Dissertation zur Erlangung des Doktorgrades

der Fakultät für Chemie und Pharmazie

der Ludwig-Maximilians-Universität München

Folding properties, molecular recognition and aggregation behavior of aromatic oligoamide helical capsules in water

Binhao Teng

aus WenZhou, Zhejiang, China

2023

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, 28.06.2023

.....……

Binhao Teng

Dissertation eingereicht am 28.06.2023

1. Gutachter: Prof. Dr. Ivan Huc

2. Gutachter: Prof. Dr. Franz Bracher

Mündliche Prüfung am 23.06.2023

Table of contents

List of publications

Accepted:

B. Teng, J. Atcher, L. Allmendinger, C. Douat, Y. Ferrand, I. Huc*. Combining local conformational preferences and solvophobic effects in helical aromatic oligoamide foldamers. *Org. Biomol. Chem.*, **2023**, 21, 3525.

(See chapter 4)

Submitted:

B. Teng, L. Allmendinger, C. Douat, Y. Ferrand, I. Huc*. Endothermic molecular recognition within helical aromatic oligoamide capsules in water. *Chem. Sci*. **2023**, submitted.

(See chapter 5)

B. Teng, P. Mandal, L. Allmendinger, C. Douat, Y. Ferrand, I. Huc*. Controlling aromatic helix dimerization in water by tuning charge repulsions. *Chem. Sci*. **2023**, submitted.

(See chapter 6)

1 Abstract

A foldamer is a discrete chain molecule (oligomer) in chemistry that folds into a conformationally ordered state in solution. Local conformation preference is primarily responsible for the well-folded structure, however, hydrophobic effects also play a significant part in the folding process. As the secondary and tertiary structures of biomacromolecules have shown biological activity, the distinct aggregation patterns of foldamers that led to complex structures are significant. It may also result in the development of novel chemical or biological activities. Narrow helical foldamers have been intensively studied over the past decade, but a strategy for rapidly achieving a large cavity helix is still urgently required. The new building blocks not only alter the geometry of foldmers but also their aggregation behaviors; additionally, the novel features brought by the enormous diameter require additional investigation.

This study introduces a novel class of aromatic δ-amino acid foldamers based on 7-amino-2-quinolinecarboxylic acid (Q^H). It was proved that these more flexible units may be folded well-ordered with more than two helical turns in water, with the stability increasing as the length of the main chain grew. The hydrophobic effects have a significant impact on the folding process; hence, the Q^H unit could not be folded in DMSO. In addition, a number of unique dimerization patterns (head-to-head, parallel double helix, and potential tetrameric helix) were discovered by altering the side chains, which may be attributed to electron repulsion and desolvation effect. Thus, these observations enable designs that enlarge the capsule cavity without modifying the capsule's backbone. In aqueous solution, the study discusses the discovery of a binding affinity between water-soluble guests and amphiphilic hosts. Due to the release of highly ordered water in the cavity, an unusual endothermic binding process that was entirely driven by entropy was observed. This particular molecular recognition can be utilized for molecular separation and reaction containers.

In conclusion, our findings constitute a significant reference for the backbone flexible sequence helical folding and the construction of large cavity foldamers for guest binding in aqueous media. In addition, various aggregation patterns and the first capsule crystal grown in water were observed. The solid state and cavity binding affinity research could serve as a starting point for future designs and help develop several different capsules for specific purposes.

2 Introduction

Molecular recognition in vivo could be achieved through the use of complementary shapes to carry out diverse biological functions, such as enzyme catalysis, signal transduction, antigen recognition, and substance transport, without the need for harsh conditions or transition metal catalysis, as in organic chemistry. Some enzymes can accelerate their substrate-to-product conversion by millions of times. Orotidine 5'-phosphate decarboxylase is an extreme example, as it permits a process that would ordinarily take millions of years to occur in milliseconds.[1,2] High specificity and structural diversity are necessary for achieving efficient and diverse catalytic processes. How can such structural variety be accomplished? The solution is given by the nature: folding. Using a combination of noncovalent interactions, various functional groups are positioned in distinct locations to produce various biological functions. To fully comprehend the folding process and establish novel structures and functionalities, a series of artificial backbone structures capable of folding into well-defined configurations have been constructed. Foldamers are a significant class of these chemicals.

Foldamers have been utilized extensively in chemical catalysis, biomolecular simulation, functional materials, etc., in recent years. Water serves as the most prevalent chemical environment in biological systems, hence more and more researchers have been working on it. To imitate molecular recognition in vivo, it is necessary to have large cavities and higher-order structures. This thesis describes capsule structure procedures while concentrating primarily on aqueous solutions. The significant hydrophobic effect and intense hydrogen bond competition make the folding and aggregation behaviors of foldamers in water completely distinct from those in organic solvents. This study gives an introduction to foldamer science with an emphasis on aromatic amides as building blocks.

2.1 Introduction to binding cavity

In biochemistry and molecular biology, a "binding site" refers to an area on a macromolecule, such as a protein, that binds specifically to another molecule. Most biomolecular binding cavities are only partially open, which facilitates binding and release. The binding site of a protein always possesses a pocket-like structure and catalyzes with great efficiency.[3] To imitate natural functional binding sites, a series of hemi-opened artificial cavities were created. The lab of Wang-Mei reported the synthesis and structural analysis of the S6-Corona[3]arene [3]pyridazines macrocycle (Figure 1a), which can form stable 1:1 complexes with dicationic guest species by ionic bonding in water.[4] And based on the architectures of Pillar[n]arenes (Figure 1b), a number of distinct substances exhibited affinity in various chemical environments due to π-π stacking, hydrophobic effect, etc.[5] Bowl-type structures are another kind of cavity system. Sealing one side of the macrocycle increases the number of binding sites in the cavity. Professor Rebek and his colleagues used resorcinarene to construct the cavitands; the various wall constructions were stabilized by hydrogen bonds to prevent flipping and rotation (Figure 1c).[6] Self-assembly was also utilized in the formation of cavities. Through π-π stacking of distinct layer macrocycles, the group of Bing Gong showed the tubic type cavity. Using guests of certain lengths, the tube's length could be precisely regulated and unspecific polymerization could be prevented (Figure 1d).[7]

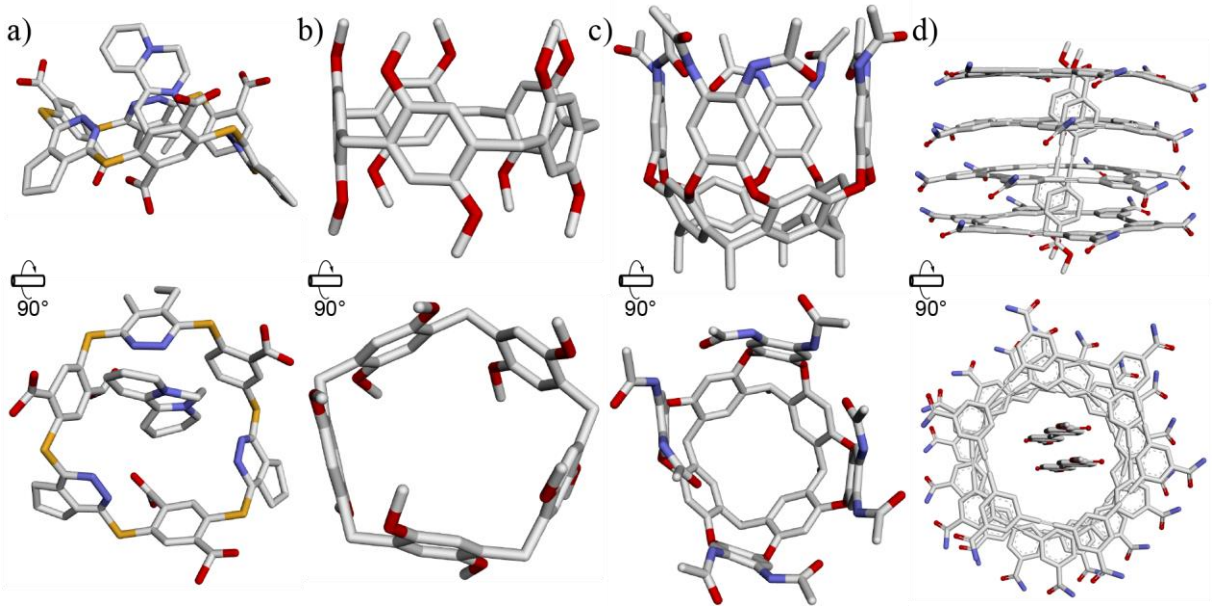


Figure 1. Overview of different hemi-opened cavities. a) crystal structure of S6-Corona[3] arene [3] pyridazines macrocycle. b) crystal structure of Pillar[5]arenes macrocycle. c) crystal structure of bowl shape cavitand. d) minimized model for tubic cavity.

Meanwhile, totally sealed cavities were constructed by connecting two bowl structures via halogen bond interactions (Figure 2a).[8] Self-assembly is an additional method for building a completely isolated system. Jonathan Nitschke and colleagues use in situ carbonyl-amine condensation to generate imines that chelate with metal ions and self-assemble into a cage structure that is separated from the bulky solvent environment (Figure 2b).[9] It might be utilized to stabilize active intermediates and prevent their separation. Additionally, unnatural aromatic amides were employed to develop cavities. Conformational bias allows the aromatic peptide to fold into a well-ordered helical shape, resulting in a large cavity inside. Two hemi-capsule strands can form a double helix cavity by capping one helical terminal (Figure 2c).[10] And upon capping both ends, a single helical aromatic amide capsule was formed (Figure 2d).[11] By hydrogen bonding and the solvent's phobic effect, a number of polar compounds are able to form internal bonds.

As a result of the variable geometry and rationally designable features of cavities, an increasing number of researchers are working on foldamers. In addition to aromatic amides, various other chemical units and linkages (e.g., hydrazide,[12] urea,[13] and alkyne[14] groups) were employed to construct self-assembly and a well-defined structure, which incorporated varied chemical qualities into the binding cavities.

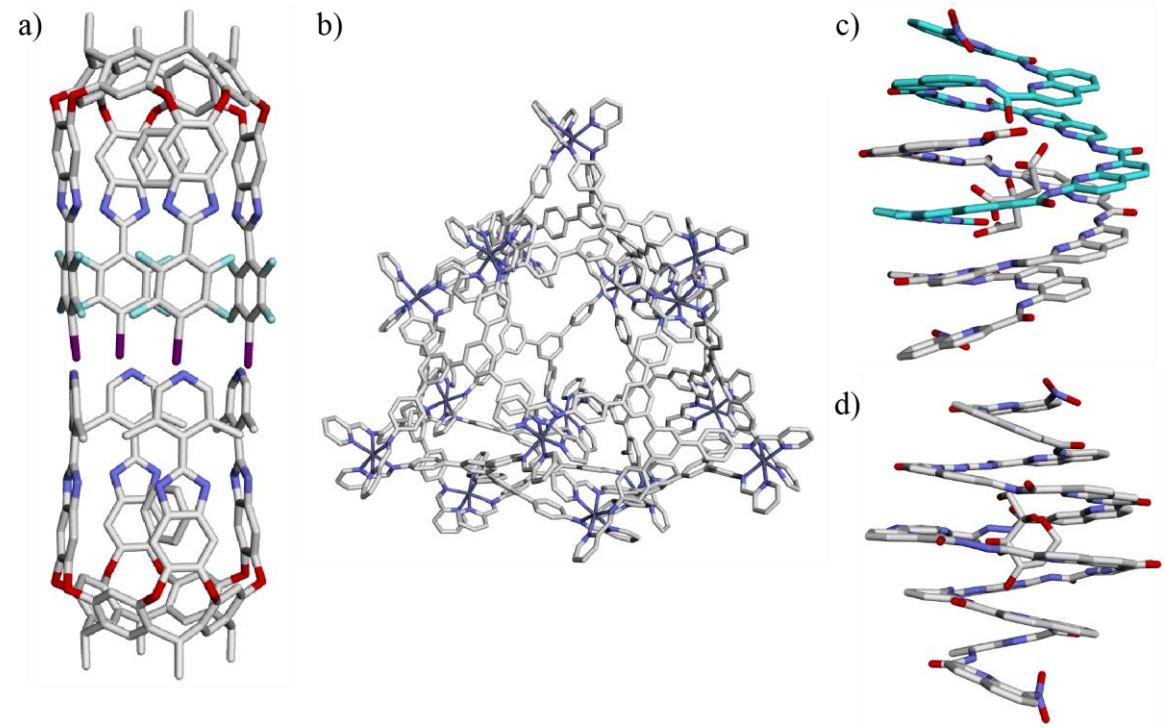


Figure 2. Overview of different sealed cavities. a) Optimized structure of double bowl cavity. b) Crystal structure of metal coordinated cage. c) Crystal structure of a double helix aromatic oligoamide. The different strands were colored in white and cyan, respectively d) Crystal structure of a single helix aromatic oligoamide.

2.2 Application of supramolecular cavities

The expression of biological processes such as enzyme catalysis,[15] cargo transport,[16] and immunological response[17] in confined nanospaces, which have vastly different physical and chemical conditions than those of bulk solution, is fundamental to living organism. In past decades, the fabrication of artificial nanoconfined containers through the use of molecular engineering has become a prominent focus for many molecular scientists.[18–20] These molecular cavities, which have specific structures that are stereochemically well-defined, have shown to be effective platforms for constructing restricted nanospaces.[21] The synthetic freedom afforded by molecular containers enables us to control cavity sizes, functional group decoration, and electronic atmosphere distribution, hence enabling the creation of adaptable binding pockets whose characteristics and functionalities are tuned to meet specific requirements (Figure 3).[22]

The fundamental mechanism of enzyme catalysis is stabilization of transition states and energy barrier reduction.[23] At the binding sites, the stereoelectronic distributions display complementarity with the transition state of the substrate to advance the reaction toward the product. Simulating the catalytic properties of enzymes with simple synthetic receptors remains challenging. It has been proven that mechanical distortions affect the reactivity enhancement of a substrate within a molecular container.[24] When the amide is encapsulated in the cavity, the solid-state analysis demonstrates that these containers favor the twisted conformation of the amide bond over the planar one. The encapsulated amides can be twisted by up to 34° which considerably increases their hydrolysis rates in the presence of bases (Figure 4a).

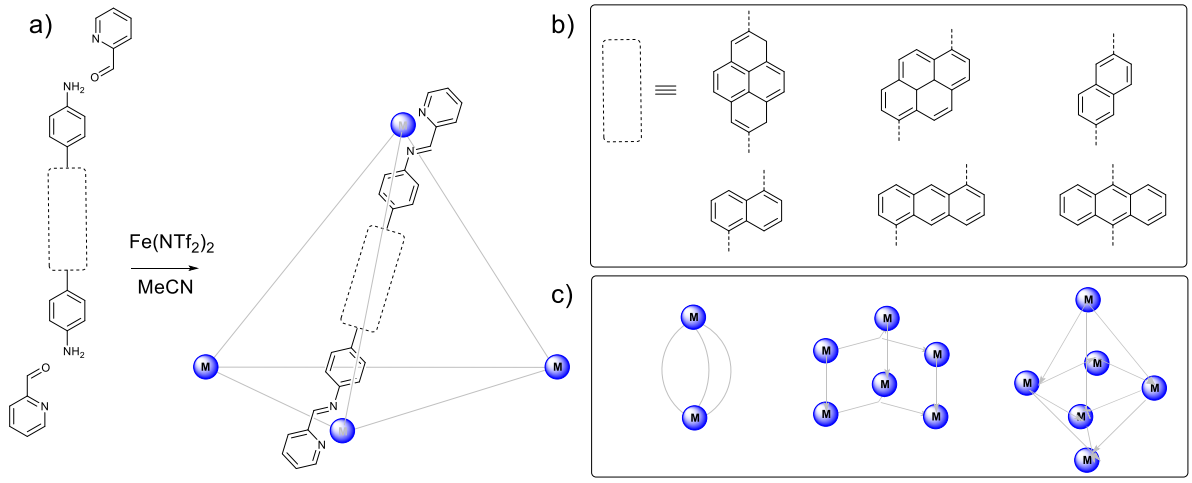


Figure 3. a) Self-assembly of M_4L_6 cages from various ligands combinding dynamic covalent and metal coordination bonds. b) The different ligands can define the sizes, shapes, and apertures of the binding cavities. c) The diverse combinations of building blocks result in various metal-organic containers.

An artificial Diels-Alderase reactor also have been reported;[25] the container polarizes the dienophile, hence activating its reactivity upon binding. Additionally, it can accommodate and stabilize the transition state (Figure 4b). In comparison to the uncatalyzed reaction in bulk solution, the cavities are capable of accelerating the reaction rate by up to a factor of 1,000. In addition to boosting the rate of the reaction, this container also alters the regio- and chemo-selectivities.

Additionally, the cavities can be used to stabilize reactive species. Once contained within the cavity, a reactive species is sheltered from the bulk environment, providing the foundation for its stability (Figure 4c). Hao Li's research has shown that synthetic containers may isolate the naphthalene diimide radical anion and prevent its dimerization in water; as a result, the radical anion is shielded against disproportionation in a noncovalent method.[26] Additionally, encapsulation makes possible the investigation of an unstable molecule. Mass spectrometry study of sulfur clusters has been a technological difficulty due to their structural instability, which leads to their instantaneous and complete fragmentation. By encapsulating sulfur clusters in a container, fragmentation can be avoided (Figure 4d). A coordination container is capable of enclosing S_6, S_8, and S_{12} clusters, the complex could be observed by ESI-MS.[27]

Highly ordered covalent organic frameworks of supramolecular cavities make the nanospace confinement possible in solid state study. Consequently, these extended frameworks have been utilized actively as porous materials, and their limited nanospaces have been used for applications in gas storage,[28] molecular separation,[29] and catalysis[30].

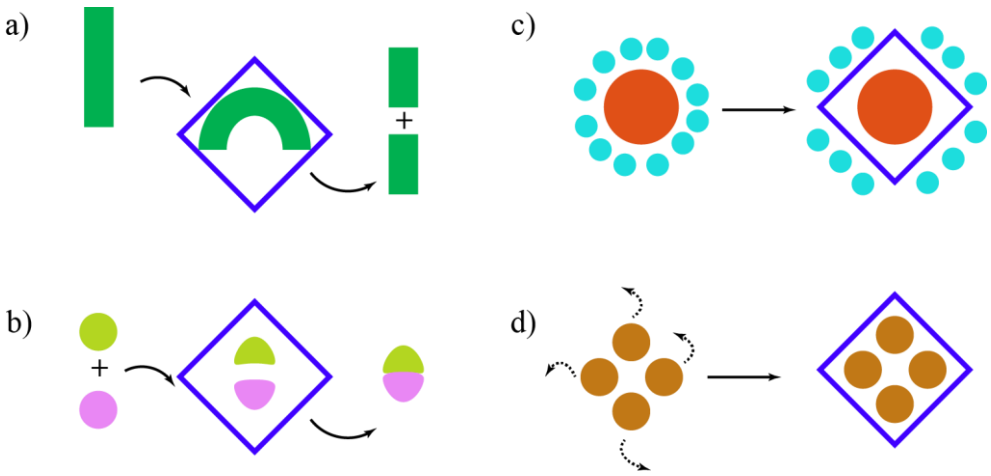


Figure 4. Representation of mechanism for supramolecular cavities catalyze reactions and stabilize reactive species. a) Twisted starting material to enhance cleavage reaction. b) Stabilize transition states to accelerate reactions. c) Isolated from solvent and get rid of disproportionation. d) Lock the cluster and avoid degregation.

2.3 Supramolecular Chemistry in water

Supramolecular chemistry in water is a rapidly expanding field of study because noncovalent interactions in aqueous environments are crucial for gaining a deeper knowledge of and exerting control over the most fundamental processes in nature. Due to its numerous atypical features compared to organic solvent, water presents both obstacles and opportunity.[31–33] Water is being utilized as a reaction medium since it is the most plentiful solvent on the earth and its use has minimal environmental impact.

In solution phase, the hydrogen bonds between distinct water molecules construct an infinitely dynamic network.[34] This extremely advantageous procedure is the primary cause of not only the bias of a number of its physical properties, but also the hydrophobic effect. Any operation that requires network disruption will incur enormous energy costs.[35] On the other hand, polar molecules undergo intense hydration by water and participate in the hydrogen-bonding network, which has a significant effect on the characteristics of the solvated species. Related to these, the fundamental problems of supramolecular chemistry in water are the solubility of receptors in water and the considerable participation of water molecules in the noncovalent process.

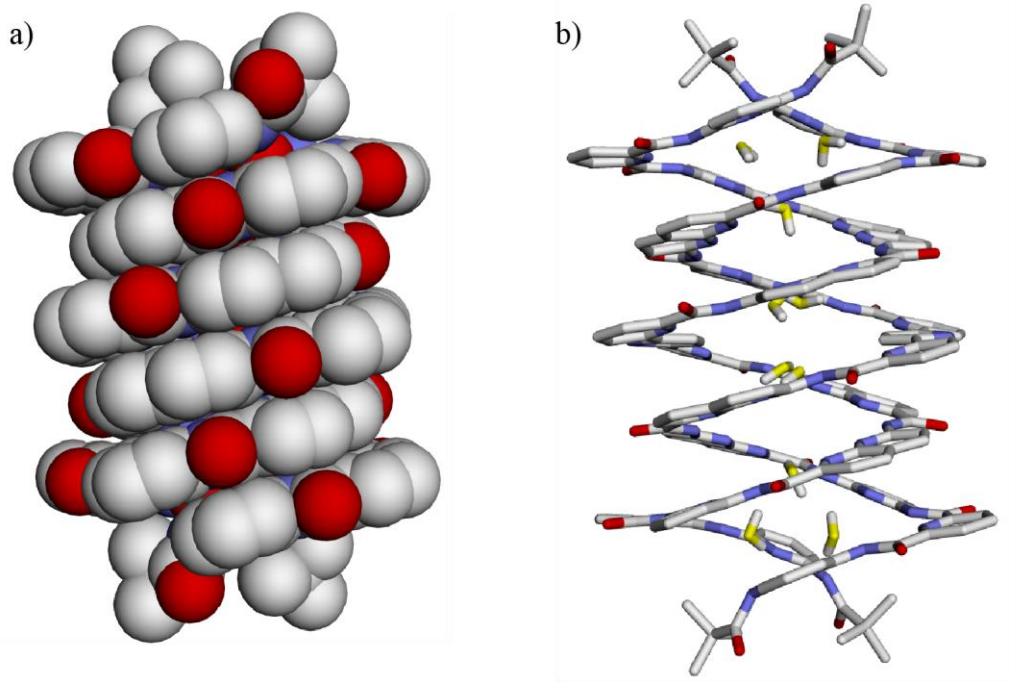


Figure 5. Crystal structure of double helix capsule: a) Side view shown in CPK representation. b) Side view with 10 encapsulated water molecules which highlight in yellow. Side chains have been omitted for clarity.

Multiple hydrogen bond interactions and the hydrophobic effect allow natural receptors such as enzymes and antibodies to form a robust and selective host–guest complex.[36–38] These natural systems serve as a source of ideas for the rational design of synthetic receptors, which may be utilized to develop a knowledge of the binding forces that result in the formation of complexes. The majority of biological activities in nature occur in aqueous medium, but the significant proportion of synthetic receptors were restricted to organic solvents.[39] To introduce synthetic receptors into water, it is necessary to overcome several obstacles. The water solubility of receptors drastically restricts the building blocks that may be employed in their synthesis. Introducing water-soluble side chains is a viable solution.[40–42] In addition, the competing impact of water during the binding process is another challenge. Frequently, the cavity was filled with several water molecules (Figure 5).[43] However, these water molecules in the cavities may also constitute the driving force behind the binding process through the "normal hydrophobic effect"[44] and "high-energy water".[45]

In enzyme mimic catalysis, water molecules not only participate in the binding process, but also play a crucial role in the catalytic activity itself. Protonation of the single hydrogen-bonded water determines the catalytic activity (Figure 6).[46] Multiple guests are encapsulated within big water-soluble receptors, which is a significant characteristic.[47] This facility permits the study of molecule interactions in limited spaces and the conduct of chemical reactions between them in aqueous solutions. A cavity guides and catalyzes synthesis, while simultaneously protecting the process from water.[48] Steric factors, which invariably add chiral control, severely regulate the incorporation and subsequent chemical changes in limited spaces.[49]

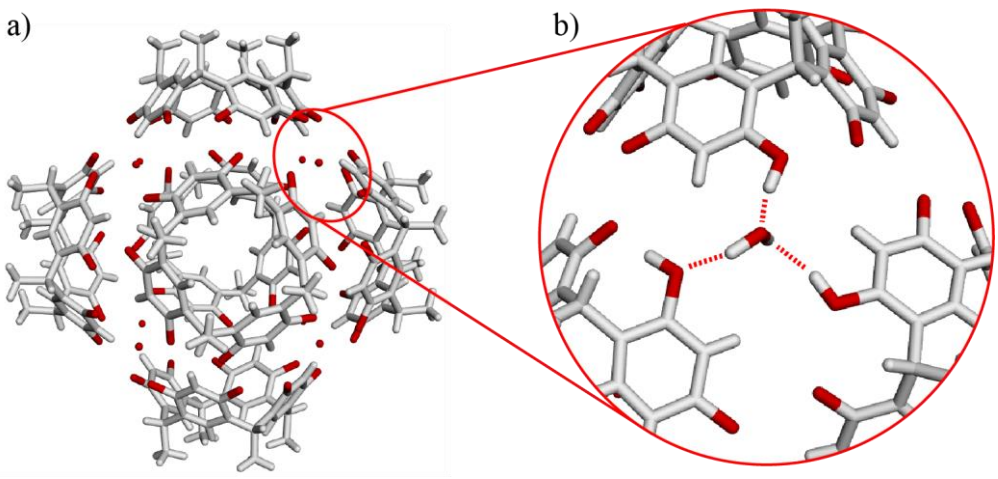


Figure 6. a) crystal structure of hexameric resorcin[4]arene capsule. b) The hydrogen-bond network and can function as hydrogen-bond donor at the inside of the cavity.

2.4 Challenges of establishing a large helical capsule

Bing Gong's research demonstrate that the helical folded structure could be stabilized by hydrogen bonds.[50] Furthermore, local electrostatic attractions and repulsions involving amide groups and neighboring quinoline endocyclic nitrogen atoms have been revealed to be responsible for a significant conformational bias at rotatable bonds, resulting in well-defined curvature and subsequently encouraging helical folding.[51] The novel 7-amino-2-quinolinecarboxylic acid moiety (Q^H), with C- and N-termini at a larger angle of 120°, was developed by me to construct a helix shape with a large diameter for the formation of a capsule with a large cavity. And the water solubility was solved by cationic side chains. The amide group at position 2 has a strong conformational preference due to hydrogen bonding and electron repulsions. However, the 7-position amide is too far to interact with endocyclic nitrogen atoms; it has an equal chance of generating either cis or trans conformations (Figure 7a). In a non-polar solvent, the rotatable position 7 bonds generate various conformations (Figure 7b, c), making it difficult to construct a helical conformation exclusively. Moore's research demonstrates that the cooperative folding process may be chain-length dependent.[52,53] Thus, we think that under the influence of the hydrophobic effect, an oligomer that has sufficient length will be capable of folding into a well-folded helix. Several experiments were done in this direction to confirm this hypothesis; the outcomes are described in Chapter 4.

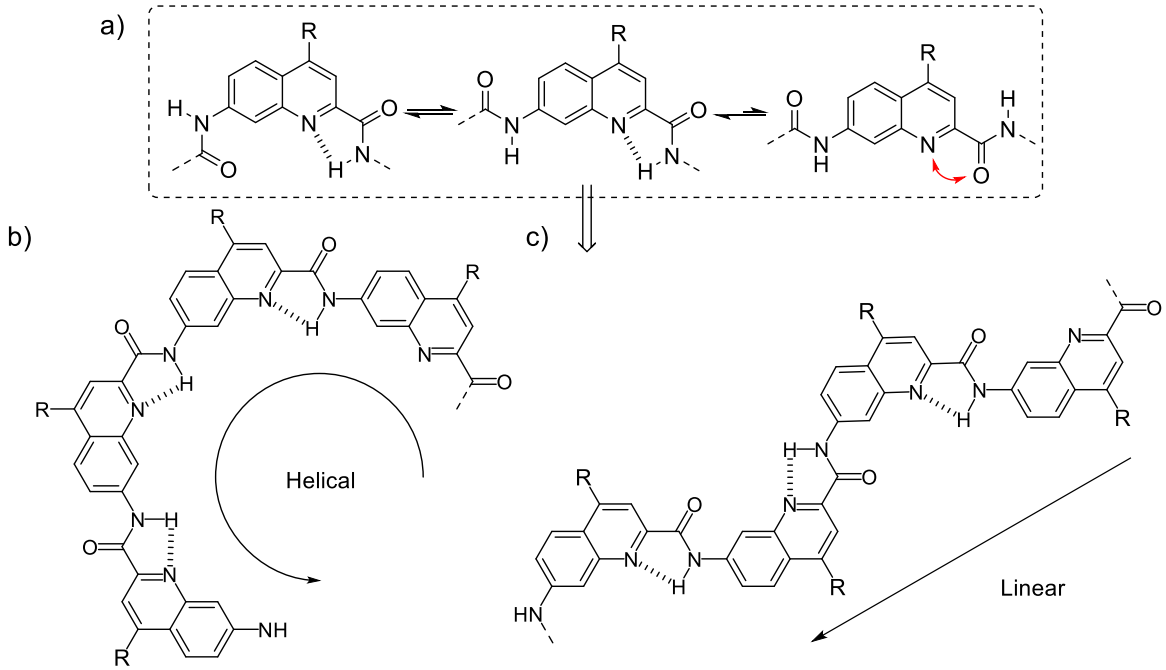
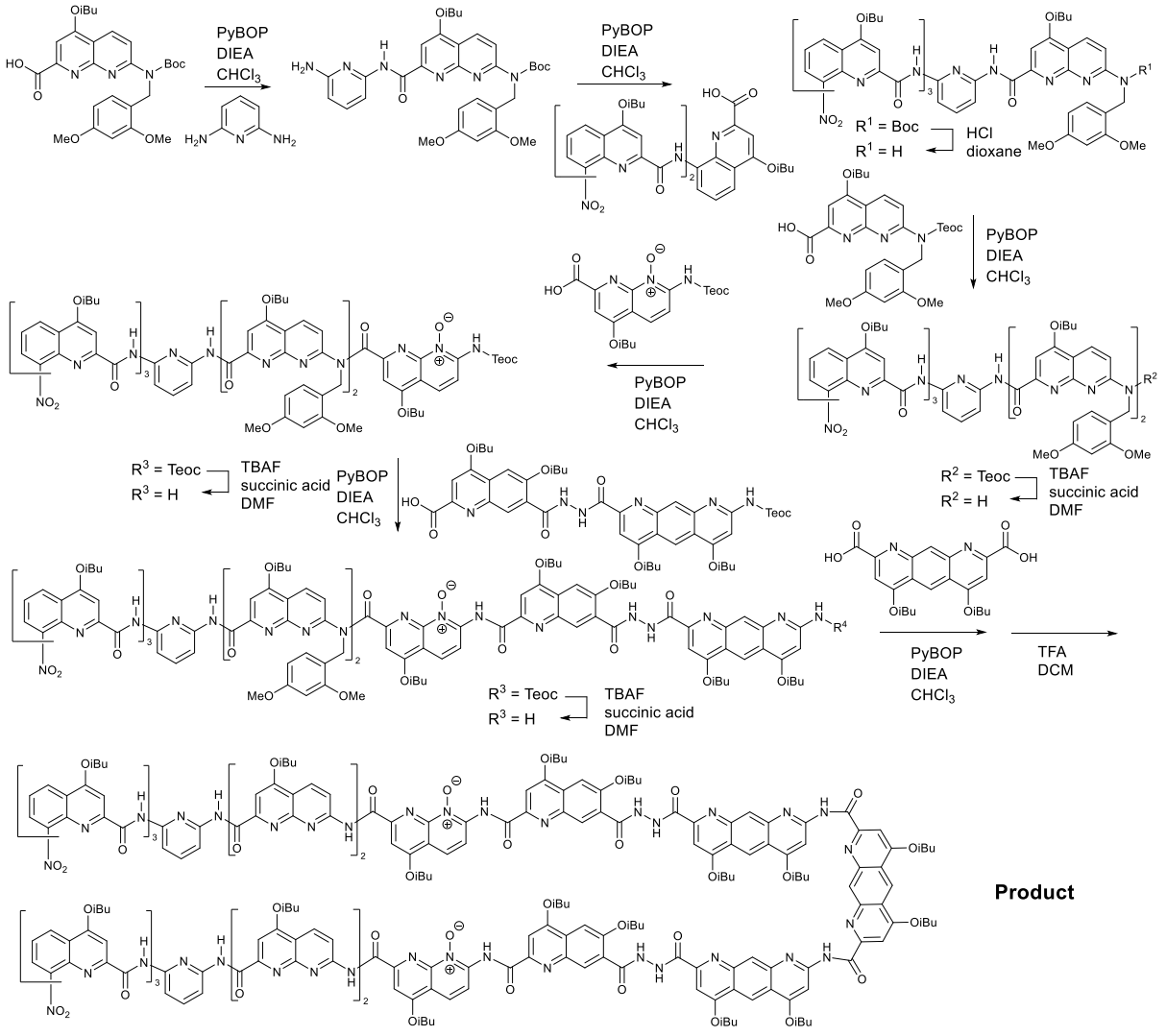


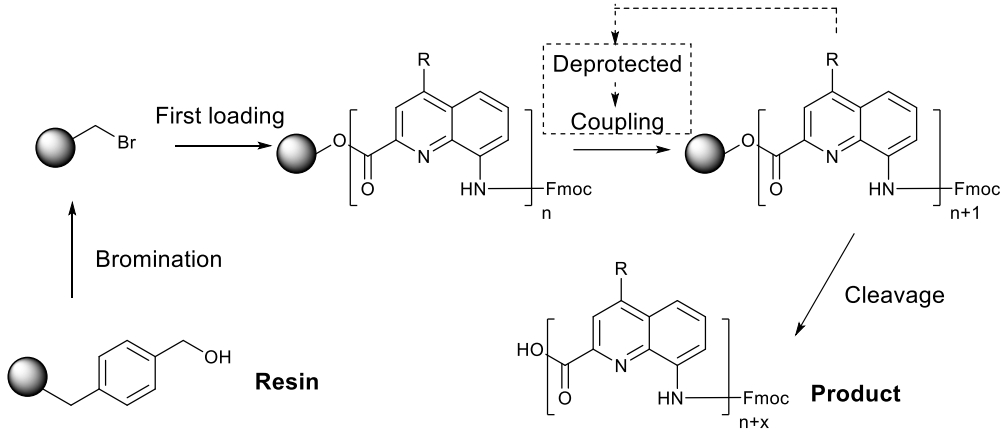
Figure 7. a) Different possible conformations for flexible 7-amino-2-quinolinecarboxylic acid moiety (Q^H) monomer. The two distinct folding order (b, c) related to Q^H conformational variation.

Although large cavity capsule foldamers have been thoroughly investigated by our lab,[54–56] an aqueous phase cavity has not yet been established. And the entire research was conducted using a solution phase-assisted strategy, which was both complex and time-consuming (Scheme 1).[54] Oligomer synthesis often involves numerous deprotection, activation, and coupling processes. Purification is required after each coupling step, which drastically restricts the scope of solution phase synthesis, frequently necessitating the employment of convergent schemes or segment doubling condensation strategies. Different segments necessitate distinct protective group units for various functions. Monomer synthesis and purification need considerable time. While these methods can produce long oligomers,[57] solution phase chemistry seldom permits the synthesis of sequences with a range of distinct side chains in arbitrary ordering. Since that, solid-phase synthesis is needed for rational foldamer capsule design.



Scheme 1. Representation of one foldamer synthetic route for Solution Phase Oligoquinoline Synthesis

As solid phase peptide synthesis for peptides has been established, reagents and coupling conditions have been tuned for the use of α-amino acids with more reactive amine functions than most aromatic amines. A solid-phase synthesis protocol designed for rapid access and rational sequence design has been developed (Scheme 2).[58,59] But limited to a few monomers that only can construct narrow helix. The new 7-amino-2-quinolinecarboxylic acid moiety (Q^H) is required to build up the central cavity part facilitating molecular recognition. When coupling onto aromatic amines, Fmoc-protected amino acids are converted in situ to their corresponding acid chlorides, whereas for aliphatic amines, conventional peptide coupling reagents are utilized. Additionally, while operating on a solid support, the folding of the molecule during synthesis becomes advantageous. In peptide synthesis, aggregation on the resin is a typical issue that hinders subsequent couplings. In contrast, the helical structure of quinoline-based oligomers allows them to protrude from the solid support, so resolving this issue, as well as the dimerization byproduct that occurred during solution phase synthesis.[60]



Scheme 2. Synthetic strategies for Solid Phase Oligoquinoline Synthesis on Wang Resin

2.5 Dimerization of supramolecule

In living organisms, biological supramolecular dimerization is a key process that enables molecules to form functional assemblies.[61] Understanding and regulating the molecular mechanisms of protein dimerization and their function leads to innovative research[62] and facilitates the development of several medicinal applications. Typically, non-covalent interactions, such as the hydrophobic effect and hydrogen bonds, maintain dimerization. Therefore, molecular glue,[63] metal ions,[64] host-guest complexes,[65] and G-quadruplex[66] have all been used to increase protein stability or to construct artificial quaternary structures (Figure 8).

