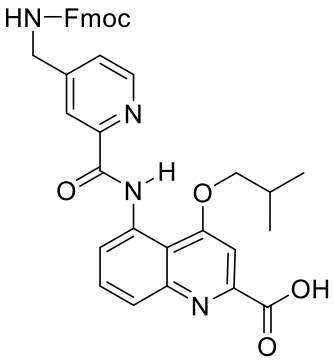


Compound 20. Compound **19** (2.3 g, 5.4 mmol) and LiI (3.6 g, 26.8 mmol) were suspended in degassed EtOAc (40 mL) under N_2. The reaction mixture was refluxed overnight. The solvent was removed in vacuo and the remaining solid was washed with aqueous citric acid solution (5%, w/w), then with aqueous $Na_2S_2O_3$ solution (5%, w/w) and finally with water. The product was obtained after drying in vacuo (1.75 g, 79%). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 13.10 (bs, 1H), 10.12 (s, 1H), 9.40 (s, 1H), 8.80 (s, 1H), 8.04 (d, *J* = 8.1 Hz, 1H), 7.95-7.87 (m, 3H), 7.82-7.75 (m, 3H), 7.43 (t, *J* = 7.4 Hz, 2H), 7.35 (t, *J* = 7.4 Hz, 2H), 4.51 (d, *J* = 7.0 Hz, 2H), 4.36 (t, *J* = 7.0 Hz, 1H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 166.7, 154.6, 152.6, 143.8, 141.9, 140.8, 134.4, 129.9, 129.7, 129.6, 127.8, 127.2, 125.3, 124.3, 120.3, 118.6, 66.3, 48.6, 46.6. **HRMS** (ESI+) calcd. for $C_{25}H_{18}N_2O_4$ $[M+H]^+$ 411.1339, found 411.1337.

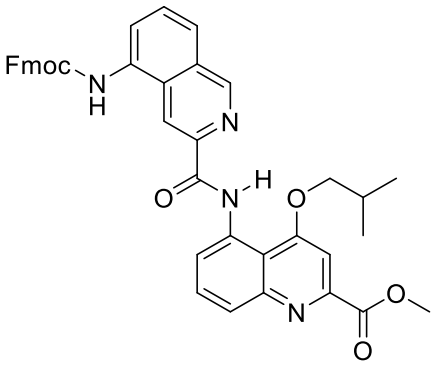


Compound 21. Compound **10** (1 g, 2.35 mmol) was dissolved in dry MeOH (7 mL) under N_2 atmosphere. Then Pd/C (38 mg) was added and N_2 was replaced by H_2. The reaction was stirred at r.t. overnight under H_2 atmosphere. After completion of the reaction, which was confirmed by TLC, the reaction mixture was filtered and concentrated. Compound **11** was obtained quantitatively and used directly in the next step. Compound **11** (506 mg, 1.85 mmol), compound **14** (574 mg, 1.53 mmol) and PyBOP (1.6 g, 3.07 mmol) were dissolved in dry $CHCl_3$ under N_2. DIPEA (1.1 mL, 6.34 mmol) was added and the reaction mixture was stirred at r.t. for 48 h. The solution was diluted with $CHCl_3$, and washed with citric acid solution (5%, w/w), then with $NaHCO_3$ solution (5%, w/w). The organic

phases were combined and concentrated. The product was precipitated from diethyl ether as a pale-yellow solid (830 mg, 86%). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 12.36 (s, 1H), 8.97 (dd, *J* = 5.9, 3.1 Hz, 1H), 8.67 (d, *J* = 4.9 Hz, 1H), 8.16 (s, 1H), 8.11 (t, *J* = 6.1 Hz, 1H), 7.93–7.83 (m, 4H), 7.71 (d, *J* = 7.5 Hz, 2H), 7.63 (d, *J* = 3.1 Hz, 1H), 7.52 (d, *J* = 4.6 Hz, 1H), 7.42 (t, *J* = 7.4 Hz, 2H), 7.34 (t, *J* = 7.4 Hz, 2H), 4.40 (d, *J* = 6.8 Hz, 2H), 4.36 (d, *J* = 6.0 Hz, 2H), 4.33–4.29 (m, 2H), 4.26 (t, *J* = 6.8 Hz, 1H), 3.96 (s, 3H), 2.58 (nonet, *J* = 6.8 Hz, 1H), 1.00 (d, *J* = 6.6 Hz, 6H). (Mixture of two conformers with a ratio of 1:0.15, only the data of the major species is listed here). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 165.3, 163.4, 162.3, 156.5, 151.5, 149.6, 149.5, 148.7, 148.3, 143.9, 140.8, 134.5, 130.8, 127.7, 127.1, 125.7, 125.4, 125.2, 120.5, 120.2, 118.2, 113.1, 102.1, 76.6, 65.6, 52.8, 46.8, 42.9, 26.9, 19.1. **HRMS** (ESI+) calcd. for $C_{37}H_{34}N_4O_6$ $[M+H]^+$ 631.2557, found 631.2542.

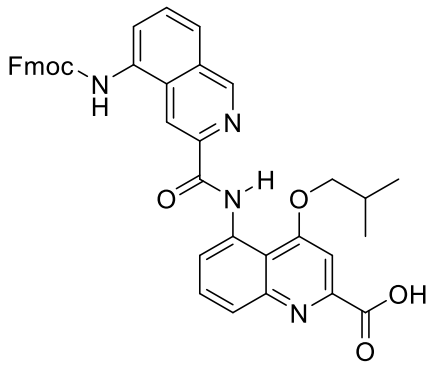


Compound 22. Compound **21** (1.08 g, 1.71 mmol) was suspended in degassed EtOAc under N_2. Then LiI (1.15 g, 8.60 mmol) was added and the mixture was refluxed overnight. The solvent was removed in vacuo and the remaining residue was washed with 5% aqueous citric acid solution, then 5% aqueous $Na_2S_2O_3$ solution and water. The product was obtained as a pale-yellow solid after drying in vacuo (800 mg, 76%). **^{1}H NMR** (500 MHz, DMSO-d_6) δ 12.42 (s, 1H), 11.57 (s, 1H), 8.70–8.65 (m, 2H), 8.10–8.07 (m, 2H), 7.90 (d, *J* = 7.5 Hz, 2H), 7.76–7.69 (m, 4H), 7.57 (t, *J* = 8.2 Hz, 1H), 7.45–7.38 (m, 2H), 7.34 (td, *J* = 7.4, 1.3 Hz, 2H), 6.52 (s, 1H), 4.41–4.24 (m, 5H), 4.20–4.14 (m, 2H), 2.58 (nonet, *J* = 6.8, 1H), 1.03–1.01 (m, 6H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 164.1, 162.3, 156.5, 151.5, 149.5, 148.2, 148.1, 143.9, 140.8, 139.5, 137.5, 134.4, 129.0, 127.7, 127.3, 127.1, 125.2, 121.4, 120.2, 120.1, 112.6, 109.9, 101.5, 76.6, 65.6, 46.8, 42.9, 26.8, 19.0. **HRMS** (ESI+) calcd. for $C_{36}H_{32}N_4O_6$ $[M+H]^+$ 617.2395, found 617.2386.



Compound 23. Compound **10** (1 g, 2.35 mmol) was dissolved in dry MeOH (7 mL) under N_2 atmosphere. Then Pd/C (38 mg) was added and N_2 was replaced by H_2. The reaction was stirred at r.t. overnight under H_2 atmosphere. After completion of the reaction, which was confirmed by TLC, the reaction mixture was filtered and concentrated. Compound **11** was obtained quantitatively and used directly in the next step. Compound **20** (1.3 g, 3.2 mmol) was dissolved in thionyl chloride (13 mL) under N_2. DMF (0.1 mL) was added and the reaction mixture was stirred at r.t. for 2 h. Thionyl chloride and DMF were removed in vacuo overnight to afford the desired acyl chloride. To compound **11** (690 mg, 2.5 mmol) and DIPEA (1.3 mL, 13.6 mmol) in dry $CHCl_3$ (20 mL) was added dropwise a solution of previously prepared acyl chloride of **20** in dry $CHCl_3$ (20 mL) under N_2. The reaction mixture was stirred overnight at r.t. and the solvent was then removed in vacuo. The resulting crude was purified by silica gel chromatography (cyclohexane/EtOAc 1:1). The product was recrystallized from acetonitrile/DCM to obtain the product as a pale yellow solid (800 mg, 48%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 12.56 (s, 1H), 9.20 (s, 1H), 9.11 (d, *J* = 7.7 Hz, 1H), 8.88 (s, 1H), 8.33 (bs, 1H), 8.04 (d, *J* = 8.8 Hz, 1H), 7.88 (d, *J* = 8.1 Hz, 1H), 7.84–7.78 (m, 3H), 7.75 (t, *J* = 7.9 Hz, 1H), 7.69-7.63 (m, 3H), 7.53 (bs,

1H), 7.42 (t, *J* = 7.4 Hz, 2H), 7.34 (t, *J* = 7.4 Hz, 2H), 4.63 (d, *J* = 6.5 Hz, 2H), 4.33 (t, *J* = 6.5 Hz, 1H), 4.27 (d, *J* = 7.1 Hz, 2H), 4.10 (s, 3H), 2.76 (nonet, *J* = 7.0 Hz, 1H), 1.12 (d, *J* = 6.7 Hz, 6H). **^{13}C NMR** (126 MHz, $CDCl_3$) δ 166.1, 164.0, 163.5, 154.0, 151.3, 150.4, 148.6, 144.3, 143.7, 141.5, 134.9, 133.9, 130.8, 130.3, 129.7, 128.0, 127.3, 126.7, 125.1, 123.7, 120.2, 119.5, 114.6, 114.2, 102.0, 67.8, 53.5, 47.2, 27.5, 19.7. **HRMS** (ESI+) calcd. for $C_{40}H_{34}N_4O_6$ $[M+H]^+$ 667.2551, found 667.2539.



Compound 24. Compound **23** (600 mg, 0.9 mmol) and LiI (1.2 g, 9.0 mmol) were suspended in degassed EtOAc (50 mL) under N_2 and the reaction mixture was refluxed overnight. The solvent was removed in vacuo. The remaining residue was washed with citric acid solution (5%, w/w), $Na_2S_2O_3$ solution (5%, w/w) and water. The product was obtained as a pale yellow solid after drying in vacuo (453 mg, 77%).**^{1}H NMR** (500 MHz, DMSO-d_6) δ 12.57 (s, 1H), 10.22 (s, 1H), 9.47 (s, 1H), 9.04 (dd, *J* = 6.1, 3.0 Hz, 1H), 8.95 (s, 1H), 8.16 (d, *J* = 8.3 Hz, 1H), 7.95–7.78 (m, 8H), 7.65 (s, 1H), 7.44 (t, *J* = 7.4 Hz, 2H), 7.37 (t, *J* = 7.4 Hz, 2H), 4.54 (d, *J* = 7.0 Hz, 2H), 4.37 (dd, *J* = 12.4, 7.0 Hz, 3H), 2.65 (nonet, *J* = 6.8 Hz, 1H), 1.04 (d, *J* = 6.6 Hz, 6H). **^{13}C NMR** (126 MHz, DMSO-d_6) δ 166.1, 163.5, 162.7, 154.7, 151.6, 149.9, 149.3, 143.8, 143.0, 140.8, 134.7, 134.5, 130.7, 130.2, 130.1, 129.5, 127.8, 127.2, 125.3, 125.3, 124.8, 120.2, 118.1, 116.3, 113.0, 102.0, 76.5, 66.3, 46.6, 26.8, 19.1. **HRMS** (ESI+) calcd. for $C_{39}H_{32}N_4O_6$ $[M+H]^+$ 653.2395, found 653.2385.

5.3 Solid phase synthesis general methods

5.3.1 Loading of the resin via HBTU activation

SASRIN resin (500 mg, 0.4–0.6 mmol, 1 eq.) was swollen in 5 mL DCM for 1 h, transferred to the microwave vessel and washed 3 times with extra dry DMF. DIPEA (170 µL, 1.0 mmol, 2 eq.) was added to a mixture of Fmoc-$Q^{B/D/M}$-OH (0.45 mmol, 0.9 eq.) and HBTU (228 mg, 1.2 eq.) in extra dry DMF (5 mL). The resulting solution was shaken for 30 s before to be poured to the resin-containing reaction vessel. The reaction mixture was subjected to treatment in a microwave oven (50 °C, 20 min, 25 W). The resin was filtered and washed with DMF (5 x 2 mL) and DCM (10 x 2 mL). Capping was performed by adding a mixture of DCM/pyridine/benzoyl chloride (3:1:1, v/v/v, 5 mL) to the resin followed by shaking for 30 min at r.t., and subsequent washing with DCM (20 x 2 mL). To monitor the efficiency of the 1st loading, a small amount of resin (around 2 mg) was taken and dried in vacuo. The loading was estimated at this scale.

HMBA-AM resin (500 mg, 0.8–1.2 mmol, 1 eq.) was swollen in 5 mL DCM for 1 h, transferred to the microwave vessel and washed 3 times with extra dry DMF. DIPEA (170 µL, 1.0 mmol, 2 eq.) was added to a mixture of Fmoc-Gly-OH (134 mg, 0.45 mmol, 0.9 eq.) and HBTU (228 mg, 1.2 eq.) in extra dry DMF (5 mL) and the resulting solution was shaken for 30 s before being poured into the resin-containing reaction vessel. The reaction mixture was subjected to a treatment in a microwave oven (50 °C, 20 min, 25 W). The resin was filtered and washed with DMF (5 x 2 mL) and DCM (10 x 2 mL). Capping was performed by adding a mixture of DCM/pyridine/benzoyl chloride (3:1:1, v/v/v, 5 mL) to the resin followed by shaking for 30 min at r.t., and subsequent washing with DCM (20 x 2 mL). To monitor the efficiency of the 1st loading, a small amount of resin (around 2 mg) was taken and dried in vacuo. The loading was estimated at this scale.

5.3.2 Estimation of the loading

To a small amount of Fmoc-Gly-HMBA AM resin or Fmoc- $Q^{B/D/M}$-SASRIN resin (1–2 mg), a freshly prepared solution of DMF/piperidine (8:2 (v/v), 3.0 mL) was added. The mixture was shaken and incubated for 5 min. Then the absorption was measured at 301 nm using a NanoDrop One Microvolume UV-Vis Spectrophotometer and a Hellma quartz glass cuvette 104 (path length 10 mm). Three replicates were measured, then the loading was calculated with the following equation:

$$\text{loading}\left(\text{in}\ \frac{\text{mmol}}{\text{g}}\right) = \frac{\text{Abs}_{301\ \text{nm}} \times \text{V}}{\varepsilon_{301\text{nm}} \times l \times \text{m}}$$

$$\varepsilon_{301\text{nm}} = 7800\ \text{L/mol/cm}$$ [13]

5.3.3 Solid phase synthesis of monomer and turn units via *in-situ* activation

For manual solid phase synthesis under MW-irradiation: The Fmoc- $Q^{B/D/M}$-SASRIN resin was swallowed in DCM (5 mL) for 2 h, and then transferred into the reaction vessel followed by washing with NMP (3 x 3 mL). The deprotection of the Fmoc group was performed by using a solution of 2% DBU in NMP (3 mL, 3 min + 7 min). The resin was next filtered off and washed with DCM (3 x 2 mL) and then with anhydrous THF (5 x 2 mL). The Fmoc deprotection step was performed before each aromatic monomer coupling. The resin was then suspended in anhydrous THF (1 mL) and 2,3,5-collidine (5 eq. with respect to the resin-loading) was added to the resin supernatant. The Fmoc-protected monomer or turn unit (2 eq. with respect to the resin-loading) and PPh_3 (4 eq. with respects to the resin-loading) were successively added in a vial to be solubilized in freshly distilled $CHCl_3$ (1 mL). Trichloroacetonitrile (4.5 eq. with respect to the resin loading) was then added to the vial and the resulting acid chloride solution was shaken for 30 s before to be poured to the resin-containing reaction vessel. The reaction vessel was then placed in the microwave oven and subjected to MW irradiation for 15 min (50 °C, 50 W). The resin was then washed 3 times with anhydrous THF. This entire coupling step was then repeated once more. For the final coupling of (1*S*)-camphanic acid, the resin was suspended in anhydrous THF (1 mL) and 2,3,5-collidine (5 eq. with respect to the resin-loading) was added to the resin suspensions. A solution of (1*S*)-camphanyl chloride (2 eq. with respect to the resin-loading, purchased from Sigma-Aldrich, 98%, ee: 99%) in freshly distilled $CHCl_3$ (1 mL) was added to the supernatant and the resin was shaken at r.t. for 2 h. The resin was filtered off, and washed 3 times with dry THF, and the same process was repeated once. After coupling, the resin was vigorously washed 3 times with DMF and 3 times with DCM.

The fully automatic solid phase synthesis on Chorus PurePep ® synthesizer followed previously reported procedures.[8]

5.3.4 Fragment condensation via BOP activation

The corresponding resin (cal. 100 mg) was swallowed in DCM (5 mL) for 2 h. The deprotection of the Fmoc group was performed by using a solution of 2% DBU/NMP (3 mL, 3 min + 7 min). The resin was next filtered off and washed with DCM (3 x 2 mL) and then with anhydrous THF (5 x 2 mL). DIPEA (2 eq. with respects to the resin-loading) was added to a mixture of Fmoc-protected fragment sequences (1 eq. with respects

to the resin-loading) and BOP (1.5 eq. with respects to the resin-loading) in anhydrous chloroform (5 mL) under N_2. After activating for 1 min, the solution was added to the resin. The supernatant was shaken at r.t. overnight. The resin was filtered and washed with anhydrous chloroform (3 x 3 mL). To perform a capping step, the resin was suspended in anhydrous chloroform (3 mL) and DIPEA (4 eq. with respects to the resin-loading) was added. Acetyl chloride (2 eq. with respects to the resin-loading) was added to the supernatant and the resin was shaken at r.t. for 30 min. The resin was filtered and washed with DCM (3 x 2 mL), 20% DIPEA/NMP (3 x 2 mL) and DCM (10 x 2 mL).

5.3.5 Mini-cleavage

For SASRIN resin: To perform a mini cleavage, the resin (1–2 mg) was swollen in 1 mL TFA/DCM (1:99, v/v) solution and incubated at r.t. for 10 min. The DCM solution was washed with aqueous $NaHCO_3$ solution and the solvent was removed under reduced pressure.

For HMBA AM resin: To perform a mini cleavage, the resin (1–2 mg) was swollen in 1 mL MeOH/DCM (1:1, v/v) solution followed by the addition of 10 µL NaOMe (25% (m/m)) and incubated at r.t. for 10 min. The cleavage solution was diluted with DCM, washed with aqueous citric acid solution (5%), dried over $MgSO_4$, filtered and the solvent was finally removed under reduced pressure.

5.3.6 Full cleavage[14,15]

For SASRIN resin: A mixture of 1% TFA/DCM (2 mL) was added to the resin and stirred at r.t. for 1 min (x 10). The solution was directly quenched by sat. $NaHCO_3$ solution. The organic phase was collected, dried over $MgSO_4$, filtered and concentrated to afford the crude product as a yellow solid.

For HMBA AM resin: The resin (around 100 mg) was dried in vacuo and slowly added to 400 mL cleavage solution (preparation see below) under N_2 atmosphere. The mixture was stirred under N_2 atmosphere for 2 h before it was added to 100 mL of aqueous citric acid solution (5%). The aqueous layer was extracted with DCM (3 x 50 mL). The combined organic phases were washed with brine, dried over $MgSO_4$, filtered and the solvent was evaporated under reduced pressure. The crude product was obtained as a solid.

Preparation of the cleavage solution: 200 mL dry MeOH was added to 200 mL dry DCM under N_2 atmosphere. 2 mL NaOMe (25% (m/m)) in methanol were added and the mixture was mixed well by magnetic stirring. A sufficient amount of cleavage solution (at least 400 mL cleavage solution per 100 mg resin) was important to avoid the formation of oligomeric acid as the by-product.

5.4 Synthesis of oligomers

O_2N-Q^B<u>X</u>Q^B-T6f-$Q^BQ^MQ^B$-OMe (**2a**) Compound **2a** was synthesized on SASRIN resin (scale: 15 µmol) using SPFS procedure reported in paragraph 5.3. The crude product obtained after full-cleavage was dried in vacuo and dissolved in dry chloroform/MeOH 3:2 (3 mL) under N_2. To this solution $TMSCHN_2$ (2 M in *n*-hexane, 17.7 µL, 60 µmol) was added dropwise. The reaction mixture was stirred at r.t. for 2 h. A few drops of acetic acid were then added and stirring continued for 30 min at r.t. to quench the excess amounts of $TMSCHN_2$. Then the solution was diluted with DCM, washed with aqueous $NaHCO_3$ solution, dried over $MgSO_4$, filtered and concentrated. The crude product was purified by RP-HPLC. Compound **2a** was obtained as a yellow solid (7.2 mg, 26%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 12.58 (s, 1H), 12.33 (s, 1H), 12.15 (s, 1H), 12.03 (s, 1H), 12.01 (s, 1H), 12.00 (s, 1H), 9.10 (dd, *J* = 7.6, 1.4 Hz, 1H), 8.96 (dd, *J* = 7.6, 1.3 Hz, 1H), 8.93 (dd, *J* = 7.7, 1.3 Hz, 1H), 8.69 (dd, *J* = 7.9, 1.2

Hz, 1H), 8.50 (dd, *J* = 8.3, 1.5 Hz, 1H), 8.31 (dd, *J* = 7.5, 1.3 Hz, 1H), 8.17–8.16 (m, 1H), 8.08 (dd, *J* = 7.8, 1.3 Hz, 2H), 8.03–7.96 (m, 7H), 7.87 (dd, *J* = 8.3, 1.3 Hz, 1H), 7.79 (s, 1H), 7.78–7.71 (m, 2H), 7.66–7.61 (m, 2H), 7.59–7.55 (m, 1H), 7.44 (s, 1H), 7.39 (dd, *J* = 8.4, 7.5 Hz, 1H), 7.29–7.25 (m, 1H), 7.15–7.13 (m, 1H), 6.94 (s, 1H), 6.90 (s, 1H), 6.68 (s, 1H), 6.59 (t, *J* = 8.1 Hz, 1H), 6.35 (dd, *J* = 8.3, 1.2 Hz, 1H), 4.37–4.19 (m, 8H), 3.98 (s, 3H), 3.89 (d, *J* = 6.5 Hz, 2H), 3.86 (d, *J* = 6.1 Hz, 2H), 3.52 (s, 3H), 2.69 (nonet, *J* = 6.9 Hz, 1H), 2.43 (m, 2H), 2.30 (tt, *J* = 13.4, 6.6 Hz, 2H), 1.70 (s, 9H), 1.35–1.10 (m, 30H). **HRMS** (ESI+) calcd. for $C_{103}H_{100}N_{16}O_{18}$ $[M+Na]^+$, 1871.7294, found 1871.7320.

O_2N-Q^{B}XQ^B-T6f-Q^{B}Q^{M}Q^B-OMe (**2b**) Compound **2a** (7.2 mg, 3.8 μmol) was treated with TFA/DCM 1:1 (3 mL) at r.t. for 2 h. The solvent was removed in vacuo. The residue was washed with diethyl ether, yielding the product as a yellow solid (7.0 mg, quant.). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 12.62 (s, 1H), 12.59 (s, 1H), 11.96 (s, 1H), 11.90 (s, 2H), 11.79 (s, 1H), 10.24 (s, 1H), 9.19 (d, *J* = 7.5 Hz, 1H), 8.80 (d, *J* = 7.5 Hz, 1H), 8.68 (d, *J* = 6.9 Hz, 1H), 8.64 (d, *J* = 7.9 Hz, 1H), 8.50 (dd, *J* = 8.4, 1.5 Hz, 1H), 8.30 (d, *J* = 7.5 Hz, 2H), 8.19–8.04 (m, 5H), 7.99–7.92 (m, 4H), 7.85–7.80 (m, 4H), 7.75 (q, *J* = 7.6 Hz, 2H), 7.65–7.60 (m, 3H), 7.47 (d, *J* = 1.5 Hz, 1H), 7.38 (t, *J* = 7.9 Hz, 1H), 7.33 (t, *J* = 7.9 Hz, 1H), 6.90 (s, 1H), 6.69 (s, 1H), 6.61 (t, *J* = 8.1 Hz, 1H), 6.37 (d, *J* = 7.0 Hz, 1H), 4.77–4.62 (m, 2H), 4.44 (s, 3H), 4.38–4.33 (m, 1H), 4.23 (m, 4H), 4.00–3.85 (m, 5H), 3.58 (s, 3H), 2.50–2.29 (m, 5H), 1.36–1.13 (m, 30H). **HRMS** (ESI+) calcd. for $C_{99}H_{92}N_{16}O_{18}$ $[M+H]^+$ 1793.6848, found 1793.6892.