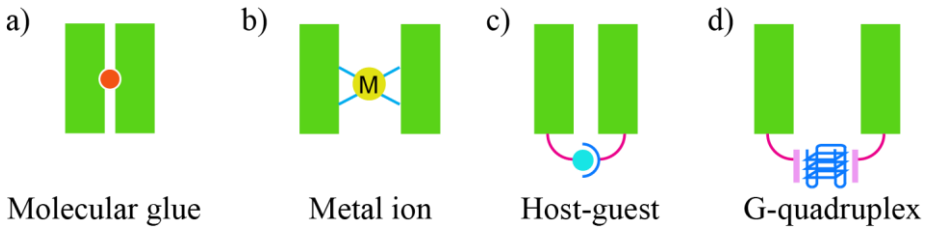


Figure 8. Representation of different methods of dimerization of proteins

Furthermore, artificial supramolecular systems have been developed to carry out the enormous dimerization complex, which facilitates rational structure design. In addition to the double helix and double bowl dimerization patterns described in the preceding section, the Huc group has recently revealed a number of unique homodimerization conformations. Connected by carboxylic acid group at –C terminals, the discrete head-to-head dimer could be established in water (Figure 9a).[67] Moreover, the incorporation of hydrogen-bonding capabilities at certain

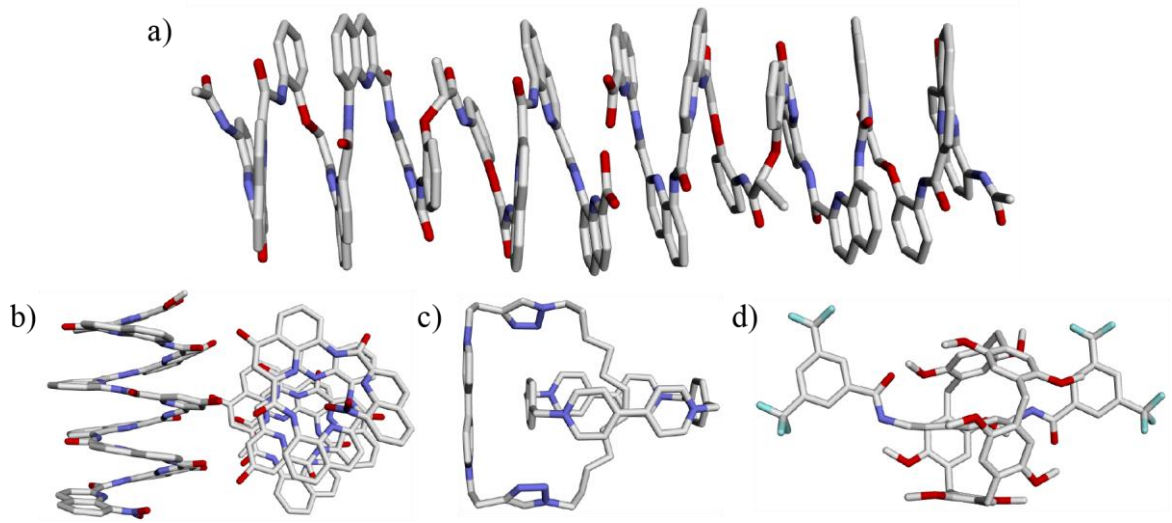


Figure 9. Overview of different dimerization pattern. a) Crystal structure of head-to-head dimer. b) Crystal structure of helix bundle. c) Crystal structure of catenane. d) Crystal structure of rotaxane.

positions on the backbones of abiotic foldamers that facilitate cooperative folding into two helix bundles in organic solvents (Figure 9b).[68] Using molecular recognition and molecular self-assembly methods, Stoddart and coworkers have achieved very efficient syntheses of mechanically interlocked molecular structures for heterodimer complexes. Catenanes are molecules with two or more interlocking rings (Figure 9c).[69] It is impossible to separate the interlocking rings without destroying the covalent bonds of the macrocycles. For rotaxanes or foldaxanes, the macrocycle or foldamer helix[70] is held in place by two large stoppers at the ends of the rod and interaction sites on the rod, which construct a stable dimer structure (Figure 9d).[71]

According to the structures of dimerization, several applications have been established. An additional strand interaction can regulate the cavity size of a helical foldamer (Figure 10a,b).[72] The Yashima group demonstrated another spring-like motion based on dimerization conformation; the reassemblable dimer is essential for this twisting in one direction in a highly homotropic allosteric manner driven by metal ions and protonation (Figure 10c,d).[73] In addition, Marc Sallé and coworkers reported redox-controlled guest release employing supramolecular dimers.[74] The dissociation of this complex upon oxidation/dimerization of the tweezers provides a unique redox-activated molecular delivery mechanism. Additionally, heterodimerization structures might do additional difficult tasks. Rotaxane-based linear motors[75] and catenane-based rotary motors[76] have been established as micromachines that transform energy from an external source into unidirectional movement on a molecular scale, which might have medical applications. Moreover, dimerization-based polymerization provides an additional option in the field of material science. The Dong group reported the synthesis of linear supramolecular polymers by stacking one single-stranded helix on top of another.[77]

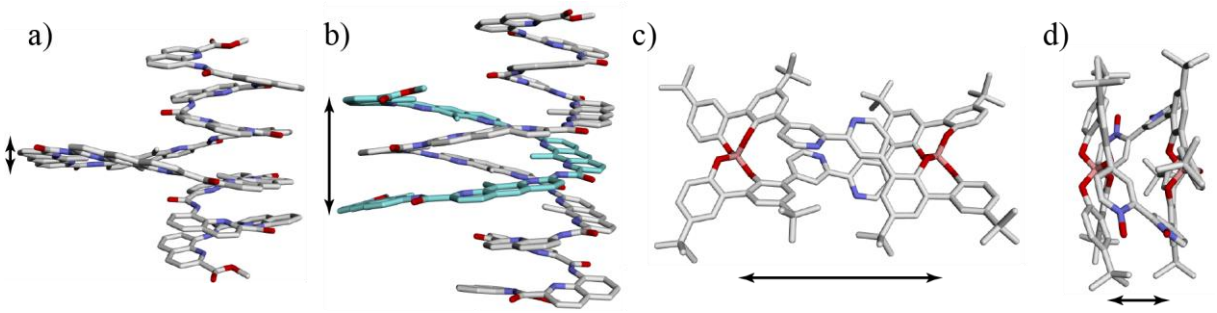


Figure 10. Representation of spring-like motion based on dimerization conformations. a) Single helix with little space. b) Extended cavity with another strand. c) Extended structure in the deprotonated condition. d) Contracted structure in the protonated condition.

3 Guiding objective: design of helical capsule and molecular recognitions

The development of water-soluble aromatic oligoamide helical capsules capable of recognizing hydrophilic guests in water served as a guiding objective throughout the research for this thesis. Although a few large cavity foldamers have been developed by our group utilizing solution phase synthesis and large diameter monomers, their application is limited by the lengthy synthesis and purification procedures. In order to produce a large cavity capsule in water employing solid-phase synthesis, a new extended-diameter monomer that can tolerate in-situ coupling conditions is required. Due to its large diameter, the new Q^H monomer can be used to construct a vast cavity in the initial helical capsule design, which consists of a cap and a cavity. And both ends can be closed by Q monomers, which will form a short, narrow helix and prevent the formation of a double helix (Figure 11). However, there is some uncertainty concerning the Q^H monomer. Because there is no conformational bias, each unit has two conformations (Figure 7a). It was difficult to be well folded in an organic solvent. In water, we hypothesize that the strong hydrophobic effect can force the sequence to fold into a helix structure, which accumulates multiple layers of helical turns and exposes a minimal solvent-accessible surface. Charged water-soluble side chains were introduced to perform solution-phase experiments and combine with neutral side chains to optimize crystallization properties, thereby facilitating X-ray analysis with high-quality crystals.

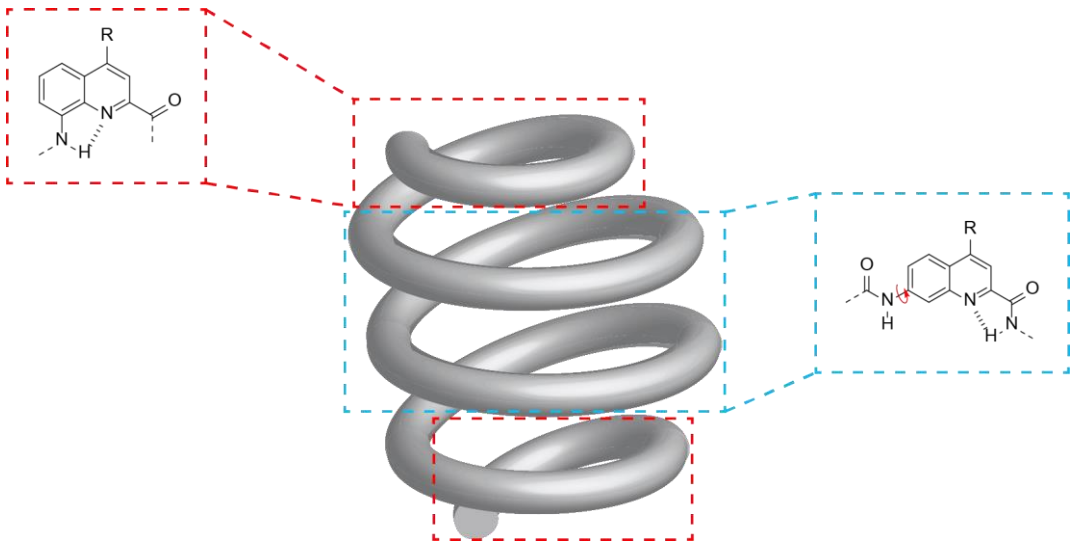


Figure 11. Design strategies for the large cavity of helical capsule foldamers. The both ends are capped by Q monomer to close the cavity and avoid dimer formation. The central helical part is constructed by Q^H monomer that establish a considerable cavity by their large diameter.

Considering the hydrophobicity of the helical aromatic oligoamide turn cavity inside, comprising hydrophilic amide proton and endocyclic nitrogen from quinoline, and the lipophilic benzene ring (Figure 12), a series of amphiphilic guests were suspected to show binding affinity combining hydrogen bonds and hydrophobic effect. And shape complementarity could be used to accomplish selectivity. Surprisingly, additional experimental results presented in chapter 5 indicate that the enormous entropy produced by the release of highly ordered water molecules also played a crucial role in the binding process. This amphiphilic cavity has the advantage of not only being able to capture hydrophobic compounds, but also water-soluble guests.

This project's typical methodology consisted of utilizing a predetermined design strategy and planning the precise sequence pertaining to the cavity geometry and hydrogen bonding positions. To avoid unnecessary titrations experiments, the molecular modeling tools (Maestro) was utilized to demonstrate the energy-minimized structure in various solvents and evaluate the shape complementarity of different sequences with distinct guests. Employing solid-phase synthesis techniques, potential candidates were synthesized and purified with semi-prep HPLC. After isolating the products, mass spectrometry and ^{1}H NMR studies were conducted to determine the initial conformational behavior of the compound. To obtain precise structural information, the X-ray crystallographic data are required. The aforementioned data is also essential for improving shape complementarity through iterative design.

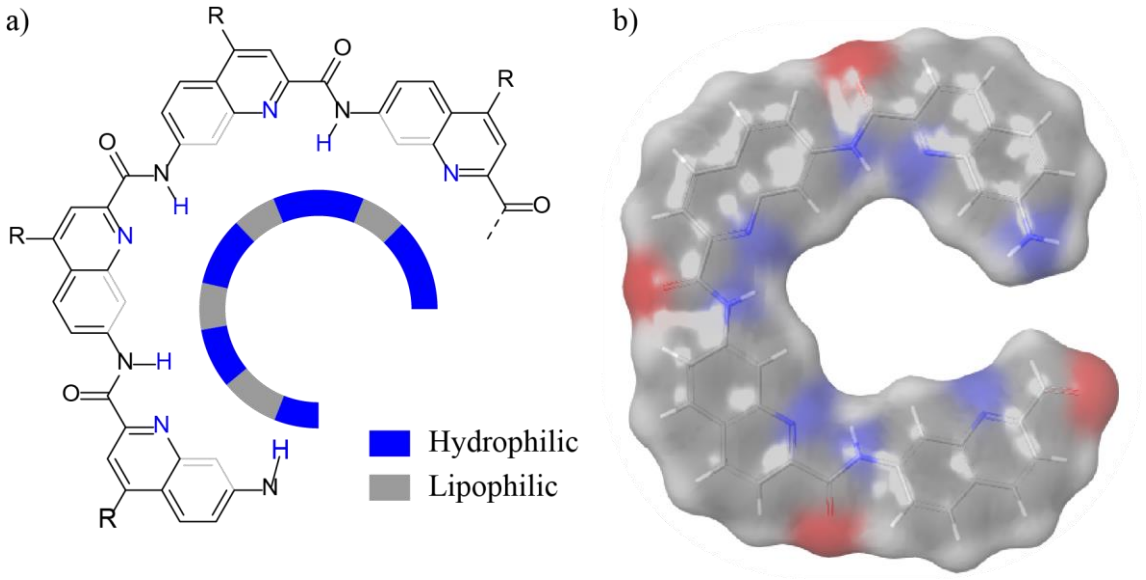


Figure 12. Representation of one helical aromatic oligoamide turn. a) The chemical formula of one helical turn. The colored circle presents the hydrophobicity of the cavity inside. The hydrophilic and lipophilic parts are highlighted in blue and grey, respectively. b) The surface present by residue charge. The carbon, oxygen and nitrogen are highlighted in grey, red and blue, respectively.

Numerous oligomers have been synthesized during the course of this research. As anticipated, NMR titration experiments revealed that a number of water-soluble guests exhibited binding affinity with the amphiphilic cavities, with the majority of them demonstrating only a single set of signals, indicating a fast exchange of binding species on the NMR timescale. By decreasing the temperature or increasing the number of binding positions, the exchange rate could be slowed down and showed two sets of signals. Although additional structural data from X-ray crystallography could not be obtained with the free host and host-guest complex, side chain optimization led to the discovery of novel aggregation patterns. (head-to-head dimer, parallel double helix, and tetrameric helix). The first aromatic oligoamide capsule crystal that grew from water was obtained from an aggregation species. The folding process, molecular recognition, and aggregation pattern have been published or submitted to peer-reviewed journals and will be discussed in the following chapters. (Note that compound numbering is not continuous and starts again at the beginning of each chapter).

4 Combining local conformational preferences and solvophobic effects in helical aromatic oligoamide foldamers

Sequential Q units can form a narrow helix depending on conformational preference and avoid double helix formation.[72,78,79] By increasing the angle between -C and –N terminals, the diameter of the helical structure can be extended and a large cavity developed. In light of this information and with the objective of generating a helical capsule, the designed amino group was relocated to position 7 (Q^H) to establish a 120° angle and optimize the helical diameter. As a reference, the same angle Q^F with conformational bias was synthesized (Figure 13a). Since the folding behavior of Q^H with a rotatable C-N bond is unknown, oligomers of varying lengths were investigated in a water/DMSO solvent mixture with varying proportions in order to comprehend the folding properties.

The established Q unit synthetic protocol, which used to introduce aminomethyl side chains, needed highly toxic KCN and was not suitable for Q^H monomers; thus, a new synthetic route had to be developed to be able to simplify the synthetic steps. A Suzuki coupling reaction was selected to introduce a boc protected side chain with a boron reagent. By getting rid of the complicated protect and deprotect strategy that was performed in the previous Q^H synthetic route, the total synthetic steps were reduced to 7 steps (for more information see supplementary information section 2.3), leading to short and fast synthesis in the course of this work.

The short and charged aminomethyl side chain was designed to provide excellent water solubility and crystal growth ability for our oligoamide foldamers. Although a well-defined helical capsule structure was characterized by NMR studies, the single helix crystal could not be obtained in water. The UV, NMR, and computational results have been summarized in a manuscript that has been published in *Organic & Biomolecular Chemistry*. The folding process in water is related to Q^H segment length, which means one helical turn (Figure 13b) is not enough to complete folding. Minimum of two helical turns are required to facilitate helical folding of the Q^H segment (Figure 13c), which is enabled by two layers of π-π stacking and hydrophobic effect. On the other hand, the Q^H segments prefer the extended state (Figure 13d) in DMSO which is proved by energy-minimized models and experimental results. A novel cooperative conformational transition was observed for these backbone flexible foldamers upon increasing water proportion in DMSO. This work develops a novel flexible backbone building

block and validates the folding process via solvent property transfer, demonstrating the possibility of various backbone units for foldamers.

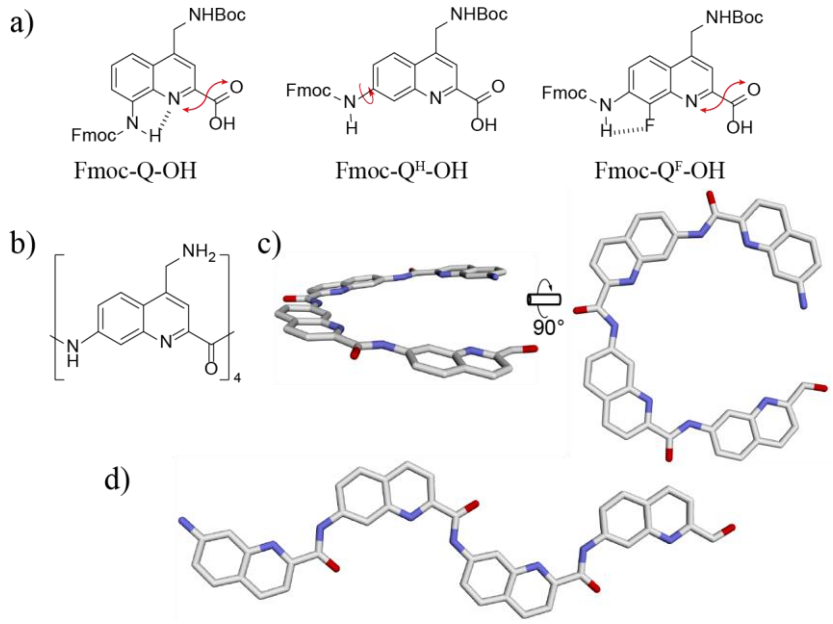


Figure 13. a) Structure of synthesized building blocks. b) Chemical formula of one minimum helical turn unit. Energy-minimized models of helical turn c) in water with side view and top view. d) in no solvent condition. Side chains have been omitted for clarity.

Contribution: I conceived the ideas for this project in collaboration with I. Huc. Synthetic routes and conditions for Q and Q^{H} monomers were developed by me with suggestions from Y. Ferrand. The synthetic route of Q^{F} monomer was established by J. Atcher. Oligomer synthesis and analysis were done by me. Energy minimized model was done by me. L. Allmendinger was involved in the design and execution of NMR experiments. The automated solid phase synthesis was done in collaboration with C. Douat. The manuscript was written by me in collaboration with I. Huc.

Organic & Biomolecular Chemistry

ROYAL SOCIETY OF CHEMISTRY

COMMUNICATION

View Article Online
View Journal | View Issue

Cite this: *Org. Biomol. Chem.*, 2023, **21**, 3525

Received 28th March 2023,
Accepted 11th April 2023

DOI: 10.1039/d3ob00473b

rsc.li/obc

Combining local conformational preferences and solvophobic effects in helical aromatic oligoamide foldamers†

Binhao Teng,[a] Joan Atcher, [a] Lars Allmendinger, [a] Céline Douat, [a] Yann Ferrand [b] and Ivan Huc *[a]

Aromatic oligoamide foldamers were designed using a newly-developed monomer so that helical folding was promoted by both local conformation preferences and solvophobic effects. Solid phase synthesis provided quick access to the desired sequences. Sharp solvent-driven conformational transitions that depended on sequence length were evidenced by both NMR and UV absorption spectroscopies.

Aromatic foldamers constitute a large ensemble of oligomers with aryl rings in their main chain that have stable folded conformations, most frequently helices. Helical folding may be driven by multiple parameters including solvophobic effects,[1–7,9] the presence of neutral or ionic guests,[10–14] steric congestion combined with aryl–aryl interactions,[15,16] or local conformational preferences.[17–21] The latter are typically mediated by hydrogen bonds and electrostatic repulsions between adjacent aryl rings or between aryl rings and amide groups, eventually restricting single bond rotations. In their absence, folding is altered.[22] Conversely, folding may be so strongly driven by local conformational preferences that it occurs in essentially any solvent.[23] The solvent then simply modulates helix stability, but neither unfolding nor important conformational transitions are observed. A benefit of high conformational stability is that it enhances crystal growth ability for the purpose of solid-state structure elucidation. In contrast, solvophobic foldamers have rarely been crystallized. Several approaches using hydrogen bonds or polar interactions have been proposed to reinforce the helical structures of solvophobic aromatic foldamers.[24–29] Here, we present an opposite strategy consisting in deliberately decreasing conformation control by removing some local conformational preferences from helical aromatic oligoamide foldamers (AOFs), while at the same time allowing for a larger role of solvophobic effects. The main motivation was to elicit and investigate conformational transitions in objects that would otherwise be rigid. Conformational transitions enrich structural behavior and allow, for example, for the dynamic control of properties such as anion recognition or self-assembly.[30,31]

Q and Q^F are typical monomers of AOF sequences in which helical folding is driven by local conformational preferences (Fig. 1).[17,23,32,33] In such AOFs, quinoline endocyclic nitrogen atoms or exocyclic fluorine atoms form hydrogen bonds with amide protons and concomitantly engage in repulsive electrostatic interactions with amide carbonyl groups. In contrast, the new 7-amino-2-quinolinecarboxylic acid monomer Q^H also has an endocyclic nitrogen atom, but no function adjacent to its 7-amino group that could set the orientation of an amide at that position. Pseudo-conjugation is therefore expected to favor two s-*cis* and s-*trans* conformers in which the quinoline

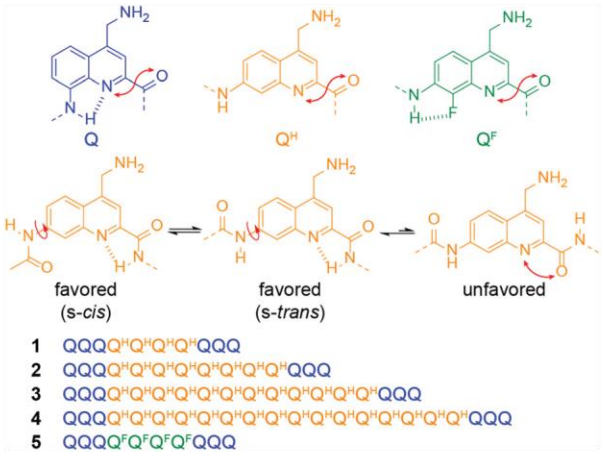


Fig. 1 Structures of color-coded monomers Q, Q^H and Q^F and composition of sequences **1–5**. When placed in an AOF sequence, Q^H may adopt two main conformations. Double-headed arrows indicate electrostatic repulsions.

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

[b]*Univ. Bordeaux, CNRS, Bordeaux Institut National Polytechnique, CBMN UMR 5248, 2 rue Escarpit, 33600 Pessac, France*

†Electronic supplementary information (ESI) available: Synthetic protocols and characterisation of new compounds. See DOI: https://doi.org/10.1039/d3ob00473b

ring of Q^H and adjacent amide groups are coplanar, leading to 2^n conformers in a sequence having n Q^H units (Fig. 1).

For this study, monomers Q, Q^H and Q^F were all equipped with a 4-aminomethyl group which, once protonated, ensures good water solubility of AOF sequences. The choice of this small polar side chain was also made in the hope it would favor crystal growth.[35] Monomer Q has previously been described,[35] yet an improved synthesis is presented in the ESI (ESI, Scheme S1†). The syntheses of monomer Q^F which had never been presented with this side chain and of new monomer Q^H are also shown in the ESI (Schemes S2 and S3†). All were produced on a 10 g scale as building blocks ready for solid phase synthesis (SPS), that is, with a free carboxylic acid in position 2, and their main chain and side chain amines protected with Fmoc and Boc, respectively.

We designed sequences **1–4** to investigate the folding behavior of Q^H monomers (Fig. 1). These sequences contain Q_4^H, Q_8^H, Q_{12}^H, or Q_{16}^H segments. Sequence **5** is analogous to **1** with Q^F units instead of Q^H. The importance of assessing any length dependence in such studies has been highlighted by Moore *et al.*[8] and we thus headed for a series instead of a single compound. Energy-minimized models show that it takes about 4.5 Q^H units to span one turn when this monomer is involved in a helix (Fig. 2a–d). In their helical conformations, the Q_n^H segments of **1–4** would thus be expected to span about 0.9, 1.8, 2.7 or 3.6 helix turns, respectively. Our expectation was that water may favor the helical conformations of **1–4** because these conformations reduce the solvent exposure of Q^H monomers. However, this may not occur in solvents that solvate the aryl amide groups such as DMSO where the Q_n^H segments may adopt any conformations from the helix (with 7-amido-quinoline bonds all in s-*trans*) to a fully extended form (Fig. 2e, with 7-amido-quinoline bonds all in s-*cis*). As an illustration, the solvent exposed surface of **2** in its helical and fully extended Q_8^H conformations were calculated to be 1720 and 2580 Å^2, respectively, giving a measure of how much is hidden from the solvent upon helix folding (Fig. S1†).

Q_n^F analogues bearing other side chains have a strong propensity to aggregate into multistranded structures through a spring-like extension and intertwining of their helices, including in water.[32,36] In anticipation that Q_n^H sequences may also aggregate and that this would complicate the investigation of their folding behavior, all sequences were flanked by Q_3 segments at both ends. Q monomers code for a narrower helix diameter than Q^F or Q^H – only 2.5 units per turn.[17,32,33] Q_n oligomers with $n > 2$ do not form multistranded helices and have been shown, when placed at the termini, to prevent double helix formation of segments having a larger diameter.[37,38] Thus, in their helical conformations (Fig. 2a–d), sequences **1–4** may possess a cavity potentially capable of encapsulating guest

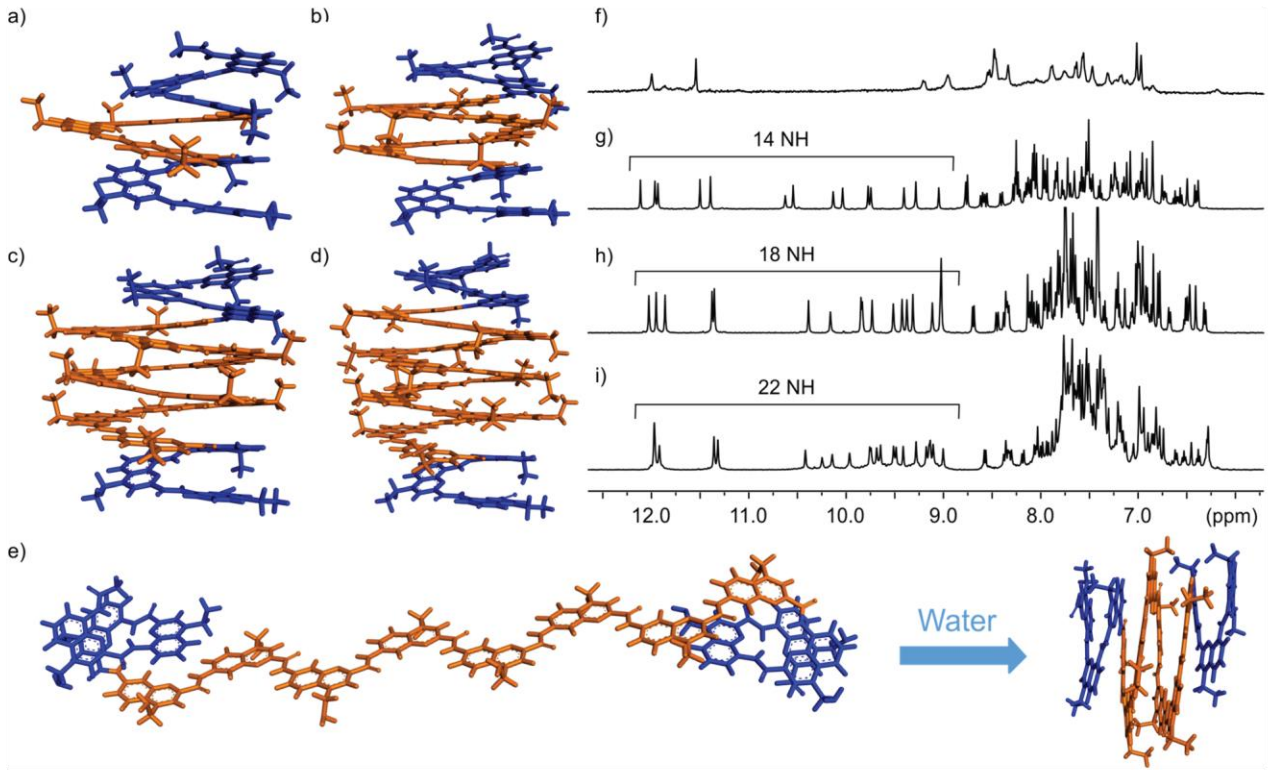


Fig. 2 Energy minimized conformations (Maestro, MMFFs force field, water as implicit solvent)[34] of **1** (a), **2** (b and e), **3** (c) and **4** (d). In (a–d), the Q_n^H segments have been arranged as a helix at the start of the minimization whereas in (e) it was placed in an all extended conformation. (e) thus illustrates two extreme conformational states. (f–i) Excerpts from the 500 MHz ^{1}H NMR spectra of compounds **1** (f), **2** (g), **3** (h), and **4** (i), all recorded at 298 K in H_2O/D_2O (90 : 10 v/v).

molecules, as in designs that we and others have presented previously.[7,37–39] This potential was an additional motivation when designing **1–4**. Yet their molecular recognition properties will be described elsewhere.

A third motivation in the design of **1–4** was to challenge the preparation of such sequences by SPS. SPS protocols are very efficient for sequences such as Q_n derivatives with various side chains,[35,40] and have recently been automated.[41] Their applicability to entire capsule-like sequences was unknown but desirable for quicker access in the context of iterative sequence improvements.[42,43] As presented in detail in the ESI,† SPS proved to be very efficient, generating sequences in good crude purity and allowing for their isolation using RP-HPLC in 15–32% isolated yield. The only exception was sequence **5** which repeatedly gave poorer yields for reasons that were not identified.

The ^{1}H NMR spectra of **2–4** at 0.5 mM in H_2O/D_2O (90 : 10 v/v) showed a single set of sharp resonances typical of folded conformations (Fig. 2g–i). In particular, distinct signals with very little overlap are clearly observed for the amide NH resonances even for the longest sequence **4**. In contrast, the signals of **1** were broader (Fig. 2f), suggesting that some conformational dynamics, or some aggregation, were at play. The spectrum of **2** did not change significantly upon varying concentration, hinting at a monomeric species or a very stable aggregate (Fig. S2†). In DMSO-d_6, a solvent that inhibits the aggregation of aromatic oligoamides and that provides weaker solvophobic effects than water, one set of ^{1}H NMR signals is also observed for all species but the patterns of signals differ from those of water (Fig. 3). Only a limited number of distinct amide NH resonances were observed, that were assigned to the NH protons of the folded Q units. Several NH signals appear under the same large peak whose intensity correlates with the length of the central Q_n^H segment (Fig. 3, green circle). The spectra also tend to become broader as the length of this segment increases. This trend is consistent with the Q_n^H segment undergoing the s-*cis*/s-*trans* equilibrium shown in Fig. 1. In the absence of folding that would make their environment different due to the ring current effects associated with aromatic stacking, the NH resonances of Q_n^H segments have, on average, a similar chemical shift value in DMSO-d_6. Equilibrium between the different conformers is fast on the NMR time scale but full averaging takes more time for the longest sequences, hence the broadening of the signals.

Upon adding water to samples in DMSO-d_6, all four ^{1}H NMR spectra were found to evolve (Fig. 3). In the case of the two shorter sequences **1** and **2**, an initial broadening led to a coalescence before a new set of signals emerged, indicating an equilibrium between two states that is neither fast nor slow on the NMR time scale. For the two longer sequences **3** and **4**, the water-dominant species emerged before the DMSO-dominant species disappeared and co-existence of the two species in slow exchange on the NMR time scale was observed at intermediate proportions of water. These results demonstrate that, for all compounds, the prevalent species in water are different

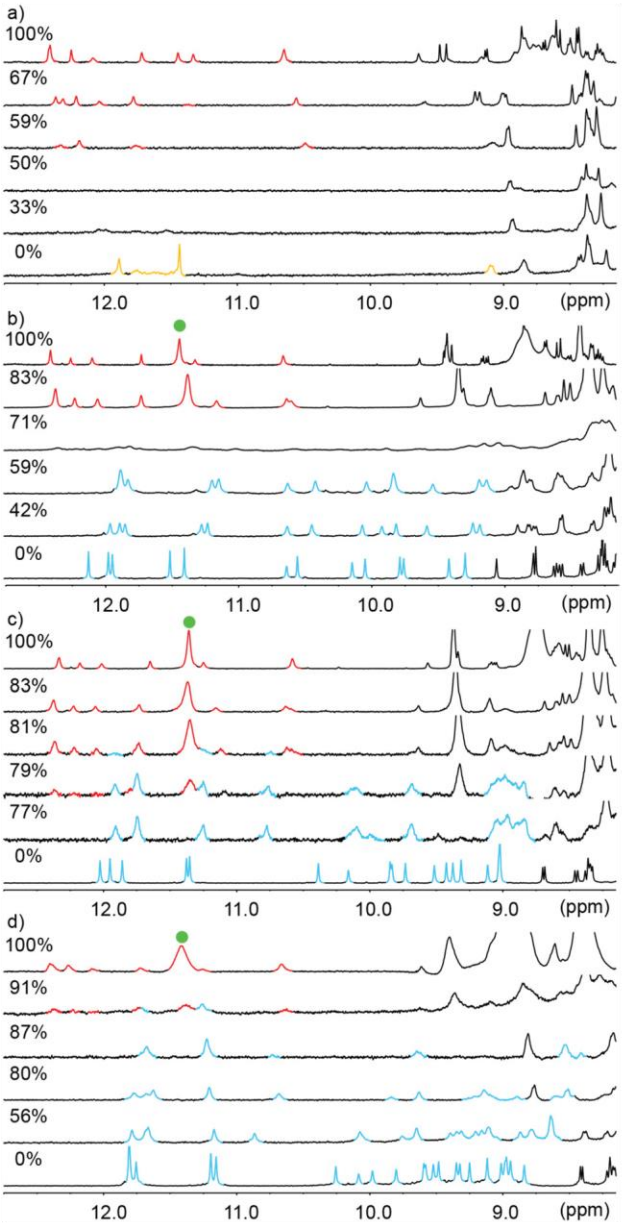


Fig. 3 Excerpts of the ^{1}H NMR (500 MHz) spectra of **1–4** in DMSO-d_6/H_2O mixtures at 298 K. (a) **1**; (b) **2**; (c) **3** and (d) **4**. The volume % of DMSO are indicated. Note: 0% DMSO means the spectrum was recorded in H_2O/D_2O (90 : 10 v/v). Red, blue and gold color amide NH signals stand for structures in which Q_n^H segments adopt unfolded, helically folded and partially folded conformations, respectively. Large peaks assigned to the NH of unfolded Q^H units are pointed out by green circles.

from those that prevail in DMSO, and that the characteristic time of exchange between these species increases with increasing the length of Q_n^H segments. A similar effect was observed upon adding CD_3OH to a DMSO-d_6 solution of **2** (Fig. S3†), albeit more CD_3OH was necessary for a full transition than with water. In contrast, the NMR spectra of reference compound **5** show that it undergoes no such transition (Fig. S4†), as would be expected in a sequence that does not contain Q^H.‡

The water-induced transition was found to require less water as the length of the Q_n^H segment was increased (Fig. 3), suggesting some cooperative phenomena. UV spectroscopy was then utilized to further investigate this cooperative conformational transition. Fig. 4a–c shows the UV spectra of oligomers **2–4** at different proportions of water and DMSO. Increasing the water content resulted in a drop in the intensity of the band centered at 357 nm (hypochromicity), consistent with the formation of stacks of aryl rings.[44] Concomitantly, hyperchromic and hypsochromic effects were observed for the band centered at 257 nm. A decrease in the A_{357}/A_{257} absorbance ratio indicates a conformational change in water compatible with the collapse shown in Fig. 2e.[45] Most of the UV spectra cross isosbestic points, confirming an equilibrium between two states. However, at high DMSO content, the spectra no longer include the isosbestic points, indicating that more complex effects are involved (*e.g.* solvent dependence of the absorbance itself).

Fig. 4d depicts the relationship between the normalized UV intensity at 357 nm and the volume fraction of DMSO in water. Oligomers **2–4** with more than one helix turn at their central Q_n^H segment exhibit a sigmoidal curve characteristic of a sharp, cooperative, folding transition.[5] The transition becomes more abrupt as the length of the chain increases. In agreement with NMR spectra, it takes less water to cause the transition of longer Q_n^H segments. The curve corresponding to **1**, shows a less clear transition that may be explained by the fact that Q_4^H is too short to span a turn. It can only undergo reduced intramolecular aromatic stacking between Q^H and Q units in its helical conformation (Fig. 2a shows two protruding, solvent exposed Q^H units). This observation is consistent with the behavior observed by NMR. The existence of a sharp folding transition induced by water suggested a possibly strong contribution of entropy and thus a dependence on temperature. We measured the UV spectra of **2–4** at temperatures ranging from 5 to 85 °C using a proportion of DMSO that was, for each compound, close to the transition point (Fig. S6–S8†). However, contrary to what has been reported for phenylacetylene oligomers,[2] no significant change of the UV spectra was observed.

Finally, to ascertain the structures of the folded conformations in solution, multiple attempts were made to grow single crystals. Unfortunately, the aminomethyl side chains did not serve this purpose well and no good diffracting crystals were obtained. We thus turned to NMR and managed to elucidate the structure of oligomer **2** using two-dimensional experiments (see ESI†). The spin systems of each quinoline ring were assigned using TOCSY, HBMC (*e.g.* correlations of H3 to the exocyclic $\underline{C}H_2$

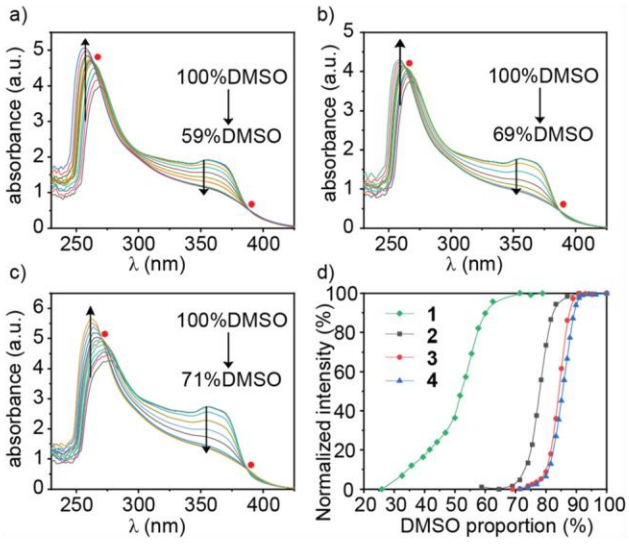


Fig. 4 UV spectra of compounds **2–4** at different proportions of water and DMSO (a–c respectively). (d) Plot of the normalized UV intensity at 357 nm *vs.* the volume percent of DMSO in water for compounds **1–4**. In the case of **1**, the normalized green curve in fact reflects a modest absolute change of the UV spectra (Fig. S5†). Isosbestic points are indicated by red circles.

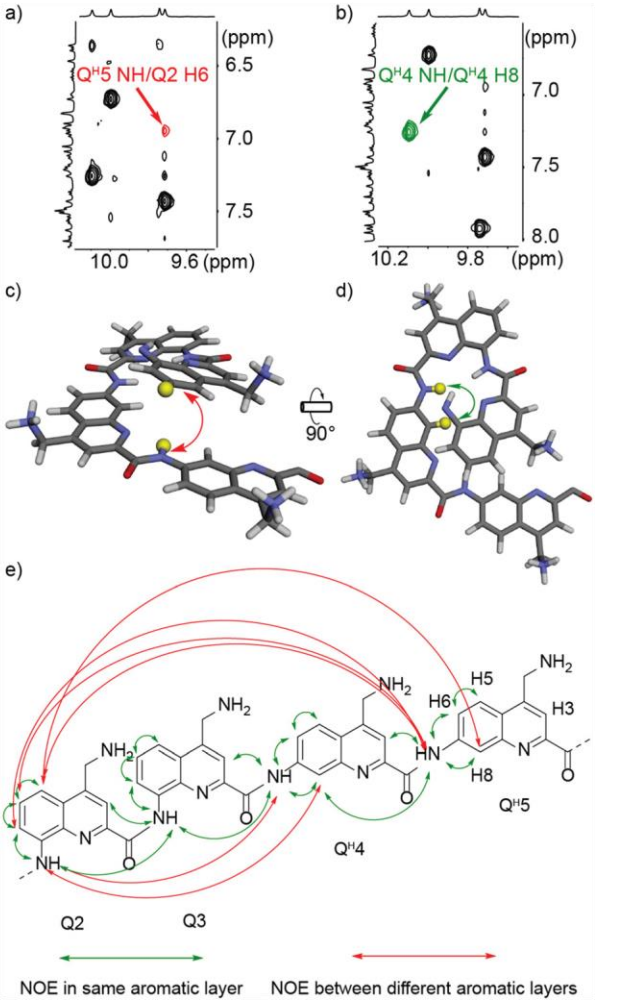


Fig. 5 Selected parts of the $^1H-^1H$ NOESY spectrum (500 MHz) of compound **2** showing in (a) the interlayer NOE correlation between Q^H5 NH/Q2 H6 (in red) and in (b) the intralayer NOE correlation between Q^H4 NH/Q^H4 H8 (in green). Energy minimized models indicating (c) Q^H5 NH/Q2 H6 and (d) Q^H4 NH/Q^H4 H8 proximity (target protons are highlighted in yellow). (e) NOE interactions observed in the NOESY spectrum of compound **2**. All spectra were recorded at 298 K in H_2O/D_2O (90 : 10 v/v).

carbon, see Fig. 5e for numbering), and NOESY (*e.g.* correlation of H3 and H5 to the exocyclic $C\underline{H}_2$ protons) experiments. NOE correlations between amide N$\underline{H}$ signals allowed for the assignment of each spin system to its position in the sequence. Finally, the NOESY spectrum revealed multiple correlations between units contiguous in the sequence or stacked above one another that can only arise in a helical structure. Representative examples are shown in Fig. 5. In particular, Q^H units could be shown to adopt the s-*trans* conformation at their position 7 (Fig. 1), which allowed to propose that the structure shown in Fig. 2b is a valid model of the conformation of **2** in water.