O_2N-Q^B<u>X</u>Q^B-T6r-Q^{B}Q^{M}Q^B-OMe (**3a**) Compound **3a** was synthesized on SASRIN resin (scale: 15 μmol) using SPFS procedure reported in paragraph 5.3. The crude after full-cleavage was dried in vacuo and dissolved in dry chloroform/MeOH 3:2 (3 mL) under N_2. $TMSCHN_2$ (2 M in *n*-hexane, 17.7 μL, 60 μmol) was added dropwise. The solution was stirred at r.t. for 2 h. A few drops of acetic acid were added and stirred at r.t. for 30 min to quench the excess $TMSCHN_2$. Then the solution was diluted with DCM, washed with aqueous $NaHCO_3$ solution. dried over $MgSO_4$, filtered and concentrated. The crude was purified by RP-HPLC. The compound **3a** was obtained as a yellow solid (13.2 mg, 47%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 12.61 (s, 1H), 12.12 (s, 1H), 12.09 (s, 1H), 12.04 (s, 1H), 12.03 (s, 1H), 11.75 (s, 1H), 10.73 (s, 1H), 9.24 (dd, *J* = 7.6, 1.3 Hz, 1H), 9.14–9.12 (m, 2H), 8.56 (dd, *J* = 8.3, 1.5 Hz, 1H), 8.40 (dd, *J* = 7.9, 1.1 Hz, 2H), 8.37 (s, 1H), 8.33 (dd, *J* = 7.5, 1.4 Hz, 1H), 8.15 (dd, *J* = 7.6, 1.3 Hz, 1H), 8.11–8.06 (m, 2H), 8.05–8.00 (m, 3H), 7.93 (s, 1H), 7.91 (dd, *J* = 8.3, 1.3 Hz, 1H), 7.83–7.77 (m, 3H), 7.74 (t, *J* = 7.8 Hz, 1H), 7.67 (dd, *J* = 7.4, 1.5 Hz, 1H), 7.62–7.59 (m, 1H), 7.55 (s, 1H), 7.47–7.43 (m, 1H), 7.35–7.29 (m, 3H), 7.16–7.12 (m, 2H), 6.80 (s, 1H), 6.70 (s, 1H), 6.59 (t, *J* = 8.0 Hz, 1H), 6.38 (dd, *J* = 8.3, 1.1 Hz, 1H), 4.49 (d, *J* = 6.0 Hz, 2H), 4.27–4.14 (m, 4H), 3.99 (d, *J* = 6.1 Hz, 2H), 3.87 (d, *J* = 6.0 Hz, 2H), 3.80 (s, 3H), 3.53 (s, 3H), 3.03–2.94 (m, 1H), 2.51–2.41 (m, 2H), 2.39–2.31 (m, 2H), 1.45 (d, *J* = 6.6 Hz, 6H), 1.29 (d, *J* = 6.5 Hz, 6H), 1.26 (d, *J* = 6.8 Hz, 12H), 1.21 (d, *J* = 6.8 Hz, 6H), 1.10 (s, 9H). **HRMS** (ESI+) calcd. for $C_{106}H_{100}N_{16}O_{18}$ $[M+Na]^+$ 1907.7294, found (HR-ESI) 1907.7255.

O_2N-Q^{B}XQ^B-T6r-Q^{B}Q^{M}Q^B-OMe (**3b**) Compound **3a** (13.2 mg, 7.0 μmol) was treated with a TFA/DCM 1:1 (3 mL) at r.t. for 2 h. The solvent was removed in vacuo. The residue was washed with diethyl ether, yielding the products as a yellow solid (12.8 mg, quant.). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 12.60 (s, 1H), 12.57 (s, 1H), 12.30 (s, 1H), 11.83 (s, 1H), 11.75 (s, 1H), 11.66 (s, 1H), 11.37 (s, 1H), 10.15 (s, 1H), 9.22 (d, *J* = 7.4 Hz, 1H), 9.01 (d, *J* = 7.4 Hz, 1H), 8.75 (s, 1H), 8.58 (dd, *J* = 8.3, 1.5 Hz, 1H), 8.47 (d, *J* = 7.9 Hz, 2H), 8.40–8.35 (m, 2H), 8.22–8.15 (m, 2H), 8.12–8.07 (m, 2H), 8.02–7.98 (m, 2H), 7.95 (dd, *J* = 8.3, 1.4 Hz, 1H), 7.90 (s, 1H), 7.86 (s, 1H), 7.86–7.77 (m, 4H), 7.70 (t, *J* = 7.9 Hz, 1H), 7.66 (s, 1H), 7.64–7.61 (m, 3H), 7.43–7.39 (m, 1H), 7.35 (t, *J* = 7.9

Hz, 2H), 6.84 (s, 1H), 6.70 (s, 1H), 6.65 (t, *J* = 7.9 Hz, 1H), 6.40 (d, *J* = 7.9 Hz, 1H), 4.63–4.59 (m, 1H), 4.57 (s, 3H), 4.42–4.38 (m, 1H), 4.27–4.19 (m, 4H), 3.98–3.85 (m, 4H), 3.57 (s, 3H), 2.91–2.82 (m, 1H), 2.57–2.48 (m, 1H), 2.48–2.41 (m, 1H), 2.40–2.31 (m, 2H), 1.38 (d, *J* = 6.8 Hz, 3H), 1.33 (d, *J* = 6.8 Hz, 6H), 1.31 (d, *J* = 6.6 Hz, 3H), 1.28 (d, *J* = 6.8 Hz, 3H), 1.26 (d, *J* = 2.4 Hz, 3H), 1.25 (d, *J* = 2.3 Hz, 6H), 1.23 (d, *J* = 1.8 Hz, 3H), 1.22 (d, *J* = 6.8 Hz, 3H). **HRMS** (ESI+) calcd. for $C_{102}H_{92}N_{16}O_{18}$ $[M+H]^+$ 1829.6848, found 1829.6872.

C*-Q^{B}XQ^B-T6r-Q^{B}Q^{M}Q^B-OMe (**4a**) Compound **4a** was synthesized on SASRIN resin (scale: 15 µmol) using SPFS procedures reported in paragraph 5.3. The crude after full-cleavage was dried in vacuo and dissolved in dry chloroform/MeOH 3:2 (3 mL) under N_2. $TMSCHN_2$ (2 M in *n*-hexane, 17.7 µL, 60 µmol) was added dropwise. The solution was stirred at r.t. for 2 h. A few drops of acetic acid were added and stirring continued at r.t. for 30 min to quench the excess amounts of $TMSCHN_2$. Then the solution was diluted with DCM, washed with aqueous $NaHCO_3$ solution, dried over $MgSO_4$, filtered and concentrated. The crude product was purified by RP-HPLC. The product was obtained as a yellow solid (13.2 mg, 43%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 12.61 (s, 1H), 12.41 (s, 1H), 12.08 (s, 1H), 12.05 (s, 1H), 12.03 (s, 1H), 11.95 (s, 1H), 10.77 (s, 1H), 9.99 (s, 1H), 9.12 (t, *J* = 7.5 Hz, 2H), 9.05 (s, 1H), 8.76 (d, *J* = 7.5 Hz, 1H), 8.49 (d, *J* = 7.8 Hz, 1H), 8.38 (s, 1H), 8.34 (d, *J* = 7.6 Hz, 1H), 8.16 (d, *J* = 6.5 Hz, 1H), 8.04–7.98 (m, 4H), 7.96 (d, *J* = 8.4 Hz, 1H), 7.92–7.89 (m, 2H), 7.86 (t, *J* = 7.8 Hz, 1H), 7.83–7.81 (m, 2H), 7.78 (t, *J* = 7.9 Hz, 1H), 7.70 (t, *J* = 7.9 Hz, 1H), 7.66 (s, 1H), 7.61 (t, *J* = 7.9 Hz, 1H), 7.56 (s, 1H), 7.32–7.11 (m, 4H), 6.93 (s, 1H), 6.81 (s, 1H), 6.70 (s, 1H), 6.61 (t, *J* = 8.0 Hz, 1H), 6.39 (d, *J* = 7.4 Hz, 1H), 4.56 (s, 1H), 4.47 (t, *J* = 8.5 Hz, 1H), 4.39 (t, *J* = 7.1 Hz, 1H), 4.34–4.27 (m, 1H), 4.26–4.19 (m, 1H), 4.18–4.11 (m, 1H), 3.99–3.93 (m, 2H), 3.90–3.84 (m, 2H), 3.80 (s, 2H), 3.73 (s, 1H), 3.69 (t, *J* = 4.8 Hz, 1H), 3.65 (s, 3H), 3.53 (s, 3H), 2.57–1.98 (m, 5H), 1.37–1.21 (m, 42H), 0.94 (s, 6H). **HRMS** (ESI+) calcd. for $C_{116}H_{114}N_{16}O_{19}$ $[M+H]^+$ 2035.8519, found 2035.8703.

C*-Q^{B}XQ^B-T6r-Q^{B}Q^{M}Q^B-OMe (**4b**) Compound **4a** (13.2 mg, 6.5 µmol) was treated with TFA/DCM 1:1 (3 mL) at r.t. for 2 h. The solvent was removed in vacuo. The residue was washed with diethyl ether, yielding the product as a yellow solid (12.8 mg, quant.). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 12.66 (s, 1H), 12.60 (s, 1H), 12.51 (s, 1H), 11.95 (s, 1H), 11.88 (s, 1H), 11.76 (s, 1H), 11.33 (s, 1H), 10.15 (s, 1H), 10.01 (s, 1H), 9.19 (d, *J* = 6.9 Hz, 1H), 8.93–8.89 (m, 1H), 8.74 (s, 1H), 8.54 (dd, *J* = 5.9, 2.7 Hz, 1H), 8.44 (s, 1H), 8.31 (ddd, *J* = 9.8, 7.8, 1.3 Hz, 2H), 8.20–8.16 (m, 2H), 8.12 (dd, *J* = 8.3, 1.3 Hz, 1H), 8.05–7.94 (m, 6H), 7.87–7.81 (m, 4H), 7.72 (dt, *J* = 13.3, 8.1 Hz, 2H), 7.66 (s, 1H), 7.59 (s, 1H), 7.56 (s, 1H), 7.47 (s, 1H), 7.35 (t, *J* = 8.0 Hz, 2H), 7.31–7.27 (m, 1H), 6.78 (s, 1H), 6.71 (s, 1H), 6.59 (t, *J* = 8.1 Hz, 1H), 6.33 (dd, *J* = 8.2, 1.2 Hz, 1H), 4.69 (dd, *J* = 8.7, 4.2 Hz, 1H), 4.48 (s, 3H), 4.45 (dd, *J* = 8.7, 6.2 Hz, 1H), 4.27–4.19 (m, 4H), 3.98–3.88 (m, 4H), 3.60 (s, 3H), 2.61–2.05 (m, 9H), 1.44–1.22 (m, 30H), 0.79 (s, 3H), 0.73 (s, 3H), 0.40 (s, 3H). **HRMS** (ESI+) calcd. for $C_{112}H_{106}N_{16}O_{19}$ $[M+H]^+$ 1979.7893, found 1979.7950.

O$_2$N-Q^{B}XQ^{B}Q^{M}PQ^{B}XQ^B-T6f-Q^{B}Q^{M}Q^{B}Q^{B}XQ^{B}Q^{M}Q^B-Gly-OMe (**5a**) Compound **5a** was synthesized on Fmoc-Gly-HMBA-AM resin (scale: 15 µmol) using SPFS procedures reported in paragraph 5.3. The crude product was purified by RP-HPLC after full-cleavage. The product was obtained as a yellow solid (12 mg, 19%). **^{1}H NMR** (500 MHz, $CDCl_3$, mixture of three conformers in a ratio of 1:0.8:0.2, only the aromatic amide signals are reported.) δ 12.27 (s, 0.2H), 11.89 (s, 0.8H), 11.77 (s, 1H), 11.53 (s, 0.8H), 11.50 (s, 0.2H), 11.49 (s, 0.2H), 11.47 (s, 0.8H), 11.46 (s, 2H), 11.44 (s, 1.8H), 11.42 (s, 1H), 11.40 (s, 1H), 11.32 (s, 0.8H), 11.31 (s, 0.2H), 11.28 (s, 1H), 11.25 (s, 0.8H), 11.24 (s, 1H), 11.21 (s, 0.2H), 11.18 (s, 0.8H), 11.17 (s, 1H), 11.16 (s, 1H), 11.14 (s, 1H),

11.07 (s, 0.2H), 11.01 (s, 0.8H), 10.98 (s, 1H), 10.97 (s, 0.8H), 10.96 (s, 1H), 10.87 (s, 0.2H), 10.86 (s, 0.2H), 10.82 (s, 0.8H), 10.82 (s, 1H), 10.80 (s, 0.8H), 10.79 (s, 1H), 10.77 (s, 0.2H), 10.70 (s, 0.8H), 10.68 (s, 1H), 10.66 (s, 0.8H), 10.65 (s, 1H). **HRMS** (ESI+) calcd. for $C_{232}H_{223}N_{37}O_{38}$ $[2M+2Na]^{2+}$ 2090.3220, found 2090.2690.

O_2N-Q^BX$Q^B$$Q^MPQ^BXQ^B$-T6f-$Q^B$$Q^M$$Q^B$$Q^BXQ^B$$Q^M$$Q^B$-Gly-OMe (**5b**) Compound **5a** (12 mg, 2.9 μmol) was treated with a TFA/DCM 1:1 (3 mL) at r.t. for 2 h. The solvent was removed in vacuo. The residue was washed with diethyl ether, yielding the product as a yellow solid (11.5 mg, quant.). **^{1}H NMR** (500 MHz, CD_2Cl_2) δ 12.09 (s, 1H), 11.53 (s, 1H), 11.52 (s, 1H), 11.41 (s, 2H), 11.22 (s, 2H), 11.05 (s, 1H), 10.99 (s, 2H), 10.90 (s, 1H), 10.79 (s, 2H), 10.58 (s, 1H), 10.31 (s, 1H), 10.27 (s, 1H), 10.14 (s, 1H), 9.76 (s, 1H), 8.71 (t, *J* = 3.3 Hz, 1H), 8.64 (d, *J* = 7.3 Hz, 1H), 8.34 (dd, *J* = 7.9, 1.5 Hz, 1H), 8.29 (d, *J* = 6.9 Hz, 1H), 8.26 (d, *J* = 7.6 Hz, 1H), 8.25–8.20 (m, 3H), 8.19–8.18 (m, 1H), 8.17 (s, 1H), 8.10–8.03 (m, 4H), 7.99 (d, *J* = 8.0 Hz, 1H), 7.98–7.94 (m, 3H), 7.91–7.89 (m, 2H), 7.88–7.83 (m, 3H), 7.81 (dd, *J* = 8.0, 1.3 Hz, 1H), 7.78 (dd, *J* = 7.8, 1.4 Hz, 1H), 7.73 (d, *J* = 8.1 Hz, 1H), 7.63 (s, 1H), 7.61 (d, *J* = 7.3 Hz, 1H), 7.59–7.54 (m, 2H), 7.53–7.41 (m, 7H), 7.41–7.34 (m, 5H), 7.32–7.24 (m, 3H), 7.23 (s, 1H), 7.22–7.16 (m, 6H), 7.14 (s, 1H), 7.11 (d, *J* = 5.3 Hz, 3H), 7.10–7.03 (m, 5H), 6.83 (s, 1H), 6.71 (d, *J* = 7.1 Hz, 1H), 6.59 (t, *J* = 7.8 Hz, 1H), 6.57 (s, 1H), 6.55 (d, *J* = 2.5 Hz, 2H), 6.51 (s, 1H), 6.41 (d, *J* = 2.8 Hz, 2H), 6.14 (d, *J* = 8.3 Hz, 1H), 6.06 (s, 1H), 5.86 (s, 1H), 4.48 (s, 3H), 4.42 (s, 3H), 4.32 (s, 3H), 4.20–4.16 (m, 2H), 4.15–4.05 (m, 2H), 4.03–3.96 (m, 5H), 3.93–3.88 (m, 2H), 3.83 (q, *J* = 7.0, 6.5 Hz, 2H), 3.78 (t, *J* = 7.0 Hz, 1H), 3.72 (t, *J* = 5.8 Hz, 2H), 3.64–3.60 (m, 2H), 3.50 (dd, *J* = 17.6, 6.3 Hz, 1H), 3.09 (s, 3H), 2.89 (dd, *J* = 17.5, 3.1 Hz, 1H), 2.62 (d, *J* = 9.3 Hz, 2H), 2.60–2.21 (m, 10H), 2.19–2.16 (m, 2H), 2.06–1.99 (m, 3H), 1.44–1.11 (m, 60H). **HRMS** (ESI+) calcd. for $C_{220}H_{199}N_{37}O_{38}$ $[2M+2H]^{2+}$ 1984.2461, found 1984.2511.

O_2N-$Q^B$$\underline{X}$$Q^B$$Q^MPQ^B$$\underline{X}$$Q^B$-T6r-$Q^B$$Q^M$$Q^B$$Q^B$$\underline{X}$$Q^B$$Q^M$$Q^B$-Gly-OMe (**6a**) Compound **6a** was synthesized on Fmoc-Gly-HMBA-AM resin (scale: 15 μmol) using SPFS procedures reported in paragraph 5.3. The crude product was purified by RP-HPLC after full-cleavage. The product was obtained as a yellow solid (14 mg, 22%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 11.58 (s, 1H), 11.44 (s, 2H), 11.30 (d, *J* = 1.6 Hz, 2H), 11.25 (s, 1H), 11.14 (s, 1H), 10.97 (s, 1H), 10.96 (s, 2H), 10.82 (s, 1H), 10.77 (s, 2H), 10.64 (s, 1H), 10.63 (s, 1H), 10.06 (s, 1H), 8.69 (s, 1H), 8.28 (dd, *J* = 8.1, 1.4 Hz, 1H), 8.21 (dd, *J* = 7.4, 1.3 Hz, 1H), 8.17 (dd, *J* = 8.0, 1.3 Hz, 1H), 8.11–8.05 (m, 5H), 8.04 (s, 1H), 8.02 (s, 1H), 7.97–7.94 (m, 2H), 7.92–7.89 (m, 3H), 7.88–7.85 (m, 2H), 7.84 (d, *J* = 6.8 Hz, 1H), 7.81–7.79 (m, 2H), 7.78–7.76 (m, 3H), 7.74 (t, *J* = 1.4 Hz, 2H), 7.73 (t, *J* = 1.5 Hz, 1H), 7.72–7.62 (m, 3H), 7.57–7.54 (m, 2H), 7.49 (dd, *J* = 7.5, 2.3 Hz, 2H), 7.46–7.43 (m, 2H), 7.41–7.35 (m, 3H), 7.30–7.27 (m, 2H), 7.25–7.22 (m, 2H), 7.22–7.20 (m, 2H), 7.19–7.15 (m, 2H), 7.14–7.07 (m, 4H), 7.05–6.99 (m, 5H), 6.96 (s, 1H), 6.96–6.86 (m, 3H), 6.77–6.73 (m, 2H), 6.65 (s, 1H), 6.60 (s, 1H), 6.47 (s, 1H), 6.46 (s, 1H), 6.37 (s, 1H), 6.23 (s, 1H), 6.10 (t, *J* = 7.9 Hz, 1H), 6.05 (s, 1H), 5.94 (s, 1H), 5.85 (s, 1H), 5.79 (s, 1H), 5.62 (dd, *J* = 7.9, 1.2 Hz, 1H), 4.69 (t, *J* = 2.9 Hz, 2H), 4.39–4.36 (m, 1H), 4.18–4.05 (m, 5H), 4.01–3.92 (m, 3H), 3.91–3.80 (m, 7H), 3.76 (s, 3H), 3.70 (t, *J* = 7.9 Hz, 1H), 3.67–3.62 (m, 3H), 3.54 (d, *J* = 7.1 Hz, 6H), 3.26–3.21 (m, 1H), 3.00 (s, 3H), 2.96 (dd, *J* = 17.8, 3.1 Hz, 1H), 2.53–2.12 (m, 10H), 1.69 (s, 9H), 1.52 (s, 9H), 1.43–1.07 (m, 60H), 0.78 (s, 9H). **HRMS** (ESI+) calcd. for $C_{235}H_{223}N_{37}O_{38}$ $[2M+2H]^{2+}$ 2086.3400, found (HR-ESI) 2086.3175.

O_2N-Q^BX$Q^B$$Q^MPQ^BXQ^B$-T6r-$Q^B$$Q^M$$Q^B$$Q^BXQ^B$$Q^M$$Q^B$-Gly-OMe (**6b**) Compound **6a** (14 mg, 3.3 μmol) was treated with a TFA/DCM 1/1 (3 mL) at r.t. for 2 h. The solvent was removed in vacuo. The residue was washed with diethyl ether, yielding the product as a yellow solid (13.4 mg, quant.). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 12.23 (s, 1H), 11.68 (s, 1H), 11.45 (s, 1H), 11.44 (s, 1H), 11.39 (s, 1H), 11.12 (s, 1H), 11.08 (s, 1H), 11.05 (s, 1H), 11.00

(s, 1H), 10.96 (s, 1H), 10.92 (s, 1H), 10.87 (s, 1H), 10.59 (s, 1H), 10.47 (s, 1H), 10.16 (s, 1H), 10.11 (s, 1H), 10.06 (s, 1H), 10.03 (s, 1H), 9.67 (s, 1H), 8.78 (s, 1H), 8.65 (d, *J* = 7.5 Hz, 2H), 8.38 (d, *J* = 7.3 Hz, 1H), 8.31–8.24 (m, 5H), 8.23–8.14 (m, 4H), 8.11–8.07 (m, 5H), 8.06–8.03 (m, 2H), 8.01 (d, *J* = 7.3 Hz, 1H), 7.99–7.95 (m, 3H), 7.92 (d, *J* = 6.5 Hz, 1H), 7.90–7.84 (m, 4H), 7.77 (s, 1H), 7.73 (s, 1H), 7.64–7.58 (m, 6H), 7.53–7.45 (m, 3H), 7.44–7.40 (m, 2H), 7.39–7.34 (m, 3H), 7.34–7.29 (m, 3H), 7.25–7.18 (m, 3H), 7.17–7.12 (m, 3H), 7.12–7.07 (m, 4H), 7.06–7.01 (m, 2H), 6.95–6.88 (m, 2H), 6.79 (s, 1H), 6.76 (s, 1H), 6.68 (d, *J* = 5.4 Hz, 3H), 6.63 (t, *J* = 7.5 Hz, 1H), 6.41 (s, 1H), 6.31 (s, 1H), 6.20 (s, 1H), 6.12 (d, *J* = 7.6 Hz, 1H), 5.95 (d, *J* = 4.3 Hz, 2H), 4.69 (t, *J* = 3.1 Hz, 2H), 4.47 (s, 3H), 4.40–4.37 (m, 2H), 4.32 (s, 3H), 4.27–4.23 (m, 1H), 4.18 (d, *J* = 12.1 Hz, 5H), 4.12 (t, *J* = 6.8 Hz, 1H), 4.04–3.83 (m, 8H), 3.83–3.79 (m, 1H), 3.77–3.71 (m, 2H), 3.67 (dd, *J* = 7.8, 6.0 Hz, 1H), 3.62 (t, *J* = 7.4 Hz, 1H), 3.53 (dd, *J* = 17.4, 6.1 Hz, 1H), 3.50–3.46 (m, 1H), 3.10 (s, 3H), 2.96 (dd, *J* = 17.3, 3.2 Hz, 1H), 2.73–2.16 (m, 10H), 1.43–1.11 (m, 60H). **HRMS** (ESI+) calcd. for $C_{223}H_{199}N_{37}O_{38}$ $[2M+2H]^{2+}$ 2002.2461, found 2202.2471.

Fmoc-Q^{D}<u>X</u>Q^{D}Q^{M}<u>Y</u>Q^{D}<u>X</u>Q^{D}-OH (**25**) Compound **25** was synthesized on SASRIN resin (scale: 100 μmol) using SPFS procedures reported in paragraph 5.3. The crude product was purified by RP-HPLC after full-cleavage. The product was obtained as a yellow solid (196.4 mg, 82%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 11.54 (s, 1H), 11.29 (s, 1H), 11.27 (s, 1H), 11.11 (s, 1H), 10.92 (s, 1H), 10.65 (s, 1H), 8.22 (dd, *J* = 7.6, 1.3 Hz, 1H), 8.15 (t, *J* = 3.7 Hz, 1H), 8.07 (dd, *J* = 7.5, 1.3 Hz, 1H), 8.03 (s, 1H), 7.98 (dd, *J* = 8.3, 1.3 Hz, 1H), 7.84–7.79 (m, 2H), 7.75–7.72 (m, 2H), 7.69 (dd, *J* = 8.3, 1.4 Hz, 1H), 7.64 (dd, *J* = 8.3, 1.2 Hz, 2H), 7.58 (dd, *J* = 8.3, 1.3 Hz, 1H), 7.51 (t, *J* = 7.9 Hz, 1H), 7.48–7.39 (m, 5H), 7.35–7.31 (m, 2H), 7.25 (t, *J* = 7.9 Hz, 1H), 7.22–7.17 (m, 4H), 7.13–7.11 (m, 2H), 7.08–7.00 (m, 2H), 7.00–6.94 (m, 2H), 6.84 (s, 1H), 6.79 (d, *J* = 7.6 Hz, 1H), 6.63 (d, *J* = 6.1 Hz, 2H), 6.58 (d, *J* = 2.4 Hz, 1H), 6.52–6.50 (m, 1H), 6.46 (t, *J* = 7.5 Hz, 1H), 6.05 (s, 1H), 4.19 (m, 2H), 4.06–3.97 (m, 3H), 3.92–3.77 (m, 10H), 3.73–3.65 (m, 10H), 3.61–3.58 (m, 3H), 3.53–3.50 (m, 2H), 3.49 (s, 3H), 3.48 (s, 3H), 3.41 (s, 3H), 3.39–3.36 (m, 3H), 3.33 (s, 3H), 3.32–3.24 (m, 2H), 3.23 (s, 2H), 3.10 (td, *J* = 6.6, 2.8 Hz, 2H), 3.01 (dt, *J* = 13.1, 6.6 Hz, 1H), 2.92 (dt, *J* = 12.9, 6.7 Hz, 1H), 2.23 (m, 1H), 1.62 (s, 9H), 1.53 (s, 9H), 1.32–1.28 (m, 2H), 0.19 (s, 9H). **HRMS** (ESI+) calcd. for $C_{126}H_{130}N_{16}O_{23}S_4Si$ $[M+H]^+$ 2391.8220, found 2391.7467.

Fmoc-<u>Y</u>Q^{D}<u>X</u>Q^{D}-OH (**26**) Compound **26** was synthesized on SASRIN resin (scale: 25 μmol) using SPFS procedure reported in paragraph 5.3. The crude product was purified by RP-HPLC after full-cleavage. The product was obtained as a yellow solid (22 mg, 66%).**^{1}H NMR** (500 MHz, $CDCl_3$) δ 12.24 (s, 1H), 11.75 (s, 1H), 11.70 (s, 1H), 8.99 (s, 1H), 8.45 (d, *J* = 7.1 Hz, 1H), 8.25 (s, 1H), 8.00 (d, *J* = 7.8 Hz, 2H), 7.91 (d, *J* = 8.1 Hz, 1H), 7.77 (d, *J* = 8.1 Hz, 1H), 7.73 (s, 1H), 7.70–7.60 (m, 5H), 7.36–7.29 (m, 4H), 7.24 (s, 2H), 7.12 (s, 1H), 6.98 (s, 2H), 6.57 (s, 1H), 5.40 (s, 2H), 4.25 (t, *J* = 7.9 Hz, 2H), 4.10 (s, 2H), 3.90 (t, *J* = 6.2 Hz, 5H), 3.77–3.66 (m, 4H), 3.65–3.55 (m, 5H), 3.44 (s, 4H), 3.31 (d, *J* = 10.0 Hz, 6H), 1.73 (s, 9H), 1.29–1.25 (m, 2H), 0.18 (s, 9H). **HRMS** (ESI+) calcd. for $C_{71}H_{76}N_8O_{13}S_2Si$ $[M+H]^+$ 1341.4815, found 1341.4820.

Fmoc-<u>Y</u>Q^{D}<u>X</u>Q^{D}-T6r-Q^{D}<u>X</u>Q^{D}Q^{M}-OH (**27**) Compound **27** was synthesized on SASRIN resin (scale: 50 μmol) using SPFS procedure reported in paragraph 5.3. The crude product was purified by RP-HPLC after full-cleavage. The product was obtained as a yellow solid (72.8 mg, 52%).**^{1}H NMR** (500 MHz, $CDCl_3$) δ 12.02 (s, 1H), 12.01 (s, 1H), 11.99 (s, 1H), 11.93 (s, 2H), 11.85 (s, 1H), 11.70 (s, 1H), 11.24 (s, 1H), 10.39 (s, 1H), 9.07 (s, 1H), 8.67 (d, *J* = 7.5 Hz, 1H), 8.62 (d, *J* = 7.4 Hz, 1H), 8.56 (d, *J* = 7.5 Hz, 1H), 8.49 (d, *J* = 7.5 Hz, 1H), 8.29 (d, *J* = 7.8 Hz, 1H), 8.23–8.20 (m, 2H), 8.14 (d, *J* = 7.4 Hz, 1H), 8.11–8.07 (m, 2H), 8.06–8.05 (m, 2H), 8.04–8.03 (m, 2H),

8.02 (s, 1H), 7.99 (d, *J* = 7.9 Hz, 1H), 7.94 (d, *J* = 8.3 Hz, 1H), 7.91–7.88 (m, 1H), 7.87 (s, 1H), 7.86 (d, *J* = 8.3 Hz, 1H), 7.81 (t, *J* = 7.8 Hz, 1H), 7.75 (d, *J* = 8.0 Hz, 1H), 7.73–7.68 (m, 3H), 7.56 (d, *J* = 6.0 Hz, 2H), 7.55–7.53 (m, 2H), 7.52–7.45 (m, 4H), 7.36–7.28 (m, 4H), 7.23–7.10 (m, 4H), 7.01 (s, 1H), 6.99 (d, *J* = 7.5 Hz, 1H), 6.86 (t, *J* = 7.3 Hz, 1H), 6.74 (t, *J* = 7.4 Hz, 1H), 6.71 (s, 1H), 6.46 (t, *J* = 8.0 Hz, 1H), 6.39 (s, 1H), 6.26 (d, *J* = 8.1 Hz, 1H), 5.22–5.19 (m, 1H), 4.60–4.47 (m, 2H), 4.28–4.19 (m, 3H), 4.11 (t, *J* = 6.6 Hz, 2H), 4.01 (s, 3H), 4.00–3.93 (m, 5H), 3.91–3.78 (m, 13H), 3.71–3.66 (m, 14H), 3.61–3.58 (m, 2H), 3.51–3.48 (m, 4H), 3.48 (s, 3H), 3.47 (s, 3H), 3.44 (s, 3H), 3.43–3.38 (m, 3H), 1.46 (s, 9H), 0.91 (s, 9H), 0.88 (t, *J* = 6.9 Hz, 2H), 0.14 (s, 9H). **HRMS** (ESI+) calcd. for $C_{150}H_{150}N_{20}O_{26}S_4Si$ $[2M+2H]^{2+}$ 1402.4914, found 1402.4937.

Fmoc-Q^{D}<u>X</u>Q^{D}Q^{M}-OH (**28**) Compound **28** was synthesized on SASRIN resin (scale: 25 µmol) using SPFS procedure reported in paragraph 5.3. The crude product was purified by RP-HPLC after full-cleavage. The product was obtained as a yellow solid (26.5 mg, 82%). **^{1}H NMR** (500 MHz, $CDCl_3$) δ 12.27 (s, 1H), 11.85 (s, 1H), 11.29 (s, 1H), 9.12 (d, *J* = 7.6 Hz, 1H), 8.58 (d, *J* = 8.1 Hz, 2H), 8.05 (d, *J* = 7.5 Hz, 1H), 7.99 (d, *J* = 7.4 Hz, 1H), 7.94–7.90 (m, 2H), 7.82–7.76 (m, 4H), 7.72 (t, *J* = 8.0 Hz, 1H), 7.63 (d, *J* = 7.4 Hz, 1H), 7.59 (d, *J* = 7.4 Hz, 1H), 7.57–7.54 (m, 1H), 7.44 (d, *J* = 7.1 Hz, 1H), 7.36–7.29 (m, 2H), 7.24–7.19 (m, 2H), 7.14 (t, *J* = 7.2 Hz, 1H), 6.86 (t, *J* = 7.2 Hz, 2H), 6.80 (s, 1H), 6.76 (d, *J* = 7.2 Hz, 1H), 4.18–4.14 (m, 2H), 4.08–4.04 (m, 5H), 3.90 (q, *J* = 6.6 Hz, 2H), 3.83–3.80 (m, 2H), 3.77–3.75 (m, 2H), 3.69–3.63 (m, 5H), 3.61 (t, *J* = 6.6 Hz, 2H), 3.46 (s, 3H), 3.45 (s, 3H), 3.34–3.26 (m, 2H), 1.74 (s, 9H). **HRMS** (ESI+) calcd. for $C_{70}H_{66}N_8O_{13}S_2$ $[M+Na]^{+}$ 1313.4083, found 1313.4009.

C*-Q^{D}<u>X</u>Q^{D}Q^{M}<u>Y</u>Q^{D}<u>X</u>Q^{D}-T6f-Q^{D}<u>X</u>Q^{D}Q^{M}<u>Y</u>Q^{D}<u>X</u>Q^{D}-T6f-Q^{D}<u>X</u>Q^{D}Q^{M}<u>Y</u>Q^{D}<u>X</u>Q^{D}-Gly-OMe (**7a**) Compound **7a** was synthesized on Fmoc-Gly-HMBA-AM resin (100 mg, 25 µmol) using SPFS procedure reported in paragraph 5.3. Compounds **25**, **22**, **25**, **22**, **25** and (1*S*)-camphanic chloride were coupled accordingly. The coupling of compound **25** was carried out according to the procedure reported in 5.3.4. The coupling of compound **22** and (1*S*)-camphanic chloride was carried out according to the procedures reported in 5.3.3. The sequence was purified by GPC after full-cleavage. The product was obtained as a yellow solid (18.7 mg, 10%). **^{1}H NMR** (500 MHz, $CDCl_3$, see section 7) showed four sets of signals with a ratio of 1:0.9:0.85:0.8 and is therefore not reported due its complexity. **HRMS** (ESI+) calcd. for $C_{388}H_{413}N_{57}O_{71}S_{12}Si_3$ $[3M+3H]^{3+}$ 2492.2211, found 2942.1987.

C*-Q^{D}XQ^{D}Q^{M}YQ^{D}XQ^{D}-T6f-Q^{D}XQ^{D}Q^{M}YQ^{D}XQ^{D}-T6f-Q^{D}XQ^{D}Q^{M}YQ^{D}XQ^{D}-Gly-OMe (**7b**) Compound **7a** (15 mg, 2.0 µmol) was treated with a TFA/DCM 1/1 (3 mL) at r.t. overnight. The solvent was removed in vacuo. The residue was precipitated from diethyl ether to afford the product as a yellow solid (13.7 mg, quant.). **^{1}H NMR** (500 MHz, pyridine-d_5, see section 7) showed four sets of signals with a ratio of around 1:0.9:0.85:0.8 and is therefore not reported due to its complexity. **HRMS** (ESI+) calcd. for $C_{349}H_{329}N_{57}O_{71}S_{12}$ $[3M+3H]^{3+}$ 2280.0251, found 2280.0125.

C*-Q^{D}<u>X</u>Q^{D}Q^{M}<u>Y</u>Q^{D}<u>X</u>Q^{D}-T6r-Q^{D}<u>X</u>Q^{D}Q^{M}<u>Y</u>Q^{D}<u>X</u>Q^{D}-T6r-Q^{D}<u>X</u>Q^{D}Q^{M}<u>Y</u>Q^{D}<u>X</u>Q^{D}-Gly-OMe (**8a**) Compound **8a** was synthesized on Fmoc-Gly-HMBA-AM resin (100 mg, 25 µmol) using SPFS procedure reported in paragraph 5.3. Compounds **26**, **27**, **27**, **28** and (1*S*)-camphanic chloride were coupled accordingly. The sequence was cleaved and purified by GPC. The product was obtained as a yellow solid (17 mg, 9%).**^{1}H NMR** (500 MHz, $CDCl_3$) δ 11.50 (s, 1H), 11.49 (s, 1H), 11.40 (s, 1H), 11.34 (s, 1H), 11.33 (s, 1H), 11.28 (s, 1H), 11.27 (s, 1H), 11.23 (s, 1H), 11.21 (s, 1H), 11.19 (s, 2H), 11.05 (s, 1H), 11.04 (s, 1H), 10.94 (s, 1H), 10.82 (s, 1H), 10.77 (s, 1H), 10.68 (s, 1H), 10.64 (s, 1H), 10.57 (s, 2H), 10.55 (s, 1H), 10.53 (s, 1H), 9.84 (s, 1H), 9.82 (s, 1H), 9.33 (s, 1H), 9.33 (s,

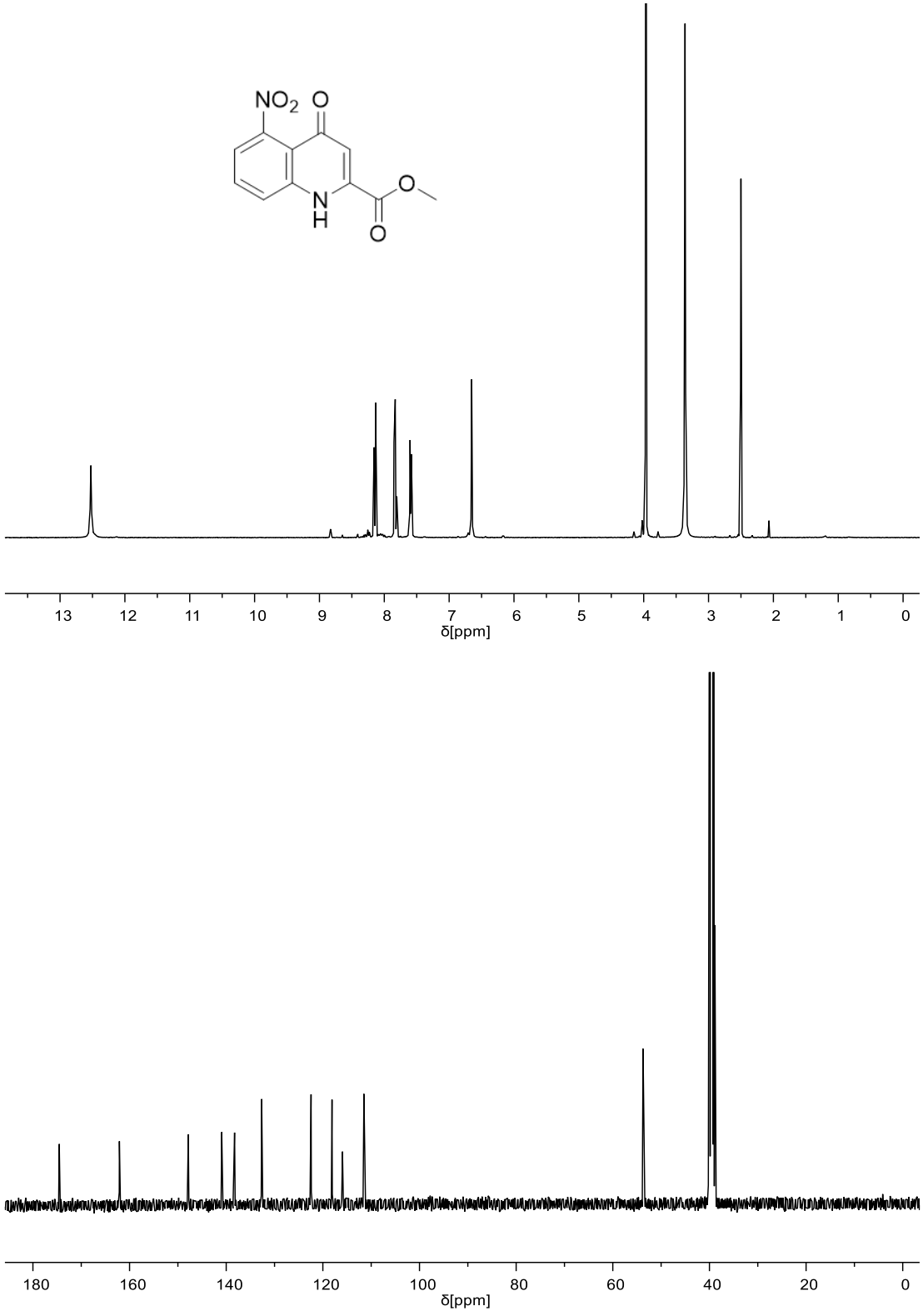
1H), 8.66 (s, 1H), 8.63 (s, 1H), 8.20 (d, *J* = 7.1 Hz, 1H), 8.15 (s, 1H), 8.12–8.09 (m, 2H), 8.07 (s, 2H), 8.05–7.98 (m, 3H), 7.97–7.91 (m, 5H), 7.90–7.83 (m, 3H), 7.78 (s, 1H), 7.77 (s, 1H), 7.75–7.68 (m, 6H), 7.67–7.61 (m, 8H), 7.60 (s, 1H), 7.59–7.56 (m, 5H), 7.55–7.52 (m, 2H), 7.51–7.46 (m, 7H), 7.46–7.40 (m, 3H), 7.38–7.33 (m, 4H), 7.32–7.31 (m, 1H), 7.31–7.30 (m, 1H), 7.28–7.21 (m, 3H), 7.21–7.15 (m, 5H), 7.15–7.11 (m, 5H), 7.10–7.09 (m, 4H), 7.08–7.05 (m, 2H), 7.04–7.01 (m, 2H), 6.99 (s, 1H), 6.96–6.91 (m, 3H), 6.86–6.80 (m, 4H), 6.70–6.63 (m, 4H), 6.55 (s, 1H), 6.46 (d, *J* = 10.5 Hz, 2H), 6.36–6.30 (m, 2H), 6.25 (s, 1H), 6.23 (s, 2H), 6.21 (s, 1H), 6.17–6.13 (m, 3H), 6.11–6.08 (m, 2H), 5.99 (s, 1H), 5.95 (s, 1H), 5.85 (s, 1H), 5.80 (d, *J* = 7.8 Hz, 1H), 5.74 (s, 1H), 4.28–4.25 (m, 1H), 4.24–4.20 (m, 4H), 4.15 (s, 4H), 4.07–3.85 (m, 25H), 3.84–3.68 (m, 28H), 3.67–3.62 (m, 42H), 3.61 (s, 3H), 3.60 (s, 3H), 3.57 (s, 3H), 3.57 (s, 3H), 3.55 (s, 3H), 3.50 (s, 3H), 3.49 (s, 3H), 3.45 (s, 3H), 3.44 (s, 3H), 3.43 (s, 3H), 3.39 (s, 3H), 3.39 (s, 3H), 3.37–3.29 (m, 6H), 3.24 (d, *J* = 27.4 Hz, 4H), 3.11 (s, 3H), 3.04 (d, *J* = 33.1 Hz, 2H), 2.99 (s, 3H), 2.79 (dd, *J* = 17.4, 2.7 Hz, 1H), 2.69–2.61 (m, 2H), 2.34–2.30 (m, 4H), 2.24–2.19 (m, 4H), 1.49 (s, 9H), 1.45 (s, 9H), 1.26 (d, *J* = 8.5 Hz, 35H), 1.05 (s, 9H), 0.64 (s, 9H), 0.48 (s, 9H), 0.18 (s, 9H), 0.07 (s, 9H), 0.06 (s, 9H). **HRMS** (ESI+) calcd. for $C_{394}H_{413}N_{57}O_{71}S_{12}Si_3$ $[3M+3Na]^{3+}$ 2538.2046, found 2538.2031.

C*-Q^{D}XQ^{D}Q^{M}YQ^{D}XQ^{D}-T6r-Q^{D}XQ^{D}Q^{M}YQ^{D}XQ^{D}-T6r-Q^{D}XQ^{D}Q^{M}YQ^{D}XQ^{D}-Gly-OMe (**8b**) Compound **8a** (15 mg, 2.0 μmol) was treated with a TFA/DCM 1:1 (3 mL) at r.t. overnight. The solvent was removed in vacuo. The residue was precipitated from diethyl ether to afford the desired product as a yellow solid (13.7 mg, quant.). **^{1}H NMR** (500 MHz, CD_2Cl_2, only the aromatic amide signals and H-bonded hydroxy group signals are reported.). δ 12.12 (s, 1H), 11.60 (s, 2H), 11.46 (s, 1H), 11.33 (s, 1H), 11.17 (s, 1H), 11.06 (s, 1H), 10.96 (s, 1H), 10.95 (s, 1H), 10.93 (s, 1H), 10.81 (s, 1H), 10.75 (s, 2H), 10.56 (s, 1H), 10.53 (s, 1H), 10.51 (s, 1H), 10.42 (s, 2H), 10.40 (s, 1H), 10.30 (s, 1H), 10.26 (s, 1H), 10.23 (s, 2H), 10.14 (s, 1H), 10.11 (s, 2H), 10.00 (s, 1H), 9.84 (s, 1H), 9.74 (s, 1H), 9.39 (s, 1H), 9.31 (s, 1H), 9.25 (s, 1H), 9.00 (s, 1H), 8.80 (s, 1H). **HRMS** (ESI+) calcd. for $C_{355}H_{329}N_{57}O_{71}S_{12}$ $[3M+3H]^{3+}$ 2304.0251, found 2034.0625.

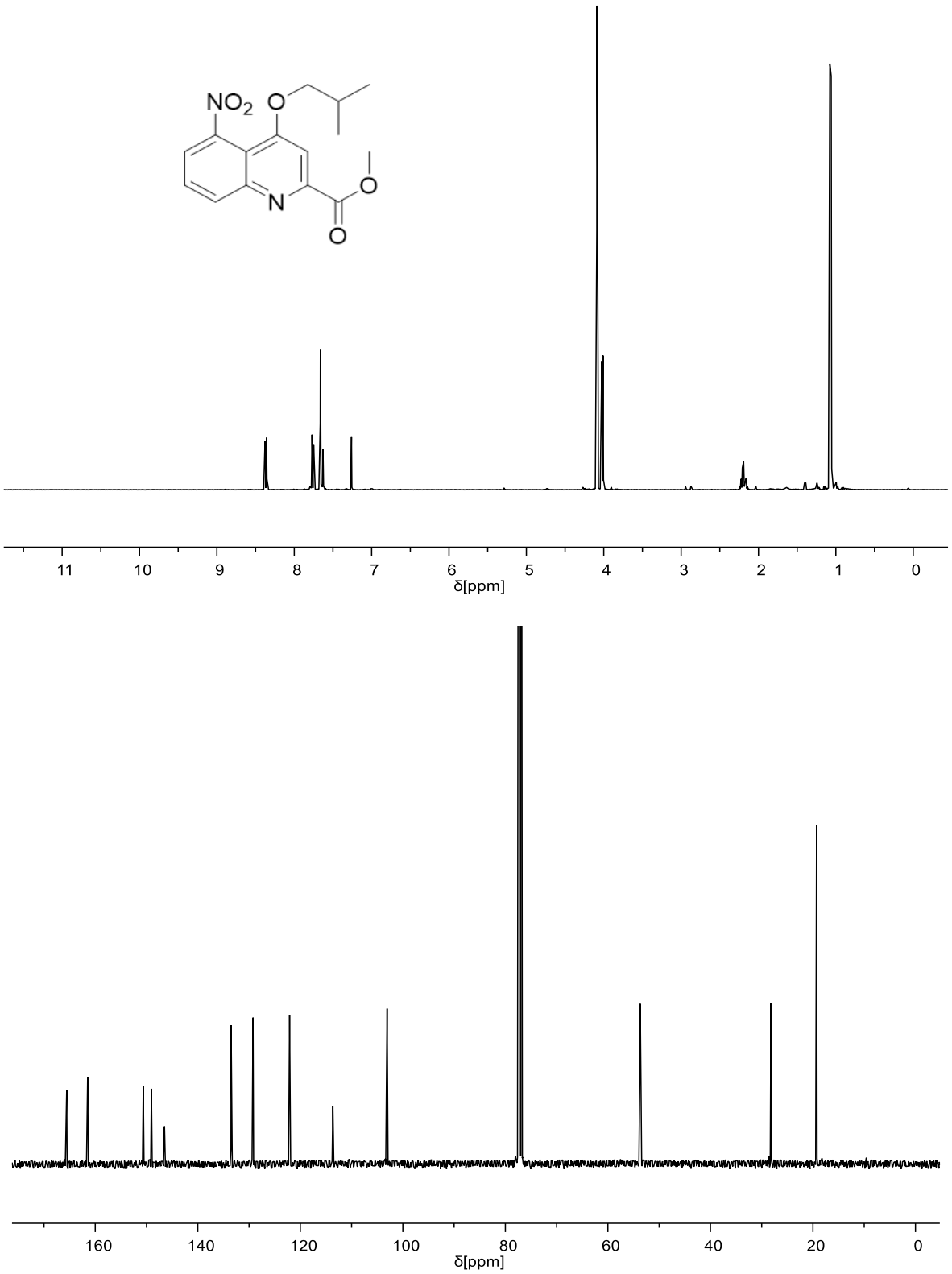
6 References

1. S. De, B. Chi, T. Granier, T. Qi, V. Maurizot and I. Huc, *Nat. Chem.*, 2018, **10**, 51-57.
2. D. Mazzier, S. De, B. Wicher, V. Maurizot and I. Huc, *Angew. Chem. Int. Ed.*, 2020, **59**, 1606-1610.
3. F. S. Menke, B. Wicher, V. Maurizot and I. Huc, *Angew. Chem. Int. Ed.*, 2023, **62**, e202217325.
4. Rigaku-Oxford-Diffraction, *CrysAlisPro Software system, version 1.171.43.130a*, 2024, Rigaku Corporation, Wroclaw, Poland.
5. Rigaku-Oxford-Diffraction, *AutoChem 6.0 software system in conjunction with Olex2 1.5-ac6*, 2024, Rigaku Corporation, Wroclaw, Poland.
6. G. M. Sheldrick, *Acta Crystallogr. Sect. A: Found. Crystallogr.*, 2008, **64**, 112-122.
7. O. V. Dolomanov, L. J. Bourhis, R. J. Gildea, J. A. Howard and H. Puschmann, *J. Appl. Crystallogr.*, 2009, **42**, 339-341.
8. V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat and I. Huc, *Chem. Eur. J.*, 2023, **29**, e202300898.
9. J. Buratto, C. Colombo, M. Stupfel, S. J. Dawson, C. Dolain, B. Langlois d'Estaintot, L. Fischer, T. Granier, M. Laguerre, B. Gallois and I. Huc, *Angew. Chem. Int. Ed.*, 2014, **53**, 883-887.
10. B. Baptiste, C. Douat-Casassus, K. Laxmi-Reddy, F. Godde and I. Huc, *J. Org. Chem.*, 2010, **75**, 7175-7185.
11. M. Vallade, P. Sai Reddy, L. Fischer and I. Huc, *Eur. J. Org. Chem.*, 2018, **2018**, 5489-5498.
12. B. Teng, J. Atcher, L. Allmendinger, C. Douat, Y. Ferrand and I. Huc, *Org. Biomol. Chem.*, 2023, **21**, 3525-3530.
13. O. Al Musaimi, A. Basso, B. G. de la Torre and F. Albericio, *ACS Comb. Sci.*, 2019, **21**, 717-721.
14. J. Hansen, F. Diness and M. Meldal, *Org. Biomol. Chem.*, 2016, **14**, 3238-3245.
15. S. Wang, B. Wicher, C. Douat, V. Maurizot and I. Huc, *Angew. Chem. Int. Ed.*, 2024, e202405091.