Conclusions

In summary, a new flexible monomer Q^H was developed to quickly access water soluble AOF sequences with capsule-like shapes *via* SPS. This monomer promotes a solvent dependence of folding behavior that is uncommon in AOFs. Sharp transitions between unfolded and helically folded states were observed upon adding water to DMSO solutions. Increasing the flexibility of a molecular system is a less common path than the reverse, but the approach has been straightforward because the starting point – the rigid conformations of AOFs – was so well defined. The approach may be extended by using monomers even less helicogenic than Q^H and modulated by combining Q^H with monomers less influenced by solvent such as Q^F. Because of their capsule shape, *i.e.* their larger diameter in the middle of their sequences, molecules **1–4** are potential hosts for molecular recognition. These properties are being investigated and will be reported in due course.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by the DFG (project HU 1766/6-1). B. Teng gratefully acknowledges financial support from the China Scholarship Council. We thank D. Gill and E. Merlet for contributing some synthetic precursors.

References

‡The NMR spectrum of **5** was found to be broad in pure DMSO-d_6 for reasons that could not be clarified (see Fig. S3†). Because this reference compound was difficult to synthesize, efforts were not pursued.

1 X. Hu, A. Schulz, J. O. Lindner, M. Grüne, D. Bialas and F. Würthner, *Chem. Sci.*, 2021, **12**, 8342–8352.
2 J. C. Nelson, J. G. Saven, J. S. Moore and P. G. Wolynes, *Science*, 1997, **277**, 1793–1796.
3 H. Goto, H. Katagiri, Y. Furusho and E. Yashima, *J. Am. Chem. Soc.*, 2006, **128**, 7176–7178.
4 K. P. de Carvasal, N. Aissaoui, G. Vergoten, G. Bellot, J.-J. Vasseur, M. Smietana and F. Morvan, *Chem. Commun.*, 2021, **57**, 4130–4133.
5 H. S. Chan, S. Bromberg, K. A. Dill, C. M. Dobson and A. R. Fersht, *Philos. Trans. R. Soc., B*, 1995, **348**, 61–70.
6 Y. Zhao and J. S. Moore, in *Foldamers*, ed. I. Huc and S. Hecht, John Wiley & Sons, Ltd, 2007, pp. 75–108.
7 Y. Hua, Y. Liu, C.-H. Chen and A. H. Flood, *J. Am. Chem. Soc.*, 2013, **135**, 14401–14412.
8 M. T. Stone, J. M. Heemstra and J. S. Moore, *Acc. Chem. Res.*, 2006, **39**, 11–20.
9 K. P. Divya, S. Sreejith, C. H. Suresh, D. S. Philips and A. Ajayaghosh, *Chem. – Asian J.*, 2013, **8**, 1579–1586.
10 K.-J. Chang, B.-N. Kang, M.-H. Lee and K.-S. Jeong, *J. Am. Chem. Soc.*, 2005, **127**, 12214–12215.
11 H. Juwarker and K.-S. Jeong, *Chem. Soc. Rev.*, 2010, **39**, 3664–3674.
12 V. R. Naidu, M. C. Kim, J. Suk, H.-J. Kim, M. Lee, E. Sim and K.-S. Jeong, *Org. Lett.*, 2008, **10**, 5373–5376.
13 E. A. John, C. J. Massena and O. B. Berryman, *Chem. Rev.*, 2020, **120**, 2759–2782.
14 C. J. Massena, D. A. Decato and O. B. Berryman, *Angew. Chem., Int. Ed.*, 2018, **57**, 16109–16113.
15 S. Peddi, M. C. Bookout, G. N. Vemuri and C. S. Hartley, *J. Org. Chem.*, 2022, **87**, 3686–3690.
16 C. S. Hartley, *Acc. Chem. Res.*, 2016, **49**, 646–654.
17 H. Jiang, J.-M. Léger and I. Huc, *J. Am. Chem. Soc.*, 2003, **125**, 3448–3449.
18 X. Yang, A. L. Brown, M. Furukawa, S. Li, W. E. Gardinier, E. J. Bukowski, F. V. Bright, C. Zheng, X. C. Zeng and B. Gong, *Chem. Commun.*, 2003, 56–57.
19 A. M. Abramyan, Z. Liu and V. Pophristic, *Chem. Commun.*, 2015, **52**, 669–672.
20 Z. Liu, A. M. Abramyan and V. Pophristic, *New J. Chem.*, 2015, **39**, 3229–3240.
21 L. Yuan, H. Zeng, K. Yamato, A. R. Sanford, W. Feng, H. S. Atreya, D. K. Sukumaran, T. Szyperski and B. Gong, *J. Am. Chem. Soc.*, 2004, **126**, 16528–16537.
22 J. Wang, B. Wicher, V. Maurizot and I. Huc, *Chem. – Eur. J.*, 2021, **27**, 1031–1038.
23 T. Qi, V. Maurizot, H. Noguchi, T. Charoenraks, B. Kauffmann, M. Takafuji, H. Ihara and I. Huc, *Chem. Commun.*, 2012, **48**, 6337–6339.
24 J. M. Cary and J. S. Moore, *Org. Lett.*, 2002, **4**, 4663–4666.
25 X. Yang, L. Yuan, K. Yamato, A. L. Brown, W. Feng, M. Furukawa, X. C. Zeng and B. Gong, *J. Am. Chem. Soc.*, 2004, **126**, 3148–3162.
26 H. H. Nguyen, J. H. McAliley and D. A. Bruce, *Macromolecules*, 2011, **44**, 60–67.
27 M. Banno, T. Yamaguchi, K. Nagai, C. Kaiser, S. Hecht and E. Yashima, *J. Am. Chem. Soc.*, 2012, **134**, 8718–8728.
28 W. Hu, N. Zhu, W. Tang and D. Zhao, *Org. Lett.*, 2008, **10**, 2669–2672.
29 H.-G. Jeon, J. Y. Jung, P. Kang, M.-G. Choi and K.-S. Jeong, *J. Am. Chem. Soc.*, 2016, **138**, 92–95.

30 Y. Hua and A. H. Flood, *J. Am. Chem. Soc.*, 2010, **132**, 12838–12840.
31 Y. Liu, F. C. Parks, W. Zhao and A. H. Flood, *J. Am. Chem. Soc.*, 2018, **140**, 15477–15486.
32 J. Shang, Q. Gan, S. J. Dawson, F. Rosu, H. Jiang, Y. Ferrand and I. Huc, *Org. Lett.*, 2014, **16**, 4992–4995.
33 C. Dolain, A. Grélard, M. Laguerre, H. Jiang, V. Maurizot and I. Huc, *Chem. – Eur. J.*, 2005, **11**, 6135–6144.
34 *Maestro, Schrödinger*, LLC, New York, NY, 2021.
35 X. Hu, S. J. Dawson, P. K. Mandal, X. de Hatten, B. Baptiste and I. Huc, *Chem. Sci.*, 2017, **8**, 3741–3749.
36 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.
37 C. Bao, B. Kauffmann, Q. Gan, K. Srinivas, H. Jiang and I. Huc, *Angew. Chem., Int. Ed.*, 2008, **47**, 4153–4156.
38 Y. Ferrand and I. Huc, *Acc. Chem. Res.*, 2018, **51**, 970–977.
39 W. Wang, C. Zhang, S. Qi, X. Deng, B. Yang, J. Liu and Z. Dong, *J. Org. Chem.*, 2018, **83**, 1898–1902.
40 B. Baptiste, C. Douat-Casassus, K. Laxmi-Reddy, F. Godde and I. Huc, *J. Org. Chem.*, 2010, **75**, 7175–7185.
41 V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat and I. Huc, *Chem. – Eur. J.*, 2023, DOI: **10.1002/chem.202300898**.
42 G. Lautrette, B. Wicher, B. Kauffmann, Y. Ferrand and I. Huc, *J. Am. Chem. Soc.*, 2016, **138**, 10314–10322.
43 N. Chandramouli, Y. Ferrand, G. Lautrette, B. Kauffmann, C. D. Mackereth, M. Laguerre, D. Dubreuil and I. Huc, *Nat. Chem.*, 2015, **7**, 334–341.
44 D. J. Hill and J. S. Moore, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 5053–5057.
45 R. B. Prince, J. G. Saven, P. G. Wolynes and J. S. Moore, *J. Am. Chem. Soc.*, 1999, **121**, 3114–3121.

Supplementary Information

Combining local conformational preferences and solvophobic effects to drive helical folding of aromatic oligoamides

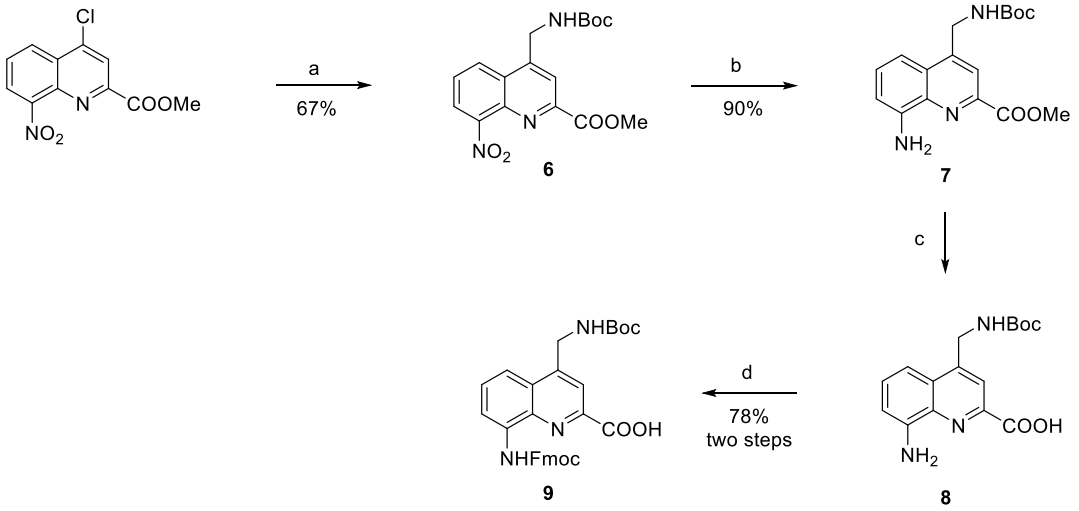
Binhao Teng, [a] Joan Atcher, [a] Lars Allmendinger, [a] Céline Douat,[a] Yann Ferrand [b] and Ivan Huc* [a]

[a.] Department of Pharmacy, Ludwig-Maximilians-Universität, Butenandtstr. 5–13, 81377, München, Germany.

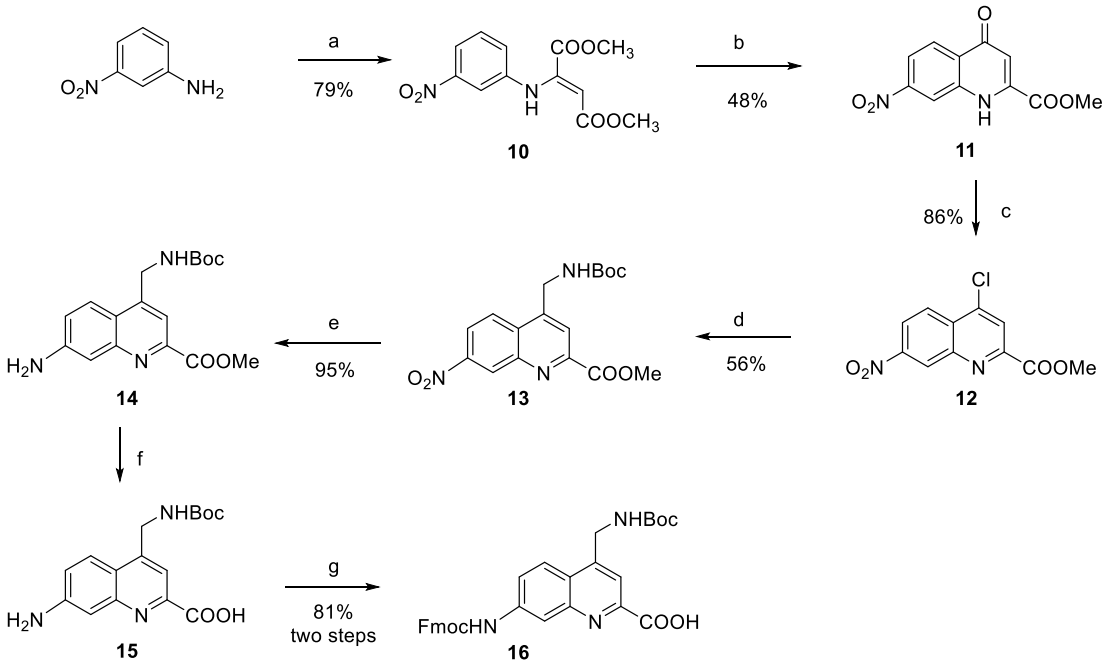
[b.] Univ. Bordeaux, CNRS, Bordeaux Institut National Polytechnique, CBMN UMR 5248, 2 rue Escarpit, 33600 Pessac, France.

Table of contents

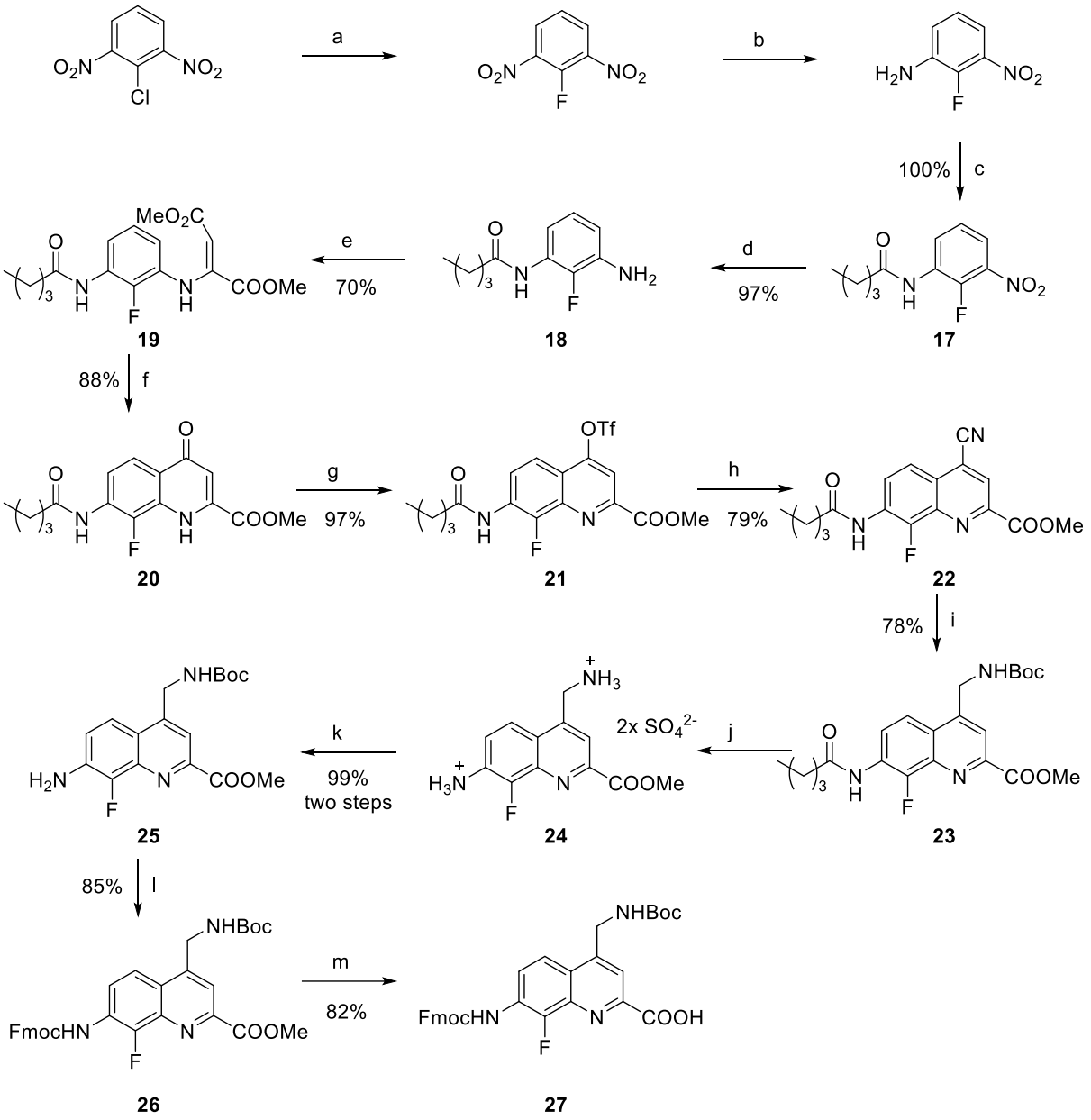
1 Supplementary Schemes and Figures



Scheme S1. Synthesis of **Q** monomer: (a) KF_3BCH_2NHBoc, $Pd(OAc)_2$, SPhos, K_2CO_3, toluene, H_2O, 85 °C; (b) Pd/C, H_2, THF, MeOH; (c) LiOH, THF, H_2O, 0 °C; (d) Fmoc-Cl, $NaHCO_3$, dioxane, H_2O, 0 °C to r.t..



Scheme S2. Synthesis of $\mathbf{Q^H}$ monomer: (a) Dimethyl acetylenedicarboxylate, MeOH; (b) PhOPh, reflux; (c) $POCl_3$, DMF; (d) KF_3BCH_2NHBoc, $Pd(OAc)_2$, SPhos, K_2CO_3, toluene, H_2O, 85 °C; (e) Pd/C, H_2, THF, MeOH; (f) LiOH, THF, H_2O, 0 °C; (g) Fmoc-Cl, $NaHCO_3$, dioxane, H_2O, 0 °C to r.t..



Scheme S3. Synthesis of $\mathbf{Q^F}$ monomer: (a) KF, dry DMSO; 100 °C; (b) Fe, MeOH, AcOH, 110 °C; (c) DIEA, valeric anhydride, 40 °C; (d) Pd/C, H_2, AcOEt; (e) Dimethyl acetylenedicarboxylate, MeOH; (f) PhOPh, reflux; (g) Tf_2O, Pyridine, DCM; (h) KCN, $Pd(PPh_3)_4$, toluene, 100 °C; (i) Boc_2O, NH_4HCOO, Pd/C, THF; (j) H_2SO_4, MeOH, reflux; (k) Boc_2O, $NaHCO_3$, DMF; (l) Fmoc-Cl, $NaHCO_3$, dioxane, H_2O, 0 ^oC to r.t.; (m) LiI, AcOEt.

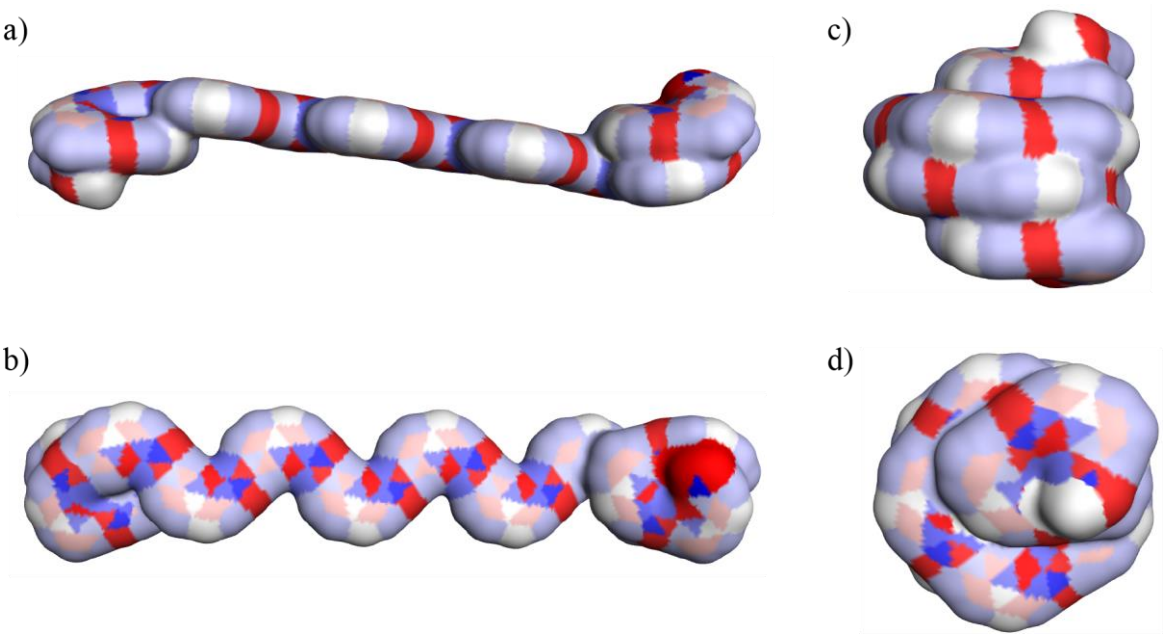


Figure S1. Solvent accessible surface of two conformations: (a) side view and (b) top view of extended **2**. Solvent accessible surface of: (c) side view and (d) top view of folded **2**. Different elements have been marked by different colors (hydrogen: white, carbon: grey, oxygen: red, nitrogen: blue). Solvent-accessible surface area was calculated by discovery studio software with a radius of 1.4 Å. Short side chains have been removed for the surface estimation. The calculated solvent accessible surface for **2** of extended and folded state were found to be 2578.08 $Å^2$ and 1717.48 $Å^2$, respectively. Thus, one can estimate that a folded **2** presents to the solvent about 67% of the surface presented by extended **2**.

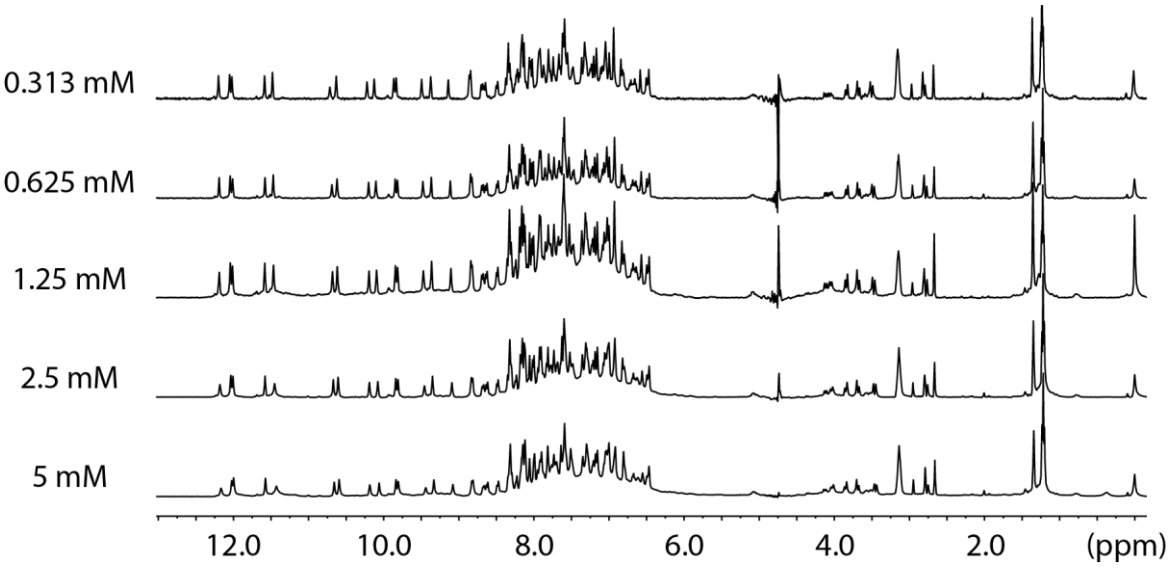


Figure S2. Variable concentration ^{1}H NMR spectra of compound 2. The spectra were recorded with a water suppression method (500 MHz, 10% D_2O in H_2O, 25 °C).

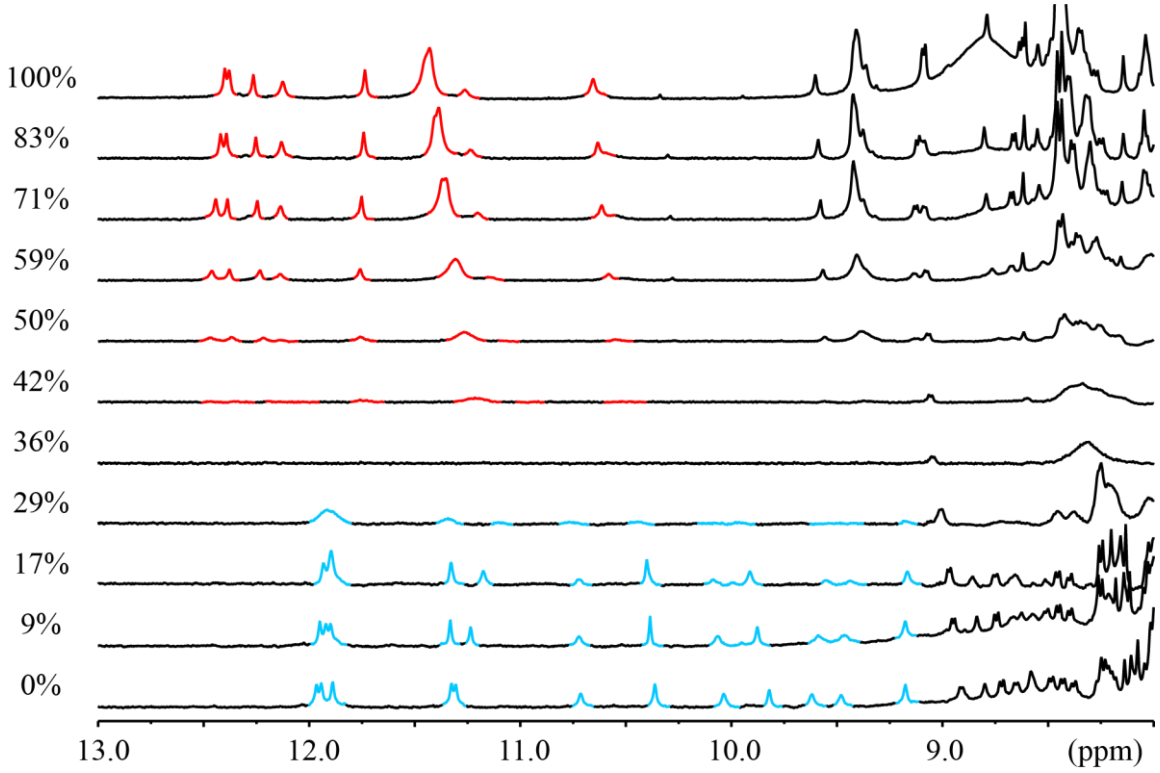


Figure S3. DMSO-d_6/CD_3OH variation ^{1}H NMR spectra of 2. Excerpts of ^{1}H NMR (500 MHz, water suppression) titration at 25 °C in DMSO-d_6/CD_3OH mixtures with different DMSO percentages. Note: 0% DMSO means that ^{1}H NMR spectrum (500 MHz) of compound **2** was performed in CD_3OH. Red and blue marked amide proton signals represent unfolded and folded conformations, respectively.

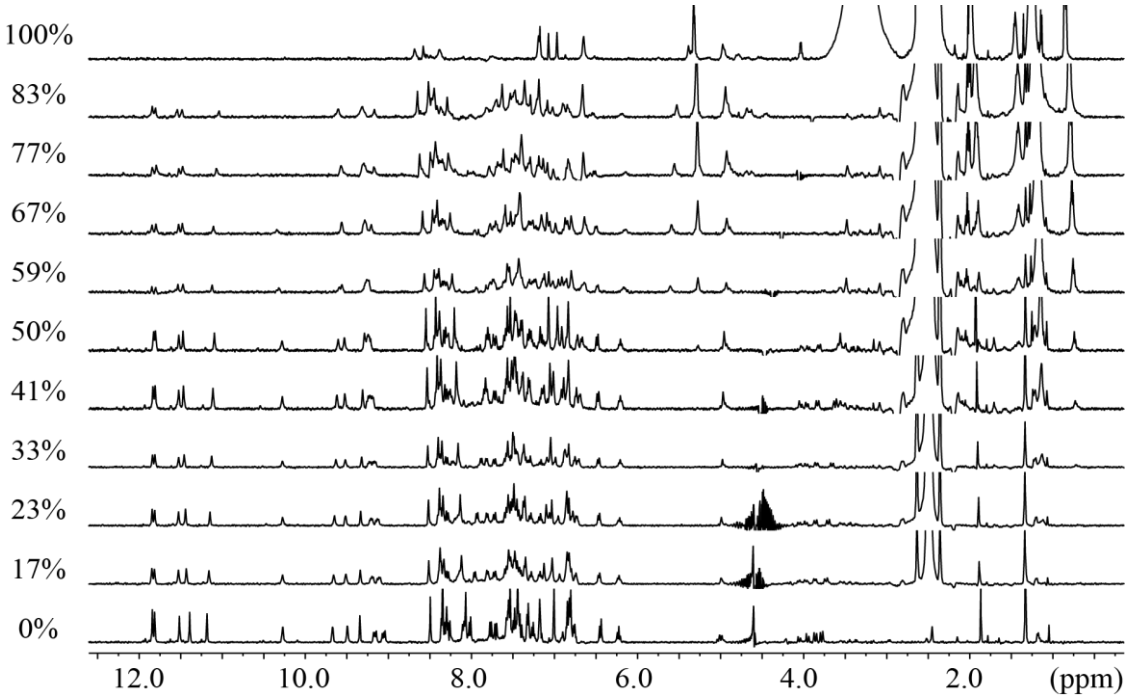


Figure S4. DMSO-d_6/H_2O variation ^{1}H NMR spectra of 5. ^{1}H NMR (500 MHz, water suppression) spectra at 25°C in DMSO-d_6/H_2O mixtures with different DMSO percentages. Note: 0% DMSO means that ^{1}H NMR (500 MHz) spectra of compound **5** were recorded in 90%H_2O+10%D_2O.

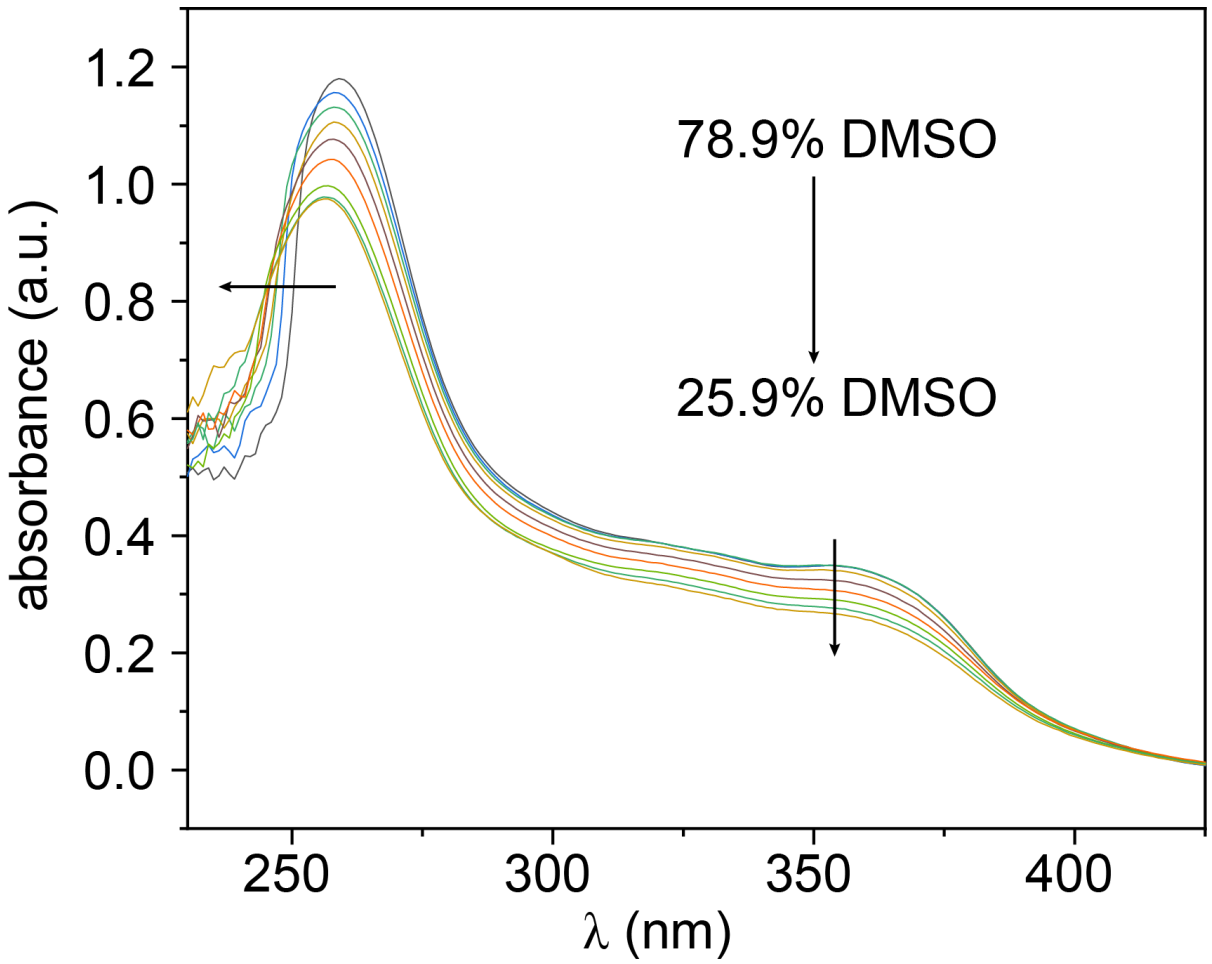


Figure S5. Water/DMSO proportion variation UV spectra of compound 1. UV spectra in different proportions of water/DMSO.

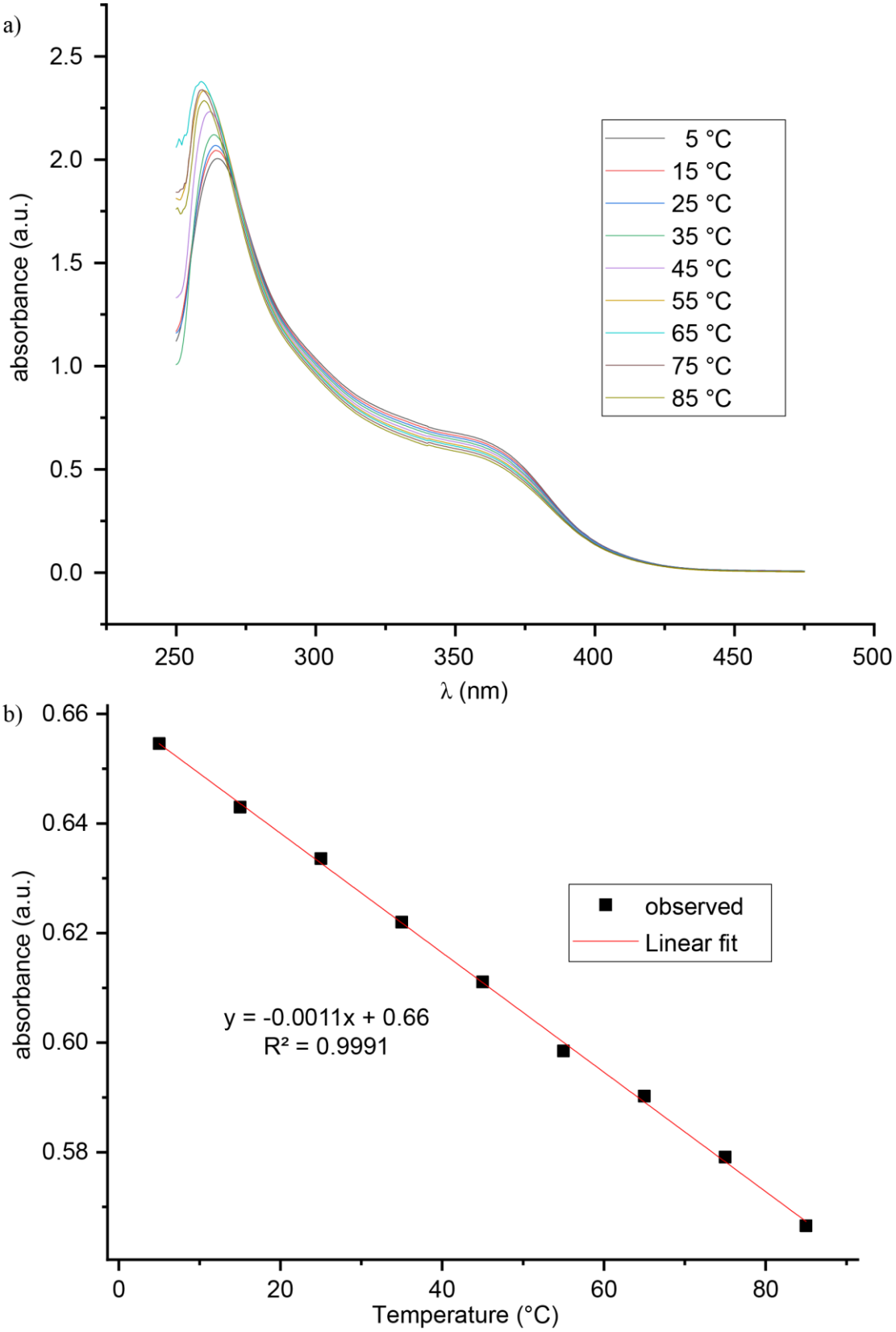


Figure S 6. Temperature variation UV of 2. (a) Variable temperature UV spectra of compound **2** in DMSO:H_2O (75.5:24.5, *v*/*v*). (b) Plot of the absorbance at 357 nm *vs* temperature.

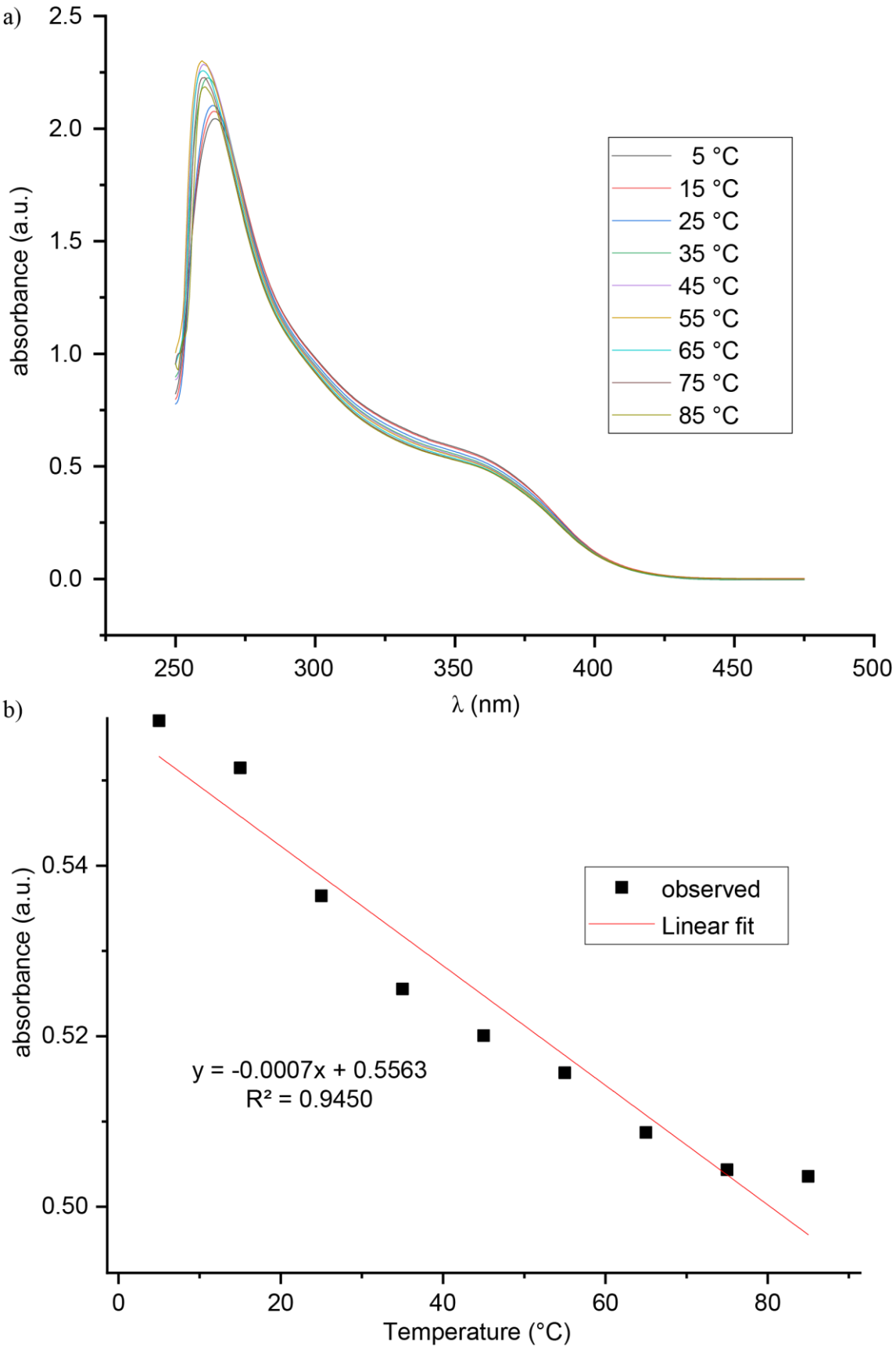


Figure S 7. Temperature variation UV of 3. (a) Variable temperature UV spectra of compound **3** in DMSO:H_2O (81.6:18.4, *v*/*v*). (b) Plot of the absorbance at 357 nm *vs* temperature.