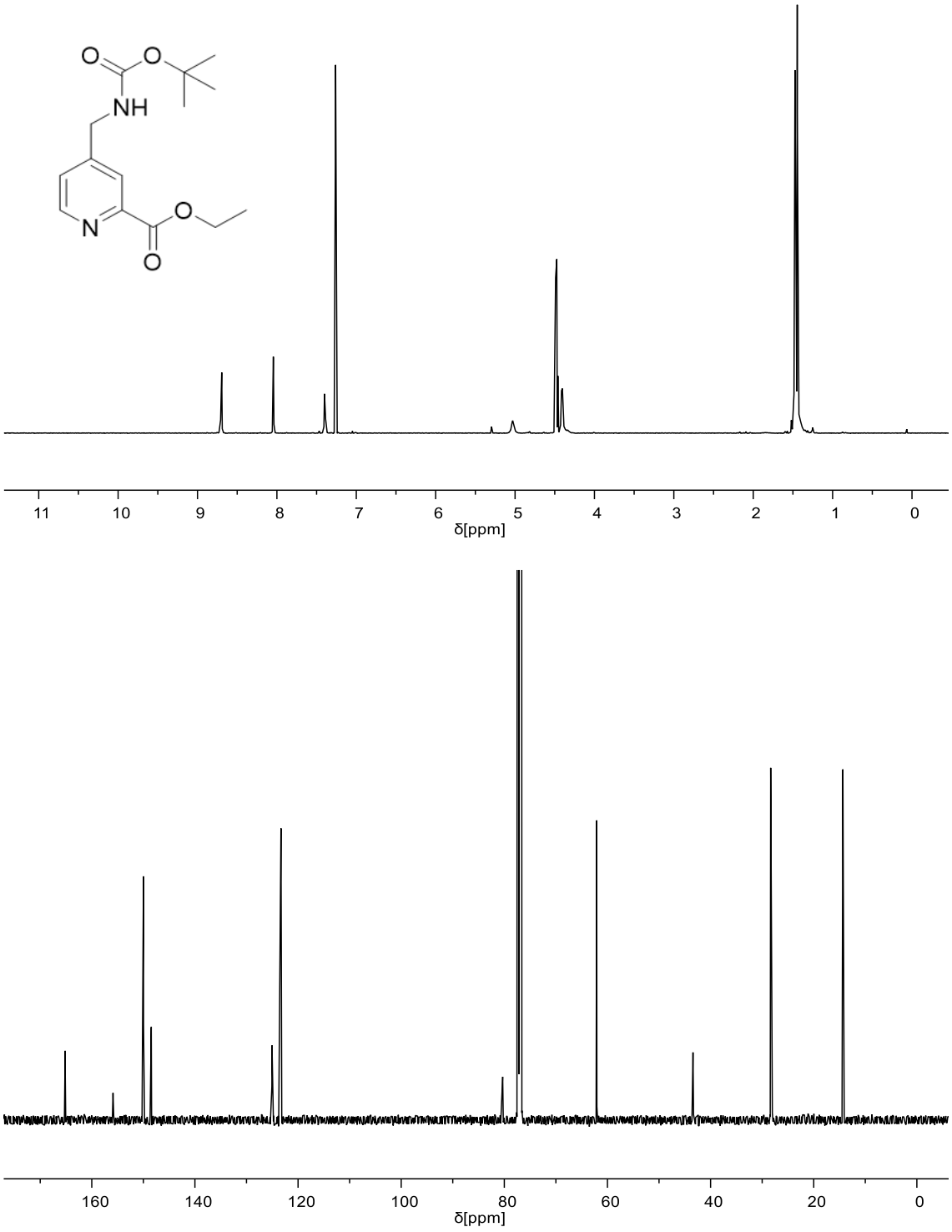
7 NMR spectra of new compounds



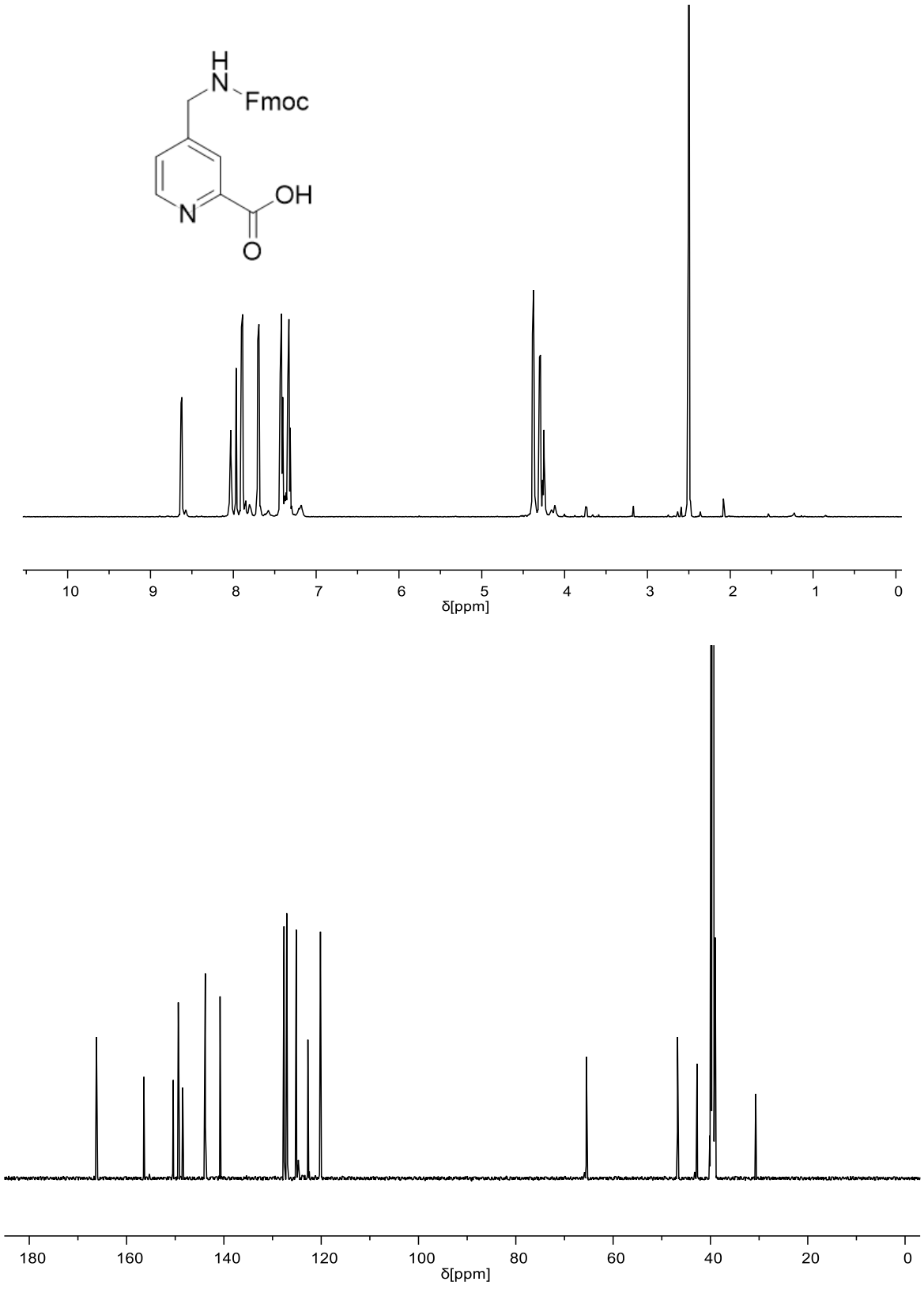
^{1}H NMR spectrum (400 MHz, DMSO-d_6) and ^{13}C NMR (101 MHz, DMSO-d_6) of compound **9**.



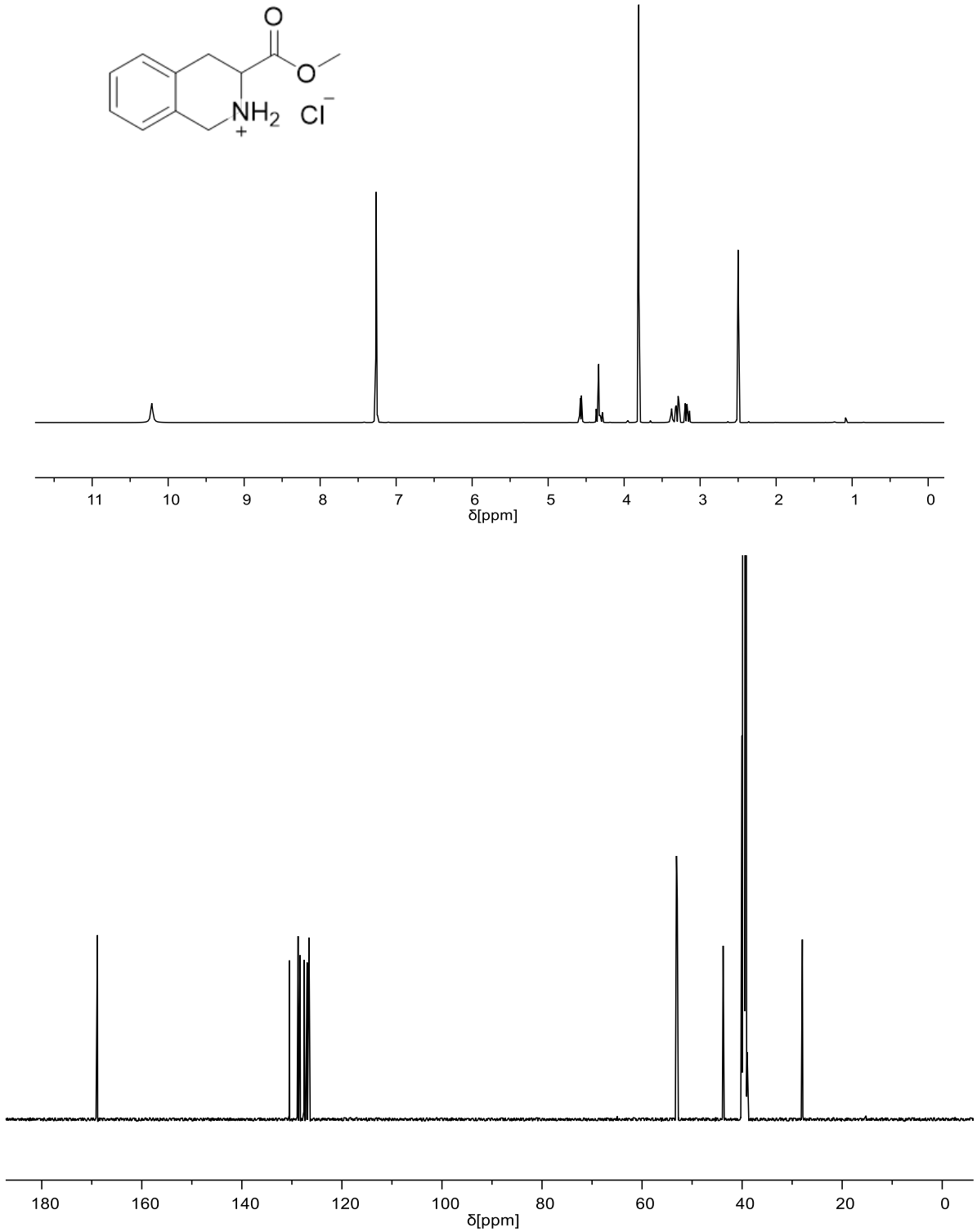
^{1}H NMR spectrum (400 MHz, $CDCl_3$) and ^{13}C NMR (101 MHz, $CDCl_3$) of compound **10**.



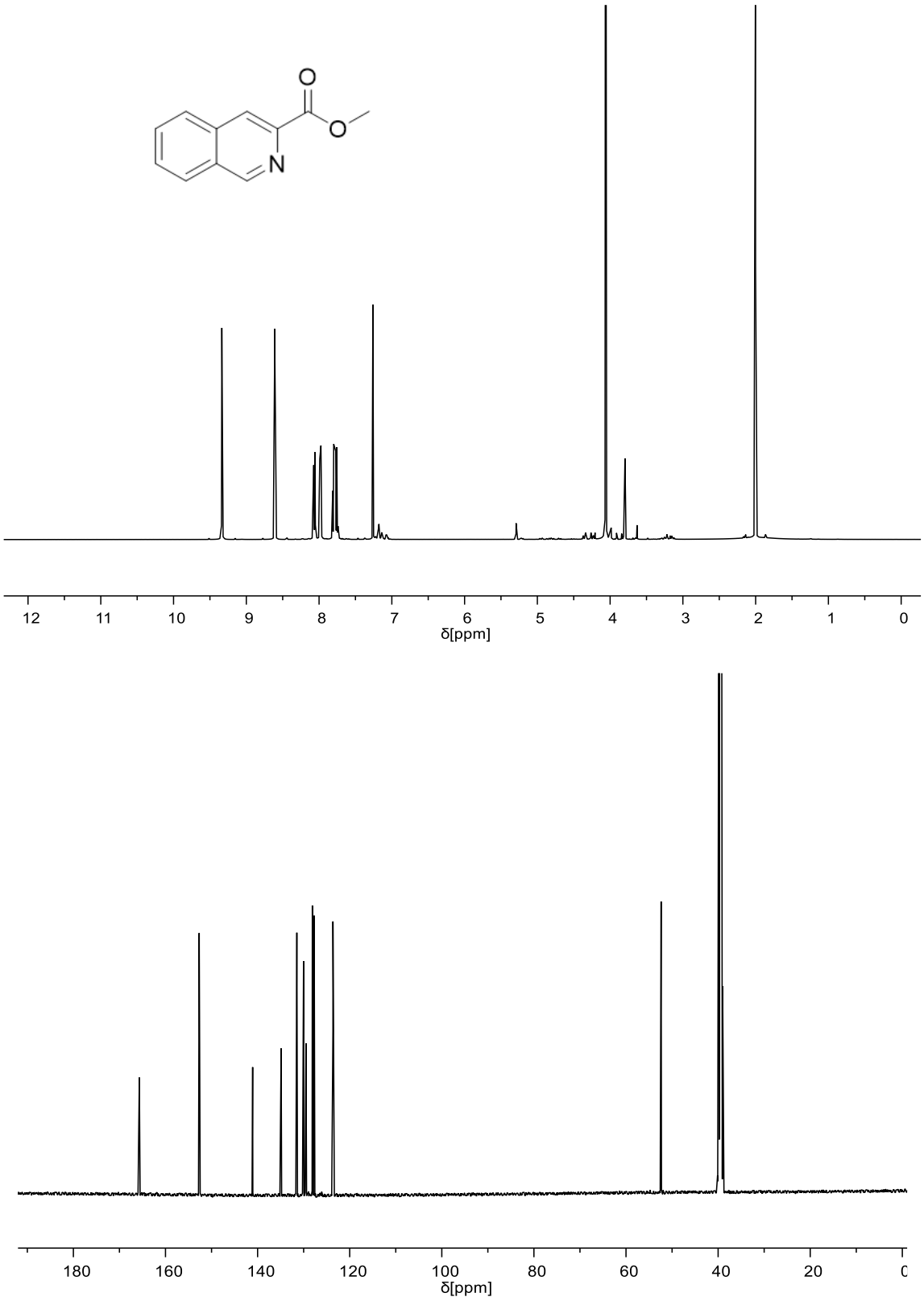
^{1}H NMR spectrum (500 MHz, $CDCl_3$) and ^{13}C NMR (126 MHz, $CDCl_3$) of compound **12**.



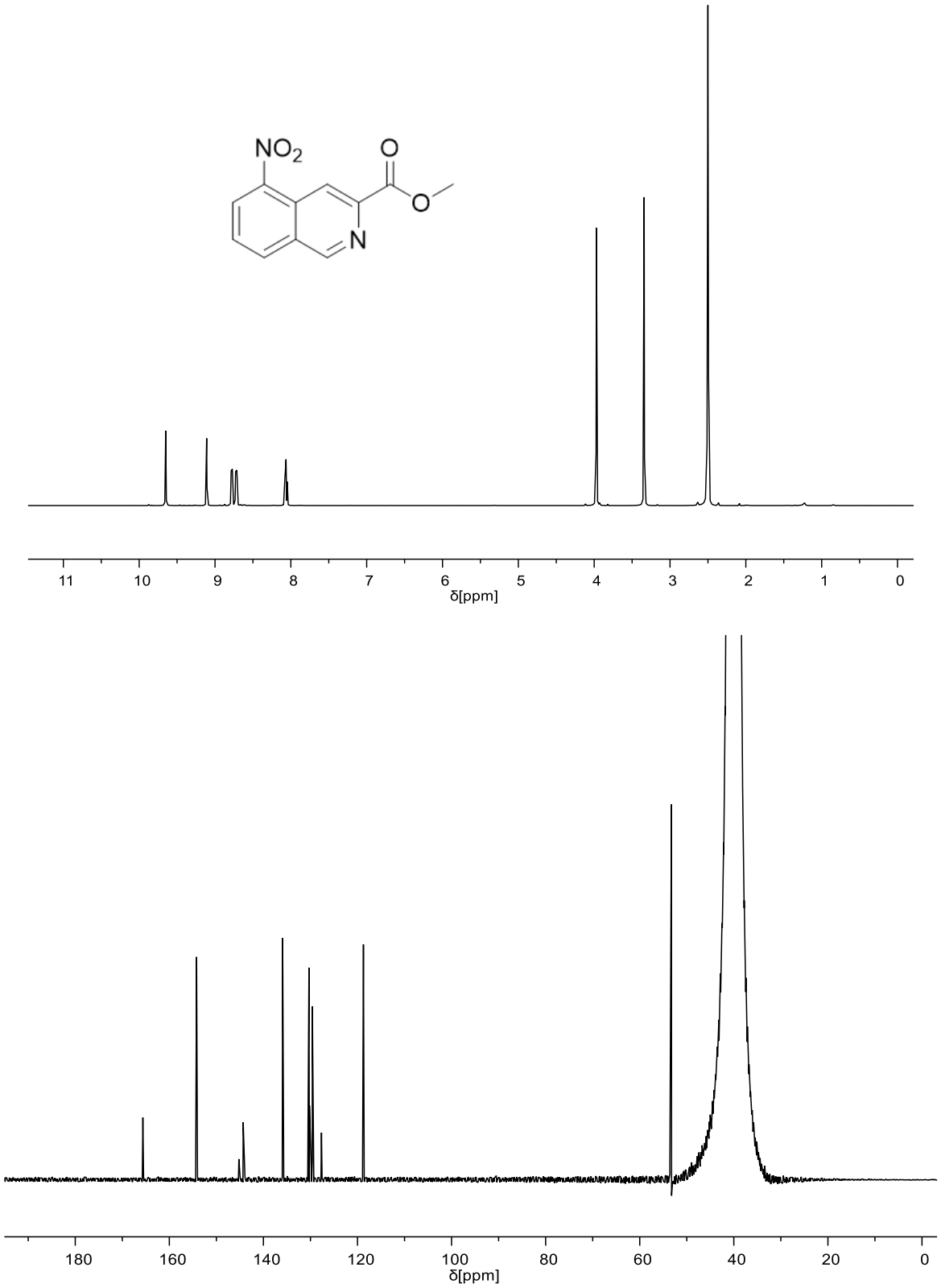
^{1}H NMR spectrum (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6) of compound **14**.



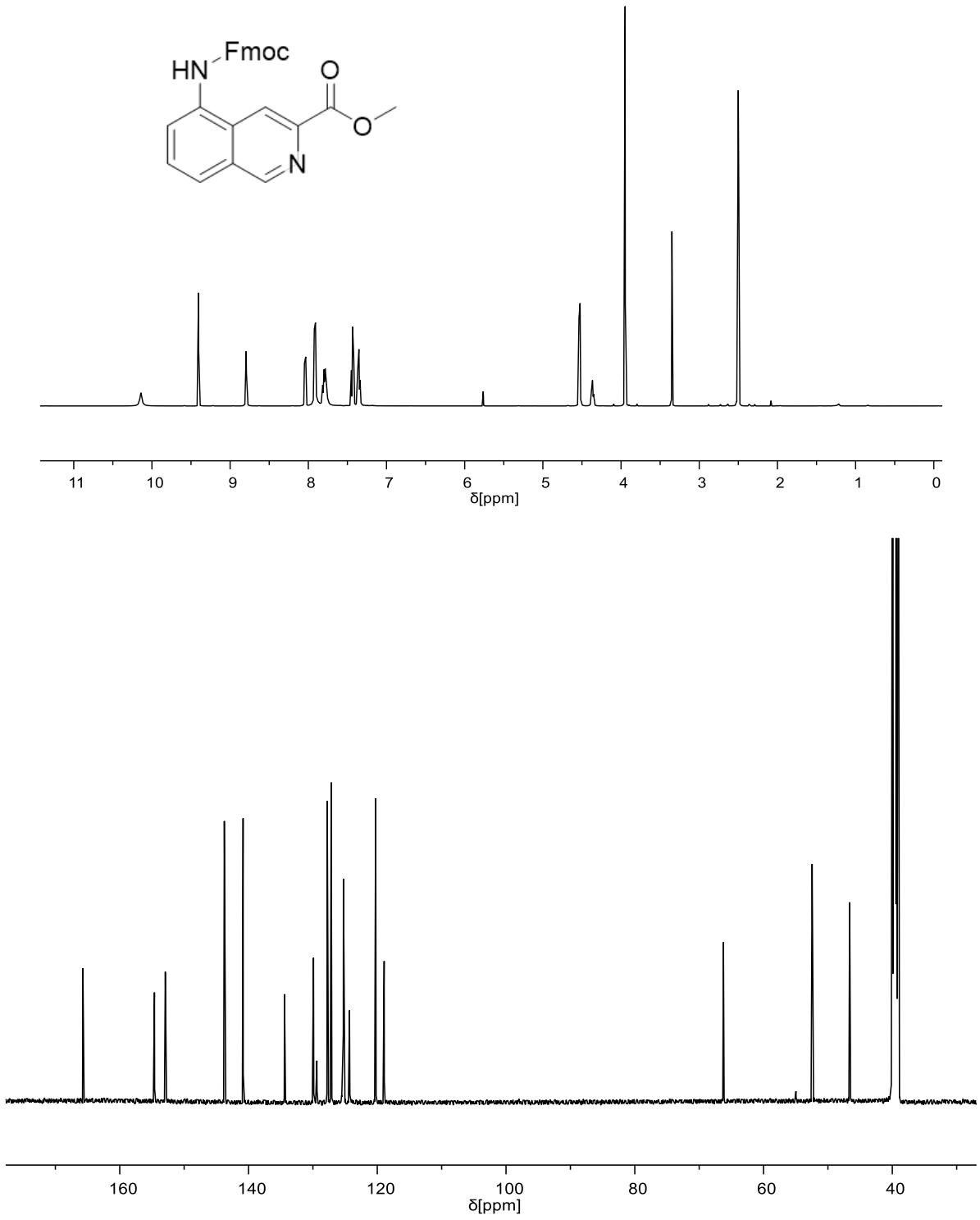
^{1}H NMR spectrum (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6) of compound **15**.



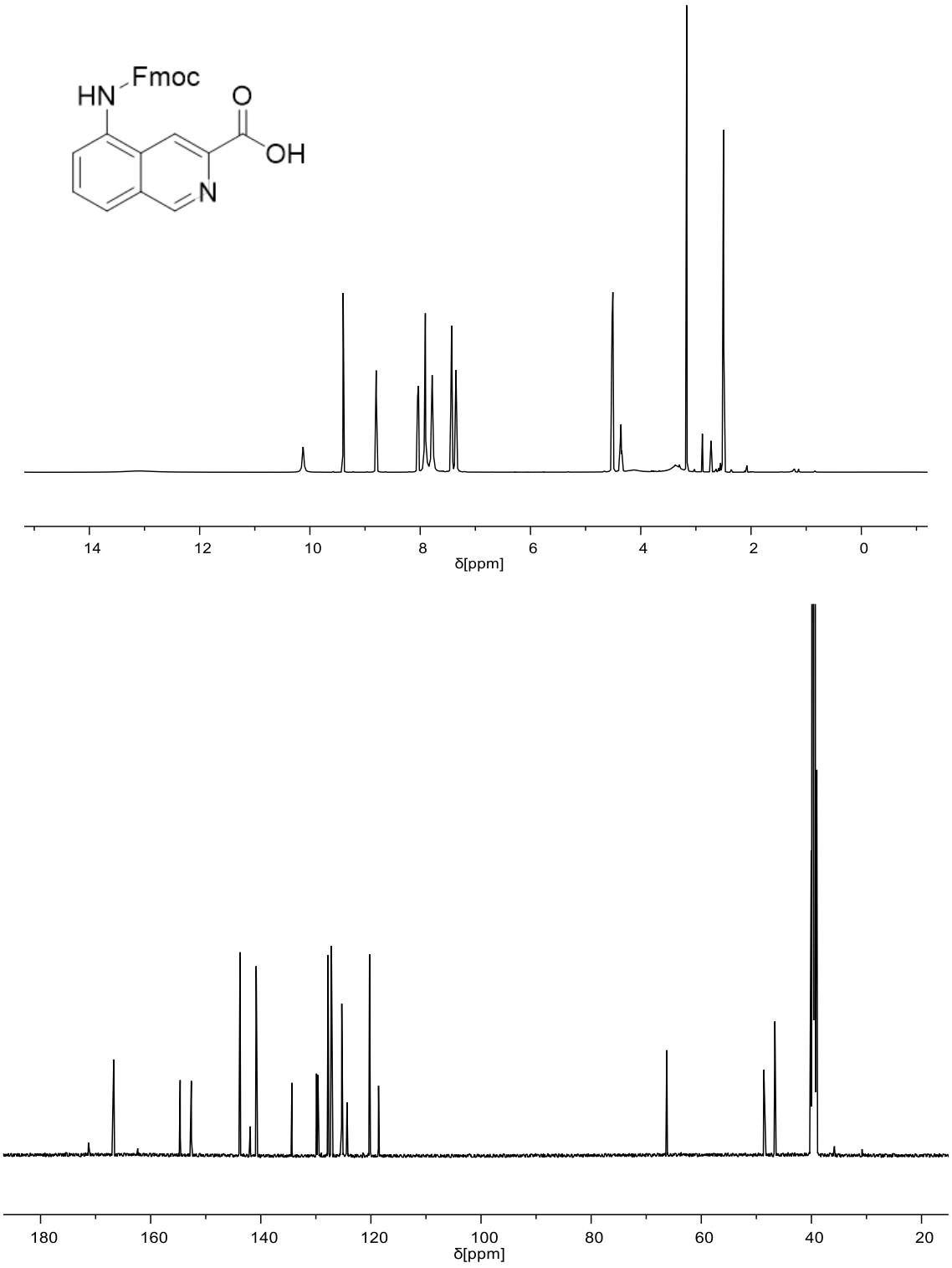
^{1}H NMR spectrum (500 MHz, $CDCl_3$) and ^{13}C NMR (126 MHz, DMSO-d_6) of compound **16**.



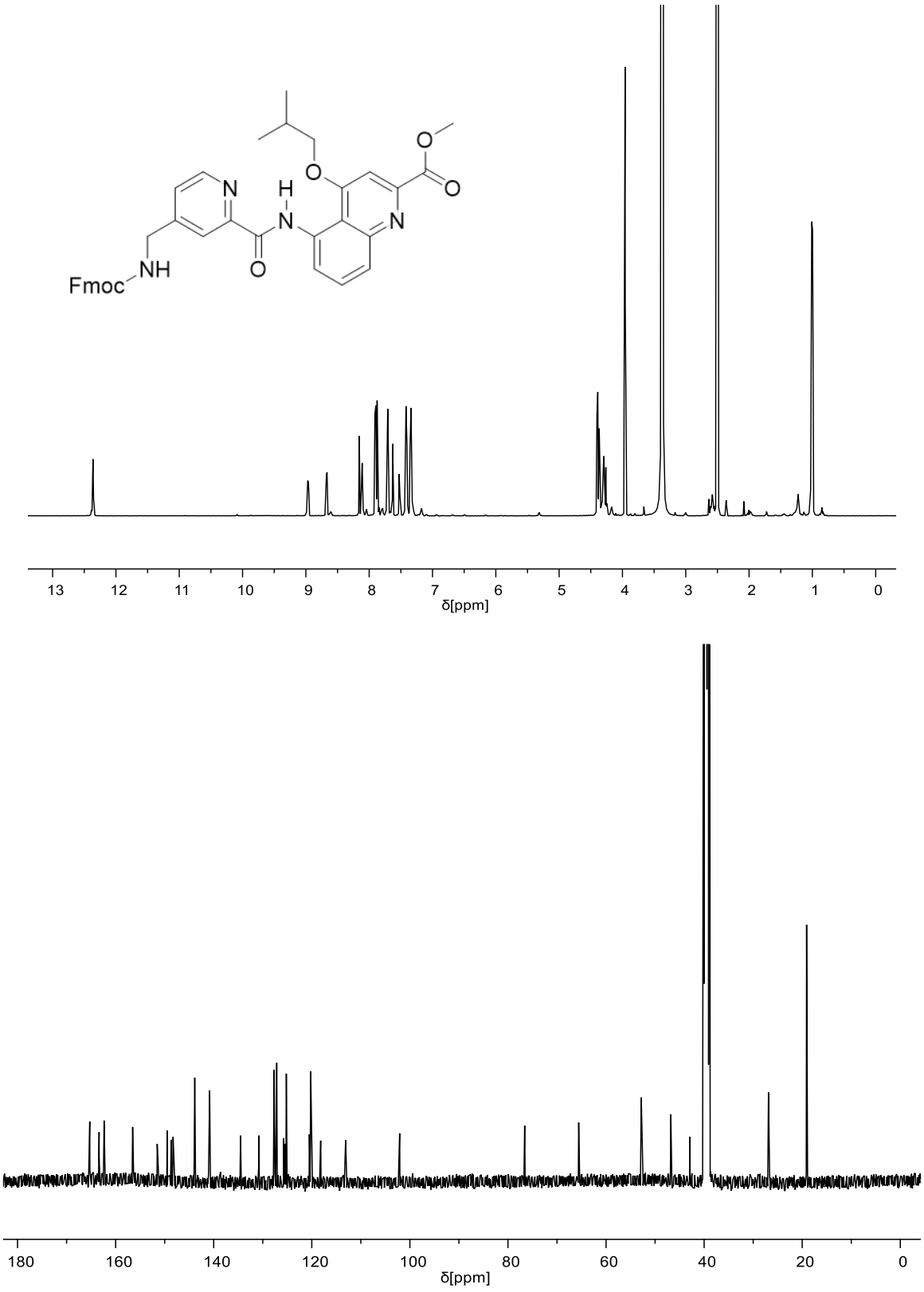
^{1}H NMR spectrum (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6) of compound **17**.



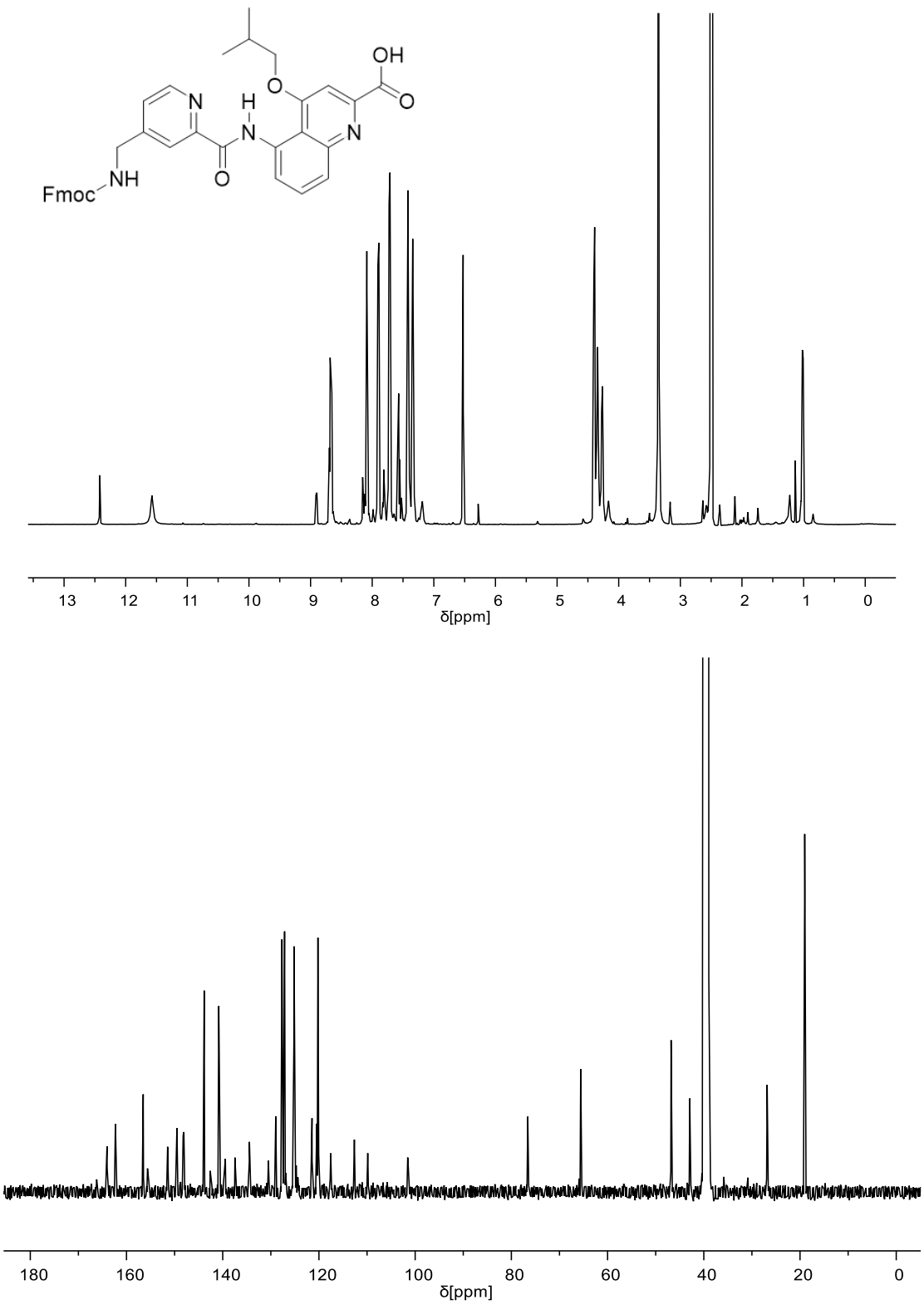
^{1}H NMR spectrum (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6) of compound **19**.



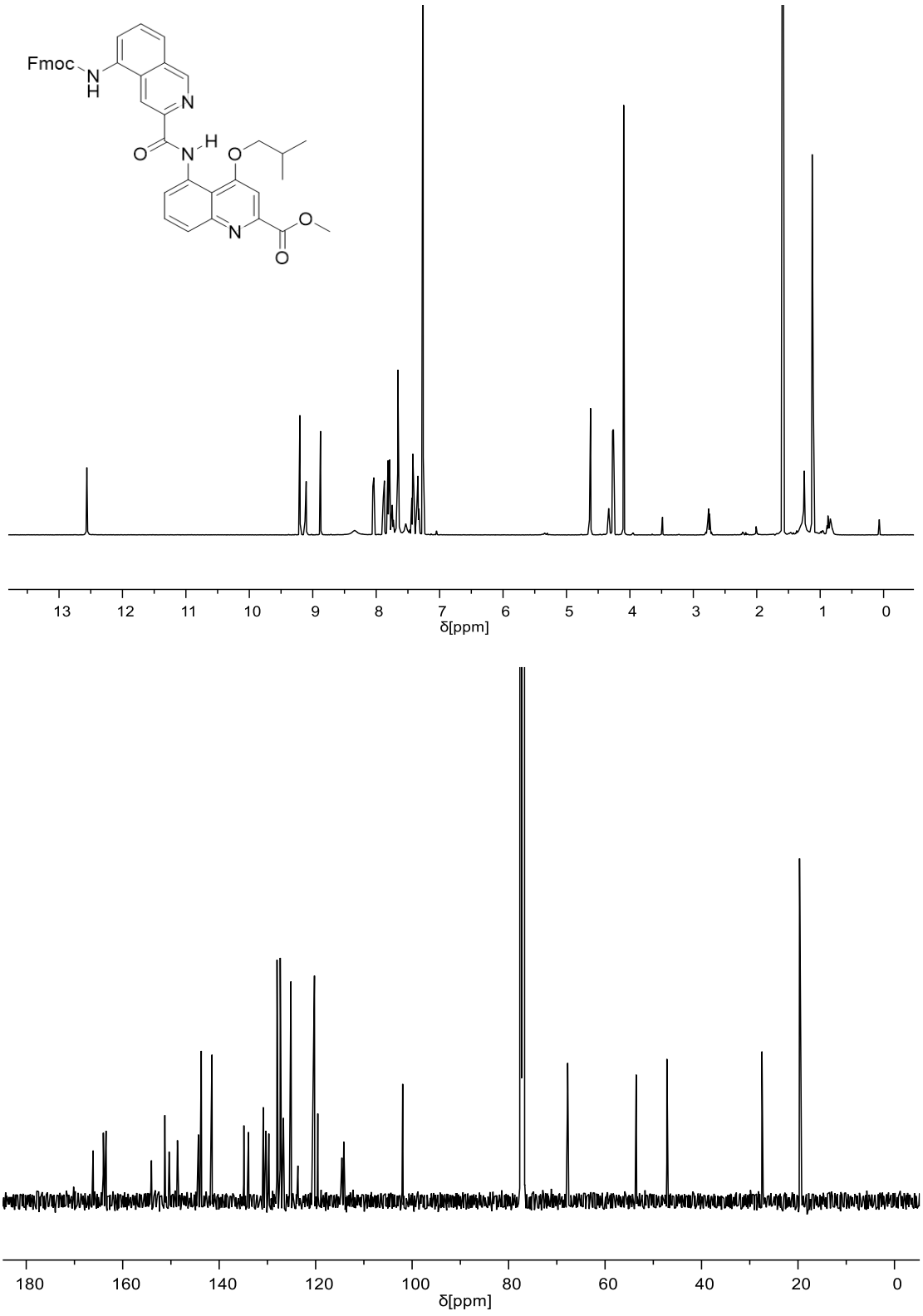
^{1}H NMR spectrum (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6) of compound **20**.



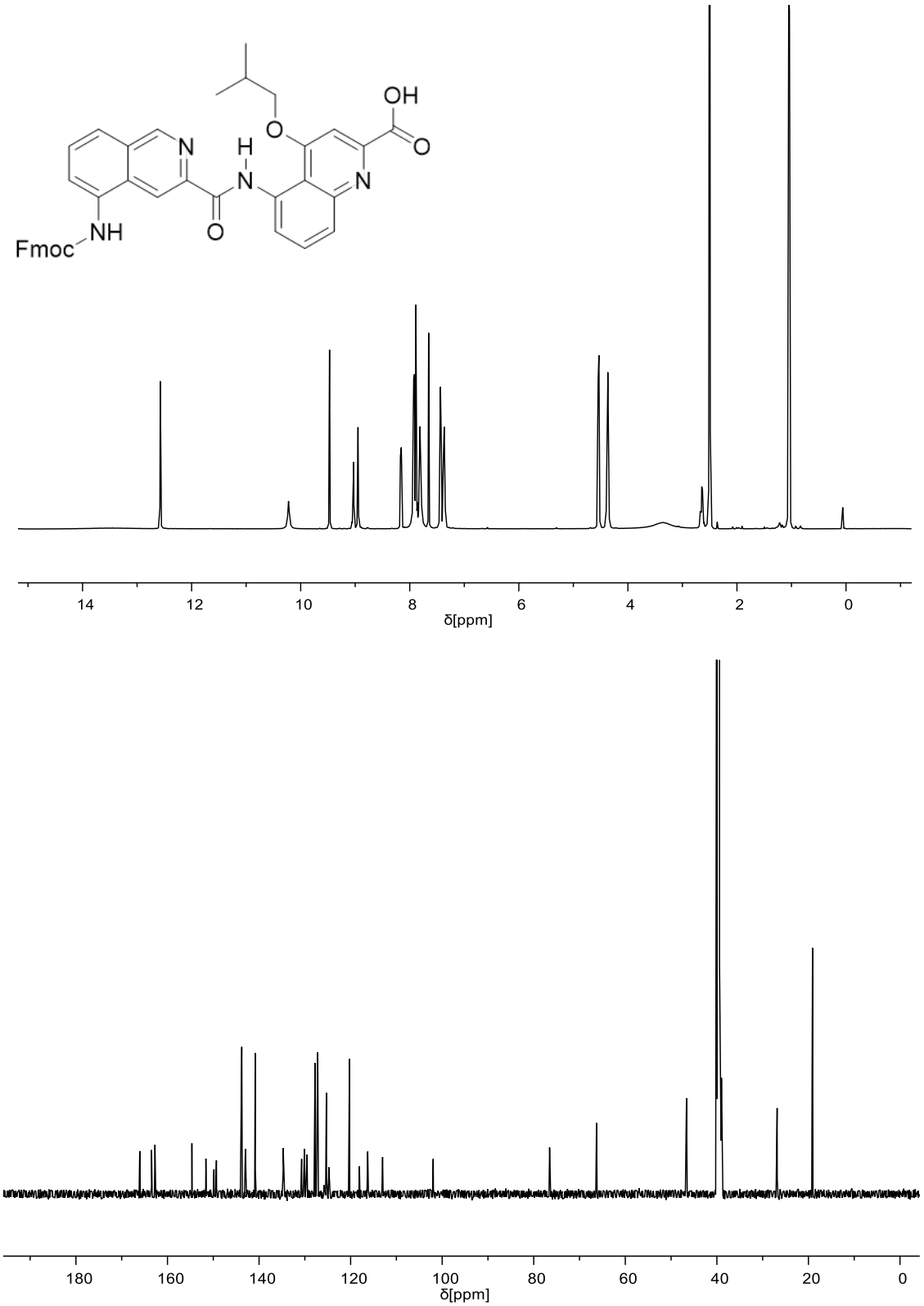
^{1}H NMR spectrum (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6) of compound **21**.



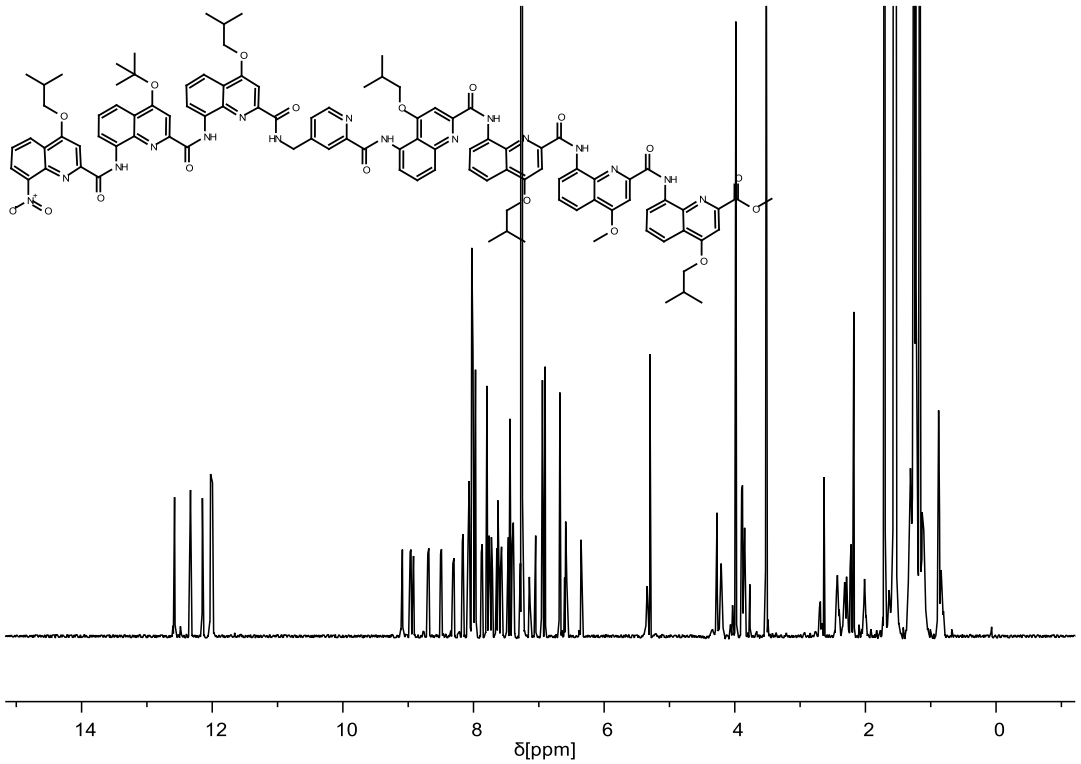
^{1}H NMR spectrum (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6) of compound **22**.



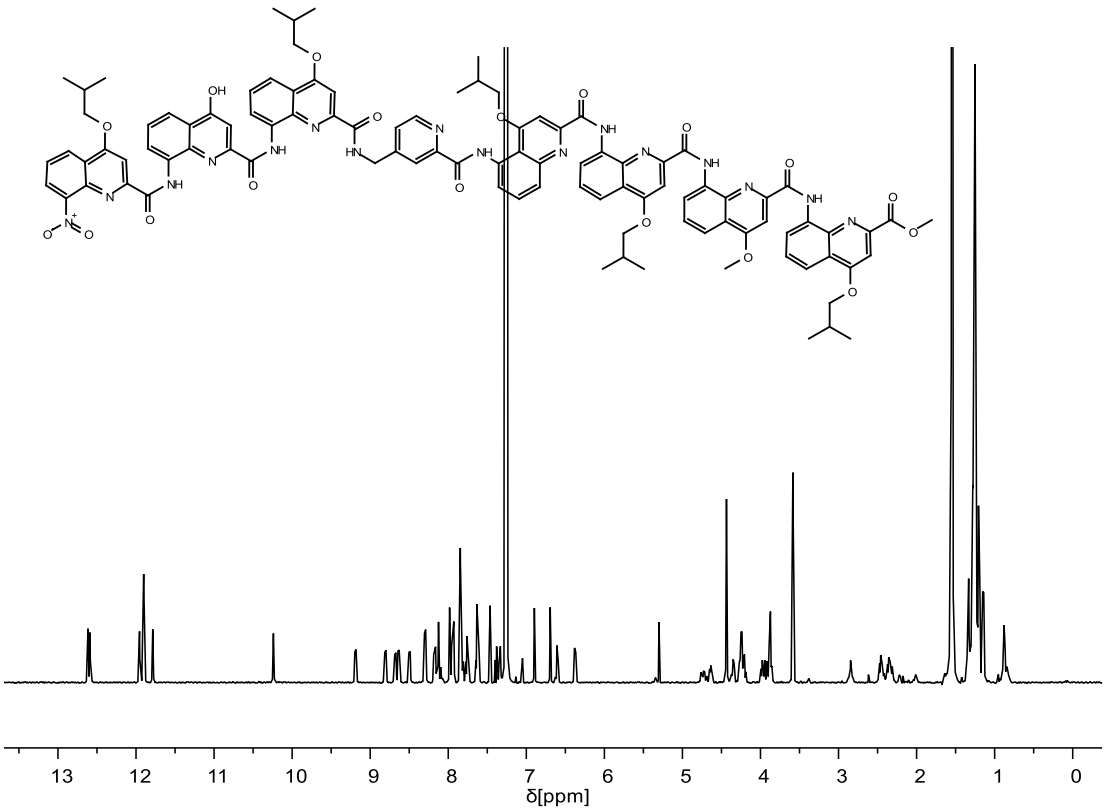
^{1}H NMR spectrum (500 MHz, $CDCl_3$) and ^{13}C NMR (126 MHz, $CDCl_3$) of compound **23**.



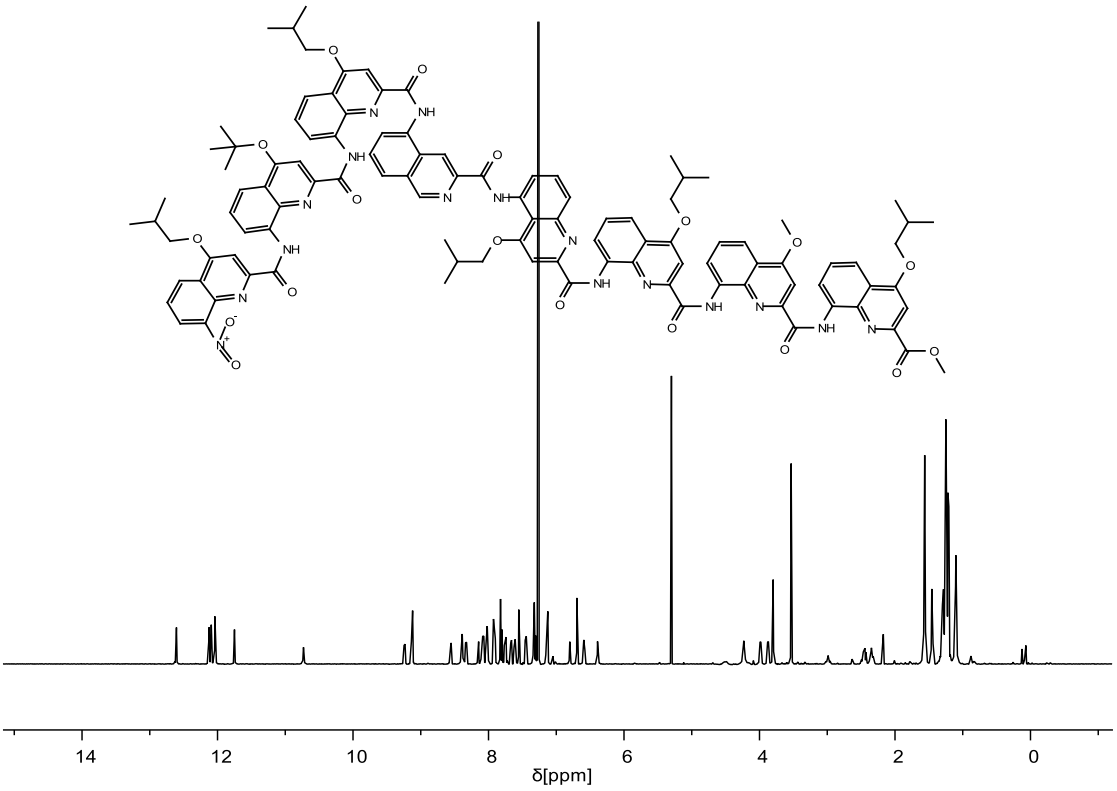
^{1}H NMR spectrum (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6) of compound **24**.