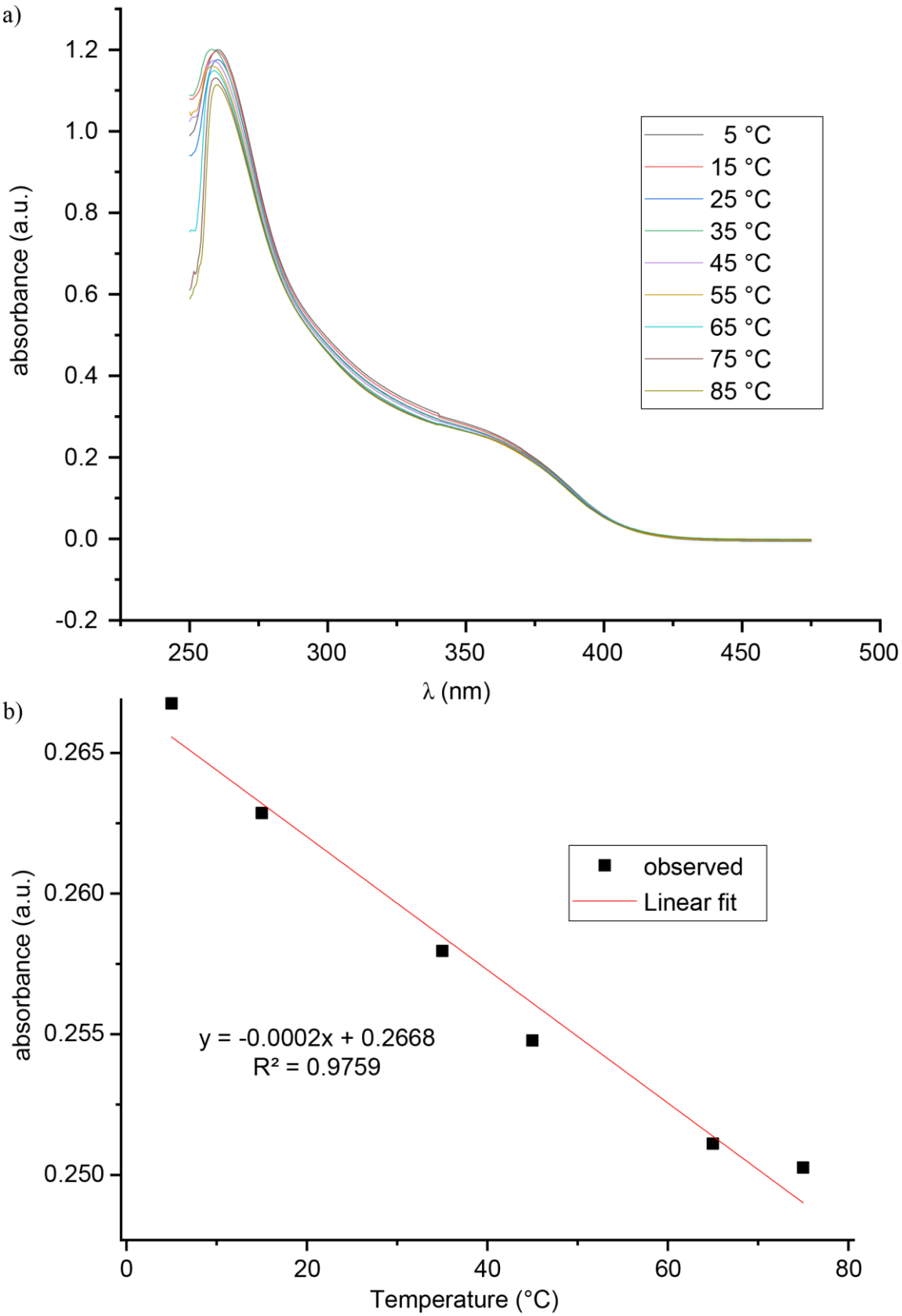


Figure S 8. Temperature variation UV of 4. (a) Variable temperature UV spectra of compound **4** in DMSO:H_2O (83.3:16.7, *v*/*v*). (b) Plot of the absorbance at 357 nm *vs* temperature.

2 Materials and Methods

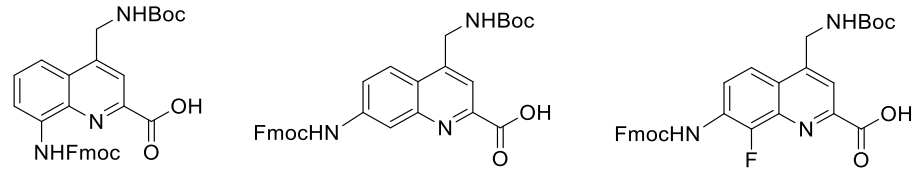


Figure S9. Fmoc-acid building blocks used in this study. For a detailed procedure to Fmoc-Q(Boc)-OH, Fmoc-Q^H(Boc)-OH and Fmoc-Q^F(Boc)-OH, see section 2.3.

2.1 General

General. Commercial reagents (Suppliers: Abcr, Fisher Scientific, Merck, Sigma-Aldrich, TCI, BLDpharm or VWR) were used without further purification unless otherwise stated. Wang resin LL (100–200 mesh) was purchased from Sigma-Aldrich. Peptide grade *N*,*N*-dimethylformamide (DMF) was purchased from Carlo Erba. Anhydrous chloroform, triethylamine (TEA) and *N*,*N*-diisopropylethylamine (DIEA) were obtained via distillation over CaH_2 prior to use. Anhydrous tetrahydrofuran (THF) and dichloromethane (DCM) were obtained via an MBRAUN SPS-800 solvent purification system. Ultrapure water was obtained via a Sartorius arium® pro VF 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). Nuclear magnetic resonance (NMR) spectra were recorded on an Avance III HD 300 MHz Bruker BioSpin spectrometer or an Avance III HD 500 MHz Bruker BioSpin spectrometer equipped with a broad band observe 5-mm BB-H&FD CryoProbe™ Prodigy. Measurements were performed at 25 °C unless stated otherwise. Water suppression was performed with excitation sculpting method. Processing was done with MestReNova (v.12.0.0-20080) NMR processing software from Mestrelab Research. Chemical shifts are reported in ppm and calibrated via residual solvent signals or 3-(trimethylsilyl)propionic-2,2,3,3-d_4 acid sodium salt (TMSP) when water suppression was applied. Signal multiplicities are abbreviated as s, singlet; d, doublet; t, triplet; q, quartet, and m, multiplet. Signals were assigned using ^{1}H-^{13}C HMQC and ^{1}H-^{13}C HMBC spectra. Electrospray ionization (ESI) mass spectra were recorded on Bruker microTOF II and Thermo Finnigan LTQ FT Ultra spectrometers. Electron ionization (EI) mass spectra were recorded on a Thermo Q Exactive GC Orbitrap or a Finnigan MAT 95 sector mass

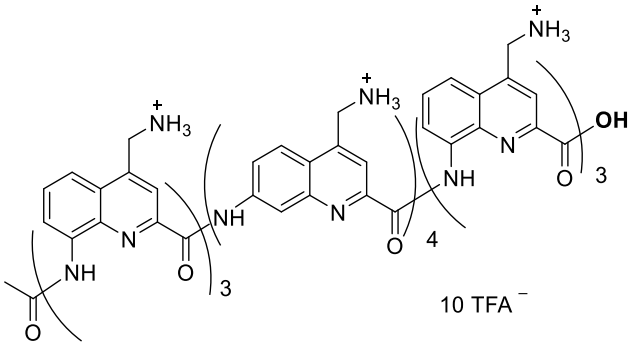
spectrometer. Analytical and semi-preparative reversed phase (RP) high performance liquid chromatography (HPLC) was performed on a Thermo Fisher Scientific Ultimate 3000 HPLC System using Macherey-Nagel Nucleodur C18 Gravity columns (4 × 100 mm, 5 µm and 10 × 250 mm, 5 µm) and Macherey-Nagel Nucleodur C8 Gravity columns (4 × 50 mm, 5 µm and 10 × 100 mm, 5 µm) with different gradients of 0.1% TFA water and 0.1% TFA acetonitrile. All ultraviolet–visible (UV/Vis) absorbance measurements were done with a Jasco V-750 spectrophotometer instrument using a 1 cm quartz cuvette. Measurements were performed at 20 °C if not stated otherwise. Microwave-assisted solid phase foldamer synthesis (SPFS) was performed via a CEM® Discover Bio manual microwave peptide synthesizer. The temperature within the reactor vessel was monitored with an optical fiber probe. Automated SPFS was done on a PurePep® Chorus synthesizer (Gyros Protein Technologies) by applying an induction heating.

2.2 Procedure for manual solid phase foldamer synthesis

Oligomers were synthesized according to previously reported SPFS protocols,[1] hereafter referred to as manual synthesis method. The automated synthesis protocol has been reported in previous work.[2]

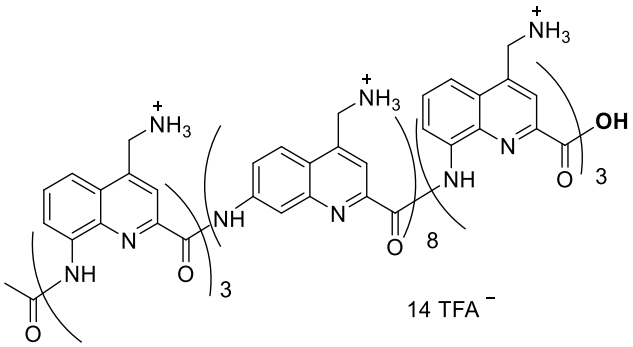
Final acetylation of the N-terminal aromatic amine: H_2N-$(Q)_n$-Wang resin and DIEA (6.0 equiv.) were suspended in anhydrous THF (1.25 mL) then acetyl chloride (4.0 equiv.) in anhydrous THF (1.25 mL) was added. The reaction vessel was then placed under microwave irradiation (25 W, ramp to 50 °C over 5 min, then hold at 50 °C for 15 min). The resin was filtered off and washed with anhydrous THF (2 x 3 mL). The coupling step was repeated twice using the same conditions and number of equivalents of coupling reagents.

Resin cleavage: The resin-bound foldamer was placed in a syringe equipped with a filter, washed with DMF (3 x 3 mL) and DCM (3 x 3 mL), and dried by passing N_2 flow through it. It was then suspended in a solution of TFA/*i*Pr_3SiH/H_2O (95:2.5:2.5, *v*/*v*/*v*). The resin was next shaken for at least 2 h at room temperature. The resin was filtered off and rinsed one time with TFA. The foldamer was precipitated from the TFA cleavage solution by adding cold Et_2O and centrifugation to obtain a crude precipitate.



Compound 1: Compound **1** was synthesized starting from LL-Wang resin (0.44 mmol g^{-1}, 25 µmol scale) according to the manual synthesis method. Loading of the first monomer: 0.31 mmol g^{-1} (70%). Final acetylation was carried out via the general acetylation method. After precipitation in cold Et_2O, the crude mixture was purified by semi prep RP-HPLC to give the **1** as a yellow solid (15 mg, 19%).

^{1}H NMR (500 MHz, DMSO-d_6) δ 12.40 (s, 2H), 12.24 (s, 1H), 12.12 (s, 1H), 11.70 (s, 1H), 11.45 (s, 1H), 11.32 (s, 1H), 10.64 (s, 2H), 9.62 (s, 1H), 9.47 (s, 1H), 9.42 (s, 1H), 9.12 (d, *J* = 7.6 Hz, 2H), 8.86 (s, 4H), 8.71 – 8.54 (m, 7H), 8.53 – 8.32 (m, 8H), 8.27 (m, 4H), 8.14 – 7.43 (m, 29H), 7.40 (d, *J* = 9.2 Hz, 1H), 7.23 (s, 1H), 4.99 – 4.34 (m, 27H).
HRMS: calcd. for $C_{112}H_{95}N_{30}O_{12}$ $[M+H]^+$ 2051.7740; found 2051.8066.

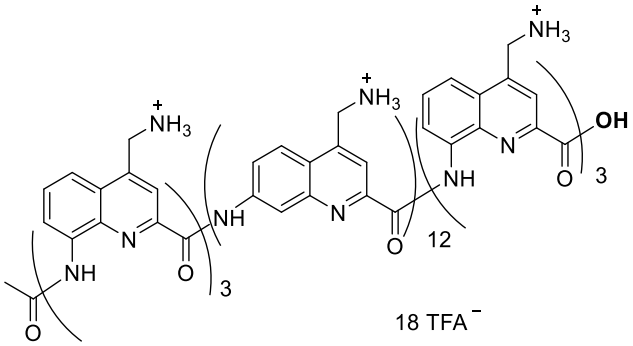


Compound 2: Compound **2** was synthesized starting from LL-Wang resin (0.44 mmol g^{-1}, 17 µmol scale) according to the manual synthesis method. Loading of the first monomer: 0.27 mmol g^{-1} (62%). Final acetylation was carried out via the general acetylation method. After precipitation in cold Et_2O, the crude mixture was purified by semi prep RP-HPLC to give the **2** as a yellow solid (16 mg, 21%).

^{1}H NMR (500 MHz, H_2O/D_2O (9:1, *v*/*v*)) δ 12.19 (s, 1H), 12.04 (s, 1H), 12.01 (s, 1H), 11.58 (s, 1H), 11.47 (s, 1H), 10.69 (s, 1H), 10.62 (s, 1H), 10.20 (s, 1H), 10.10 (s, 1H), 9.85 (s, 1H), 9.82 (s, 1H), 9.48 (s, 1H), 9.37 (s, 1H), 9.11 (s, 1H), 8.83 (d, *J* = 9.3 Hz, 2H), 8.66 (m, 3H), 8.49 (d,

J = 9.7 Hz, 2H), 8.39 – 8.27 (m, 6H), 8.25 (d, *J* = 9.0 Hz, 3H), 8.15 (m, 10H), 8.09 – 7.98 (m, 6H), 7.98 – 7.42 (m, 39H), 7.39 – 7.12 (m, 18H), 7.12 – 6.41 (m, 29H).

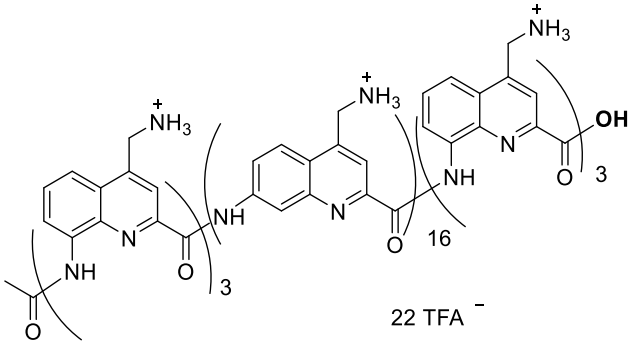
HRMS: calcd. for $C_{156}H_{132}N_{42}O_{16}$ $[M+2H]^{2+}$ 1424.5398; found 1424.5300.



Compound 3: Compound **3** was synthesized starting from LL-Wang resin (0.44 mmol g^{-1}, 23 µmol scale) according to the manual synthesis method. Loading of the first monomer: 0.31 mmol g^{-1} (70%). Final acetylation was carried out via the general acetylation method. After precipitation in cold Et_2O, the crude mixture was purified by semi prep RP-HPLC to give the **3** as a yellow solid (42 mg, 32%).

^{1}H NMR (500 MHz, H_2O/D_2O (9:1, *v*/*v*)) δ 12.03 (s, 1H), 11.95 (s, 1H), 11.86 (s, 1H), 11.38 (s, 1H), 11.35 (s, 1H), 10.39 (s, 1H), 10.16 (s, 1H), 9.85 (s, 1H), 9.83 (s, 1H), 9.73 (s, 1H), 9.52 (s, 1H), 9.43 (s, 1H), 9.38 (s, 1H), 9.31 (s, 1H), 9.11 (s, 1H), 9.03 (d, *J* = 4.6 Hz, 3H), 8.69 (d, *J* = 8.1 Hz, 1H), 8.45 (d, *J* = 10.7 Hz, 1H), 8.40 – 8.30 (m, 3H), 8.16 – 8.00 (m, 4H), 7.99 – 7.87 (m, 6H), 7.88 – 7.78 (m, 6H), 7.78 – 7.57 (m, 20H), 7.56 – 7.30 (m, 18H), 7.26 – 7.10 (m, 5H), 7.08 – 6.75 (m, 18H), 6.68 (d, *J* = 9.1 Hz, 1H), 6.55 – 6.44 (m, 3H), 6.41 (s, 1H), 6.31 (d, *J* = 9.0 Hz, 1H), 4.19 (d, *J* = 14.6 Hz, 1H), 4.01 (d, *J* = 14.9 Hz, 1H), 3.70 – 3.55 (m, 2H), 3.47 (m, 1H).

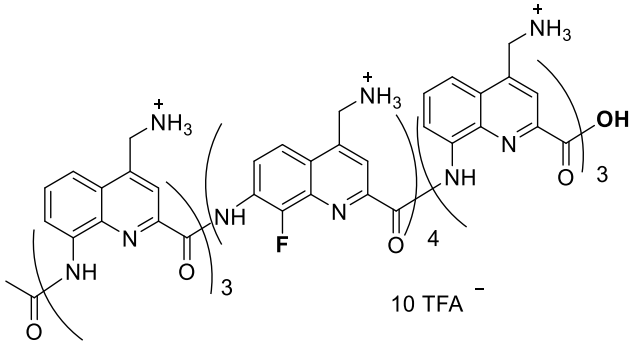
HRMS: calcd. for $C_{200}H_{168}N_{54}O_{20}$ $[M+2H]^{2+}$ 1822.6889; found 1822.7448.



Compound 4: Compound **4** was synthesized on Cl-MPA-Protide resin (0.17 mmol g^{-1}, 24 µmol scale) according to the automated synthesis method. Loading of the first monomer: 0.12

mmol g^{-1} (70%). Final acetylation was carried out via the general acetylation method. After precipitation in cold Et_2O, the crude mixture was purified by semi prep RP-HPLC to give the **4** as a yellow solid (25 mg, 15%).

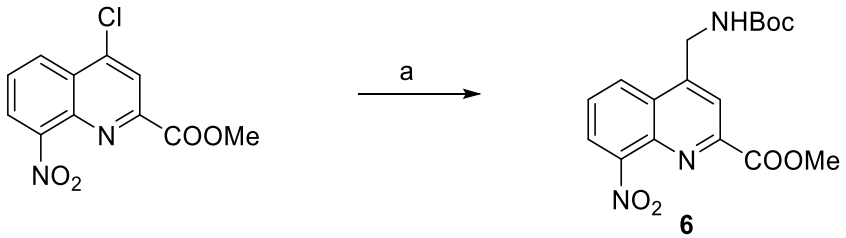
^{1}H NMR (500 MHz, H_2O/D_2O (9:1, *v*/*v*)) δ 11.81 (s, 2H), 11.76 (s, 1H), 11.19 (s, 1H), 11.15 (s, 1H), 10.26 (s, 1H), 10.08 (s, 1H), 9.98 (s, 1H), 9.80 (s, 1H), 9.60 (s, 1H), 9.58 (s, 1H), 9.52 (s, 1H), 9.48 (s, 1H), 9.35 (s, 1H), 9.32 (s, 1H), 9.25 (s, 1H), 9.12 (s, 1H), 9.01 (s, 1H), 8.98 (s, 1H), 8.97 (s, 1H), 8.94 (s, 1H), 8.84 (s, 1H), 8.41 (d, *J* = 8.1 Hz, 1H), 8.25 – 8.10 (m, 3H), 8.02 (d, *J* = 10.0 Hz, 1H), 7.92 – 7.74 (m, 6H), 7.73 – 7.50 (m, 26H), 7.50 – 7.40 (m, 14H), 7.40 – 7.27 (m, 15H), 7.20 (m, 18H), 7.11 – 6.92 (m, 10H), 6.92 – 6.53 (m, 18H), 6.53 – 6.07 (m, 8H), 3.44 (dd, *J* = 11.7, 4.3 Hz, 3H), 3.39 – 3.27 (m, 3H).
HRMS: calcd. for $C_{244}H_{205}N_{66}O_{24}$ $[M+3H]^{3+}$ 1480.8944; found 1480.9345.



Compound 5: Compound **5** was synthesized starting from LL-Wang resin (0.40 mmol g^{-1}, 15 µmol scale) according to the manual synthesis method. Loading of the first monomer: 0.33 mmol g^{-1} (76%). Final acetylation was carried out via the general acetylation method. After precipitation in cold Et_2O, the crude mixture was purified by semi prep RP-HPLC to give the **5** as a yellow solid (5 mg, 10%).

^{1}H NMR (500 MHz, H_2O/D_2O (9:1, *v*/*v*)) δ 11.94 (s, 1H), 11.92 (s, 1H), 11.62 (s, 1H), 11.49 (s, 1H), 11.28 (s, 1H), 10.37 (s, 1H), 9.77 (s, 1H), 9.60 (s, 1H), 9.45 (s, 1H), 9.31 – 9.22 (m, 1H), 9.20 – 9.11 (m, 1H), 8.60 (s, 1H), 8.49 – 8.31 (m, 5H), 8.23 – 8.14 (m, 3H), 8.12 (d, *J* = 8.4 Hz, 1H), 7.87 (d, *J* = 8.6 Hz, 1H), 7.81 (d, *J* = 10.5 Hz, 1H), 7.70 – 7.47 (m, 9H), 7.42 (d, *J* = 8.2 Hz, 2H), 7.36 (t, *J* = 8.2 Hz, 1H), 7.30 – 7.24 (m, 2H), 7.11 (s, 1H), 6.98 – 6.83 (m, 6H), 6.55 (d, *J* = 9.3 Hz, 1H), 6.33 (t, *J* = 8.7 Hz, 1H).
HRMS: calcd. for $C_{112}H_{92}F_4N_{30}O_{12}$ $[M+2H]^{2+}$ 1062.3718; found 1062.3538.

2.3 Monomer synthesis procedures



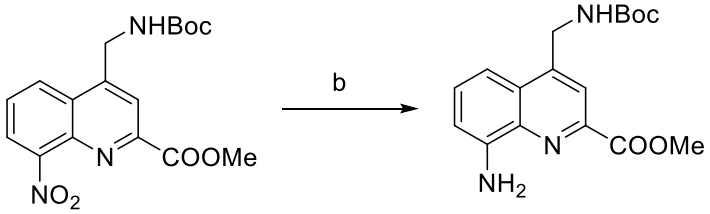
6

Compound 6. Methyl 4-chloro-8-nitroquinoline-2-carboxylate[3] (0.954 g, 0.0036 mol), KF_3BCH_2NHBoc (0.744 g, 0.0037 mol), $Pd(OAc)_2$ (0.036 g, 0.00016 mol), SPhos (0.148 g, 0.00036 mol) and K_2CO_3 (1.248 g, 0.011 mol) were added in a sealed tube and the mixture was purged 3 times with argon. Then toluene (9.6 mL) and water (2.4 mL) were added and the reaction mixture was heated to 85 °C for 18 h. After completion of the reaction that was confirmed by TLC, phosphate buffer (pH = 7) was added then the crude mixture was extracted with DCM. The combined organic layers were washed with water and brine, dried (Na_2SO_4) and then concentrated under reduced pressure. The resulting residue was purified by column chromatography to give 0.86 g (67%) of **6** as a white solid.

^{1}H NMR (500 MHz, DMSO-d_6) δ 8.53 (dd, J = 8.6, 1.2 Hz, 1H), 8.39 (dd, J = 7.5, 1.2 Hz, 1H), 8.13 (s, 1H), 7.93 (dd, J = 8.6, 7.5 Hz, 1H), 7.77 (t, J = 6.0 Hz, 1H), 4.77 (d, J = 6.0 Hz, 2H), 3.96 (s, 3H), 1.43 (s, 9H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 164.6, 155.86, 149.0, 148.9, 148.6, 137.8, 128.0, 127.6, 127.5, 123.9, 119.7, 78.6, 54.9, 53.0, 28.1.

HRMS: calcd. for $C_{17}H_{19}N_3O_6Na$ $[M+Na]^+$ 384.1166; found 384.1180.



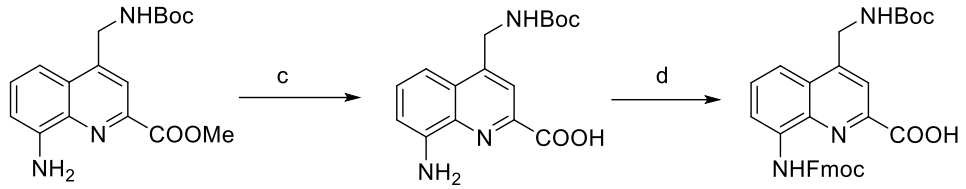
6 **7**

Compound 7. Quinoline **6** (2 g, 0.006 mmol) was dissolved in a mixture of THF (133 mL)/ MeOH (200 mL) under N_2. Then 10 wt.% Pd/C (1 g, 35% *w*/*w*) was added and N_2 was replaced by H_2. The mixture was stirred for 3 h at room temperature and the reaction was checked by TLC. After completion of the reaction, the mixture was filtered and concentrated then the residue was precipitated in diethyl ether to give 1.65 g (90%) of **7** as a yellow powder.

^{1}H NMR (500 MHz, DMSO-d_6) δ 7.92 (s, 1H), 7.63 (t, J = 6.1 Hz, 1H), 7.44 (dd, J = 8.3, 7.7 Hz, 1H), 7.22 (dd, J = 8.3, 1.1 Hz, 1H), 6.95 (dd, J = 7.7, 1.1 Hz, 1H), 6.14 (s, 2H), 4.58 (d, J = 6.0 Hz, 2H), 3.94 (s, 3H), 1.43 (s, 9H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 165.5, 155.9, 146.8, 146.4, 143.3, 136.2, 130.3, 127.7, 124.9, 117.7, 109.2, 108.6, 78.3, 52.5, 28.2.

HRMS: calcd. for $C_{17}H_{22}N_3O_4$ $[M+H]^+$ 332.1605; found 332.1600.

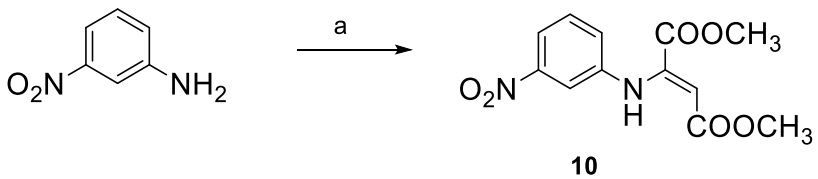


Compound 9. The protocol was modified from the reference (X. Hu, et al., *Chem. Sci.*, 2017, **8**, 3741–3749.).[3] Compound **7** (3.43 g, 0.01 mol) was dissolved in dioxane (206 mL) then LiOH·H_2O (0.65 g, 0.015 mol) in H_2O (65 mL) was added slowly at 0 °C. The mixture was stirred for 1 h at 0 °C, then neutralized by 0.1 mol/L HCl. Then H_2O (80 mL) and $NaHCO_3$ (2.61 g, 0.031 mol) were added and cooled to 0 °C. Fmoc-Cl (3.48 g, 0.013 mol) was added slowly as a solution in dioxane. The resulting mixture was stirred at 0 ° for 1 h and let to stir overnight at room temperature. After completion of the reaction that was confirmed by TLC, the solution was acidified to pH=4 using a saturated citric acid solution. Then water was added and the aqueous layers were extracted with DCM. The combined organic layers were washed with water and brine, dried (Na_2SO_4) and then concentrated under reduced pressure. The residue was precipitated in acetonitrile, filtered and washed with a small amount of acetonitrile to give 4.36 g (78%) of **9** as a grey powder.

^{1}H NMR (500 MHz, DMSO-d_6) δ 13.60 (s, 1H), 10.49 (s, 1H), 8.10 (s, 1H), 7.93 (dt, J = 7.5, 0.9 Hz, 2H), 7.84 (d, J = 8.4 Hz, 1H), 7.78 (dd, J = 7.4, 1.0 Hz, 2H), 7.72 (t, J = 6.7 Hz, 2H), 7.44 (t, J = 7.5 Hz, 3H), 7.36 (td, J = 7.4, 1.2 Hz, 2H), 4.70 (d, J = 6.0 Hz, 2H), 4.62 (d, J = 6.9 Hz, 2H), 4.46 (t, J = 6.8 Hz, 1H), 1.43 (s, 9H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 165.5, 155.9, 153.5, 148.4, 144.9, 143.7, 140.8, 136.5, 136.3, 129.5, 127.8, 127.3, 127.2, 125.2, 120.3, 118.1, 117.6, 116.5, 116.1, 78.4, 66.4, 46.6, 28.2.

HRMS: calcd. for $C_{31}H_{30}N_3O_6$ $[M+H]^+$ 540.2129; found 540.2138.



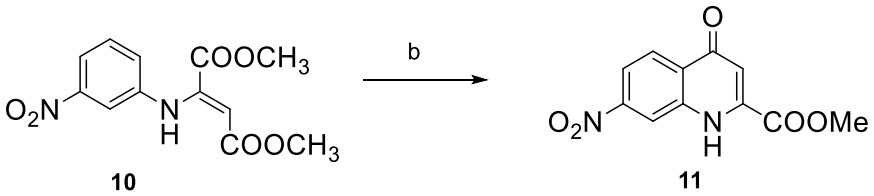
Compound 10. The protocol was modified from the reference (N. D. Heindel, et al., *J. Org. Chem.*, 1967, **32**, 4155–4157.).[4] 3-Nitroaniline (20 g, 0.145 mol) was dissolved in MeOH (350

mL) under N_2. The solution was cooled to 0 ˚C in an ice-bath and dimethyl acetylenedicarboxylate (26.7 mL, 0.217 mol) was added dropwise. The mixture was stirred at 0 ˚C for 30 min and then at room temperature for 23 h, during which time the product precipitated. The suspension was cooled to 0 ˚C to introduce additional precipitation. The product was then filtered off and washed with cold MeOH to give 32 g (79%) of **10** as a white solid.

^{1}H NMR (500 MHz, DMSO-d_6) δ 9.70 (s, 1H), 7.86 (ddd, J = 8.3, 2.2, 0.9 Hz, 1H), 7.81 (t, J = 2.2 Hz, 1H), 7.55 (t, J = 8.1 Hz, 1H), 7.33 (ddd, J = 8.1, 2.3, 0.9 Hz, 1H), 5.57 (s, 1H), 3.71 (s, 3H), 3.67 (s, 3H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 167.2, 164.2, 148.1, 144.5, 142.1, 130.2, 125.6, 117.4, 113.7, 98.1, 53.2, 51.3.

HRMS: calcd. for $C_{12}H_{12}N_2O_6$ $[M]^+$ 280.0690; found 280.0692.



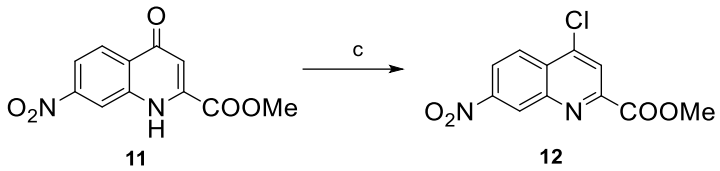
10 11

Compound 11. The protocol was modified from the reference (N. D. Heindel, et al., *J. Org. Chem.*, 1967, **32**, 4155–4157.).[4] Diphenyl ether (680 mL) was placed in a 1 L round bottom flask and heated using a heating mantle to its boiling point at 260 ˚C. Then compound **10** (32 g, 0.114 mol) was added to the boiling solvent by means of a glass funnel. The mixture was kept boiling for 20 min; no condenser was used. The reaction mixture was left to cool down to room temperature. The product was filtered off, and the solid was washed several times with cyclohexane then dried under reduced pressure. The product was recrystallized from hot DMF to give 13.5g (48%) of **11** as a golden solid.

^{1}H NMR (500 MHz, DMSO-d_6) δ 12.49 (s, 1H), 8.86 (d, J = 2.3 Hz, 1H), 8.27 (d, J = 8.9 Hz, 1H), 8.07 (dd, J = 8.9, 2.3 Hz, 1H), 6.73 (s, 1H), 3.99 (s, 3H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 176.9, 162.2, 149.7, 139.8, 139.3, 128.7, 127.2, 117.5, 115.9, 111.5, 53.8.

HRMS: calcd. for $C_{11}H_8N_2O_5$ $[M]^+$ 248.0428; found 248.0437.

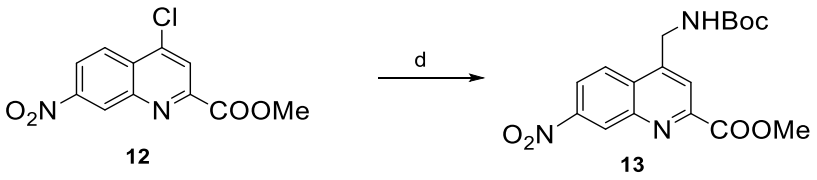


Compound 12. Compound **11** (8.86 g, 0.036 mol) was dissolved in dry DMF (93 mL) then $POCl_3$ (3.66 mL, 0.039 mmol) was added slowly. The reaction mixture was stirred for 15 min at 80 °C. After completion of the reaction that was confirmed by TLC, the mixture was poured into ice and neutralised with saturated $NaHCO_3$ solution. Then the mixture was extracted with DCM. The combined organic layers were washed with water and brine, dried (Na_2SO_4) and then concentrated under reduced pressure to give 8.18 g (86%) of **12** as a brown solid.

^{1}H NMR (500 MHz, DMSO-d_6) δ 9.01 (dd, J = 2.3, 0.6 Hz, 1H), 8.59 (dd, J = 9.2, 2.3 Hz, 1H), 8.54 (dd, J = 9.2, 0.6 Hz, 1H), 8.44 (s, 1H), 4.01 (s, 3H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 163.8, 150.0, 149.1, 146.6, 143.2, 129.6, 126.4, 126.0, 123.6, 123.3, 53.2.

HRMS: calcd. for $C_{11}H_7ClN_2O_4$ $[M]^+$ 266.0089; found 266.0086.

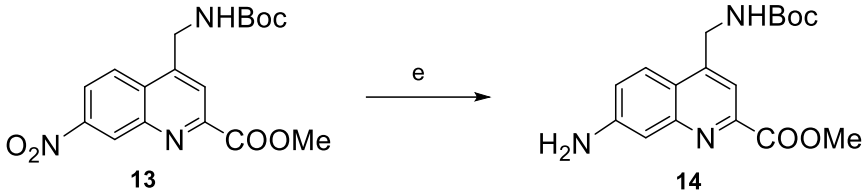


Compound 13. Compound **12** (0.954 g, 0.0036 mol), KF_3BCH_2NHBoc (0.744 g, 0.0037 mol), $Pd(OAc)_2$ (0.036 g, 0.00016 mol), SPhos (0.148 g, 0.00036 mol) and K_2CO_3 (1.248 g, 0.011 mol) were added in a sealed tube and the mixture was purged 3 times with argon. Then toluene (9.6 mL) and water (2.4 mL) were added and heated to 85 °C for 30 h. After completion of the reaction that was confirmed by TLC, pH = 7 phosphate buffer was added then the mixture extracted with DCM. The combined organic layers were washed with water and brine, dried (Na_2SO_4) and then concentrated under reduced pressure. The resulting residue was purified by column chromatography. The recovered quinoline derivative **12** was used to repeat the reaction again and the product that was obtained after column was precipitated in DCM/diethyl ether to give 0.729 g (56%) of **13** as white solid.

^{1}H NMR (500 MHz, DMSO-d_6) δ 8.96 (d, J = 2.3 Hz, 1H), 8.52 (d, J = 9.2 Hz, 1H), 8.46 (dd, J = 9.2, 2.4 Hz, 1H), 8.15 (s, 1H), 7.78 (t, J = 5.9 Hz, 1H), 4.77 (d, J = 6.0 Hz, 2H), 3.99 (s, 3H), 1.43 (s, 9H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 165.3, 156.3, 150.2, 148.9, 148.8, 146.5, 130.6, 126.7, 126.5, 122.2, 121.1, 79.0, 53.5, 41.1, 28.7.

HRMS: calcd. for $C_{17}H_{19}N_3O_6Na$ $[M+Na]^+$ 384.1166; found 384.1170.

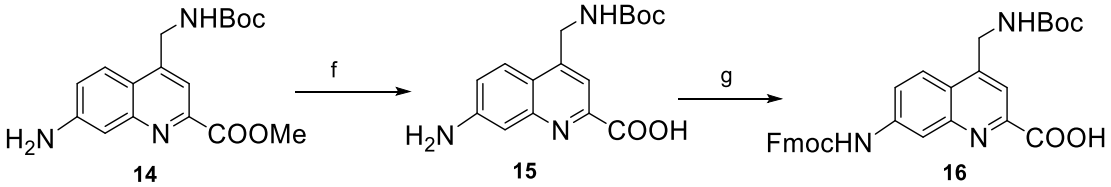


Compound 14. Compound **13** (2.91 g, 0.008 mmol) was dissolved in mixture of THF (260 mL) and MeOH (390 mL) under N_2. Then 10 wt.% Pd/C (1 g, 35% *w*/*w*) was added and N_2 was replaced by H_2. The mixture was stirred for 3 h at room temperature. After completion of the reaction that was confirmed by TLC, the mixture was filtered and concentrated then the residue was precipitated in diethyl ether to give 2.54 g (95%) of **14** as a yellow powder.

^{1}H NMR (500 MHz, DMSO-d_6) δ 7.90 (d, *J* = 9.0 Hz, 1H), 7.61 (s, 1H), 7.59 (t, *J* = 6.1 Hz, 1H), 7.14 (dd, *J* = 9.0, 2.3 Hz, 1H), 7.04 (d, *J* = 2.3 Hz, 1H), 5.95 (s, 2H), 4.56 (d, *J* = 6.1 Hz, 2H), 3.91 (s, 3H), 1.43 (s, 9H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 166.4, 156.3, 151.0, 145.0, 147.5, 146.7, 124.4, 121.4, 119.8, 114.2, 107.6, 78.7, 52.8, 40.8, 28.7.

HRMS: calcd. for $C_{17}H_{22}N_3O_4$ $[M+H]^+$ 332.1605; found 332.1608.



Compound 16. Compound **14** (2.54 g, 0.0077 mol) was dissolved in dioxane (153 ml) then LiOH H_2O (0.483 g, 0.0112 mol) in water (48 mL) was added slowly at 0 °C. The mixture was stirred for 1 h at 0 °C, then neutralized by 0.1 mol·L^{-1} HCl. Then water (14 ml) and $NaHCO_3$ (1.93 g, 0.023 mol) were added and cooled down to 0 °C by ice-bath. Fmoc-Cl (2.58 g, 0.01 mol) is added slowly as a solution in dioxane. The resulting mixture is stirred at 0 °C for 1 h and allowed to warm to room temperature overnight. After completion of the reaction that was confirmed by TLC, the solution was acidified to pH=4 using a saturated citric acid solution. Then water was added and the aqueous layers were extracted with DCM. The combined organic layers were washed with water and brine, dried (Na_2SO_4) and then concentrated under reduced

pressure. The residue was precipitated in acetonitrile then precipitated in DCM/diethyl ether to give 3.30 g (81%) of **16** as a yellow powder.

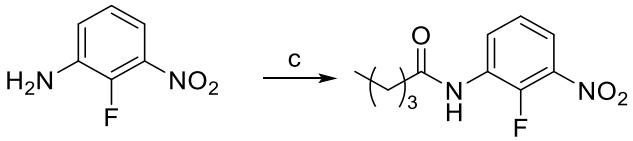
^{1}H NMR (500 MHz, DMSO-d_6) δ 10.20 (s, 1H), 8.33 (s, 1H), 8.14 (d, J = 9.1 Hz, 1H), 7.93 (dt, J = 7.6, 0.9 Hz, 2H), 7.85 (s, 1H), 7.79 (dd, J = 7.5, 1.1 Hz, 3H), 7.65 (t, J = 6.1 Hz, 1H), 7.44 (tt, J = 7.5, 0.8 Hz, 2H), 7.37 (td, J = 7.5, 1.2 Hz, 2H), 4.65 (d, J = 6.1 Hz, 2H), 4.60 (d, J = 6.5 Hz, 2H), 4.36 (t, J = 6.5 Hz, 1H), 1.42 (s, 9H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 166.9, 156.3, 153.9, 149.1, 148.2, 147.4, 144.2, 141.3, 141.1, 128.2, 127.6, 125.6, 124.6, 123.2, 121.9, 120.7, 116.9, 116.0, 78.8, 66.3, 47.1, 40.9, 28.7.

HRMS: calcd. for $C_{31}H_{30}N_3O_6$ $[M+H]^+$ 540.2129; found 540.2138.

2-fluoro-3-nitroaniline

2-fluoro-3-nitroaniline (CAS: 21397-11-5) was synthesized as previously described[5] (3 steps, only two are depicted in the scheme S1).



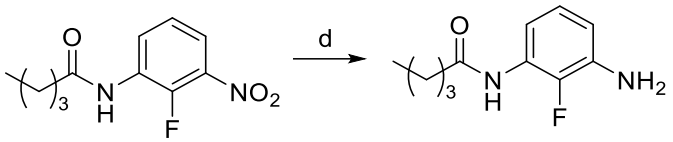
17

Compound 17: Dry DIEA (12.5 mL, 90.0 mmol, freshly distilled over CaH_2) was added to a suspension of 2-fluoro-3-nitroaniline (12.77 g, 81.8 mmol) in pentanoic anhydride (17.8 mL, 90.0 mmol) under inert atmosphere of N_2. The mixture was stirred at 40 ˚C for 19 h, after which complete conversion of the starting material was observed by TLC. The mixture was diluted with DCM, washed twice with 5% aqueous citric acid and several times with saturated aqueous $NaHCO_3$, dried over $MgSO_4$ and concentrated under reduced pressure, to obtain 19.66 g (quantitative) of compound **17** as a pale white solid.

^{1}H NMR: (300 MHz, $CDCl_3$) δ 8.68 (ddd, J = 8.4, 6.6, 1.7 Hz, 1H), 7.73 (ddd, J = 8.6, 7.1, 1.7 Hz, 1H), 7.51 (s, 1H), 7.26 (td, J = 8.3, 1.8 Hz, 1H), 2.46 (t, J = 8.5, 2H), 1.73 (ddt, J = 8.5, 7.3, 6.4 Hz, 2H), 1.50 – 1.35 (m, 2H), 0.96 (t, J = 7.3 Hz, 3H).

^{13}C NMR: (75 MHz, $CDCl_3$) δ 171.9, 145.6 (d, $J_{C\text{-}F}$ = 260.3 Hz), 137.4, 128.9 (d, $J_{C\text{-}F}$ = 9.3 Hz), 126.9 (d, $J_{C\text{-}F}$ = 2.3 Hz), 124.5 (d, $J_{C\text{-}F}$ = 5.2 Hz), 119.8 (d, $J_{C\text{-}F}$ = 3.1 Hz), 37.6, 27.5, 22.4, 13.9.

HRMS: calcd. for $C_{11}H_{14}FN_2O_3$ $[M+H]^+$ 241.0983; found 241.0985.



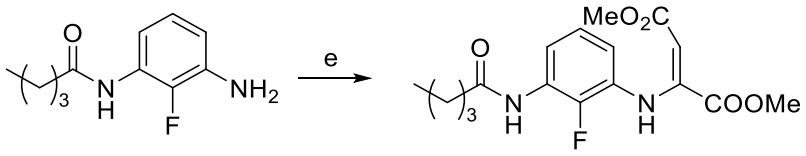
17 **18**

Compound 18: Compound **17** (10.92 g, 45.5 mmol) dissolved in 100 mL of degassed AcOEt was placed into a two-necked 500 mL round bottom flask equipped with a teflon-coated magnetic stirring bar. One of the necks was closed with a septum, N_2 was flushed and 250 mg of Pd/C (10%) catalyst was added to the solution. The reaction flask was placed under a H_2 atmosphere by alternately purging the flask under vacuum, shaking it and filling the flask with H_2 (3 times). The reaction mixture was vigorously stirred under H_2 atmosphere (balloon pressure) at room temperature for 24 h, after which complete conversion of the starting material was observed by TLC. The solution was filtered through Celite® and solvent was removed under reduced pressure, to give 9.29 g (97%) of compound **18** as a pale white solid.

^{1}H NMR: (300 MHz, $CDCl_3$) δ 7.65 (t, *J* = 7.4 Hz, 1H), 7.30 (s, 1H), 6.88 (td, *J* = 8.2, 1.7 Hz, 1H), 6.51 (td, *J* = 8.4, 1.6 Hz, 1H), 3.71 (s, 2H), 2.38 (t, *J* = 7.6 Hz, 2H), 1.80 – 1.62 (m, 2H), 1.40 (dq, *J* = 14.6, 7.3 Hz, 2H), 0.94 (t, *J* = 7.3 Hz, 3H).

^{13}C NMR: (75 MHz, $CDCl_3$) δ 171.5, 142.1 (d, $J_{C\text{-}F}$ = 234.9 Hz), 134.2 (d, $J_{C\text{-}F}$ = 11.4 Hz), 126.7 (d, $J_{C\text{-}F}$ = 8.1 Hz), 124.4 (d, $J_{C\text{-}F}$ = 4.4 Hz), 112.2, 111.5, 37.7, 27.7, 22.5, 13.9.

HRMS: calcd. for $C_{11}H_{16}FN_2O$ $[M+H]^+$ 211.1241; found 211.1243.

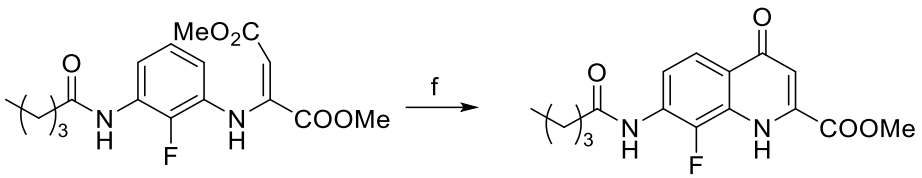


18 **19**

Compound **19**: Compound **18** (9.03 g, 42.9 mmol) was dissolved in 130 mL of degassed MeOH under inert atmosphere of N_2. The solution was cooled down to 0 ˚C in an ice-bath and dimethyl acetylenedicarboxylate (5.8 mL, 47.2 mmol) was added dropwise under stirring. The mixture was stirred at 0 ˚C for 30 min and then at room temperature for 22 h, during which time the product precipitated. Volatiles were partially evaporated and the resulting concentrated

suspension was cooled down to 5 ˚C to force the precipitation. The product was filtered off and washed with cold MeOH, to give 10.65 g (70%) of compound **19** as a white solid.

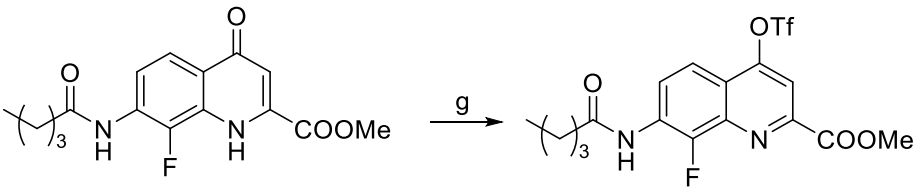
^{1}H NMR: (300 MHz, $CDCl_3$) δ 9.53 (s, 1H), 8.06 (t, *J* = 7.7 Hz, 1H), 7.34 (s, 1H), 7.00 (td, *J* = 8.2, 1.7 Hz, 1H), 6.59 (td, *J* = 8.0, 1.5 Hz, 1H), 5.54 (s, 1H), 3.75 (s, 3H), 3.72 (s, 3H), 2.40 (t, *J* = 7.6 Hz, 2H), 1.79 – 1.62 (m, 2H), 1.48 – 1.32 (m, 2H), 0.94 (t, *J* = 7.3 Hz, 3H).
^{13}C NMR: (75 MHz, $CDCl_3$) δ 171.6, 169.9, 164.2, 147.1, 144.9 (d, $J_{C\text{-}F}$ = 241.6 Hz), 128.3 (d, $J_{C\text{-}F}$ = 10.0 Hz), 127.2 (d, $J_{C\text{-}F}$ = 8.9 Hz), 124.2 (d, $J_{C\text{-}F}$ = 4.7 Hz), 117.1 (d, $J_{C\text{-}F}$ = 19.9 Hz), 95.3, 53.0, 51.5, 37.7, 27.6, 22.4, 13.9.
HRMS: calcd. for $C_{17}H_{21}FN_2O_5Na$ $[M+Na]^+$ 375.1326; found 375.1389.



19 **20**

Compound 20: Diphenyl ether (415 mL) was placed in a 1 L round bottom flask and heated to its boiling point at 260 ˚C with a heating mantle and without any reflux condenser or stirring bar. Then compound **19** (16.60 g, 47.1 mmol) was added to the boiling solvent by means of a glass funnel. The addition was done stepwise during 10 min to have control over the MeOH bubbling produced each time a portion of the starting material is added. The mixture was kept boiling for 20 min leaving the reaction flask open to allow MeOH vapors out. Then the heating mantle was removed and the reaction mixture was left to cool down to approximately 40 ˚C, observing the partial precipitation of the product. The cooled suspension was transferred to a bigger flask and cyclohexane (700 mL) was added to force the precipitation. The product was filtered off, washed thoroughly with more cyclohexane and dried under reduced pressure, obtaining 13.33 g of compound **20** (88%) as a light-gray solid.

^{1}H NMR: ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.94 (s, 1H), 8.35 (dd, *J* = 9.1, 7.1 Hz, 1H), 8.07 (dd, *J* = 9.1, 1.7 Hz, 1H), 7.52 (s, 1H), 7.26 (s, 1H), 6.92 (d, *J* = 1.8 Hz, 1H), 4.05 (s, 3H), 2.48 (t, *J* = 7.6 Hz, 2H), 1.84 – 1.67 (m, 2H), 1.52 – 1.34 (m, 2H), 0.97 (t, *J* = 7.3 Hz, 3H).
^{13}C NMR: (126 MHz, $CDCl_3$) δ 178.3, 171.7, 163.3, 140.5 (d, $J_{C\text{-}F}$ = 242.8 Hz), 136.1, 129.5 (d, $J_{C\text{-}F}$ = 6.7 Hz), 128.9 (d, $J_{C\text{-}F}$ = 11.0 Hz), 123.0, 122.3 (d, $J_{C\text{-}F}$ = 4.2 Hz), 117.3, 112.5, 54.1, 37.7, 27.5, 22.5, 13.9.
HRMS: calcd. for $C_{16}H_{18}FN_2O_4$ $[M+H]^+$ 321.1245; found 321.1339.



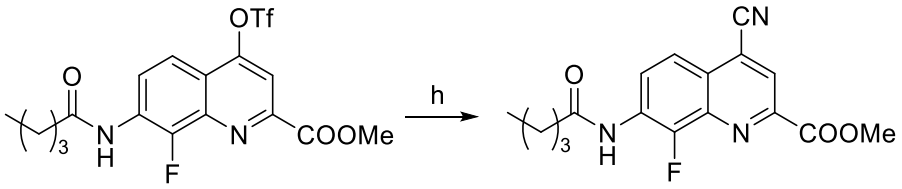
20 **21**

Compound 21: Dry pyridine (11.0 mL, 136 mmol) was added to a suspension of compound **20** (8.69 g, 27.1 mmol) in dry DCM (100 mL) under inert atmosphere of N_2. The slurry was cooled to –10 ˚C in an ice/NaCl-batch, and trifluoromethanesulfonic anhydride (Tf_2O, 6.8 mL, 40.4 mmol) was added dropwise under stirring. The resulting mixture was stirred at –10 ˚C for 2 h, after which complete conversion of the starting material was observed by TLC. Then the cold reaction mixture was poured into saturated aqueous NH_4Cl. The product was extracted with DCM, washed with 0.1 M aqueous HCl and once with saturated aqueous NH_4Cl, dried over $MgSO_4$ and concentrated under reduced pressure, obtaining 11.95 g (97%) of compound **21** as a white-gray solid.

^{1}H NMR: ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.99 (dd, J = 9.4, 6.8 Hz, 1H), 8.16 (s, 1H),7.90 (dd, J = 9.4, 1.8 Hz, 1H), 7.71 (d, J = 3.6 Hz, 1H), 4.10 (s, 3H), 2.54 (t, J = 7.6 Hz, 2H), 1.88 – 1.69 (m, 2H), 1.55 – 1.33 (m, 2H), 0.99 (t, J = 7.3 Hz, 3H).

^{13}C NMR: (75 MHz, $CDCl_3$) δ 172.1, 164.3, 153.4 (d, $J_{C\text{-}F}$ = 4.5 Hz), 149.8, 145.9 (d, $J_{C\text{-}F}$ = 254.8 Hz), 139.6 (d, $J_{C\text{-}F}$ = 11.4 Hz), 128.7 (d, $J_{C\text{-}F}$ = 8.1 Hz), 124.9, 120.9, 119.2, 116.6 (d, $J_{C\text{-}F}$ = 5.4 Hz), 111.8, 53.9, 37.8, 27.5, 22.5, 13.9.

HRMS: calcd. for $C_{17}H_{16}F_4N_2O_6SNa$ $[M+Na]^+$ 475.0557; found 475.0617.

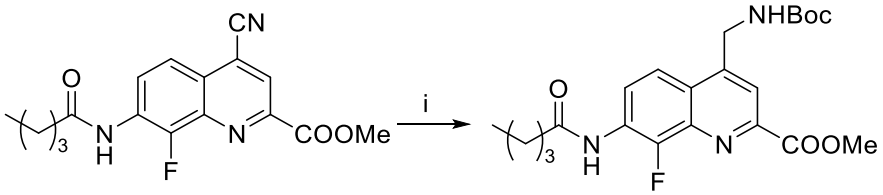


21 **22**

Compound 22: Compound **21** (4.00 g, 8.84 mmol) and KCN (3.46 g, 53.1 mmol) were suspended in 40 mL of dry THF under inert atmosphere of N_2. The slurry was sonicated until the suspended particles of solid became very small. Then the solvent was evaporated leaving a slightly wet solid residue in the reaction flask. Dry toluene (150 mL) and $Pd(PPh_3)_4$ (817 mg, 0.707 mmol) were added under inert atmosphere of N_2. The resulting suspension was stirred at

100 ˚C for 7 h, after which complete conversion of the starting material was observed by NMR. Volatiles were evaporated under reduced pressure and the solid residue was dissolved in AcOEt, washed with saturated aqueous $NaHCO_3$ and once with saturated aqueous NaCl, dried over $MaSO_4$ and concentrated under reduced pressure. The crude residue was suspended in MeOH, filtered off, washed with more MeOH and dried under reduced pressure, obtaining 2.31 g (79%) of compound **22** as a pale white solid.

^{1}H NMR: (300 MHz, $CDCl_3$) δ 9.05 (dd, *J* = 9.3, 6.8 Hz, 1H), 8.45 (s, 1H), 8.03 (dd, *J* = 9.3, 1.7 Hz, 1H), 7.77 (s, 1H), 4.11 (s, 3H), 2.55 (dd, *J* = 8.2, 6.9 Hz, 2H), 1.78 (ddt, *J* = 8.5, 7.4, 6.4 Hz, 2H), 1.53 – 1.38 (m, 2H), 0.98 (t, *J* = 7.3 Hz, 3H).
^{13}C NMR: (75 MHz, $CDCl_3$) δ 172.1, 164.2, 148.2, 146.0 (d, $J_{C\text{-}F}$ = 255.8 Hz), 137.6 (d, $J_{C\text{-}F}$ = 11.2 Hz), 128.7 (d, $J_{C\text{-}F}$ = 8.1 Hz), 125.7, 124.5, 123.7, 120.6 (d, $J_{C\text{-}F}$ = 5.5 Hz), 120.1 (d, $J_{C\text{-}F}$ = 3.9 Hz), 114.8, 53.9, 37.8, 27.4, 22.5, 13.9.
HRMS: calcd. for $C_{17}H_{16}FN_3O_3Na$ $[M+Na]^+$ 352.1068; found 352.1135.

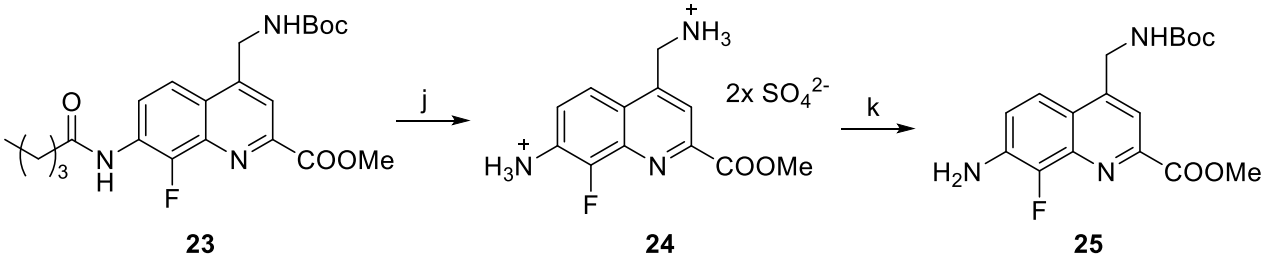


22 **23**

Compound 23: Pd/C (10%) (2.27 g) was added to a solution of compound **22** (3.25 g, 9.86 mmol), di-*tert*-butyl dicarbonate (8.61 g, 39.4 mmol) and ammonium formate (3.73 g, 59.1 mmol) in dry THF (190 mL), under inert atmosphere of N_2. After 4 h stirring at 40 ˚C, the complete conversion of the starting material was observed by TLC. The reaction mixture was filtered through Celite® using DCM to completely wash out the product from the filter. The crude product was concentrated under reduced pressure and purified by flash chromatography (SiO_2) using acetone/DCM as eluent (from 5% to 10% acetone) to give 3.32 g (78%) of compound **23** as a white solid.

^{1}H NMR: (300 MHz, $CDCl_3$) δ 8.80 (dd, *J* = 9.4, 7.1 Hz, 1H), 8.12 (s, 1H), 7.81 (dd, *J* = 9.4, 1.7 Hz, 1H), 7.68 (d, *J* = 3.4 Hz, 1H), 5.07 (s, 1H), 4.85 (d, *J* = 6.2 Hz, 2H), 4.07 (s, 3H), 2.52 (t, *J* = 7.6 Hz, 2H), 1.85 – 1.69 (m, 2H), 1.48 (s, 9H), 1.43 (m, 2H), 0.98 (t, *J* = 7.3 Hz, 3H).
^{13}C NMR: (75 MHz, $CDCl_3$) δ 172.1, 165.8, 155.9, 148.5, 146.6 (d, $J_{C\text{-}F}$ = 253.5 Hz), 146.3, 137.7 (d, $J_{C\text{-}F}$ = 10.1 Hz), 126.8 (d, $J_{C\text{-}F}$ = 8.2 Hz), 124.6, 123.3, 118.5 (d, $J_{C\text{-}F}$ = 4.9 Hz), 118.4, 80.6, 53.5, 41.6, 37.7, 28.5, 27.5, 22.5, 14.0.

HRMS: calcd. for $C_{22}H_{29}FN_3O_5$ $[M+H]^+$ 434.2086; found 434.2147.



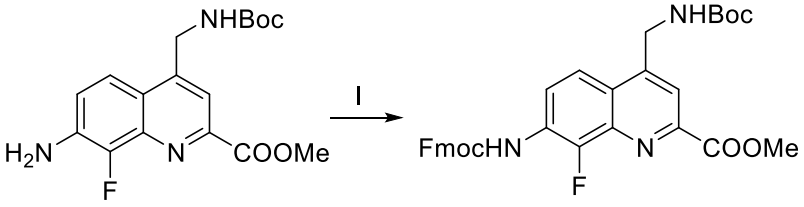
23 **24** **25**

Compound 25: Concentrated sulfuric acid (1.65 mL, 30.9 mmol) was added to a suspension of compound **23** (5.15 g, 11.9 mmol) in 110 mL of MeOH at room temperature under stirring. The mixture was refluxed for 16 h, during which time it turned from a white slurry to a red solution. The crude mixture was left to cool down to room temperature observing the formation of a red precipitate. The complete precipitation was forced by adding some Et_2O and the supernatant was separated by decantation. The precipitate was washed twice with more Et_2O, separating the supernatant by decantation, and dried under reduced pressure to obtain compound **24** as a red solid. Then di-*tert*-butyl decarbonate (2.65 g, 12.1 mmol), $NaHCO_3$ (2.00 g, 23.8 mmol) and 75 mL of dry DMF were added under inert atmosphere of N_2. The resulting solution was stirred at room temperature for 26 h, after which complete conversion of the starting material was observed by NMR. The mixture was diluted with AcOEt, washed several times with saturated aqueous $NaHCO_3$ to completely remove the DMF, dried over $MgSO_4$ and concentrated under reduced pressure, to obtain 4.11 g (99%) of compound **25** as a yellow solid.

^{1}H NMR: (300 MHz, DMSO-d_6) δ 7.78 (d, J = 9.1 Hz, 1H), 7.70 (s, 1H), 7.62 (t, J = 6.1 Hz, 1H), 7.32 (dd, J = 9.1, 8.1 Hz, 1H), 5.94 (s, 2H), 4.58 (d, J = 6.0 Hz, 2H), 3.93 (s, 3H), 1.42 (s, 9H).

^{13}C NMR: (126 MHz, DMSO-d_6) δ 165.7, 155.9, 147.2, 146.9, 142.2 (d, $J_{C\text{-}F}$ = 242.8 Hz), 138.3 (d, $J_{C\text{-}F}$ = 8.3 Hz), 136.5 (d, $J_{C\text{-}F}$ = 10.9 Hz), 121.7 (d, $J_{C\text{-}F}$ = 5.3 Hz), 119.2 (d, $J_{C\text{-}F}$ = 4.1 Hz), 119.1, 114.4, 78.3, 52.5, 40.5, 28.2.

HRMS: calcd. for $C_{17}H_{21}FN_3O_4$ $[M+H]^+$ 350.1511; found 350.1583.



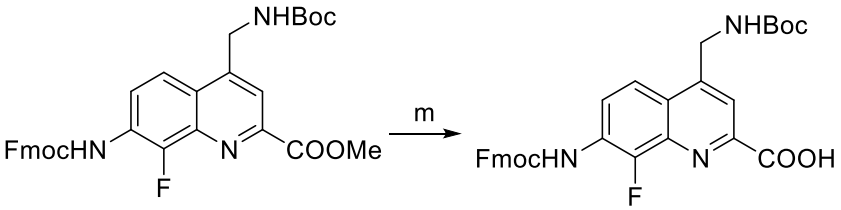
25 **26**

Compound 26: A suspension of compound **25** (4.11 g, 11.8 mmol) in 350 mL of water/1,4-dioxane (3:4, *v*/*v*) was cooled down to 0 ˚C in an ice-bath. Then a solution of 9-fluorenylmethoxycarbonyl chloride (Fmoc-Cl, 4.57 g, 17.6 mmol) in 50 mL of dioxane was added dropwise under stirring, during approximately 1 h. The reaction mixture was stirred at 0 ˚C for 2 h and then at room temperature for 5 h, after which some more Fmoc-Cl (1.52 g, 5.88 mmol) dissolved in 10 mL of dioxane was added dropwise under stirring. After 16 h stirring at room temperature, again some more Fmoc-Cl (1.52 g, 5.88 mmol) dissolved in 10 mL of dioxane was added dropwise under stirring. 1 h later the complete conversion of the starting material was finally observed by NMR. The reaction mixture was diluted with DCM, washed with water, dried over $MgSO_4$ and concentrated under reduced pressure. The crude product was purified by flash chromatography (SiO_2) using acetone/DCM as eluent (from 0% to 10% acetone) to give 5.74 g (85%) of compound **26** as a white solid.

^{1}H NMR: (300 MHz, $CDCl_3$) δ 8.53 (s, 1H), 8.12 (s, 1H), 7.84 (s, 1H), 7.80 (dd, *J* = 7.4, 1.0 Hz, 2H), 7.65 (dd, *J* = 7.3, 1.0 Hz, 2H), 7.44 (tdd, *J* = 7.5, 1.2, 0.6 Hz, 2H), 7.35 (td, *J* = 7.4, 1.3 Hz, 2H), 7.27 (s, 1H), 5.04 (s, 1H), 4.85 (d, *J* = 6.2 Hz, 2H), 4.63 (d, *J* = 6.6 Hz, 2H), 4.33 (t, *J* = 6.6 Hz, 1H), 4.08 (s, 3H), 1.50 (s, 9H).

^{13}C NMR: (75 MHz, $CDCl_3$) δ 165.8, 155.9, 153.1, 148.7, 146.4 (d, $J_{C\text{-}F}$ = 253.9 Hz), 146.3 (d, $J_{C\text{-}F}$ = 2.3 Hz), 143.6, 141.5, 137.8 (d, $J_{C\text{-}F}$ = 9.5 Hz), 128.1, 127.4, 126.6 (d, $J_{C\text{-}F}$ = 8.5 Hz), 125.1, 124.3, 122.0, 120.3, 118.7 (d, $J_{C\text{-}F}$ = 4.5 Hz), 118.3, 80.6, 67.8, 53.4, 47.1, 41.6, 28.5.

HRMS: calcd. for $C_{32}H_{31}FN_3O_6$ $[M+H]^+$ 572.2191; found 572.2229.



26 **27**

Compound 27: LiI beads (778 mg, 5.82 mmol) were added to a solution of compound **26** (1.11 g, 1.94 mmol) in degassed AcOEt (110 mL, degassed by freeze-thaw). The solution was protected from light and heated to reflux temperature under inert atmosphere of N_2. After stirring at reflux temperature for 19 h, the mixture was cooled down to room temperature, washed with 5% aqueous $Na_2S_2O_3$ and 2% aqueous citric acid, dried over $MgSO_4$ and concentrated under reduced pressure, washed twice with Et_2O and dried *in vacuo*, to obtain 891 mg (82%) of compound **27** as a white solid.

^{1}H NMR (500 MHz, DMSO-d_6) δ 13.55 (s, 1H), 10.04 (s, 1H), 8.03 (s, 1H), 7.99 (d, J = 7.4 Hz, 2H), 7.92 (d, J = 7.6 Hz, 2H), 7.80 (d, J = 7.5 Hz, 2H), 7.70 (t, J = 6.1 Hz, 1H), 7.44 (t, J = 7.4 Hz, 2H), 7.36 (t, J = 7.4 Hz, 2H), 4.68 (d, J = 6.0 Hz, 2H), 4.51 (d, J = 6.9 Hz, 2H), 4.35 (t, J = 6.9 Hz, 1H), 1.43 (s, 9H).
^{13}C NMR: (75 MHz, DMSO-d_6) δ 166.2, 155.9, 153.8, 148.9, 148.3 (d, $J_{C\text{-}F}$ = 256.5 Hz), 147.3, 143.7, 140.8, 137.5 (d, $J_{C\text{-}F}$ = 9.1 Hz), 127.7, 127.1, 125.8 (d, $J_{C\text{-}F}$ = 9.8 Hz), 125.3, 124.9, 124.5, 120.2, 118.8 (d, $J_{C\text{-}F}$ = 4.8 Hz), 117.9, 78.4, 66.4, 46.6, 40.6, 28.2.
HRMS: calcd. for $C_{31}H_{29}FN_3O_6$ $[M+H]^+$ 558.2035; found 558.2076.

3 NMR structure elucidation of 2 in water

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

2D NOESY was recorded with a phase-sensitive pulse sequence with water suppression employing an excitation sculpting element from the Bruker pulse program library (noesyesfpgpphrs). Data acquisition was performed with 1K (F2) x 512 (F1) data points and a mixing time of 0.5 s. The recycling delay was 1.5 s and 16 transients per increment were applied at a sweep width of 8 kHz in both dimensions resulting in an acquisition time of 0.1204 s. The special acquisition parameters regarding the water suppression element of the pulse sequence were adopted from the optimized parameter set of the respective one-dimensional experiment. A 90° shifted sinesquare multiplication, an exponential window of 1.0 Hz as well as a gaussian window of 1 Hz in both dimensions prior to FT was applied.

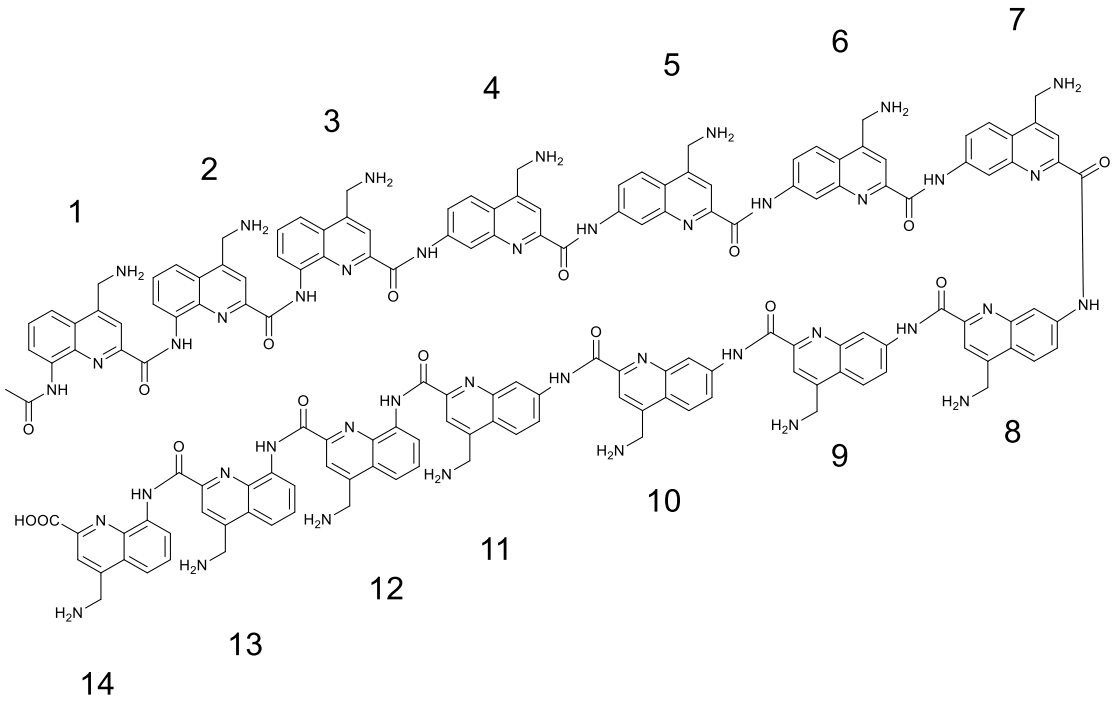
2D COSY was recorded using with water suppression using watergate W5 pulse sequence with gradients (cosygpphppw5) from the Bruker pulse program library. Data acquisition was performed with 2048 (F2) x 256 (F1) data points in States-TPPI mode. The recycling delay was 1.0 s and 128 transients per increment were applied at a sweep width of 8 kHz in both dimensions resulting in an acquisition time of 0.1204 s. The TOCSY mixing time was set to 80 ms. Special acquisition parameters regarding the water suppression element of the pulse sequence were adopted from the optimized parameter set of a respective one-dimensional experiment with a watergate W5 pulse sequence. A 90° shifted sinesquare multiplication was applied in both dimensions prior to FT.

2D TOCSY was recorded with a phase-sensitive pulse sequence using composite pulse scheme MLEV with water suppression employing an excitation sculpting element (mlevesgpph) from

the Bruker pulse program library. Data acquisition was performed with 2048 (F2) x 256 (F1) data points in States-TPPI mode. The recycling delay was 2.0 s and 8 transients per increment were applied at a sweep width of 8 kHz in both dimensions resulting in an acquisition time of 0.1283 s. The TOCSY mixing time was set to 80 ms. Special acquisition parameters regarding the water suppression element of the pulse sequence were adopted from the optimized parameter set of the respective one-dimensional experiment. A 90° shifted sinesquare multiplication was applied in both dimensions prior to FT.

^{1}H-^{13}C HSQC and ^{1}H-^{13}C HMBC spectra were recorded using standard pulse sequences from the Bruker pulse program library applying standard processing parameters.

a)



b)

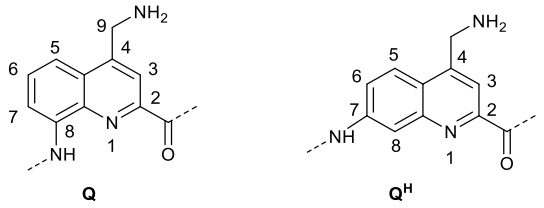


Figure S 10. a) Numbering of the units of compound **2** used in NMR assignment. b) Representative numbering of the carbon atoms in Q- and Q^H-units

Table S1. Assignment of the ^{1}H chemical shifts of compound **2** in H_2O/D_2O (9:1, *v*/*v*) at 25 °C (500 MHz).

Monomer	Atom	^{1}H (ppm)	Monomer	Atom	^{1}H (ppm)
Q1	H3	7.09	$Q^{H}8$	H3	7.50
	H5	7.38		H5	7.84
	H6	7.49		H6	8.21
	H7	7.21		H8	6.36
	NH	9.26		NH	9.00
	AcNH	1.25			
Q2	H3	8.05	$Q^{H}9$	H3	8.01
	H5	7.12		H5	8.24
	H6	6.95		H6	8.72
	H7	7.68		H8	7.92
	NH	11.94		NH	9.74
Q3	H3	6.82	$Q^{H}10$	H3	7.70
	H5	6.97		H5	8.21
	H6	6.59		H6	8.53
	H7	7.55		H8	7.63
	NH	11.90		NH	10.52
$Q^{H}4$	H3	7.90	$Q^{H}11$	H3	6.89
	H5	7.48		H5	7.81
	H6	7.21		H6	7.74
	H8	7.26		H8	6.47
	NH	10.09		NH	9.37
$Q^{H}5$	H3	8.03	Q12	H3	7.51
	H5	8.04		H5	6.69
	H6	8.38		H6	6.69
	H8	8.09		H7	6.92
	NH	9.71		NH	11.48
$Q^{H}6$	H3	8.21	Q13	H3	7.05
	H5	8.04		H5	6.39
	H6	8.38		H6	6.53
	H8	8.09		H7	8.73
	NH	10.58		NH	12.08
$Q^{H}7$	H3	7.95	Q14	H3	6.83
	H5	7.82		H5	7.48
	H6	8.58		H6	7.37
	H8	6.73		H7	7.57
	NH	9.99		NH	11.36

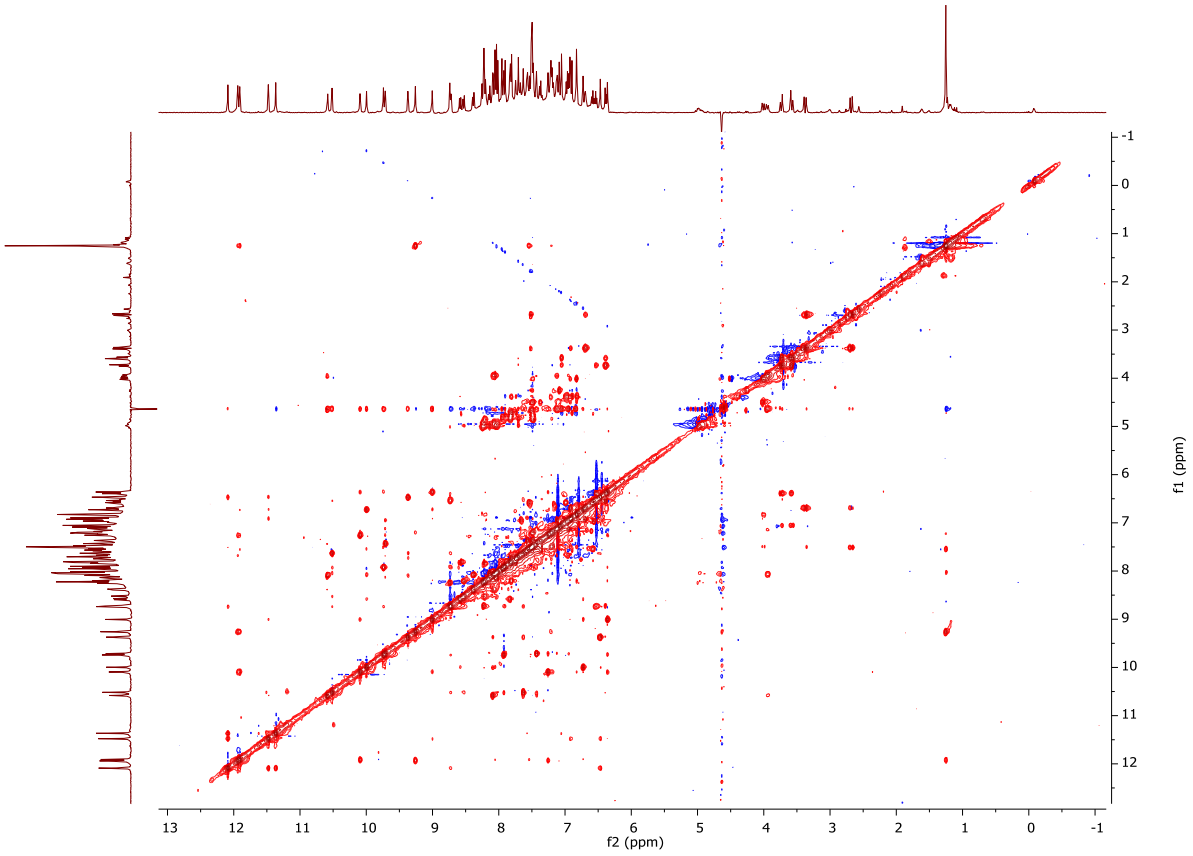


Figure S11. ^{1}H-^{1}H NOESY spectrum of **2** (0.5 mM, 500 MHz, H_2O/D_2O (9:1, *v*/*v*), 25 °C, 0.5 s mixing time).