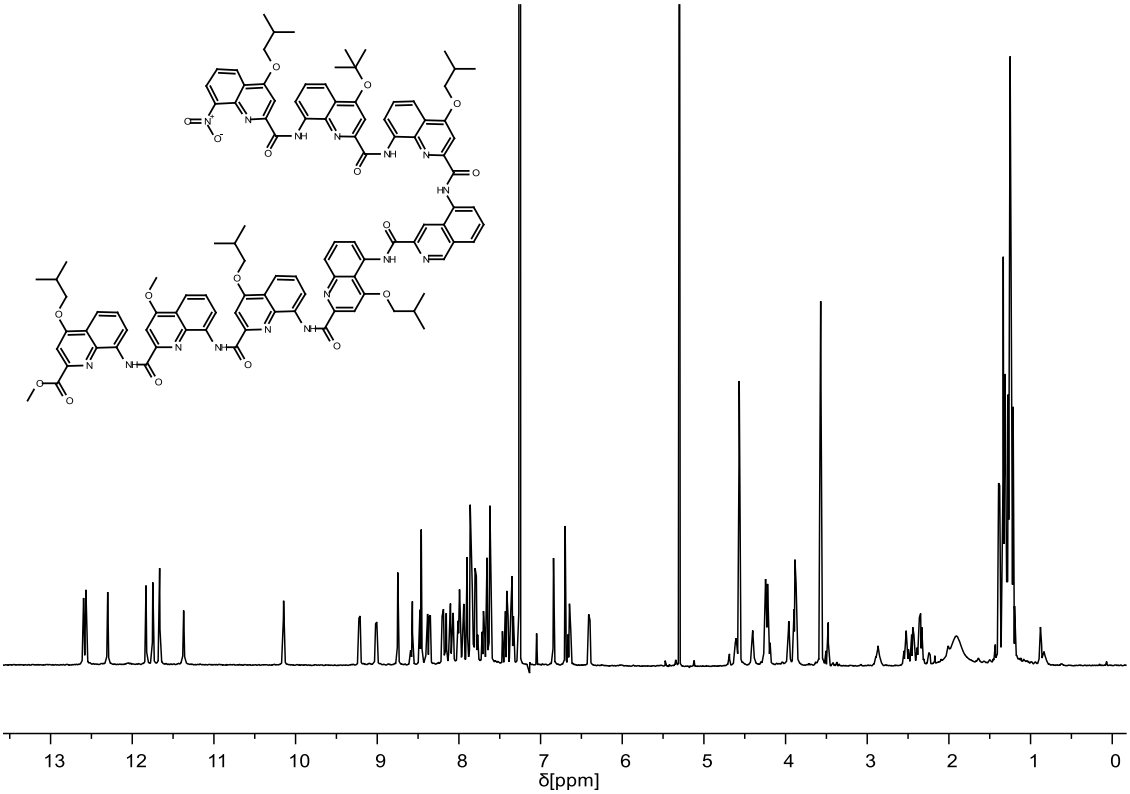
^{1}H NMR spectrum (500 MHz, $CDCl_3$) of compound **2a**.



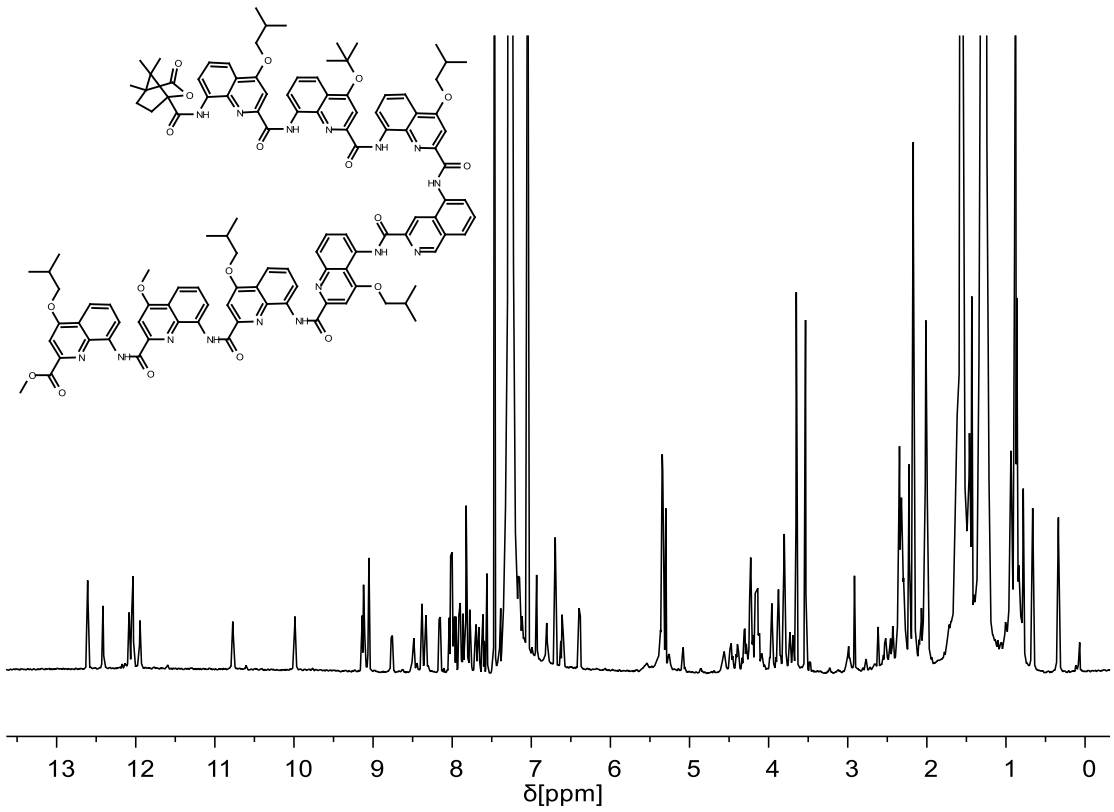
^{1}H NMR spectrum (500 MHz, $CDCl_3$) of compound **2b**.



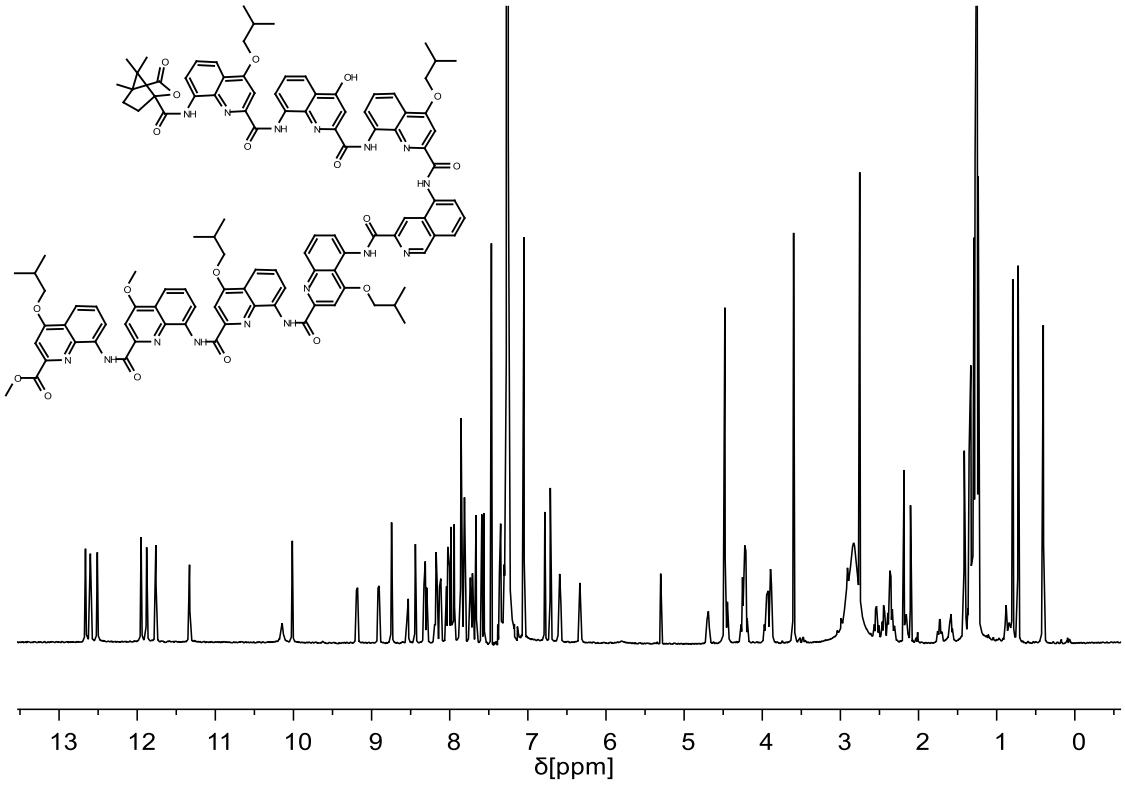
^{1}H NMR spectrum (500 MHz, $CDCl_3$) of compound **3a**.



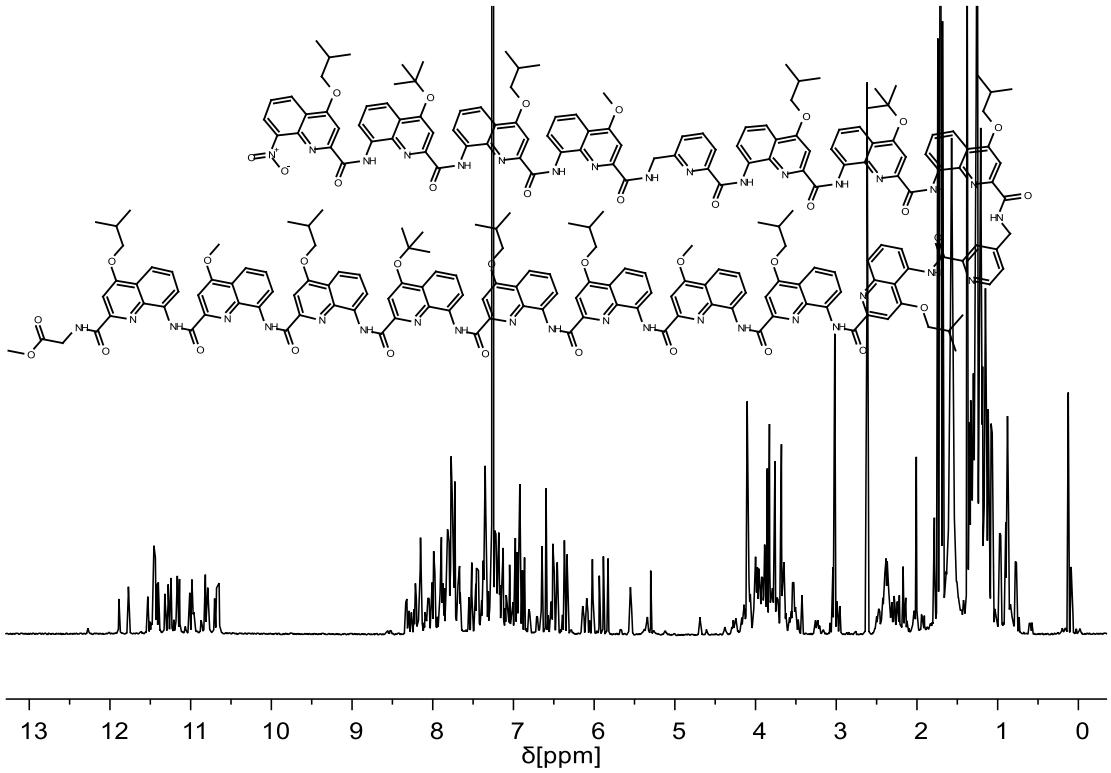
^{1}H NMR spectrum (500 MHz, $CDCl_3$) of compound **3b**.



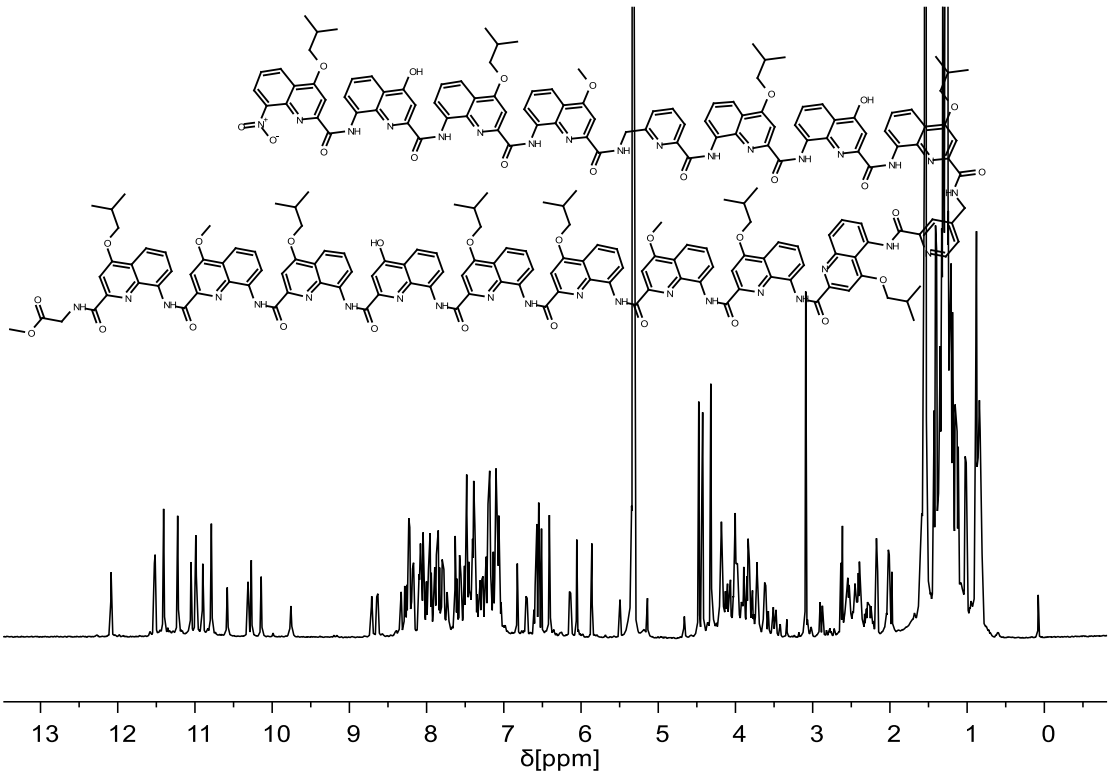
^{1}H NMR spectrum (500 MHz, $CDCl_3$) of compound **4a**.



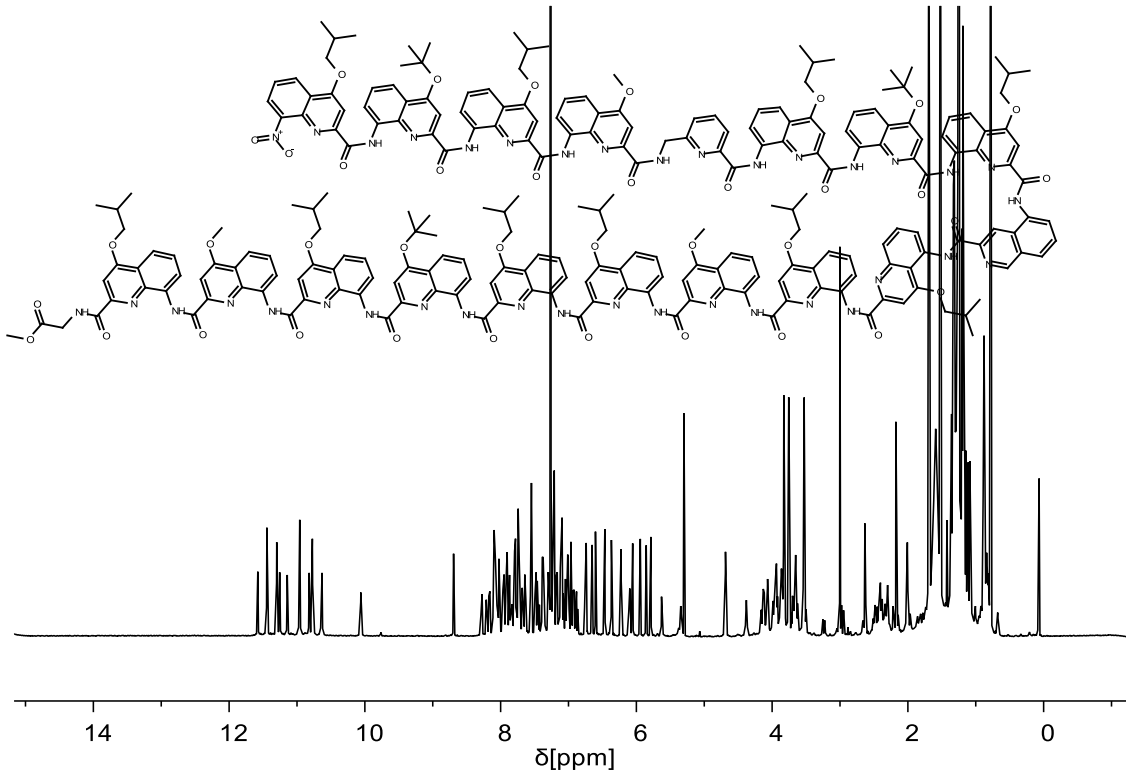
^{1}H NMR spectrum (500 MHz, $CDCl_3$) of compound **4b**.



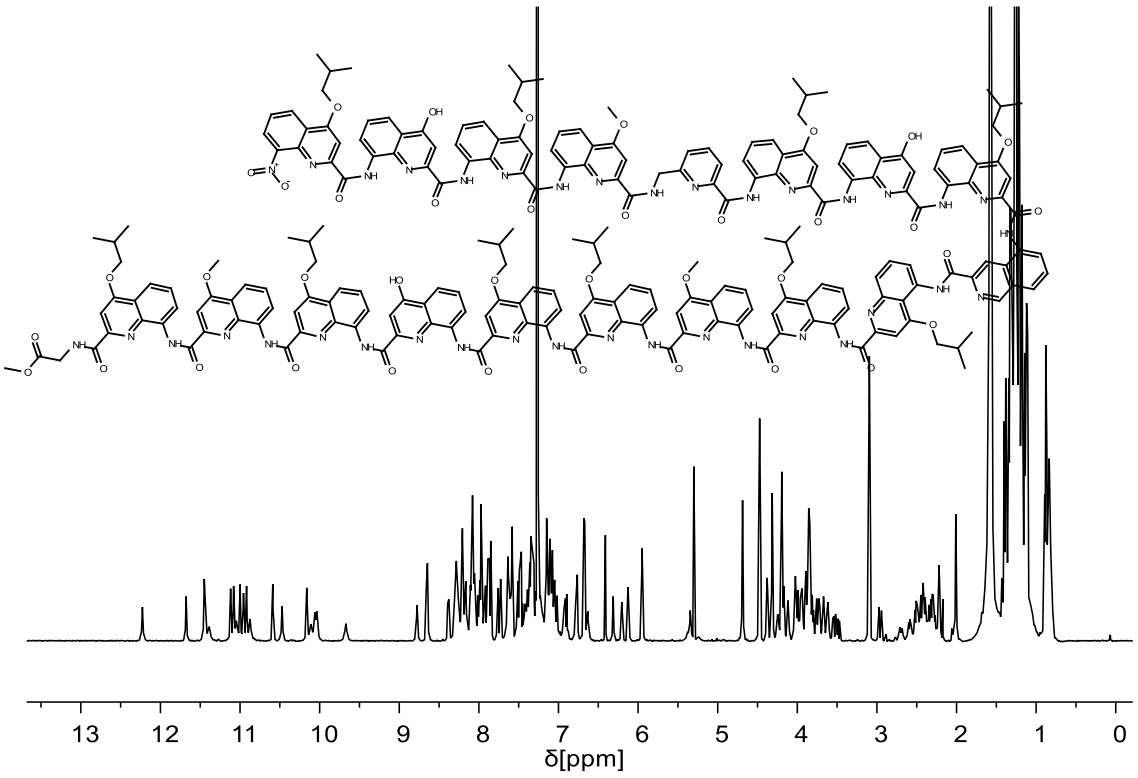
^{1}H NMR spectrum (500 MHz, $CDCl_3$) of compound **5a**.



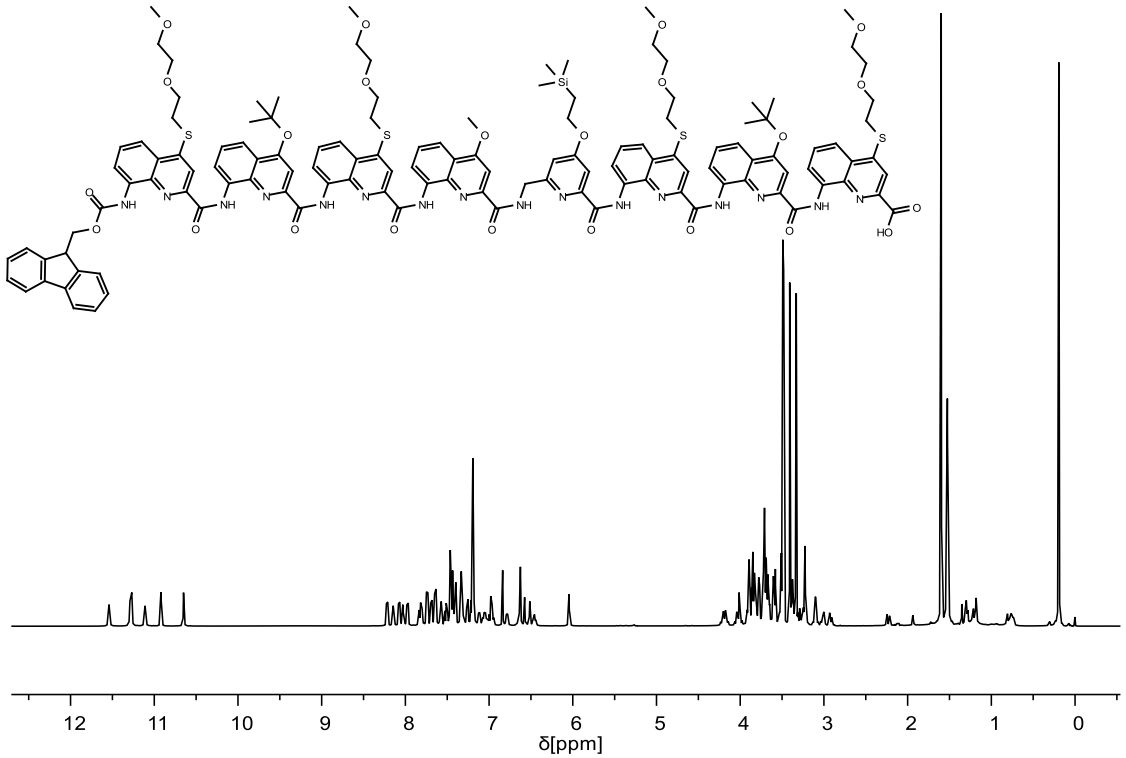
^{1}H NMR spectrum (500 MHz, CD_2Cl_2) of compound **5b**.



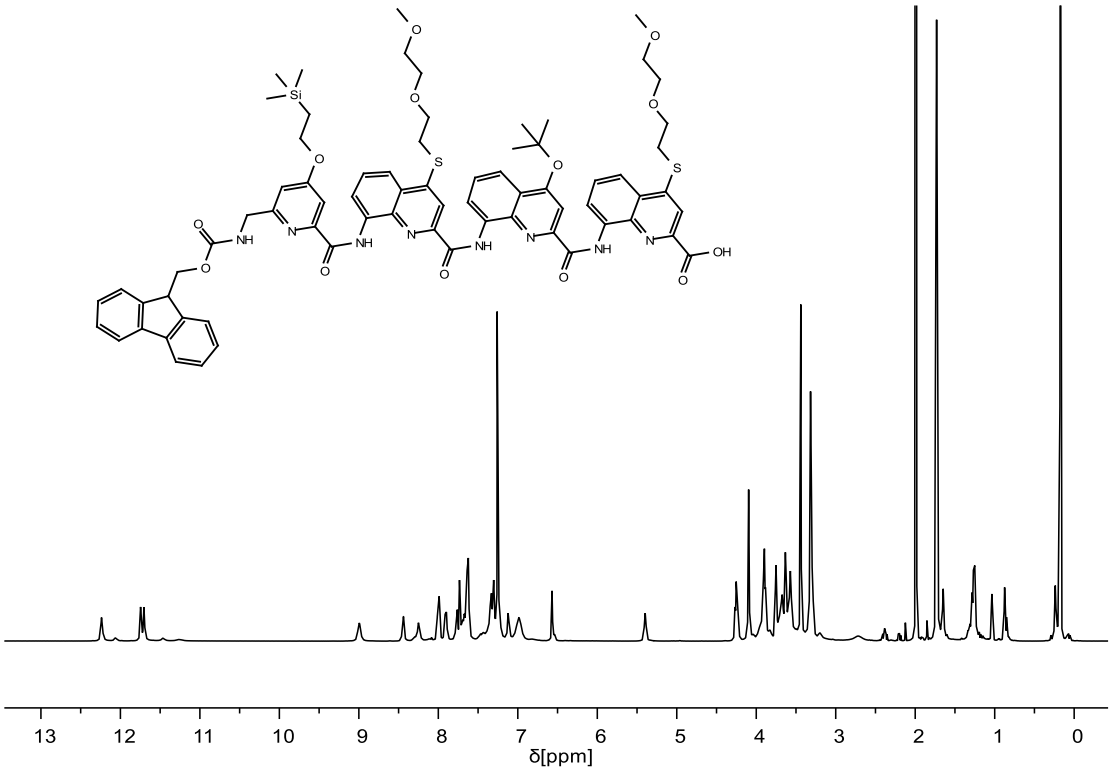
^{1}H NMR spectrum (500 MHz, $CDCl_3$) of compound **6a**.