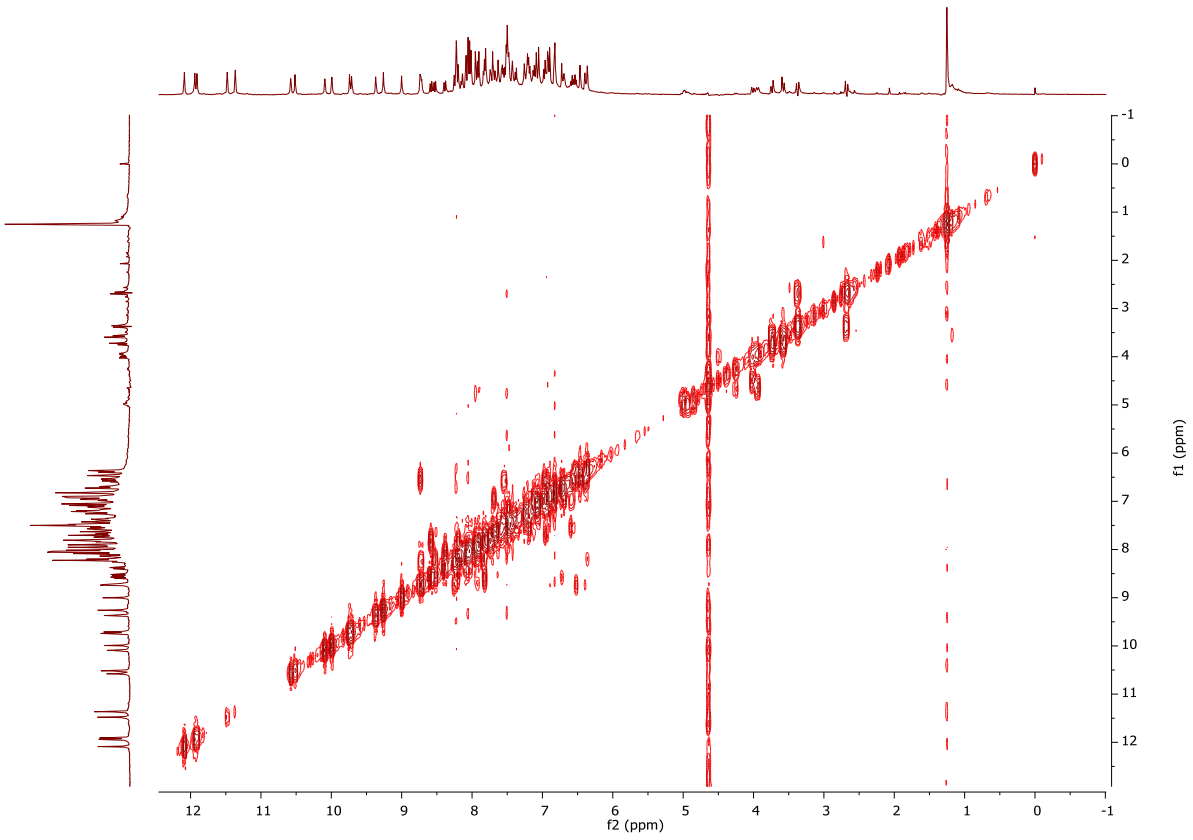


Figure S12. ^{1}H-^{1}H COSY spectrum of **2** (0.5 mM, 500 MHz, H_2O/D_2O (9:1, *v*/*v*), 25 °C).

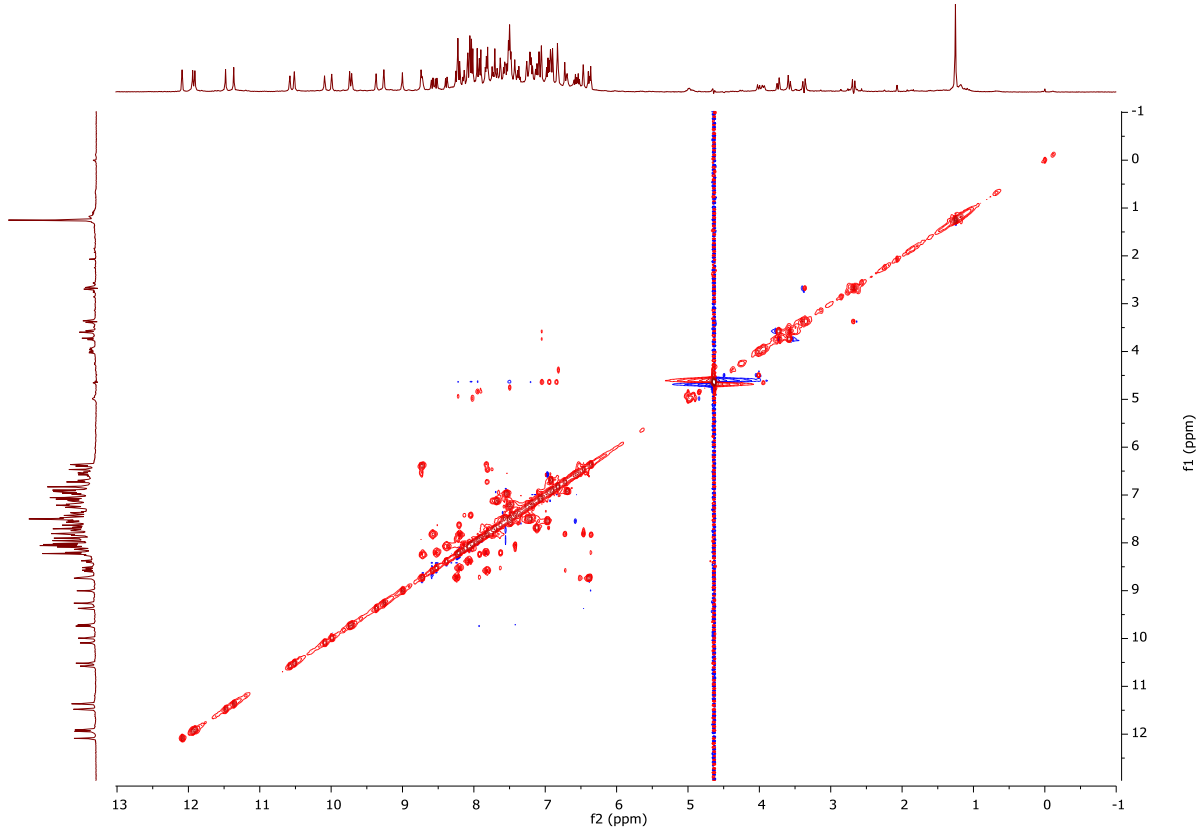


Figure S13. ^{1}H-^{1}H TOSY spectrum of **2** (0.5 mM, 500 MHz, H_2O/D_2O (9:1, *v*/*v*), 25 °C).

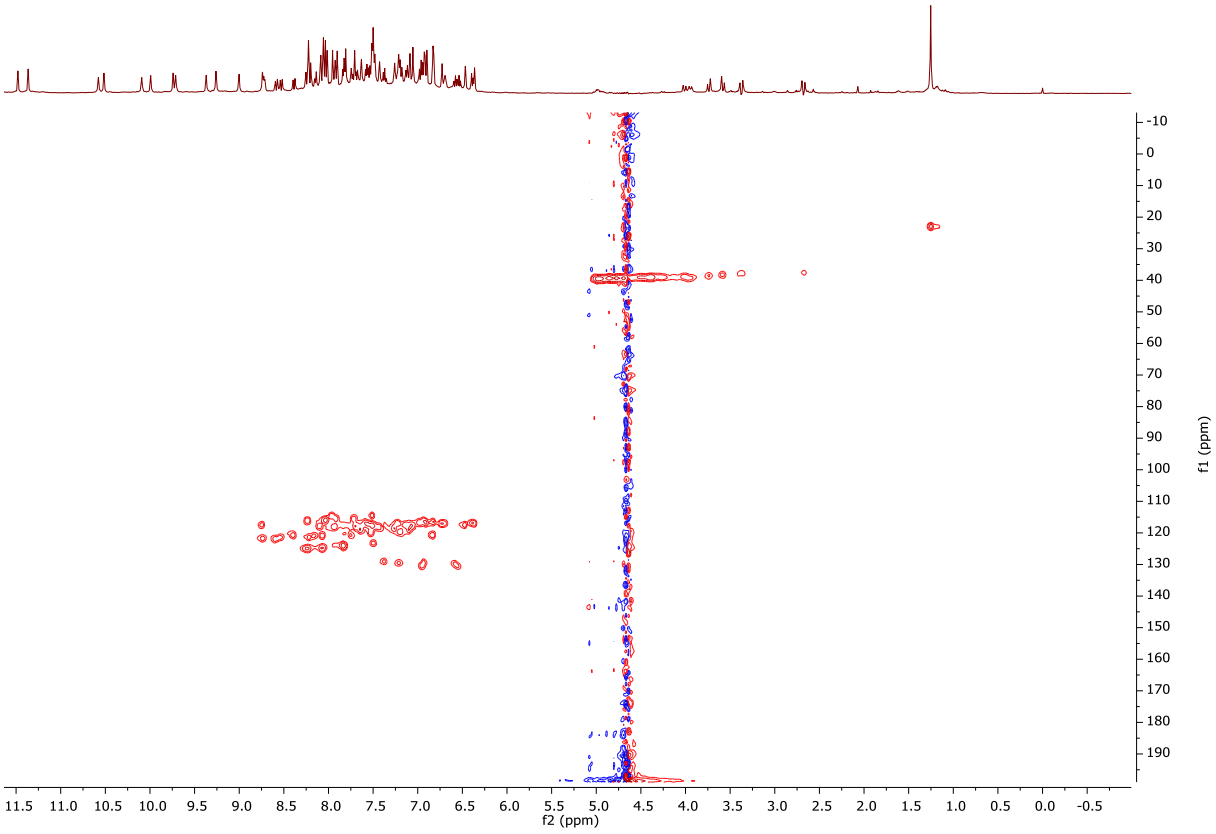


Figure S14. ^{1}H-^{13}C HSQC spectrum of **2** (0.5 mM, 500 MHz, H_2O/D_2O (9:1, *v*/*v*), 25 °C).

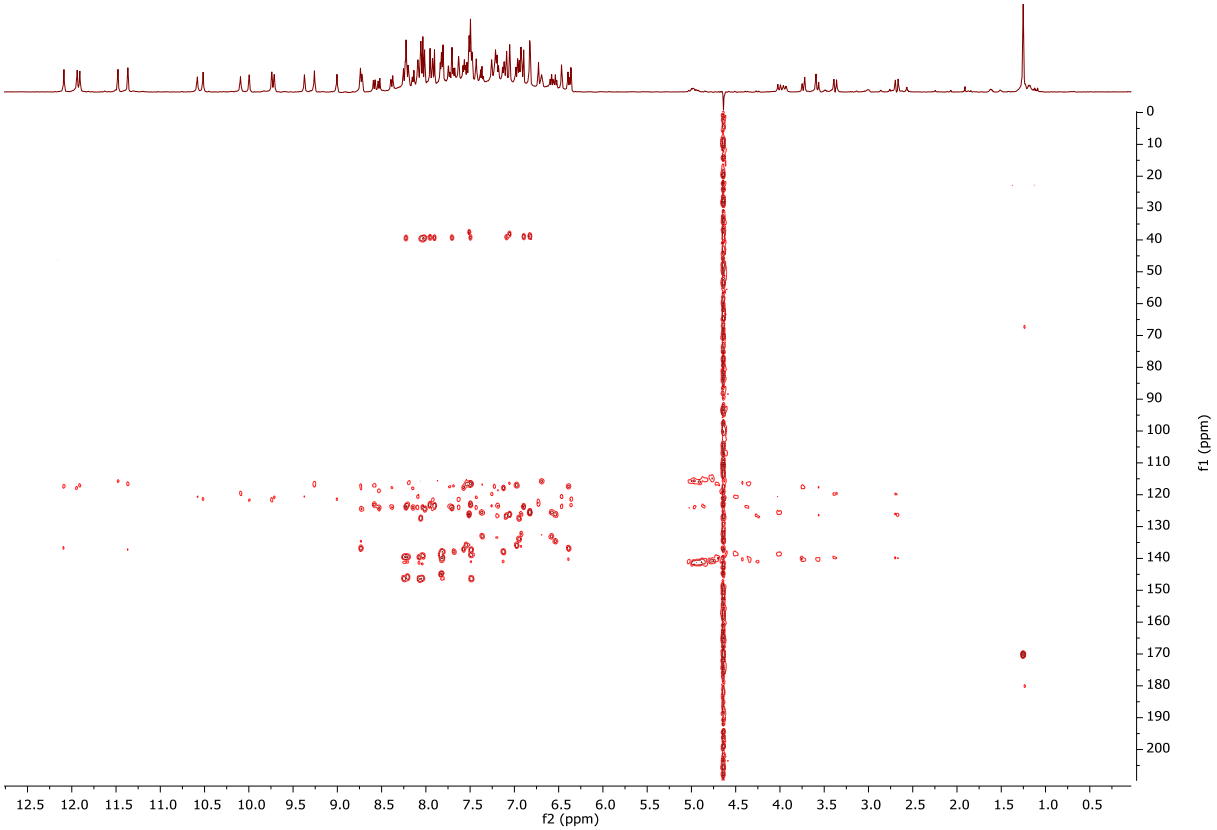


Figure S15. ^{1}H-^{13}C HMBC spectrum of **2** (0.5 mM, 500 MHz, H_2O/D_2O (9:1, *v*/*v*), 25 °C).

4 Spectra and chromatograms of new compounds

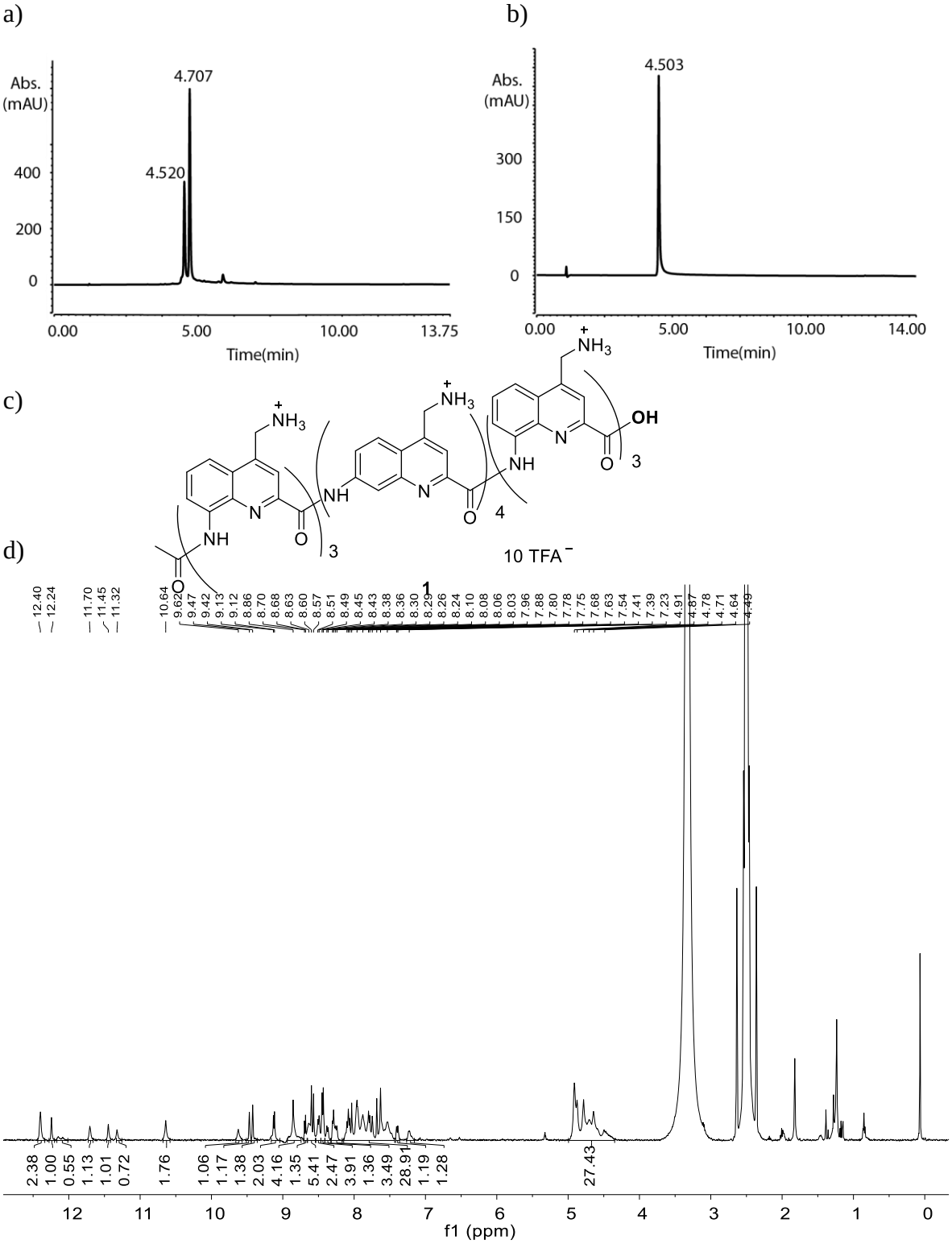


Figure S16. Analytical data of compound **1**. HPLC chromatograms (a) after cleavage from the resin (C18, 5 to 60 B% over 10 min, λ = 300 nm) and (b) after purification (C18, 5 to 60 B% over 10 min, λ = 300 nm); A: 0.1% TFA water, B: 0.1% TFA acetonitrile. (c) Chemical structure of compound **1**. (d) ^{1}H NMR spectrum (500 MHz, DMSO-d_6, 25 °C).

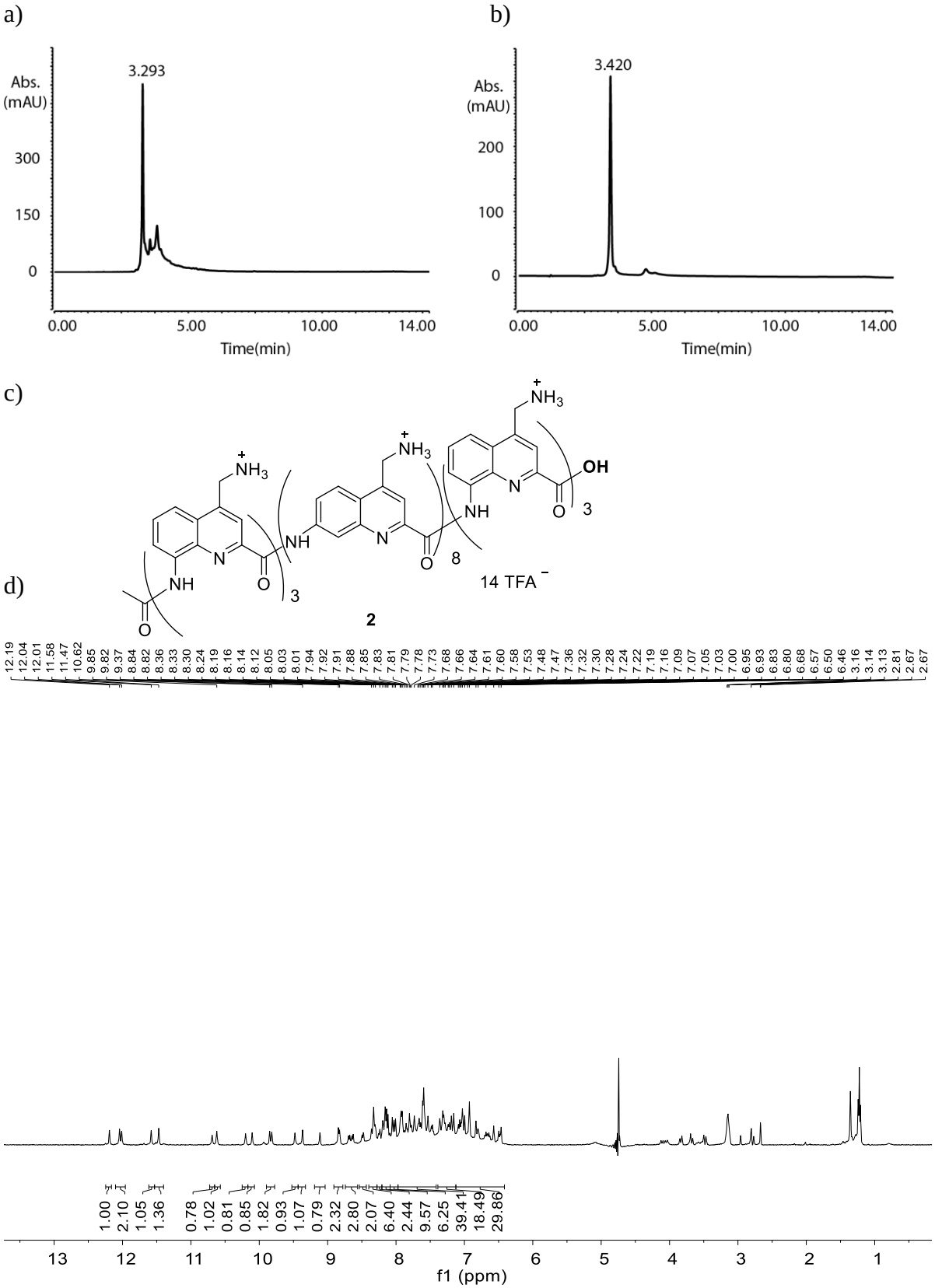


Figure S17. Analytical data of compound **2**. HPLC chromatograms (a) after cleavage from the resin (C18, 5 to 70 B% over 10 min, λ = 300 nm) and (b) after purification (C18, 10 to 40 B% over 10 min, λ = 300 nm); A: 0.1% TFA water, B: 0.1% TFA acetonitrile. (c) Chemical structure of compound **2**. (d) ^{1}H NMR spectrum with water suppression (500 MHz, H_2O/D_2O = (9:1, *v*/*v*), 25 °C).

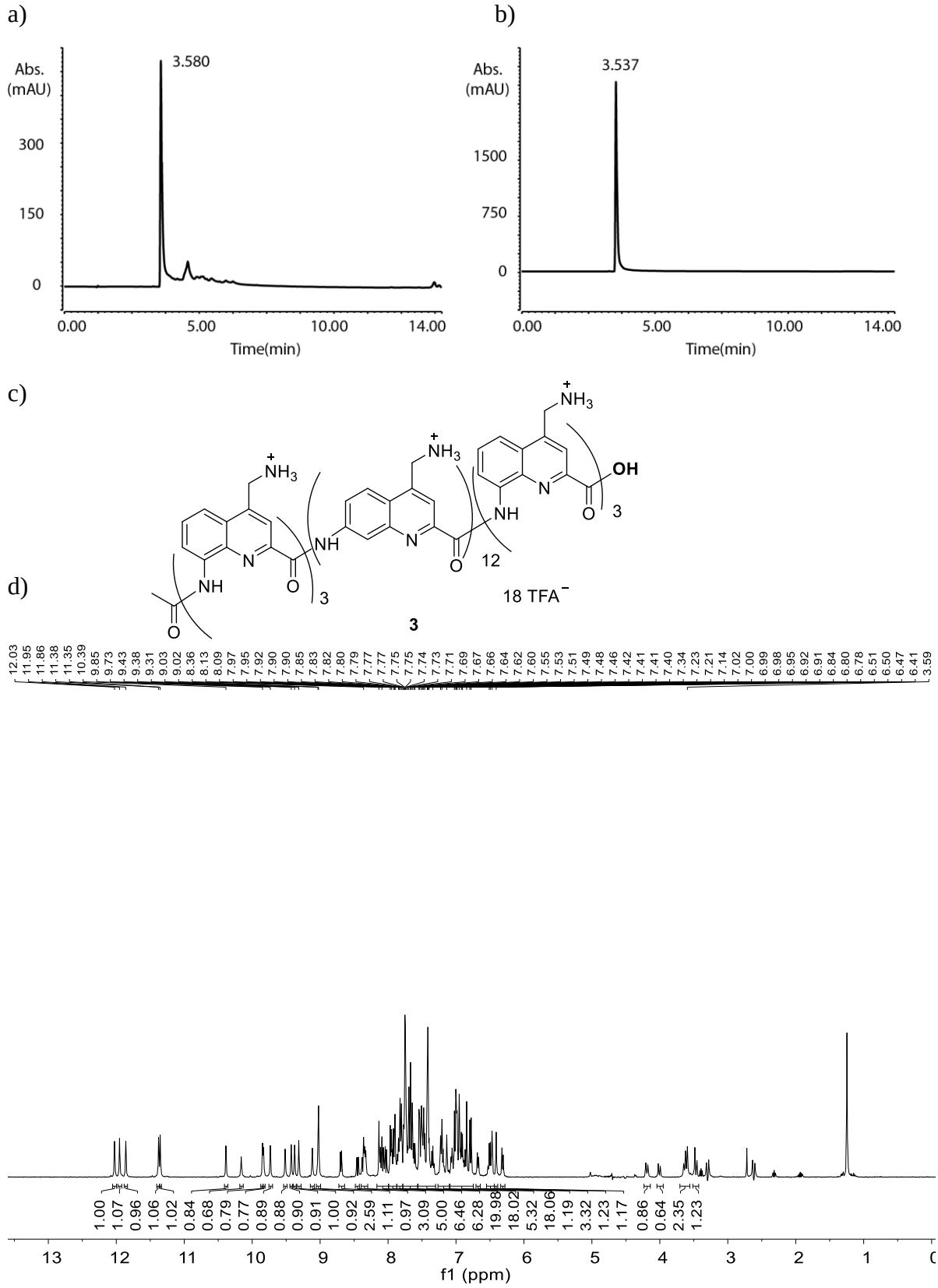


Figure S18. Analytical data of compound **3**. HPLC chromatograms (a) after cleavage from the resin (C18, 5 to 60 B% over 10 min, λ = 300 nm) and (b) after purification (C18, 5 to 60 B% over 10 min, λ = 300 nm); A: 0.1% TFA water, B: 0.1% TFA acetonitrile. (c) Chemical structure of compound **3**. (d) ^{1}H NMR spectrum with water suppression (500 MHz, H_2O/D_2O (9:1, *v*/*v*), 25 °C).

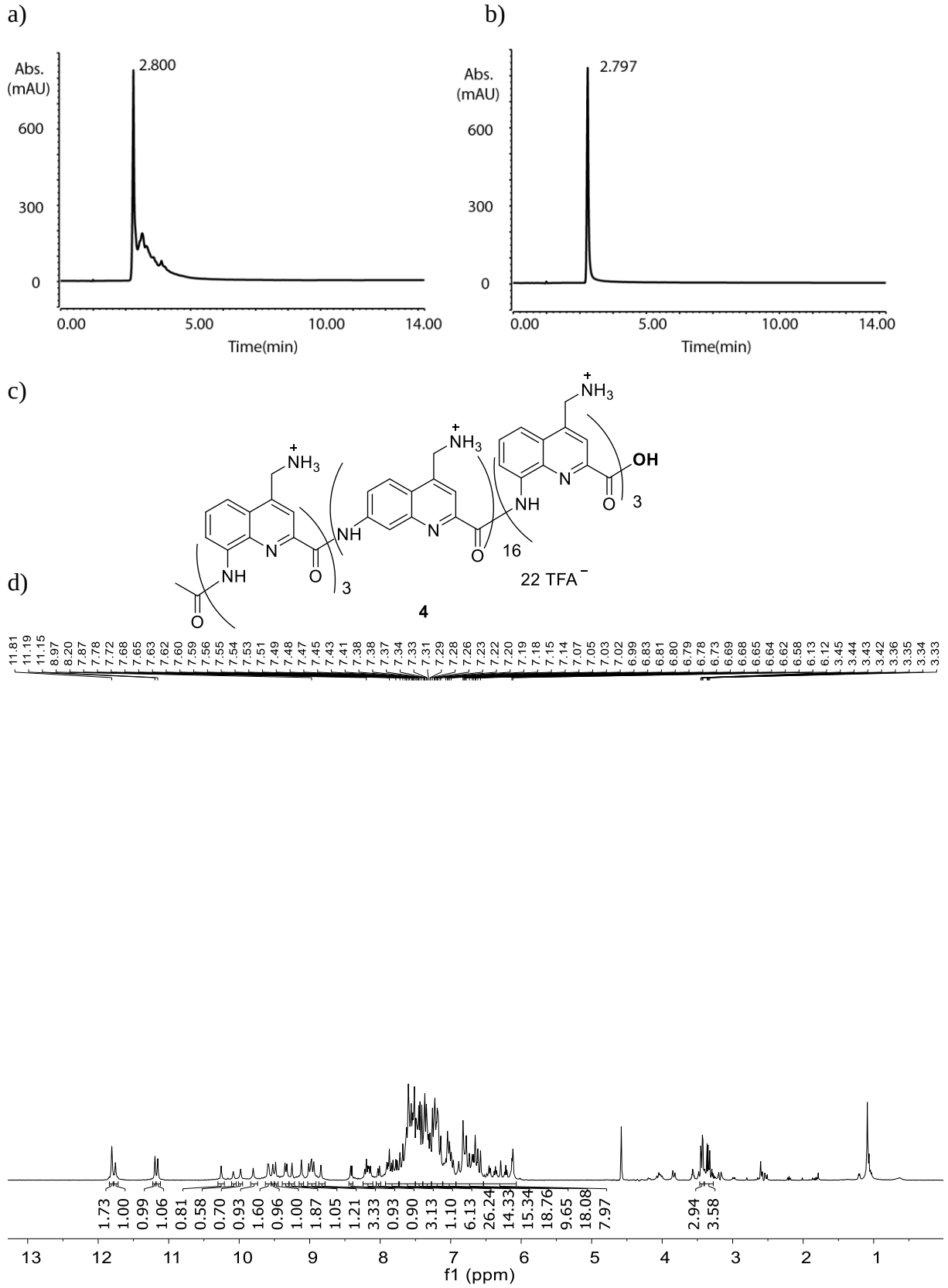


Figure S19. Analytical data of compound **4**. HPLC chromatograms (a) after cleavage from the resin (C18, 5 to 100 B% over 10 min, λ = 300 nm) and (b) after purification (C18, 5 to 100 B% over 10 min, λ = 300 nm); A: 0.1% TFA water, B: 0.1% TFA acetonitrile. (c) Chemical structure of compound **4**. (d) ^{1}H NMR spectrum with water suppression (500 MHz, H_2O/D_2O (9:1, *v*/*v*), 25 °C).

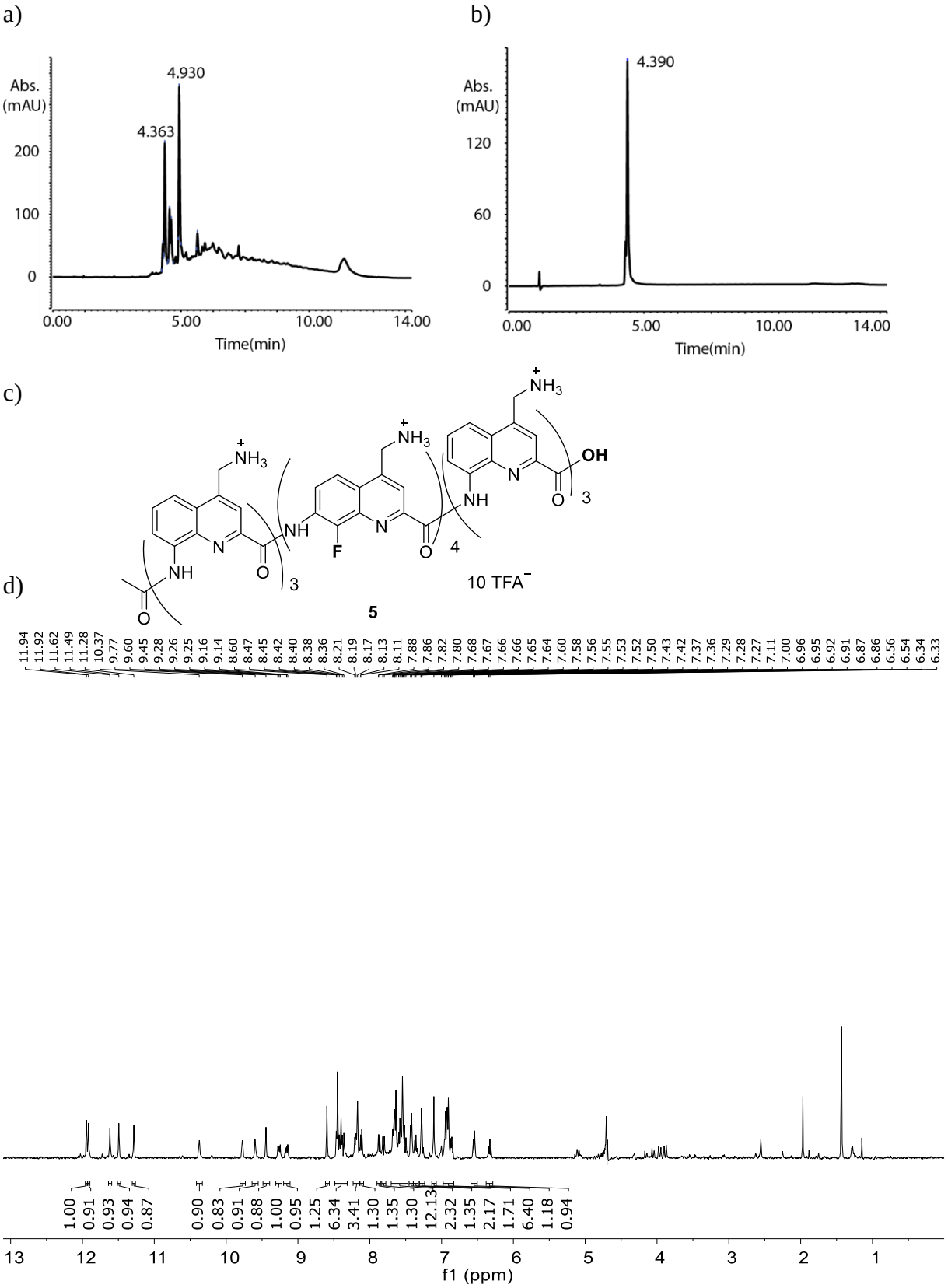


Figure S20. Analytical data of compound **5**. HPLC chromatograms (a) after cleavage from the resin (C18, 5 to 40 B% over 10 min, λ = 300 nm) and (b) after purification (C18, 5 to 40B% over 10 min, λ = 300 nm; A: 0.1% TFA water, B: 0.1% TFA acetonitrile). (c) Chemical structure of compound **5**. (d) ^{1}H NMR spectrum with water suppression (500 MHz, H_2O/D_2O (9:1, *v*/*v*), 25 °C).

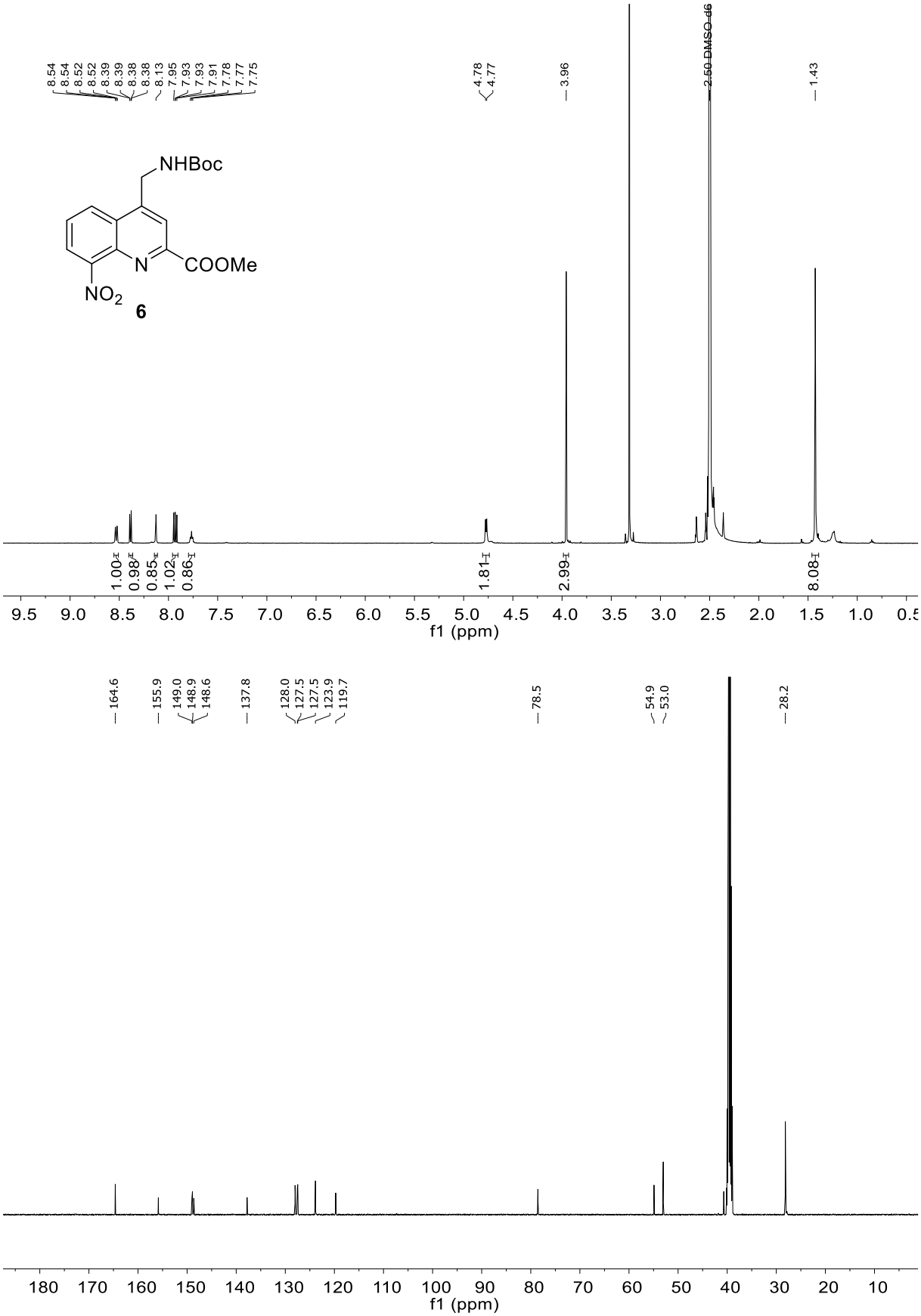


Figure S21. NMR spectra of compound **6**: ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

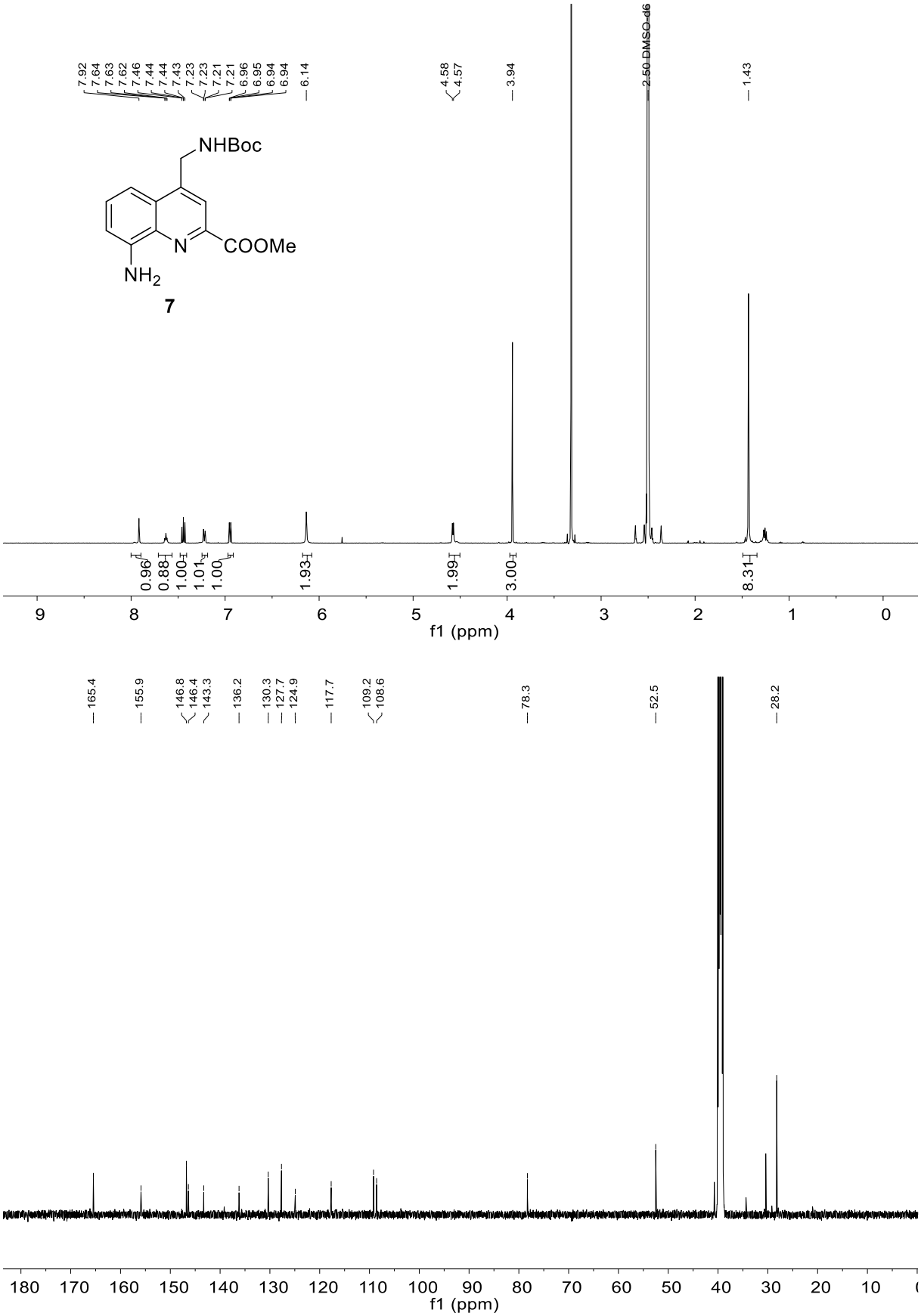


Figure S22. NMR spectra of compound **7**: ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

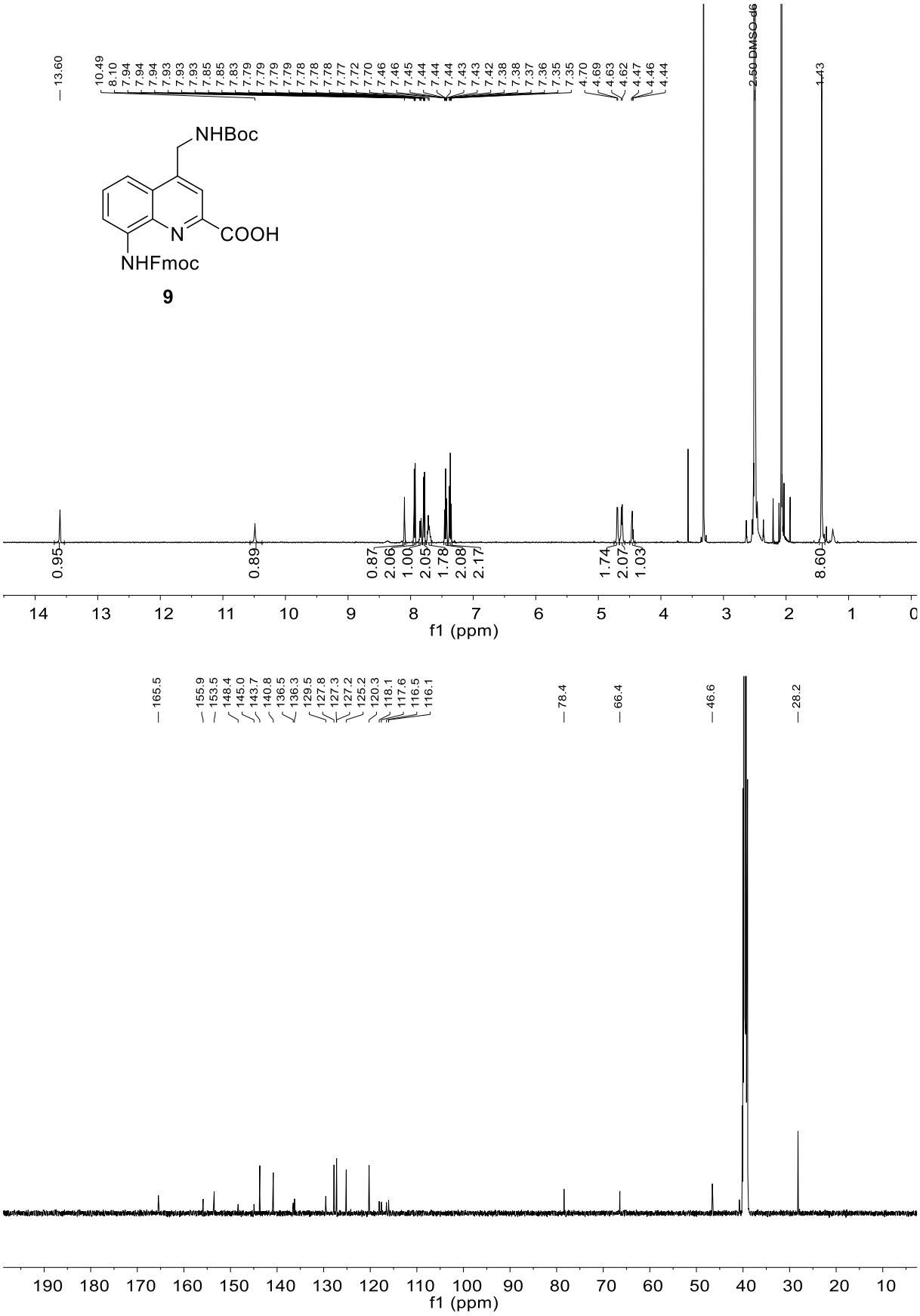


Figure S23. NMR spectra of compound **9**: ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

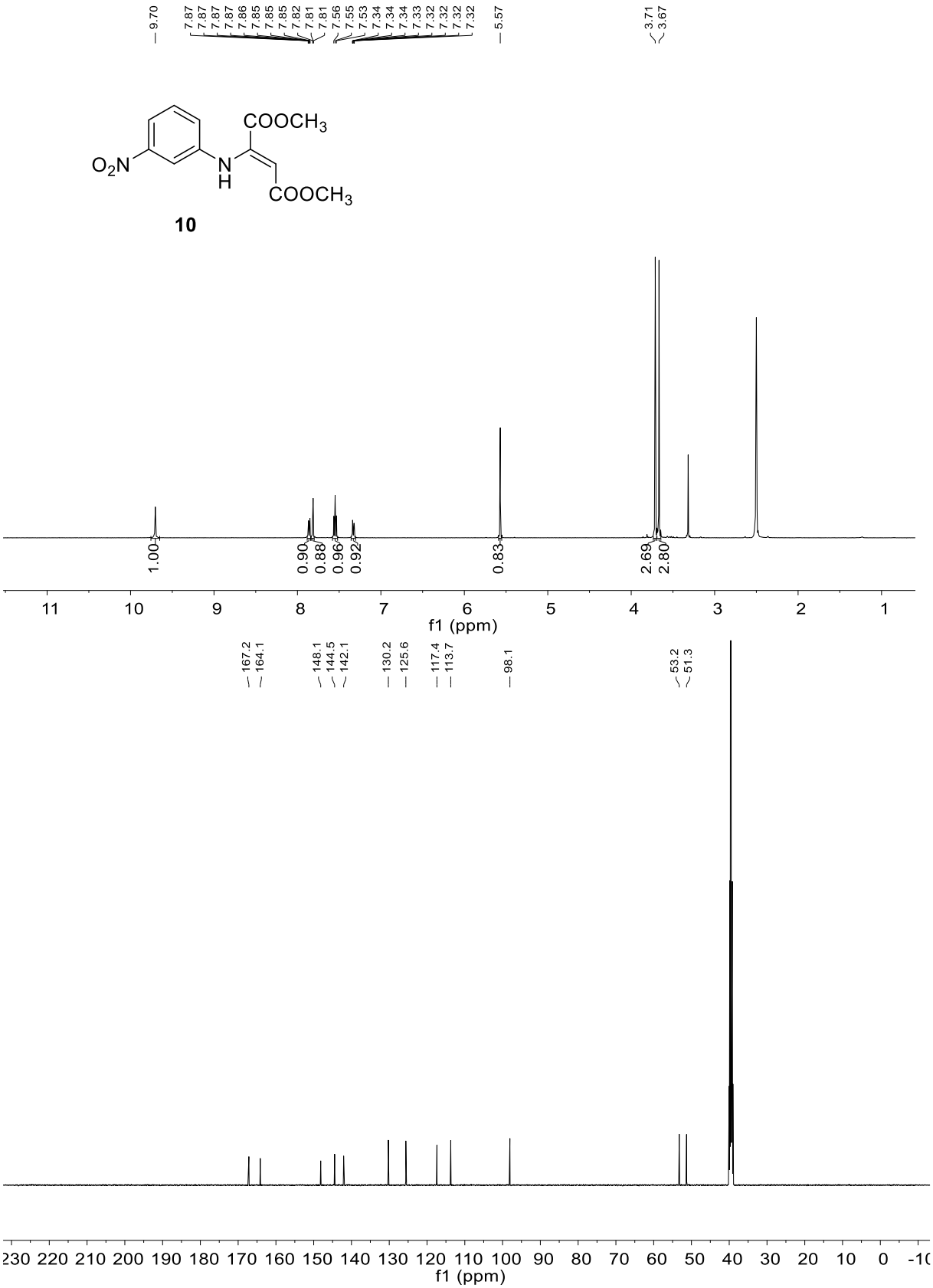


Figure S24. NMR spectra of compound **10**: ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

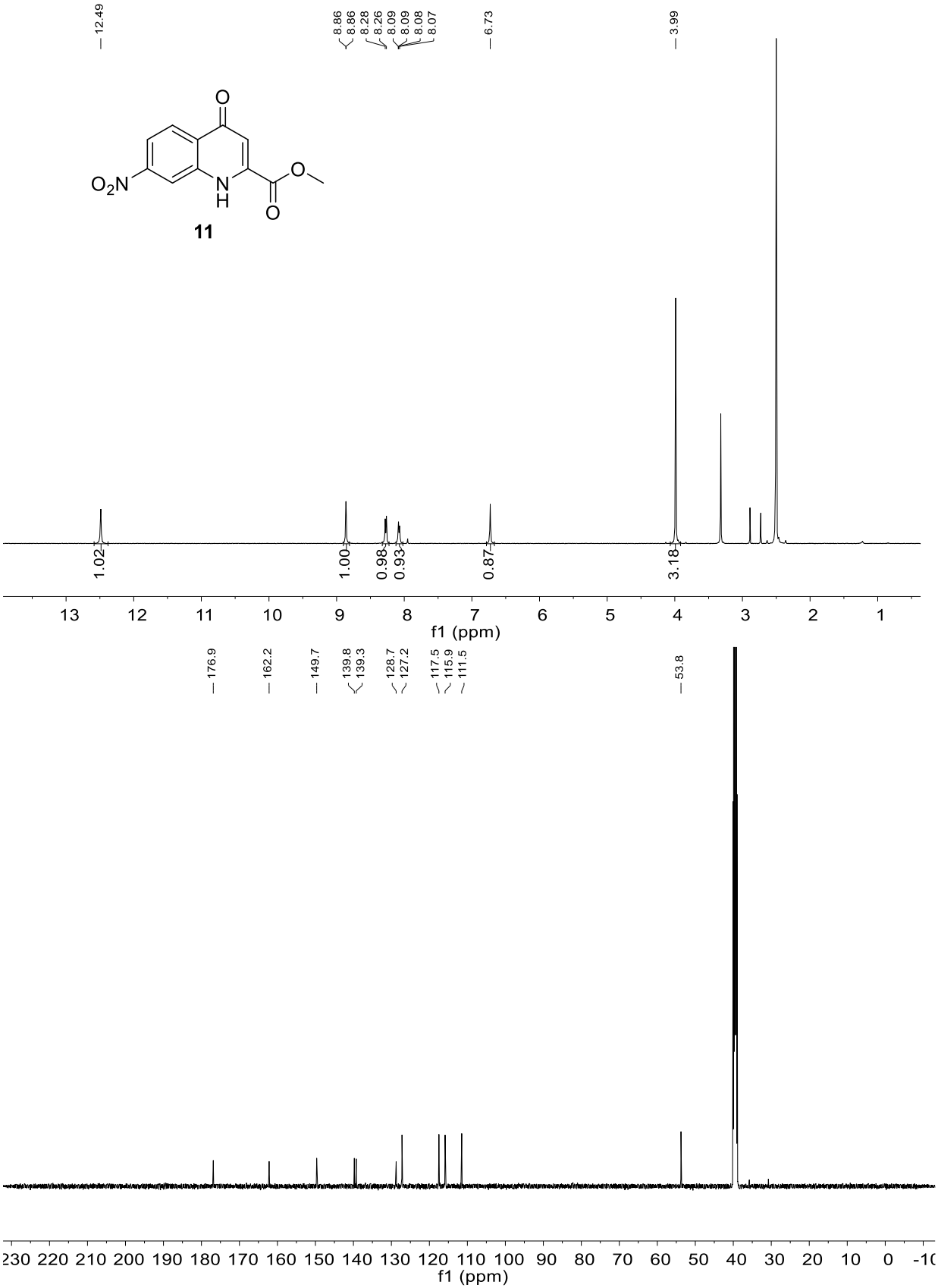


Figure S25. NMR spectra of compound **11**: ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

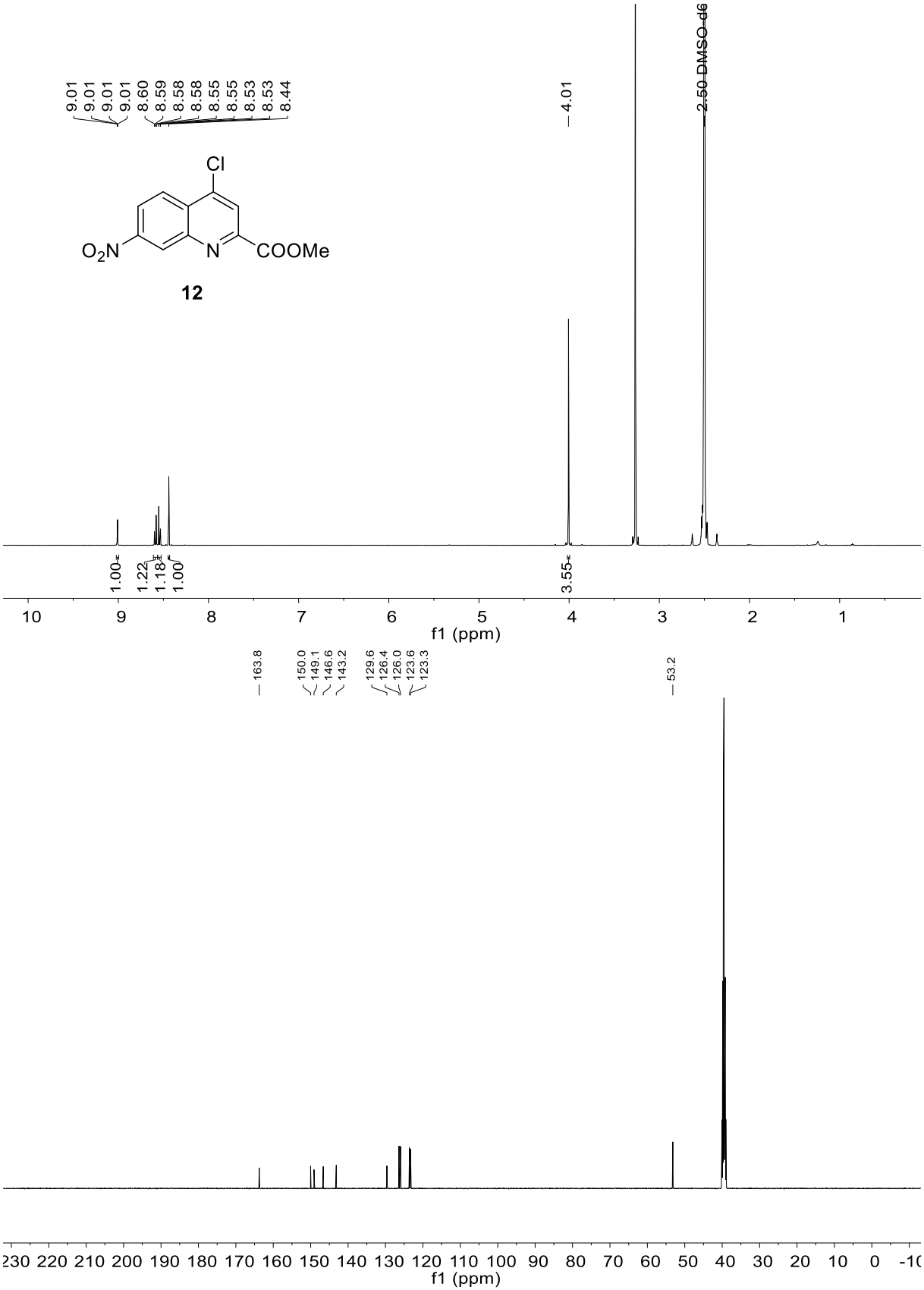


Figure S26. NMR spectra of compound **12**: ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

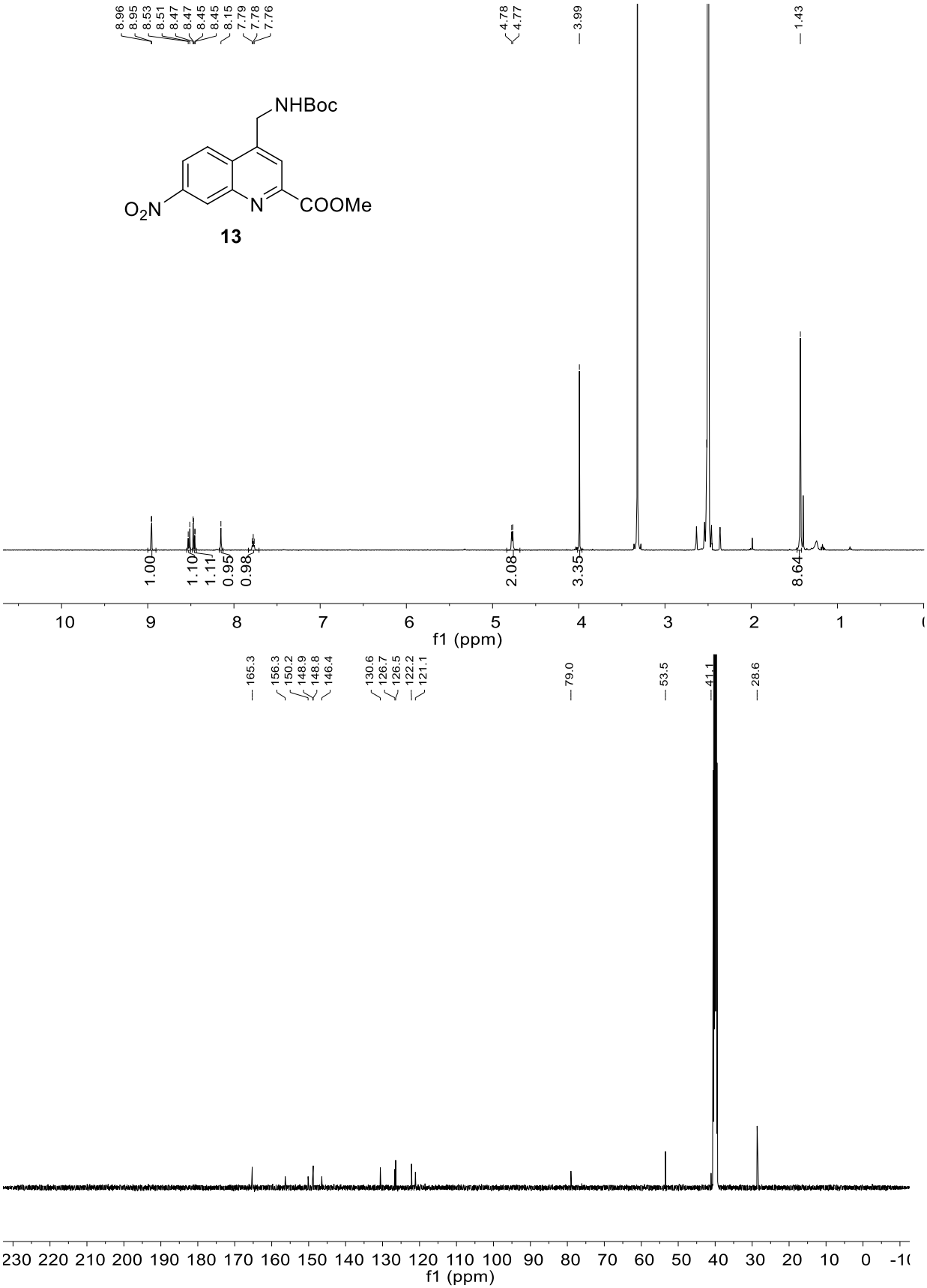


Figure S27. NMR spectra of compound **13**: ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

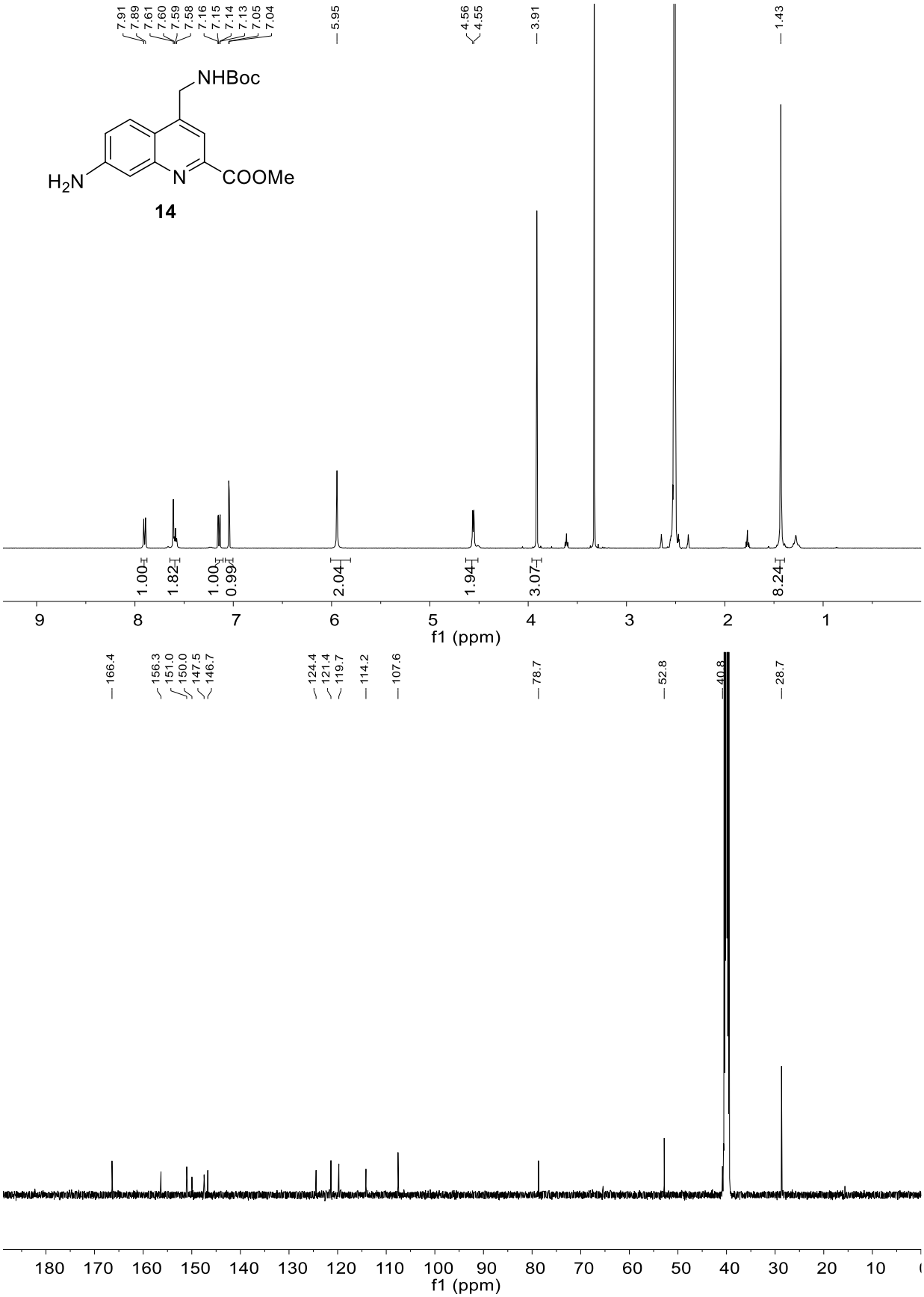


Figure S28. NMR spectra of compound **14**: ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

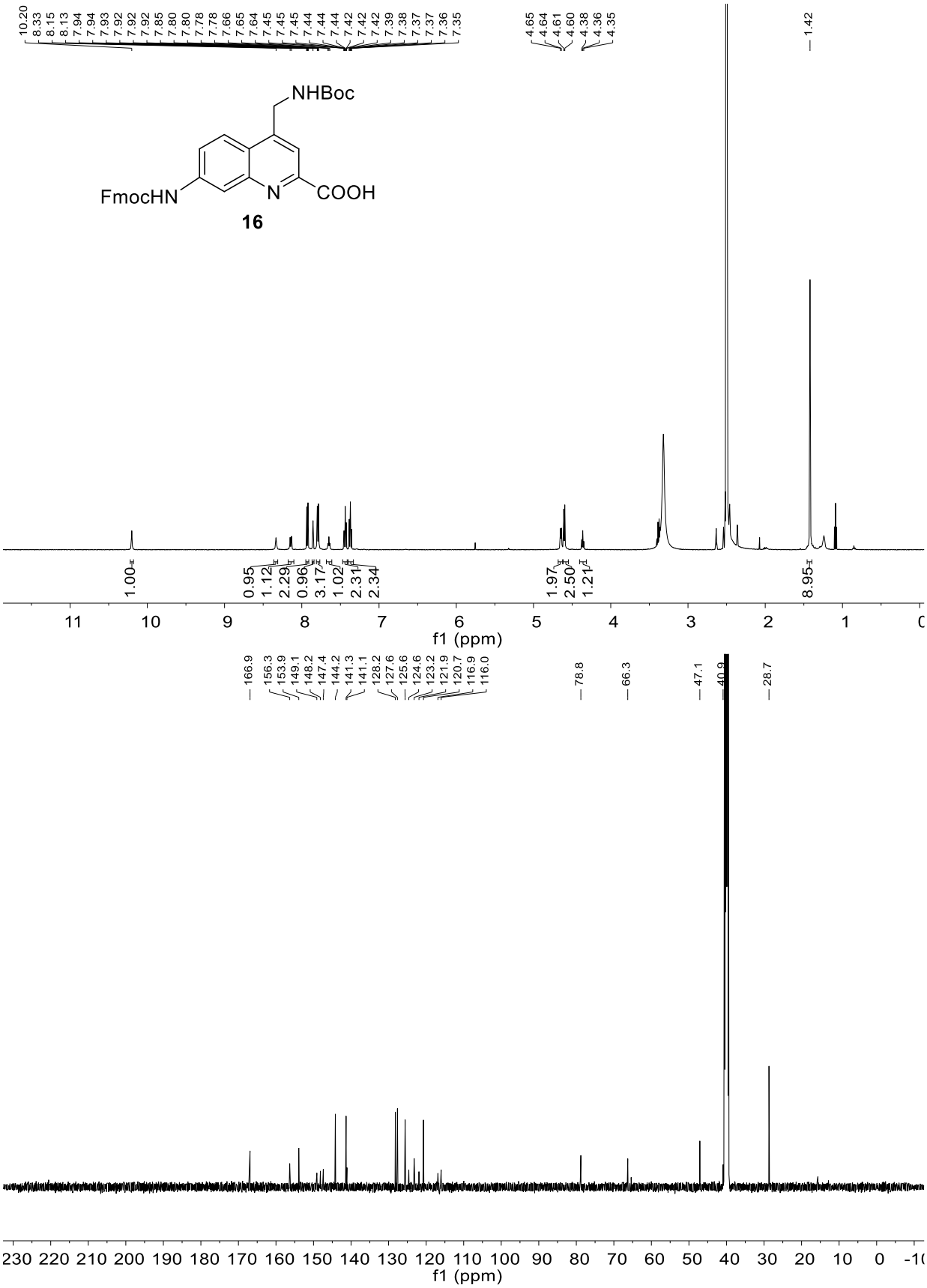


Figure S29. NMR spectra of compound **16**: ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

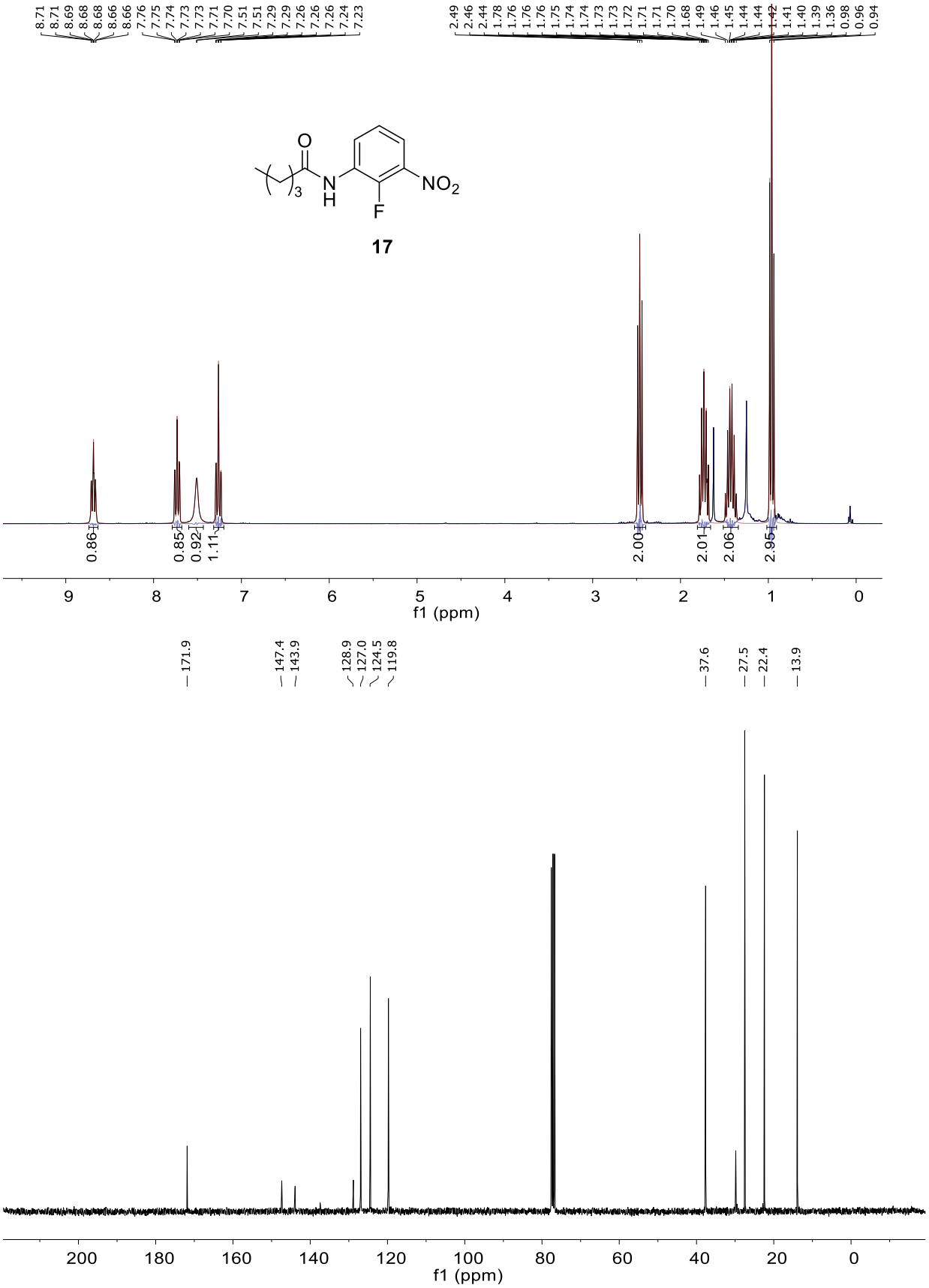


Figure S30. NMR spectra of compound **17**: ^{1}H NMR (300 MHz, $CDCl_3$) and ^{13}C NMR (75 MHz, $CDCl_3$).

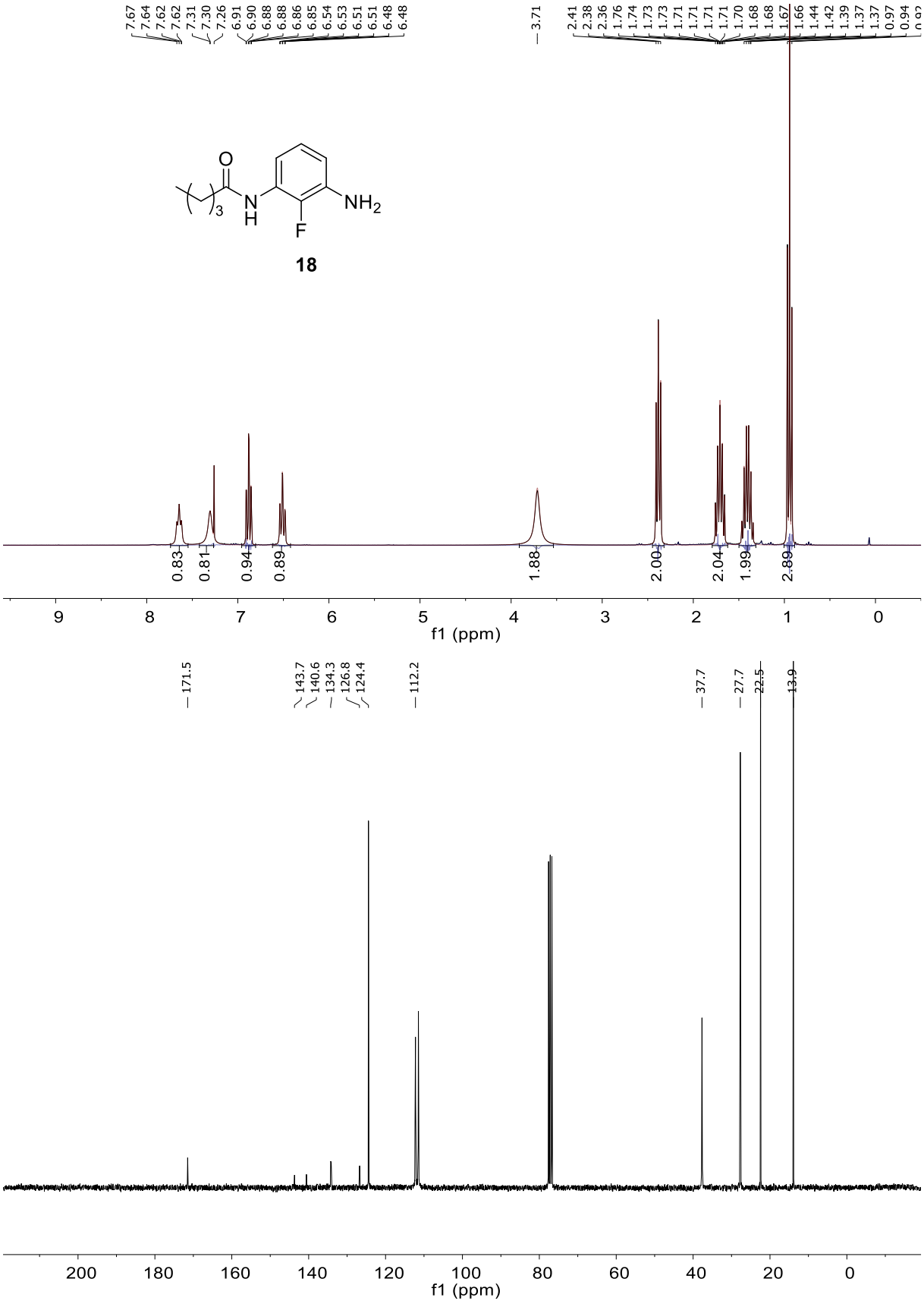


Figure S31. NMR spectra of compound **18**: ^{1}H NMR (300 MHz, $CDCl_3$) and ^{13}C NMR (75 MHz, $CDCl_3$).

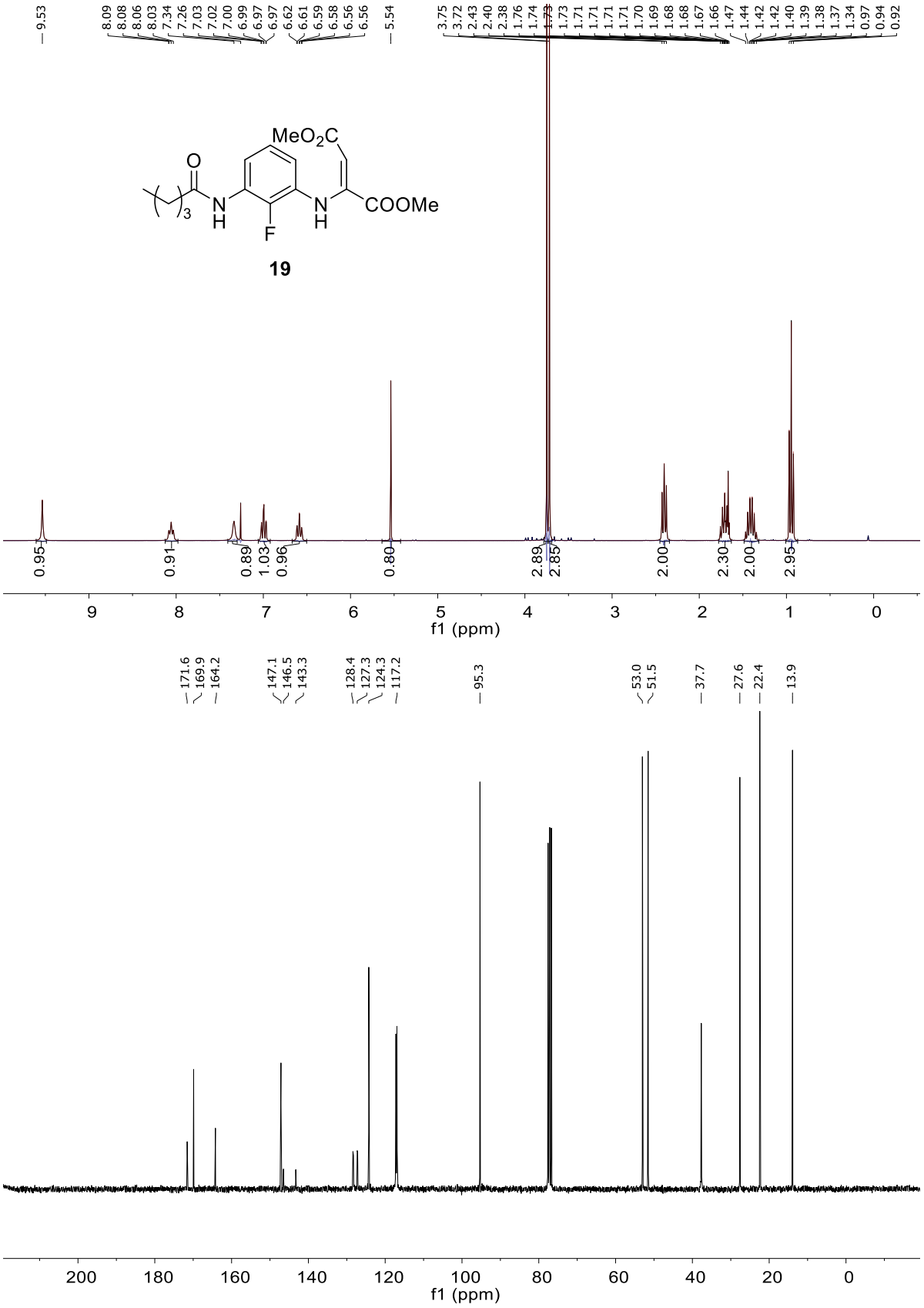


Figure S32. NMR spectra of compound **19**: ^{1}H NMR (300 MHz, $CDCl_3$) and ^{13}C NMR (75 MHz, $CDCl_3$).