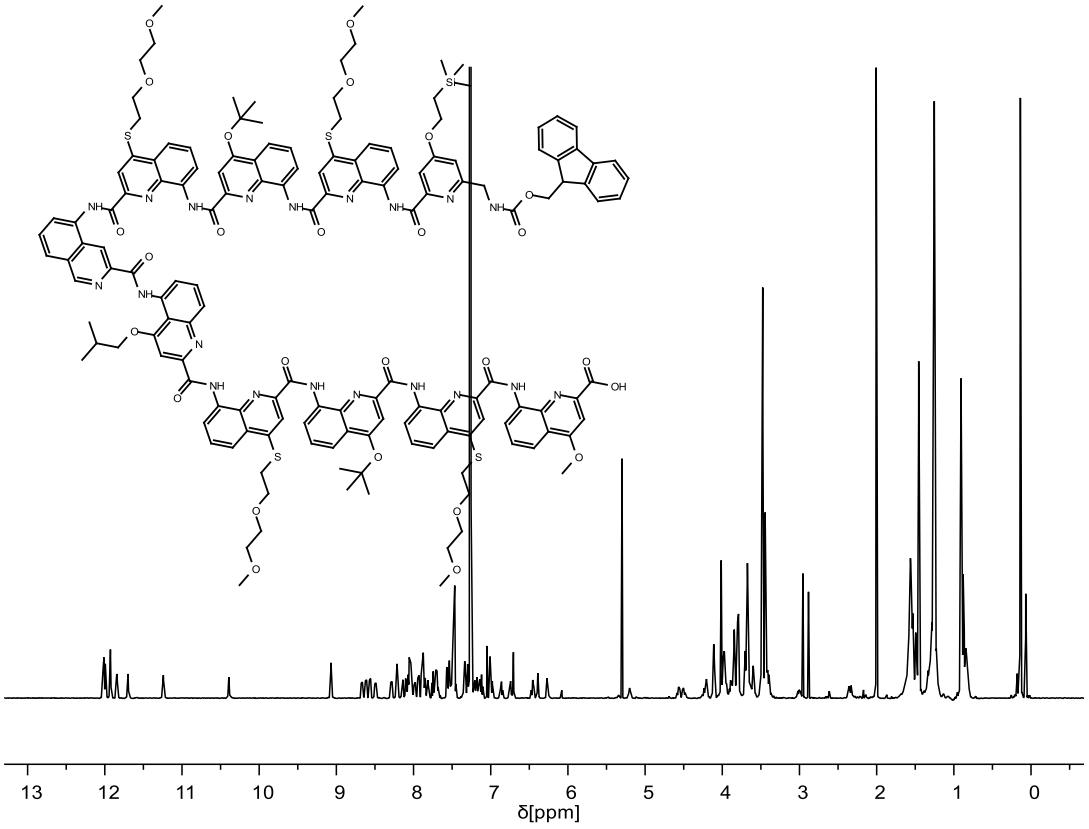
^{1}H NMR spectrum (500 MHz, $CDCl_3$) of compound **6b**.



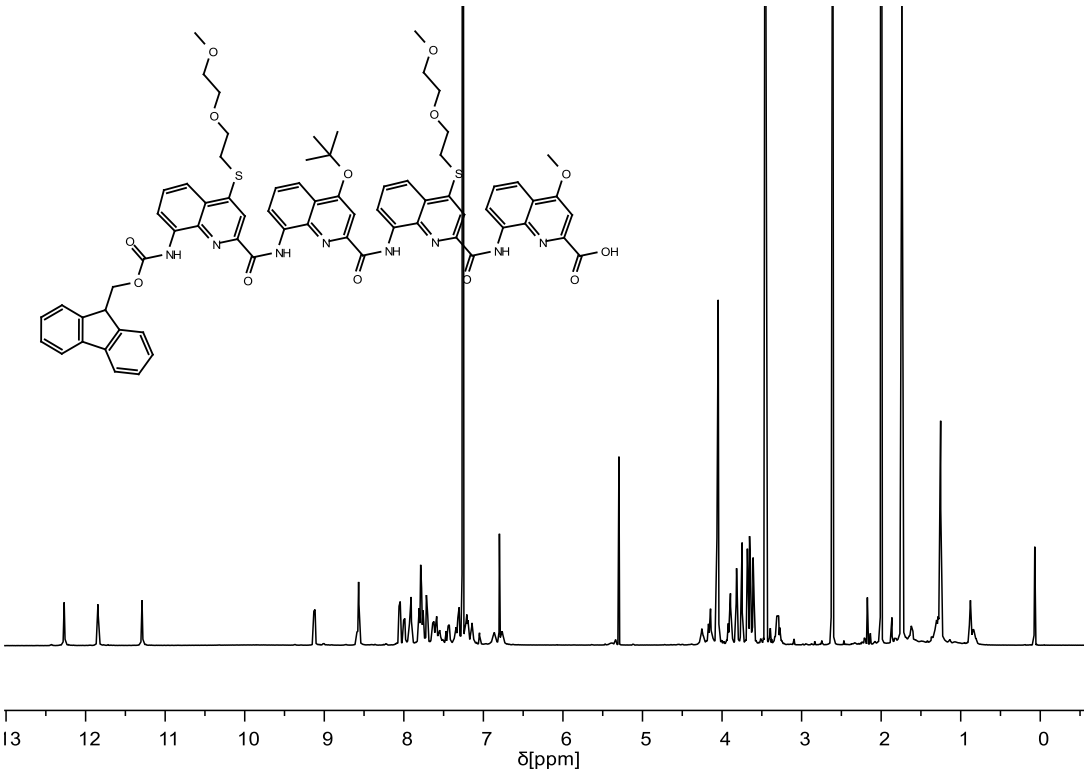
^{1}H NMR spectrum (500 MHz, $CDCl_3$) of compound **25**.



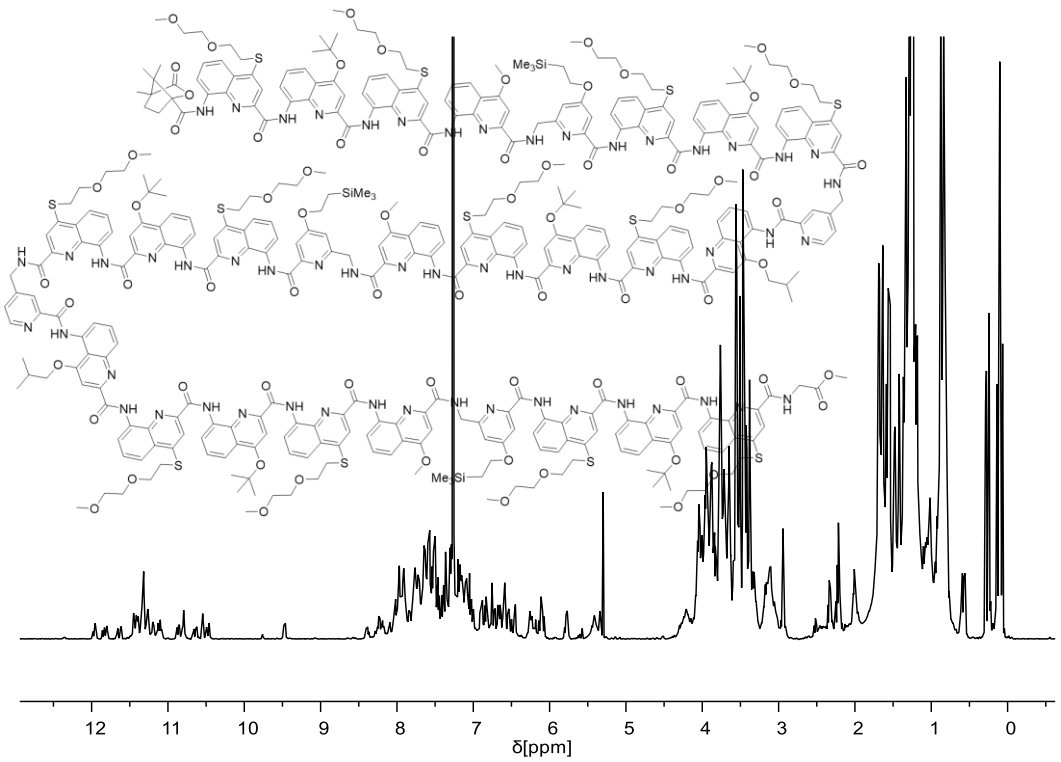
^{1}H NMR spectrum (500 MHz, $CDCl_3$) of compound **26**.



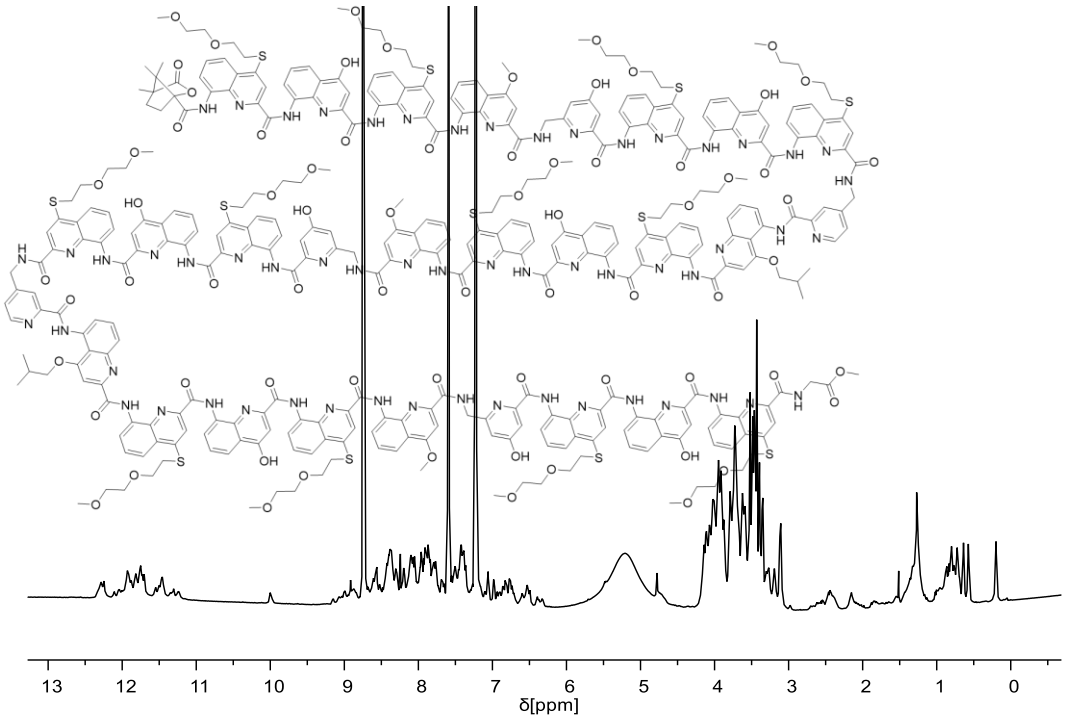
^{1}H NMR spectrum (500 MHz, $CDCl_3$) of compound **27**.



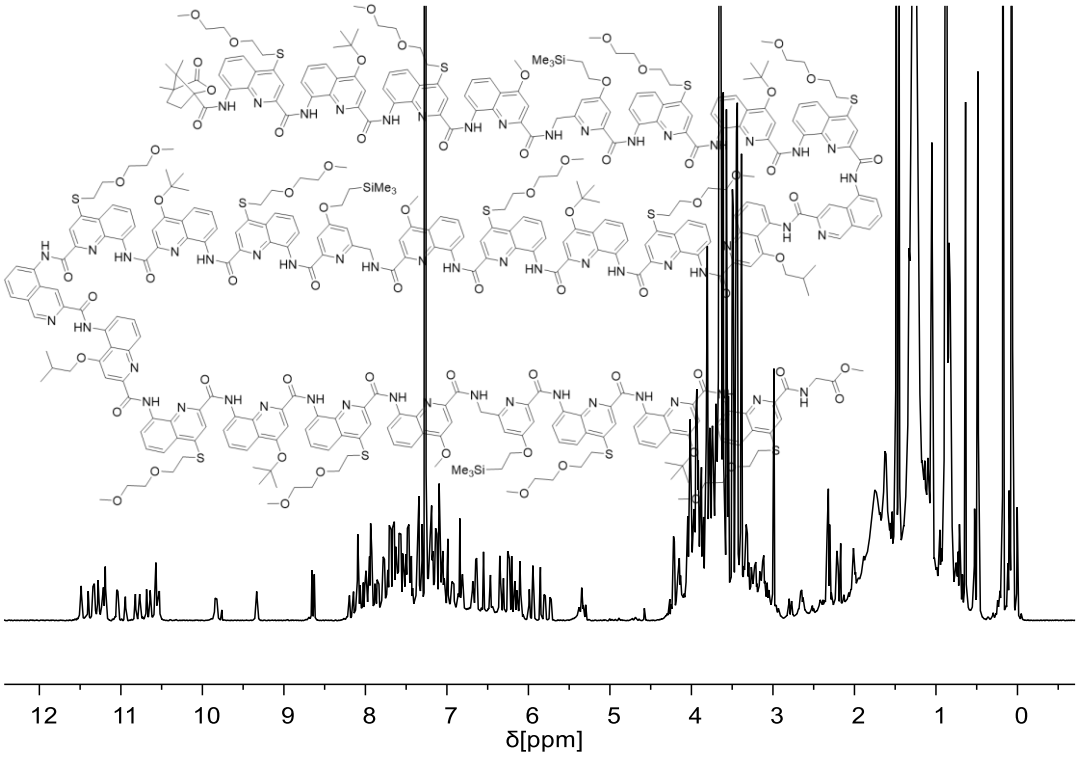
^{1}H NMR spectrum (500 MHz, $CDCl_3$) of compound **28**.



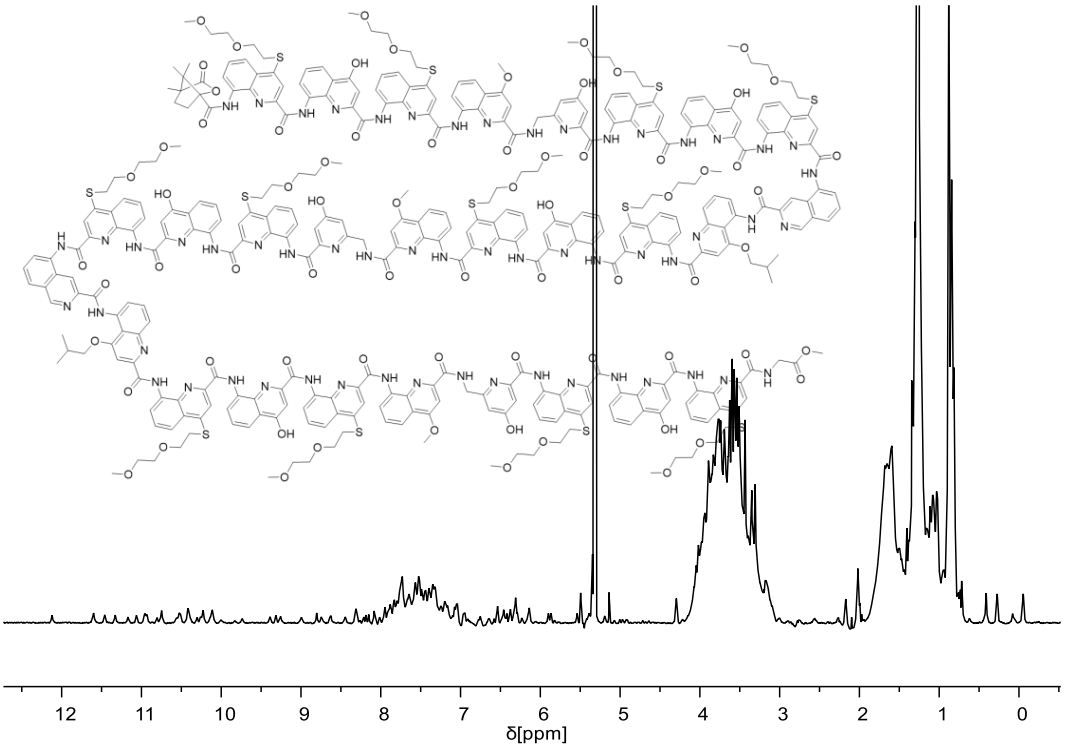
^{1}H NMR spectrum (500 MHz, $CDCl_3$) of compound **7a**.



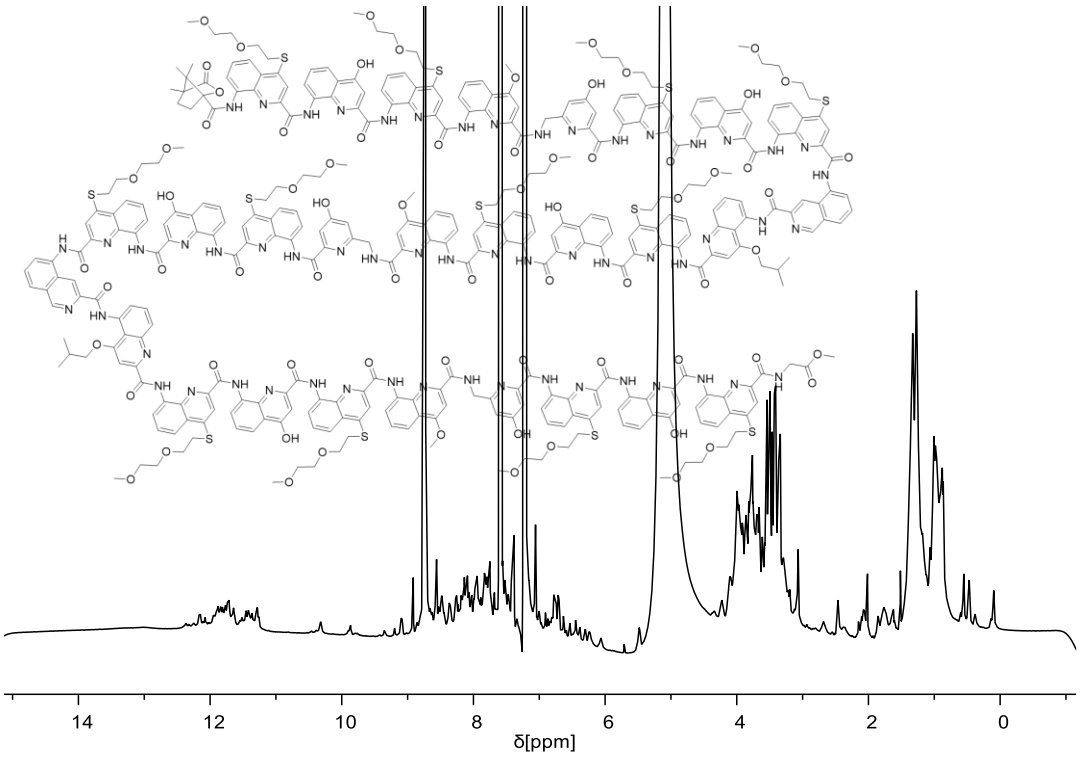
^{1}H NMR spectrum (500 MHz, pyridine-d_5) of compound **7b**.



^{1}H NMR spectrum (500 MHz, $CDCl_3$) of compound **8a**.

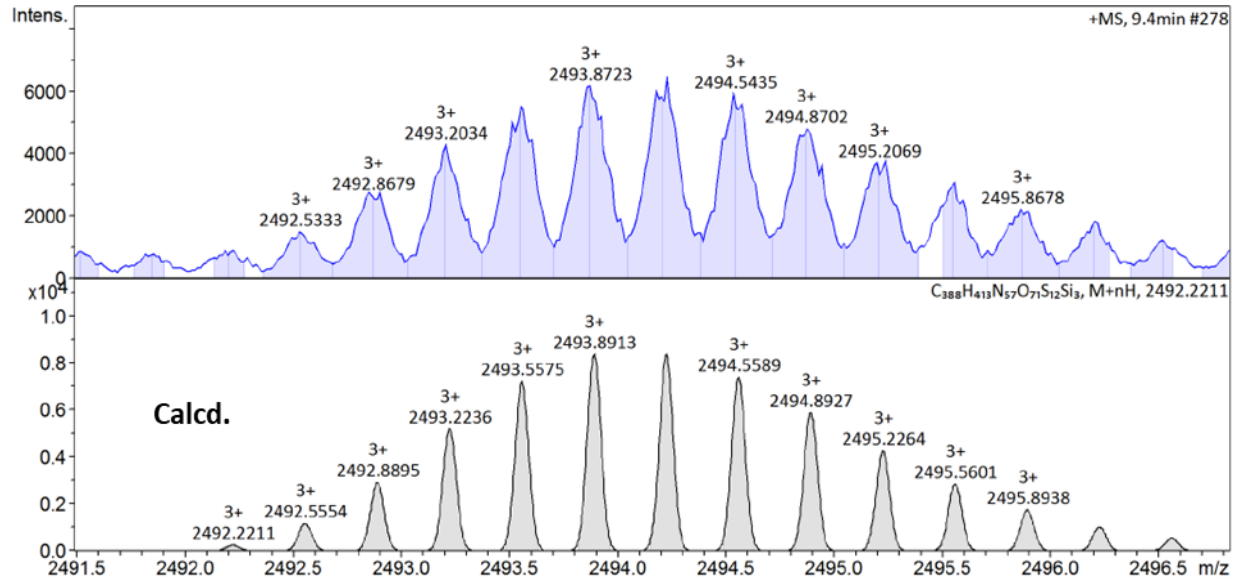


^{1}H NMR spectrum (500 MHz, CD_2Cl_2) of compound **8b**.

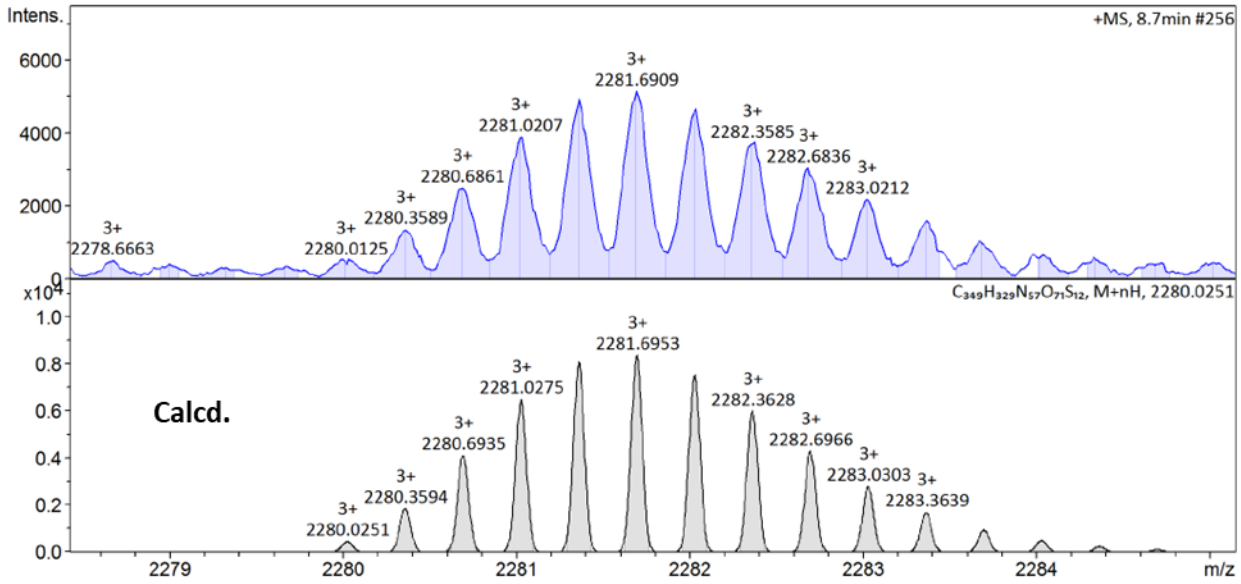


^{1}H NMR spectrum (500 MHz, pyridine-d_5) of compound **8b**.

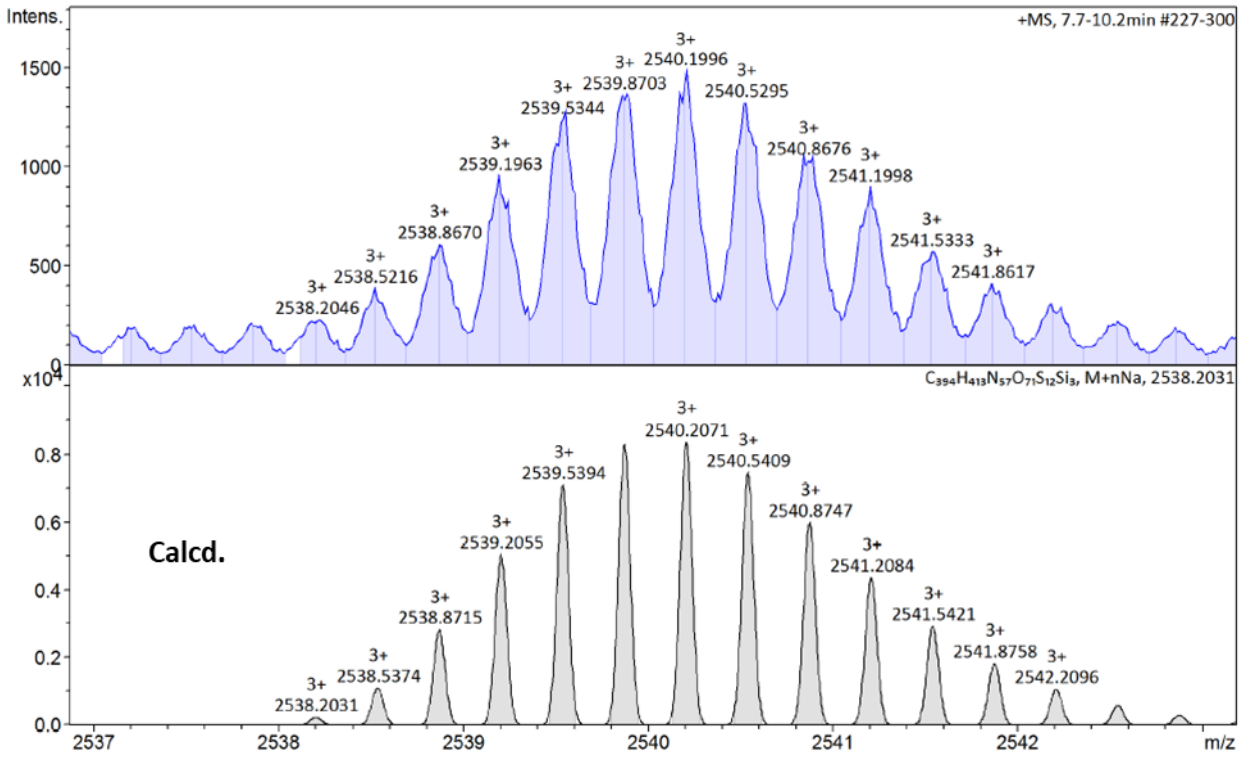
8 MS spectra of 7a, 7b, 8a and 8b



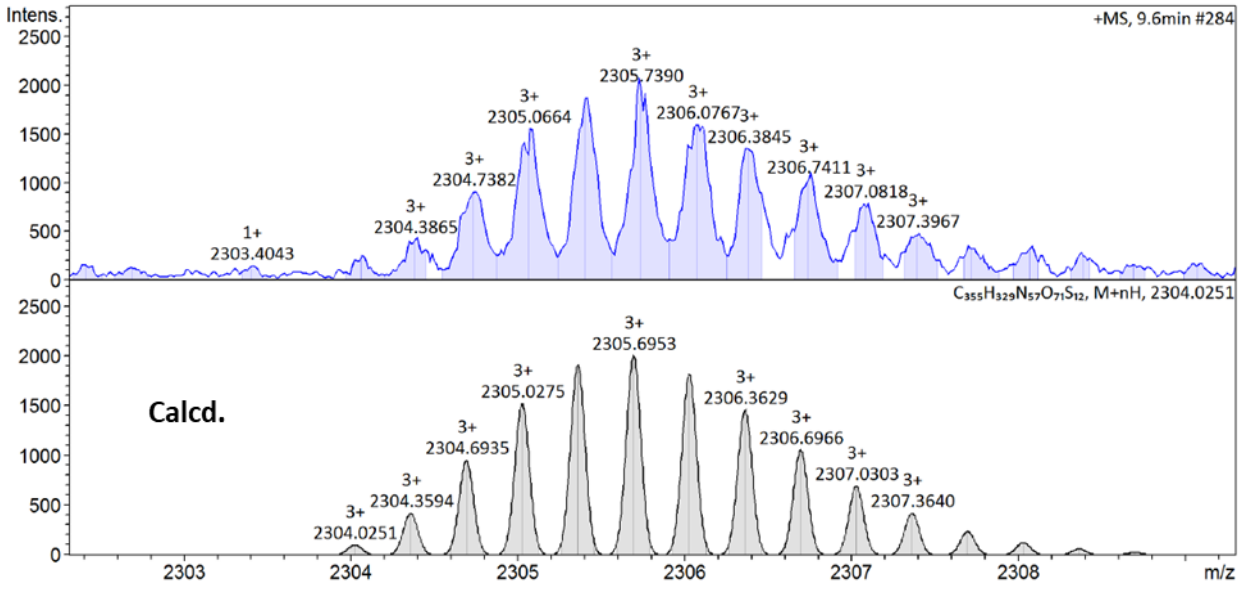
HRMS (ESI+) of compound **7a**.



HRMS (ESI+) of compound **7b**.



HRMS (ESI+) of compound **8a**.



HRMS (ESI+) of compound **8b**.

7. Summary and Prospective

7.1. Summary of publish/submitted works

This work broadened the boundaries of structural complexity in abiotic foldamers and allowed us to gain comprehensive knowledge in how to better control the interactions between foldamer helices through hydrogen bonds. First, the aggregation of previously characterized helix-turn-helix tertiary structures was investigated. Three additional hydroxy groups were introduced and positioned on one side of the helix-turn-helix structure to create a complementary array of hydrogen bond donors and acceptors that can potentially promote the formation of dimerized tertiary structures, *i.e.* genuine abiotic quaternary structure. Notably, the dimerization intermolecular interface was predicted by molecular modelling and had not been observed before. Solution studies supported the aggregation behavior; however, solid-state structures revealed a domain-swapped dimer that differed from our initial design. The conformation of the diacylhydrazine group of the turn unit twisted by approximately 90°, allowing the intramolecular interaction interface in the initial design to swap with another oligomer and form two "tilted dimer" intermolecular interaction interfaces. The domain-swapping dimer formed even though two hydroxy groups were not involved in hydrogen bonds, possibly due to the tilted dimer pattern being favored over the parallel arrangement in the intramolecular interface of the initial design. Gradually disrupting hydrogen bonds by slowly increasing the proportion of DMSO in chloroform would first lead to the dissociation of the domain-swapping dimer and the re-folding of the helix-turn-helix tertiary structure, followed by the disruption of intramolecular hydrogen bonds, resulting in "opened" monomers. This work reflected the ability of a folded structure to respond to changes and revealed a domain-swapped dimer structure which is more commonly observed in protein science.

Later, we explored the possibility of the trimerization of helix-turn-helix tertiary structures through intermolecular interfaces following a well-defined "shifted dimer" pattern. The intermolecular interaction interface between two helix-turn-helix structures was established by introducing two hydroxy groups at the N-terminus of one oligomer and one hydroxy group at the C-terminus of another oligomer. The molecular model of the dimer suggested the relative angle between two tertiary folds is close to 60°, which potentially allows the third tertiary fold to participate in the aggregation, leading to the formation of a symmetrical trimer. The molecular model of the corresponding trimer could be easily built with minimal conformational constraints. The oligomers were synthesized and the ^{1}H NMR showed two sets of signals, the proportion of which varied depending on temperatures or the amount of water in the chlorinated solvent. Other solution studies supported that they are aggregated helix-turn-helix structures in a "U" shape. Combined with molecular dynamic simulation studies, they were assigned to the initially designed trimer and unexpected hydrated dimer.

Finally, we aimed to design more complex unimolecular tertiary folds starting from existing aggregates. We selected a well-defined parallel trimer structure as our starting point and we planned to bridge the termini of adjacent helices to form a tertiary fold. Designing suitable linkers for the parallel trimer is particularly challenging because the three helix have the same handedness and orientation. By inverting the handedness and orientation of one helix, we can swap the C-terminus and N-terminus while preserving the side chain positioning. A molecular model of the heterochiral trimer was built, and its intermolecular interaction pattern is very similar to that of the homochiral trimer. Based on this model, we designed and synthesized two linkers with different flexibilities. The properties of the two linkers were first validated with two series of model sequences, Q_3-turn-Q_3 and Q_8-turn-Q_8. The solution studies and solid-state structures suggested both two linkers could promote the formation of heterochiral helix-turn-helix as expected, even though the hydrogen bond patterns slightly differed from the initial design. The final sequences are relatively long and non-polar, which posed challenges during purification. Therefore, we adopted a fragment condensation strategy: the final sequences were divided into different fragments based on the different linkers used. These fragments are relatively short and can be purified by RP-HPLC. The fragments were then coupled on the solid phase accordingly using BOP activation. The synthesized final sequences were cleaved from the solid phase and purified by size exclusion chromatography, such as gel permeation chromatography (GPC). This strategy allows high-purity, relatively large oligomers (6.9 kDa) to be obtained in a relatively easy manner. Solution studies of the final sequences indicated that only the sequence with a relatively rigid linker could fold into a well-defined species, which could be assigned to the three-helix tertiary structure as designed.

7.2. Continuing challenges and perspectives

The synthesis of all Fmoc-protected monomers has been well-optimized. Synthesizing relatively long oligomers using solid-phase foldamer synthesis is no longer challenging (which was not the case when all oligomers were synthesized in solution), especially with the use of an automated foldamer synthesizer. The purification method using RP-HPLC is also well-established, allowing oligomers to be obtained with high purities. For the oligomers that are not compatible with RP-HPLC, the fragment condensation strategy that is developed and described in Chapter 6 provides a solution to such a problem. Overall, there are currently no serious barriers to the preparation of oligomer sequences in the 5-10 kDa range.

The challenges primarily focus on sequence design. All oligoamides in this series consist of four monomers (Q, X, Y, and P), occasionally incorporating one or more linker units. The different arrangements of these four monomers, the effects of the solvent, and the constraints imposed by the linker units lead to various hydrogen bond patterns, including the formation of tilted dimers (counter-clockwise and clockwise), shifted dimers, parallel trimers, and parallel dimers. It is fascinating that such distinct structures can be obtained through different arrangements of a few monomers. However, on the

contrary, the sequence design is very delicate and small changes in sequence can lead to large changes in folding or aggregation.

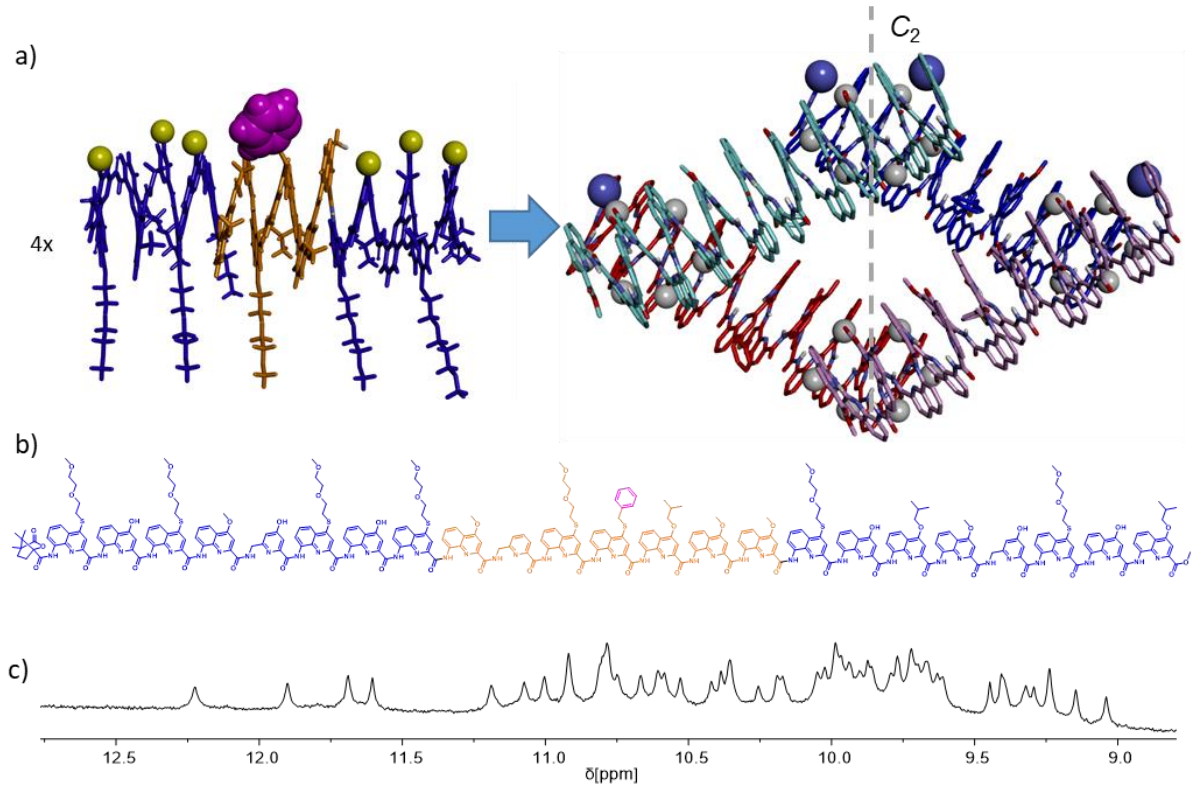


Figure 16. . a) The molecular model of the helix that has two hydrogen bond surfaces (the hydroxy proton is shown in yellow). The color code of the helix is consistent with that of the chemical structure shown in (b). The molecular model of tetramerized (a) through four clockwise tilted dimer intermolecular hydrogen bond interfaces. The N-terminus is shown as a blue ball. Four identical helices are colored red, blue, pink and cyan, respectively. (b) The chemical structure of the helix is shown in (a). c) Extract of ^{1}H NMR (500 MHz, $CDCl_3$, 25°C) of sequence shown in (b). It shows two sets of signals which could be assigned to the helices colored in cyan/blue and red/pink, respectively.

Moreover, while the influence of the solvent can be beneficial when we wish to drive the sequences to aggregate into different structures by altering the solvent, it also introduces instability and unpredictability. For example, different chlorinated solvents can quantitatively drive the formation of homochiral and heterochiral shifted dimers, respectively, while tilted dimers and parallel trimers often coexist. Controlling the quantitative formation of a particular interface is advantageous for designing self-assembled structures based on these interfaces. For instance, a tetrameric bundle stabilized by four tilted dimer interfaces could be computationally designed (Figure 16a,b). All parallel arrangements are prohibited due to steric clashes caused by the bulky groups (shown in pink in magenta in Figure 16a,b) located in the middle of the helix. The preliminary results (Figure 16c) are encouraging, and this project is currently being continued in our lab by Dr. Pronay Biswas. As we have seen in this thesis, when

targeting higher structural hierarchies and complex topologies, a lack of strong interaction pattern preference may lead to unexpected aggregation or folding. While extending the two termini of a helix can easily prevent parallel arrangements, as in the tetrameric bundle design shown in Figure 16, quantitatively preventing the formation of one of the two types of tilted dimers (counter-clockwise and clockwise) remains challenging. The research which is currently conducted in our lab by Dr. Pronay Biswas reveals that two sites of P-to-Q mutation at two termini can achieve this goal.

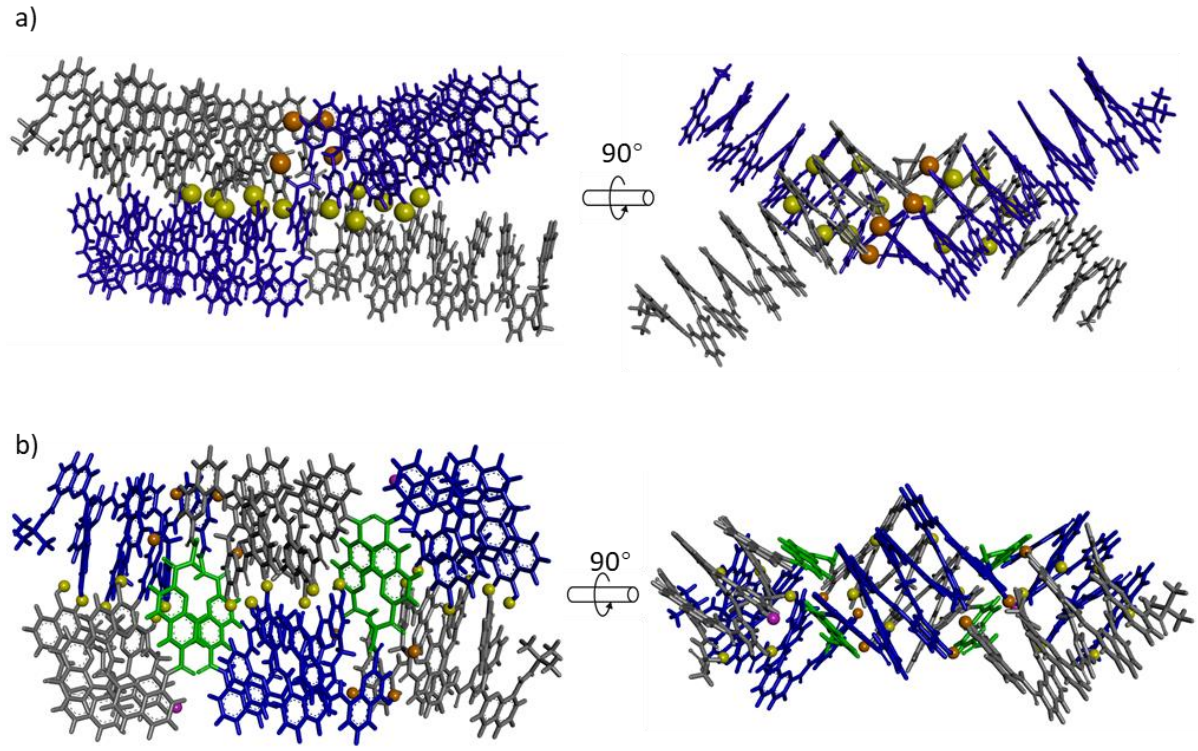


Figure 17. a) The molecular model of domain swapping dimer with two termini prolonged with oligo-Q. b) The molecular model of domain swapping dimer with N-terminus prolonged with a turn-helix motif. Two identical sequences are colored in blue and grey, respectively.

The sequence of the unexpected domain-swapped dimer contains two helices connected by a turn unit T2. In the domain-swapped dimer, the desired intramolecular hydrogen bond interface was shown not to form as designed; instead, intermolecular hydrogen bonding interfaces are formed. Since the four helices in the domain-swapped dimer point in four different directions, one possible design is to prolong the termini. The desired structure will contain four overhanging termini that facilitate modifications for further applications (Figure 17a). Another attractive design is to extend the N-terminus of the sequence with a helix-turn motif. The corresponding sequence is a helix-turn-helix-turn-helix structure. This sequence is expected to swap domains with another identical sequence three times to form a larger domain-swapping dimer (Figure 17b).

As reported in Chapter 5, the exploration of the aggregation of helix-turn-helix structures revealed the co-existence of trimers and hydrated dimers. The dimer may not be desired in projects that exploit the large cavity of the trimer. By introducing 6-bromo substituted X monomers (Figure 18a) into the sequence, modelling predicts that the formation of the hydrated dimer will be prevented due to steric hindrance, while not affecting the formation of the trimer (Figure 18b). Modification of the monomers toward the cavity will endow the aggregate with functions, such as catalysis.

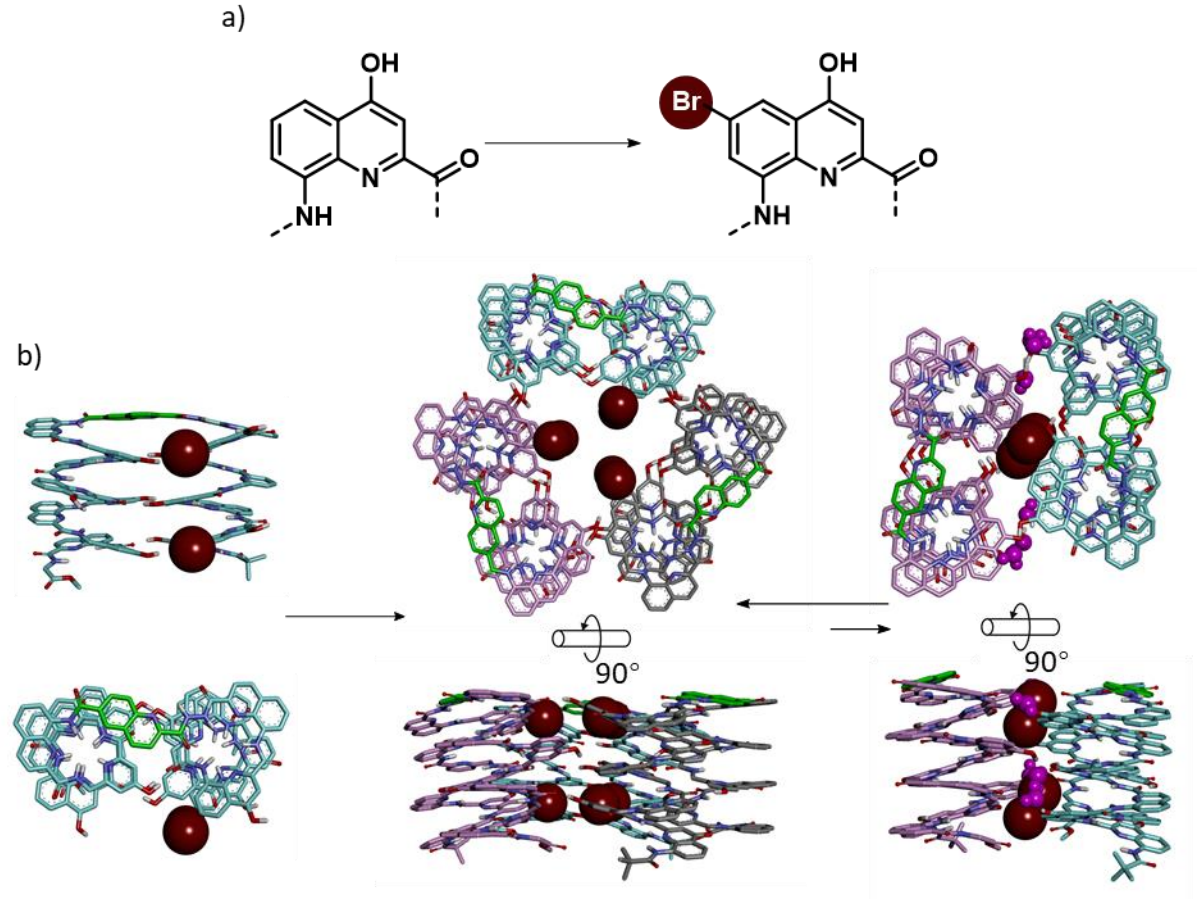


Figure 18. a) The structure of X monomer and 6-bromo substituted X monomer. b) Molecular model of helix-turn-helix structure with two 6-bromo substituted X monomers. The bromo atom is shown as a brown ball. The steric clash caused by the extra bromo atoms will prevent the formation of the dimer and favor the trimer.

8. Acknowledgements

I would like to express my deepest gratitude to my supervisor, Professor Ivan Huc, for his dedicated guidance throughout my PhD studies. I am grateful for the opportunity he provided for me to be part of such an outstanding team. His invaluable guidance, encouragement, and patience have accompanied me through my entire PhD journey.

I also want to thank all the members of the group for creating a collaborative work environment. My special thanks go to Céline Douat for her valuable advice on solid-phase synthesis, HPLC, and MS analysis. I want to thank Victor Maurizot for his guidance on project design and molecular modeling. I want to thank Lars Almmendinger for NMR spectra measurement and for his valuable advices in NMR data analysis and interpretation. Thanks to Barbara Wicher and Johannes Sigl for their help with X-ray diffraction and data analysis. I am also grateful to Lizeth Bodero and Maryam Klidsar for their assistance with using the automatic synthesizer, Thanks to Ivan Alonso for his technical support. I would like to thank Agnieszka Breyer for making campus life easier for me. I want to thank Friedericke Menke for her advice on project design and for introducing me to the lab's instruments. Thanks to Sudhakar Gaikwad for his guidance in organic synthesis. I am also thankful to Binhao Teng for his help with research and daily life. I am grateful to Arundhati Roy, Florian Sanchez, Alexandru Grozavu, Pronay Biswas, Maria Lago Silva, Mathieu Denis, Artheswari Gunanithi, and Ilaria Giussani for their support and for creating a warm and friendly environment. Additionally, I thank Susnata Pramanik, Pradeep Mandal, Daniel Bindl, Sebastian Dengler, Kathrin Aftahy, Deepak, Mahshid Alizadeh, David Konrad, Vasilii Morozov, Ryan Howard, Lingfei Wang, Wei Tuo, Robin Hess, Manuel Loos, Jiaojiao Wu, Melis Cabbar, Cyrille Rivière, Tulika Chakrabortty, Ignacio Muñoz Alonso, Yajuan Ma, Klein Hubert, Ashley Roby, and Vasileios Athanasopoulos for their support. Many thanks to Mariia Palchyk and Daniel Gill for their help with monomer synthesis, and to Lisa Gourdon-Grünewaldt for her support with IT issues. I also want to give a special thanks to Niklas Böcher and Nikolas Schneider for their help with various German language issues in my daily life.

Finally, I am deeply grateful to my friends and family, who have supported and encouraged me along the way. I am especially thankful to my parents, whose support has made this doctoral thesis possible. Special thanks go to Weijia Zhang, who has always motivated me and provided support in every aspect.