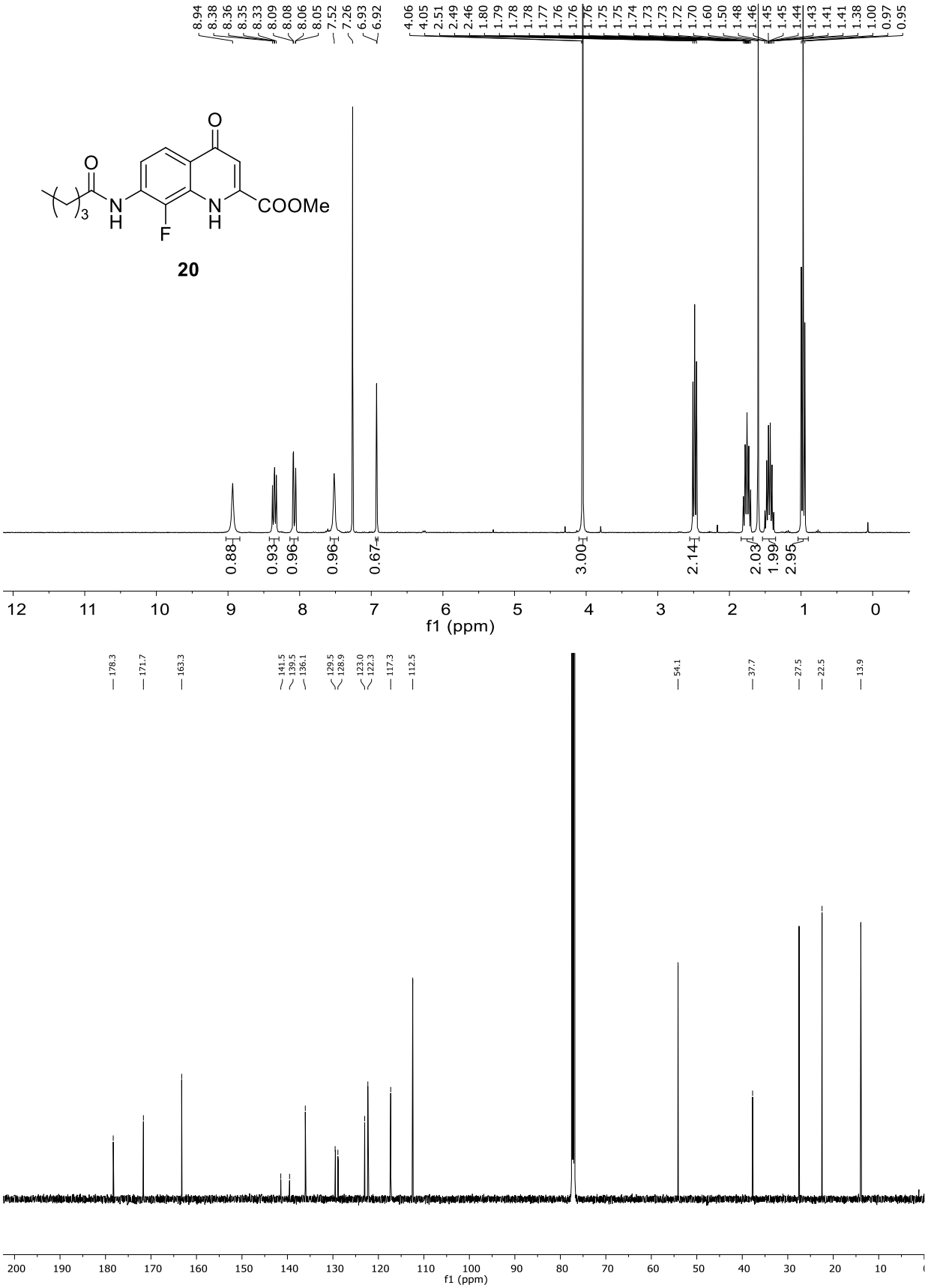


Figure S33. NMR spectra of compound **20**: ^{1}H NMR (300 MHz, $CDCl_3$) and ^{13}C NMR (75 MHz, $CDCl_3$).

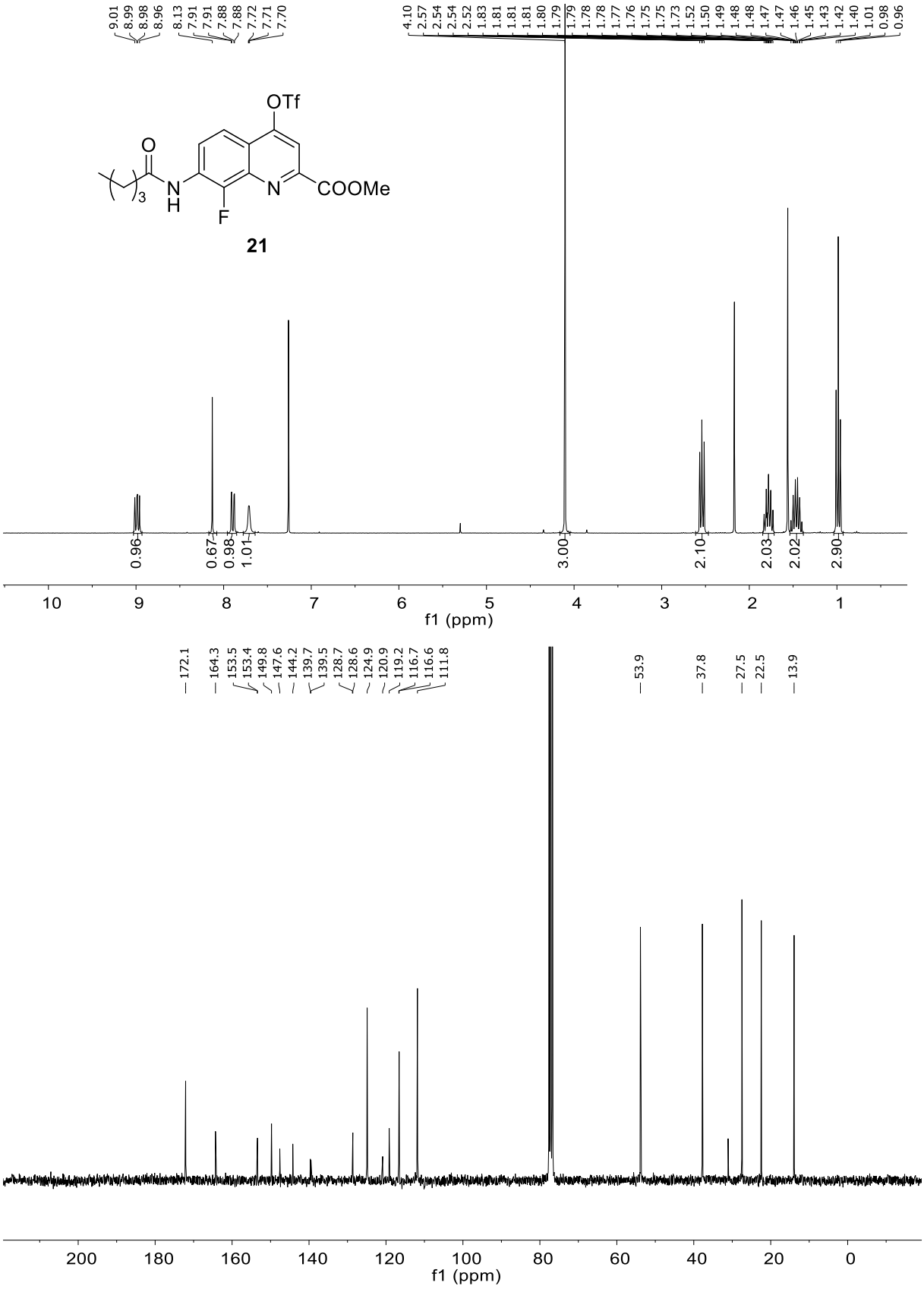


Figure S34. NMR spectra of compound **21**: ^{1}H NMR (300 MHz, $CDCl_3$) and ^{13}C NMR (75 MHz, $CDCl_3$).

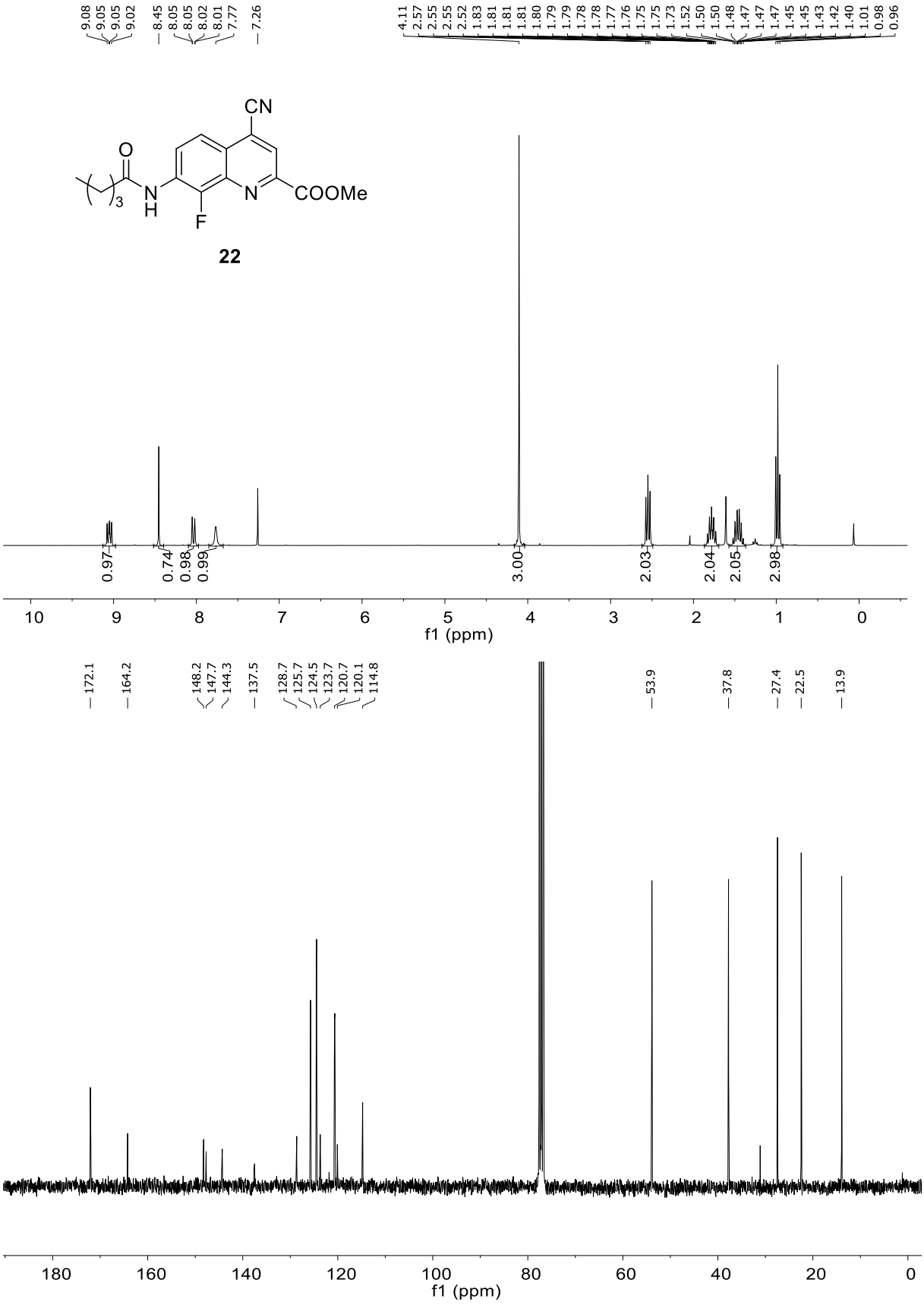


Figure S35. NMR spectra of compound **22**: ^{1}H NMR (300 MHz, $CDCl_3$) and ^{13}C NMR (75 MHz, $CDCl_3$).

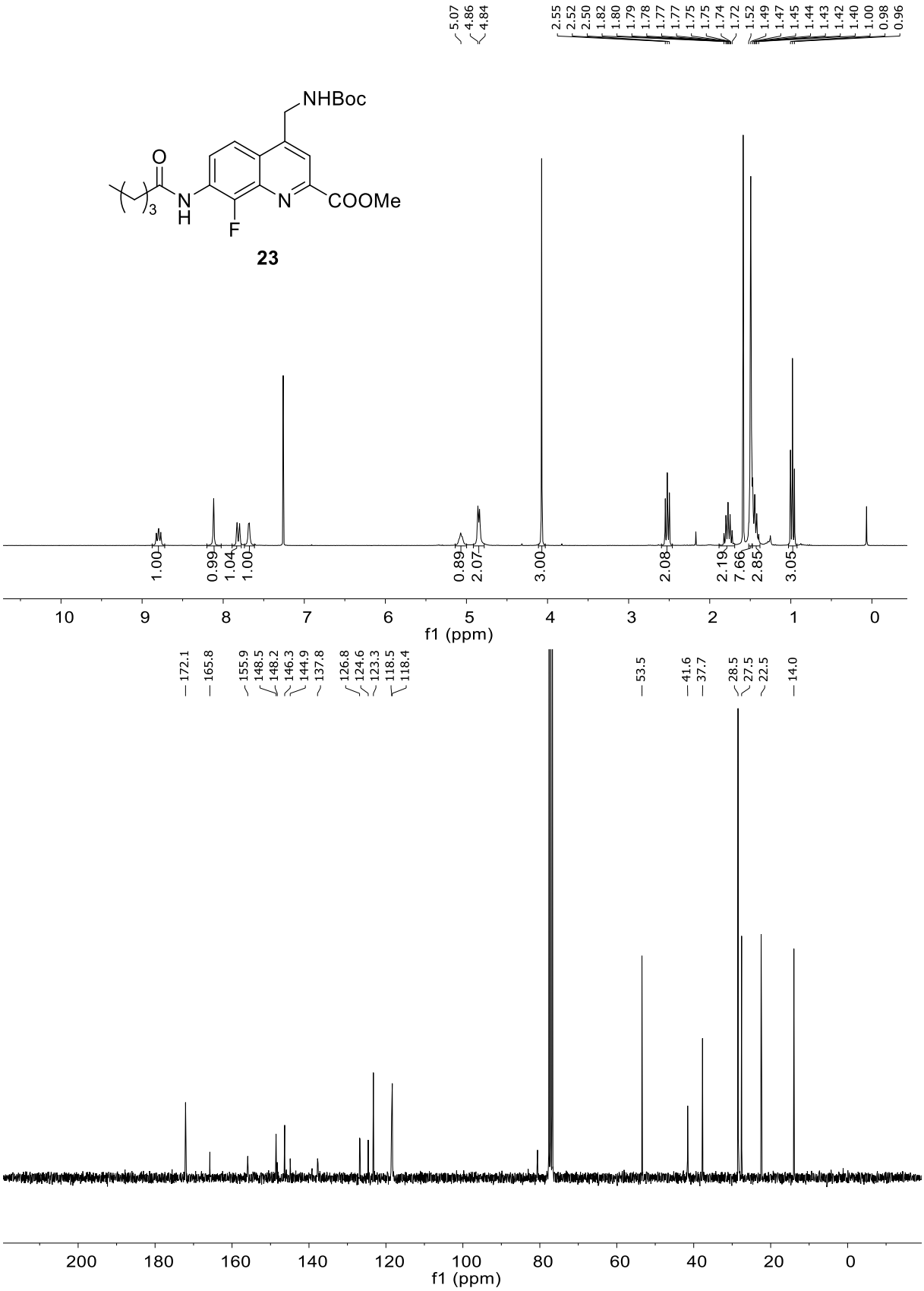


Figure S36. NMR spectra of compound **23**: ^{1}H NMR (300 MHz, $CDCl_3$) and ^{13}C NMR (75 MHz, $CDCl_3$).

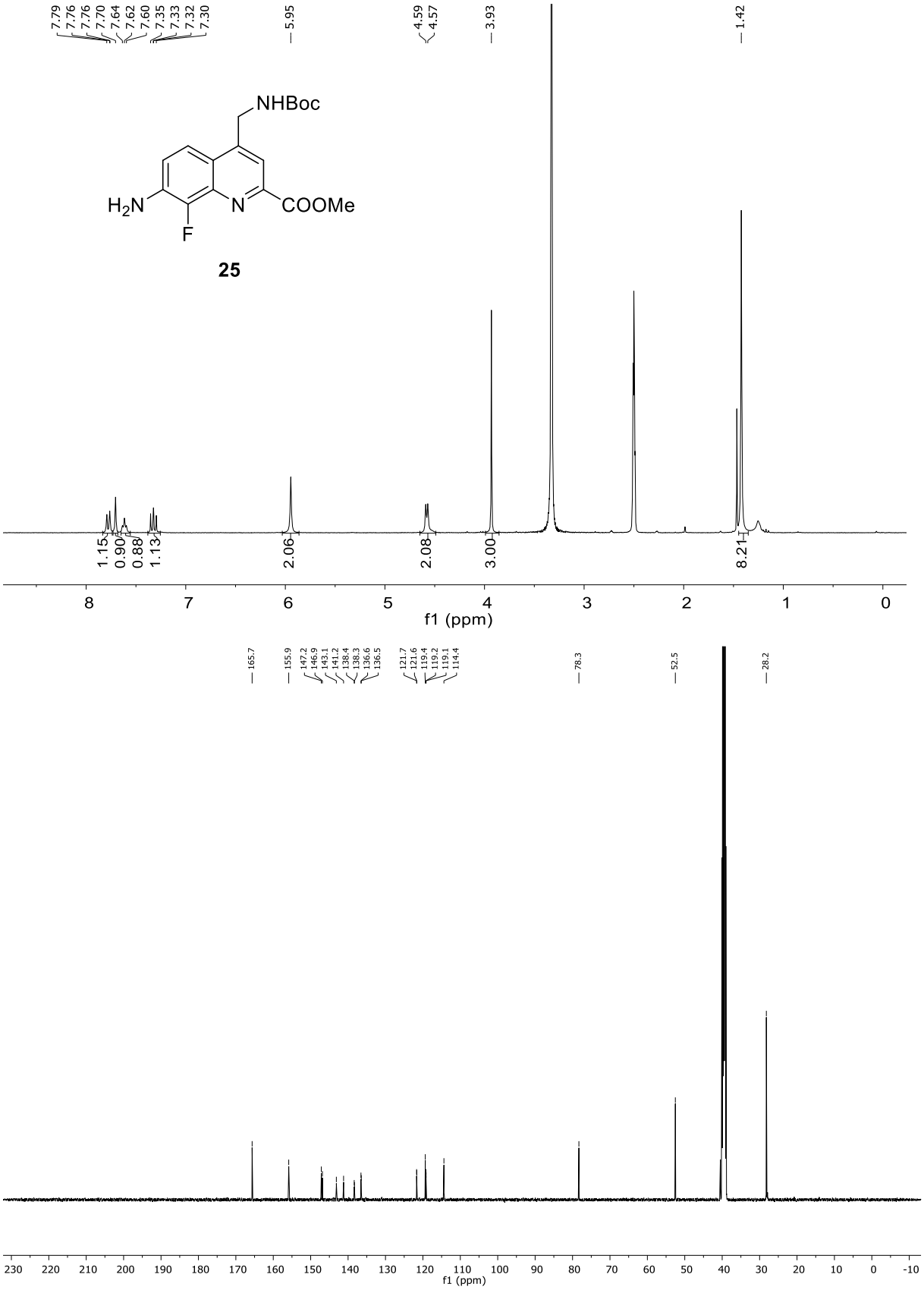


Figure S37. NMR spectra of compound **25**: ^{1}H NMR (300 MHz, DMSO-d_6) and ^{13}C NMR (75 MHz, DMSO-d_6).

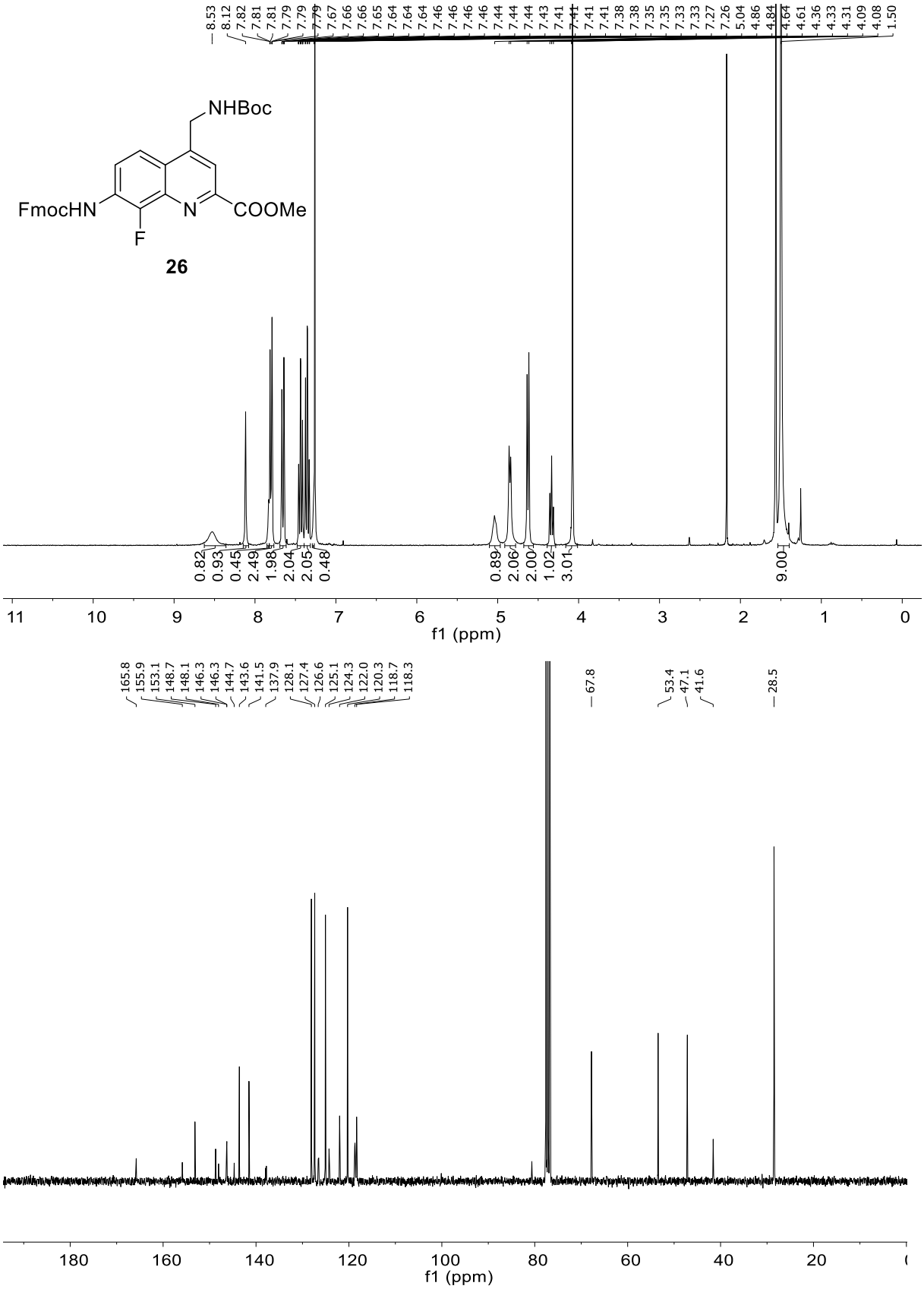


Figure S38. NMR spectra of compound **26**: ^{1}H NMR (300 MHz, $CDCl_3$) and ^{13}C NMR (75 MHz, $CDCl_3$).

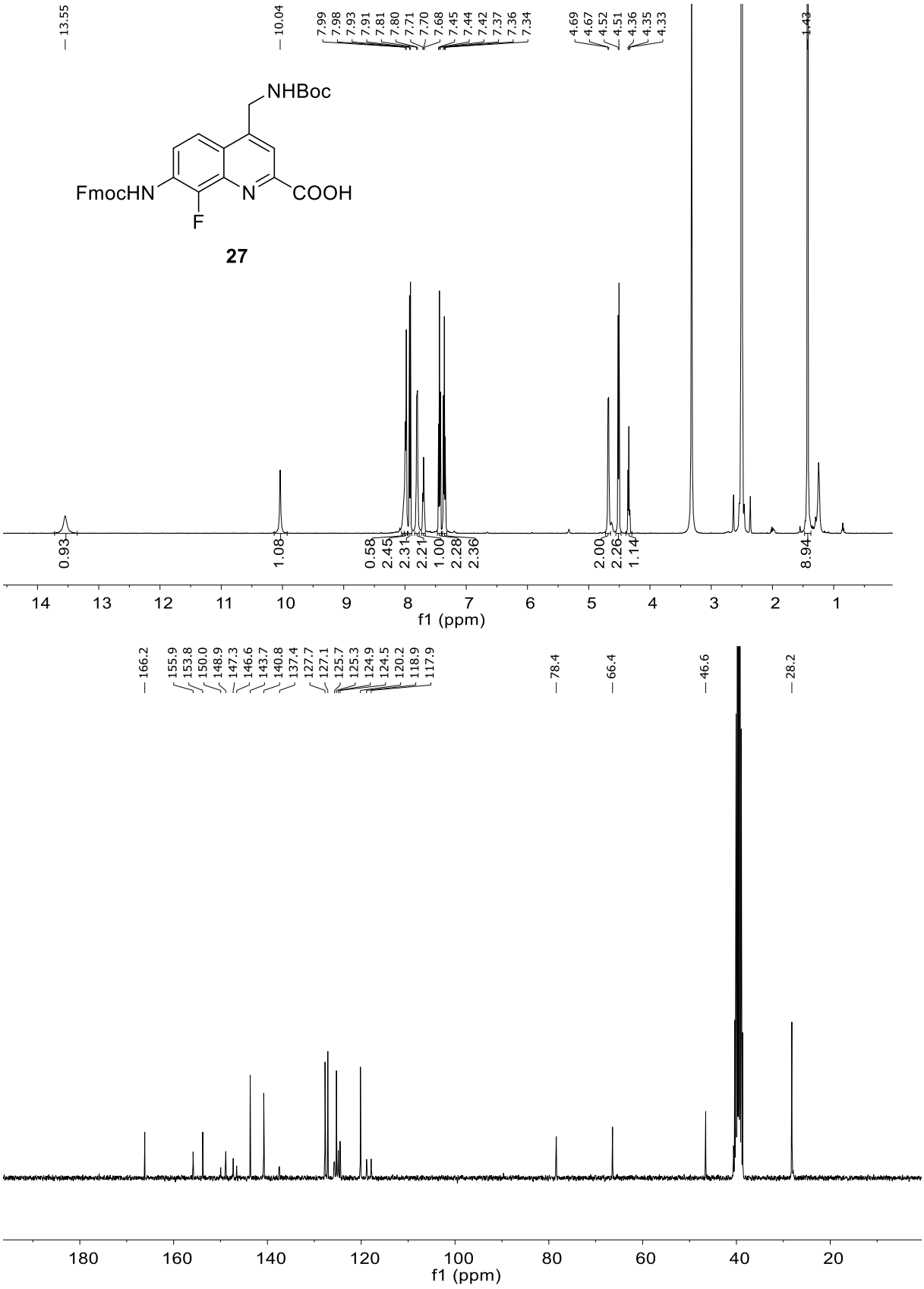


Figure S39. NMR spectra of compound **27**: ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (75 MHz, DMSO-d_6).

5 References

1 S. Dengler, P. K. Mandal, L. Allmendinger, C. Douat and I. Huc, *Chem. Sci*., 2021, **12**, 11004–11012.
2 V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat, and I. Huc, *Chem. Eur. J*, submitted.
3 X. Hu, S. J. Dawson, P. K. Mandal, X. de Hatten, B. Baptiste and I. Huc, *Chem. Sci.*, 2017, **8**, 3741–3749.
4 N. D. Heindel, I. S. Bechara, T. F. Lemke and V. B. Fish, *J. Org. Chem.*, 1967, **32**, 4155–4157.
5 J. F. K. Wilshire, *Aust. J. Chem.* 1967, **20**, 2089–2811.

5 Endothermic molecular recognition within helical aromatic oligoamide capsules in water

The rational design (Figure 14a) of a helical-folded aromatic oligoamide sequence for high-selectivity monosaccharide encapsulation in a non-polar organic solvent has been accomplished (Figure 14c).[11] Further binding research has been attempted in a polar organic solvent with hydrophilic guests (Figure 14d).[10] But the affinity completely disappeared when the higher polarity solvent was added, not to mention the more challenging work of encapsulating water-soluble guests in water, which needs to resist a strong solvation effect. In order to achieve this objective, a series of novel helical folded capsules were synthesized with different cavity sizes facilitating the shape complementary. The folded capsule structures have been described in the previous chapter. According to the cavity properties with hydrophobic –CH groups from quinoline rings and hydrophilic endocyclic nitrogen and amide groups, several amphiphilic guests were performed to determine binding affinities.

In comparison to energy-minimized models, the cap of helices containing only the Q_3 caps and Q^H cavity segments was not centered. Monomers F and H (Figure 14b) were synthesized to extend the diameter for larger guests and optimize curvature to alter the location of caps. Utilizing the same side chain introduction protocol and selective saponification of one side ester group, the building block H with a large diameter was produced. The conformational bias is driven by the hydrogen bond between hydrazine and the methoxy group and the electron repulsion between endocyclic nitrogen and carboxylic acid oxygen. The relatively small F monomer can serve as a smooth transition between the narrow helical Q and the cavity building block Q^H, so modifying the overall shape of capsules and relocating the position of the cap.

With the short and charged aminomethyl side chain, whole sequences could be well soluble in water and construct capsule structures. The NMR, ITC, computational results, and solid-state study were demonstrated in the following manuscript, which was submitted in *Chemical Science*. Several amphiphilic water-soluble guests showed strong binding affinity in water, which is considered challenging work. A unique endothermic binding process was discovered that was solely driven by releasing highly ordered water molecules to gain entropy. Conclusively, these cavities showed the possibility of binding highly solvated substrates in water and customization of the foldamer capsule shape for targeting distinct molecules.

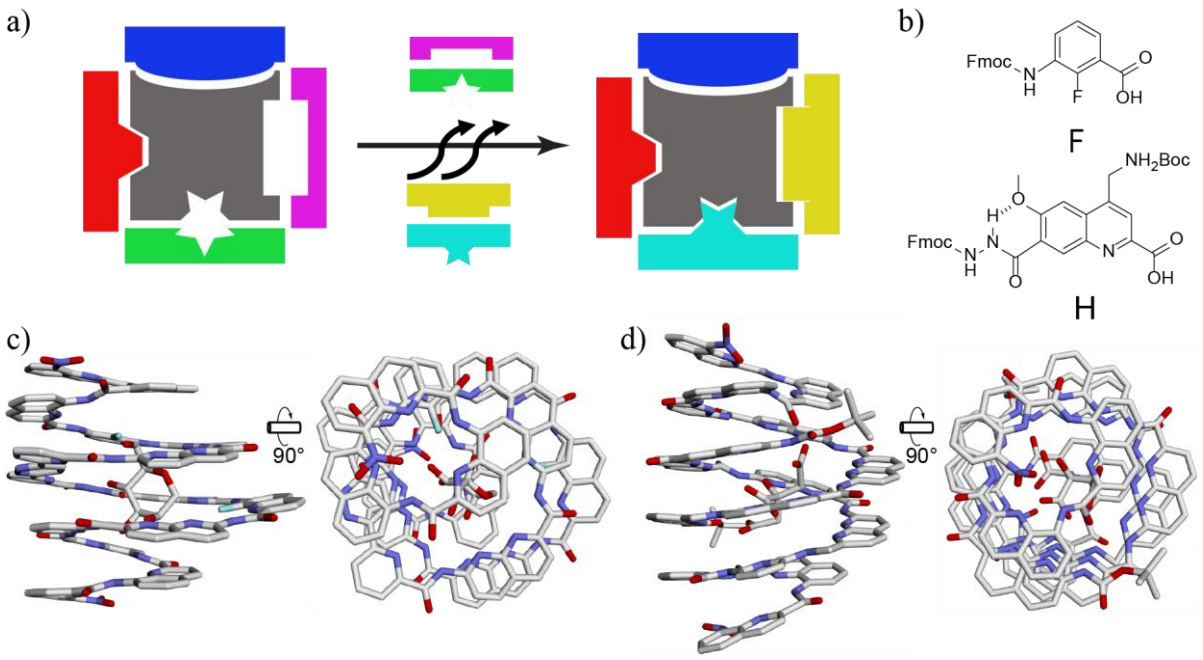


Figure 14. a) Representation of the rational design to optimize the binding cavity by replacing the sequence component facilitating shape complementary. The guest is depicted in grey and the different sequence components with various shapes are colored in different colors. b) Structure of synthesized building blocks. c) Crystal structure of a rationally designed foldamer containing the β-D-fructopyranose, allowing for high binding affinity and selectivity. d) Crystal structure of a foldamer with the citric acid in the polar organic solvent. Side chains have been omitted for clarity.

Contribution: I conceived the ideas for this project in collaboration with I. Huc. Synthetic routes and conditions for H and F monomers were developed by me with suggestions from Y. Ferrand. Oligomer synthesis and analysis were done by me. Energy minimized model and host-guest titration were done by me. L. Allmendinger was involved in the design and execution of NMR experiments. The automated solid phase synthesis was done in collaboration with C. Douat. The manuscript was written by me in collaboration with I. Huc.

ARTICLE

Endothermic molecular recognition within helical aromatic oligoamide capsules in water

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Binhao Teng,[a] Lars Allmendinger,[a] Céline Douat,[a] Yann Ferrand[b] and Ivan Huc*[a]

Helical aromatic oligoamide foldamer capsules have emerged as outstanding molecular containers in organic solvents but their potential for molecular recognition in water remained to be demonstrated. Four such water soluble capsules were designed. Monomers were prepared in solution and sequences were synthesized on solid phase. Their molecular recognition properties were assessed by ^{1}H NMR titrations. Binding in the low micromolar range was evidenced, including several cases of 1:2 associations. Hydroxymethyl-pyridine derivatives emerged as a privileged group of guests that have the peculiar property of being neither strongly hydrophilic nor strongly lipophilic (log P values close to zero). Variations of the guests structure demonstrated high selectivity of the host-guest interactions. Isothermal titration calorimetry revealed several cases of endothermal recognition, a most unusual property in water. In those cases, binding is thus driven solely by entropy which was assigned to the release of ordered water molecules from the binding cavity.

Introduction

Molecular recognition in water is at the basis of innumerable fundamental biological processes ranging from enzyme-substrate recognition to biopolymer self-assembly. It also underlies drug-target interactions and the action of pesticides, herbicides as well as depolutants. As presented in detail in a recent review, supramolecular chemistry has produced diverse synthetic hosts molecules capable of molecular recognition in water.[1] Besides the direct applications that such host molecules may have, for example to recognize protein residues,[2–4] or as sensors,[5] their main interest has been to shed light on the various non-covalent interactions involved and on the multiple roles that water molecules may play during the solvation and desolvation processes that accompany molecular recognition. An important conclusion of the study of these model systems is that recognition in water is almost systematically enthalpically favorable. Be it directed by ionic interactions,[6–10] ion-dipole interactions,[11,12] cation-π or CH-π interactions,[13,14] hydrogen bonds[15–17] or halogen bonds,[18,19] or non-classical hydrophobic effects,[20–22] binding is generally dominated by a strong favorable enthalpic term. Even when binding is driven by entropy, that is, when a strong favorable entropic term dominates, as in classical hydrophobic effects, the enthalpic term is also negative.[21,23] Examples of endothermic molecular recognition in water are very rare and include the formation of some cyclodextrin complexes[24] and some complexes between anions and bis(cyclopeptide)[25,26] a bis cyclam receptor,[27] or aromatic oligomer hosts.[19‡] In these cases, interactions within the host-guest complex and between the host-guest complex and water are altogether weaker than the sum of the interactions between the host and the guest with water. Such recognition must therefore be driven by entropy.

Here, we report our finding of endothermic recognition in water in the context of guest binding by helical aromatic oligoamide capsule hosts. We found that favorable entropic effects presumably resulting from the release of multiple water molecules from the host cavity are so strong that they allow for guest recognition even with a net loss of interactions. This occurs in particular for a class of guests that have not been the focus of much attention until now, namely neutral molecules with a partition coefficient close to zero that are neither strongly hydrophilic, strongly lipophilic, nor amphipathic (both hydrophilic and lipophilic).[§]

Aromatic foldamer helices have emerged as an important class of host molecules capable of binding lipophilic[28] and polar neutral[29–34] guests as well as ions[35–39] in organic solvents.[40–42] Among these, helical capsules constitute a peculiar class of hosts that are capped at both ends and that therefore completely surround their guest and seclude it from the surrounding medium.[37,41,43] Binding and release of the guest therefore requires the temporary opening of a window in the host structure through partial unfolding.[34] However, molecular recognition by aromatic helices has rarely been examined in water.[44] In an early attempt, it was shown that the hydrogen-bond-mediated encapsulation of a polar guest that is effective in aprotic solvents essentially vanishes in water.[45] Hydrophobic guest recognition has been highlighted in a water rich medium.[44] Stacks of shape-persistent aromatic macrocycles are structurally related to helices and have also been shown to bind

[a] Department of Pharmacy, Ludwig-Maximilians-Universität, Butenandtstr. 5–13, 81377, München, Germany. E-mail: ivan.huc@cup.lmu.de
[b] Univ. Bordeaux, CNRS, Bordeaux Institut National Polytechnique, CBMN UMR 5248, 2 rue Escarpit, 33600 Pessac, France.

†Electronic Supplementary Information (ESI) available: Synthetic protocols, crystallographic studies and characterisation of new compounds. CCDC xxxxxxx. For ESI and crystallographic data in CIF or other electronic format See DOI: 10.1039/x0xx00000x

to hydrophobic guests in water.[46] We devised that molecular recognition of aromatic helices in water would be worth exploring and may potentially display some original properties because the specific structural features of aromatic helices contrast with those of many other synthetic hosts. Specifically, despite being consisting mainly of aromatic moieties, aromatic helices only expose aromatic edges and not aromatic faces to their guests. This contrasts with the numerous hosts that operate in water thanks to cavities surrounded by aromatic faces that contribute hydrophobic effects as well as cation-π or CH-π interactions.[1] In the following, we demonstrate that aromatic helices can recognize neutral organic molecules in water and that binding is driven by a strongly favorable entropic term although the guests are not hydrophobic.

Results and discussion

Design and synthesis

Aromatic oligoamide helical capsules consist of sequences where monomers in the center code for a wide helix diameter, thereby creating a cavity, whereas those near the ends code for a narrow diameter thus closing the cavity.[42] Segments consisting of three 8-amino-2-quinoline carboxylic acid units (Q in Fig. 1) have been typically used as end-caps at the ends of capsule sequences. These Q_3 segments also help ensuring the single helical nature of the capsules by effectively hampering the aggregation of the central part of the sequences into double helices.[47] We recently introduced 4-substituted 7-amino-2-quinoline carboxylic acid units (Q^H, Fig.1) as an analogue of naphthyridine units N.[48] Like N, Qh units encode for a larger helix diameter as a result of the wider angle defined by the orientation of its amine and acid groups. However, unlike N, the position 8 of Qh exposes an aromatic CH towards the cavity – at the inner rim of the helix – instead of the endocyclic nitrogen atom of the naphthydridine. The purpose of this change was to introduce some hydrophobic patches in the cavity, following the observation that polar environments comprised of N units failed to bind polar guests through hydrogen bonding in water.[45] Another difference between N and Qh is the lack of conformational control at the aryl-NH linkage in the latter. The conformations of aryl-CO and aryl-NH bonds shown in all aromatic monomers shown in Fig. 1 are stabilized by hydrogen bonds or electrostatic repulsions involving the amide group and endocyclic nitrogen atoms or exocyclic fluoro or methoxy substituents, expect the aryl-NH bond of Qh. We have shown that this additional degree of freedom hampers folding in DMSO but that helical folding driven by hydrophobic effects occurs in water. Sequences **1**-**4** were thus previously synthesized and shown to fold into single helices in water, resulting in a single set of sharp signals in their ^{1}H NMR spectra (Fig. 2).[48] Water solubility in ensured by the 4-aminomethyl groups that diverge from the helix and are protonated in neutral or acid conditions. An earlier study also pointed to the importance of these many charged side chains prevent aggregation in water.[49]

Compounds **1**, **3** and **4** consist of Ac-$Q_3Qh_nQ_3$-OH sequences with n = 8, 2 or 16, with their Qh segments spanning

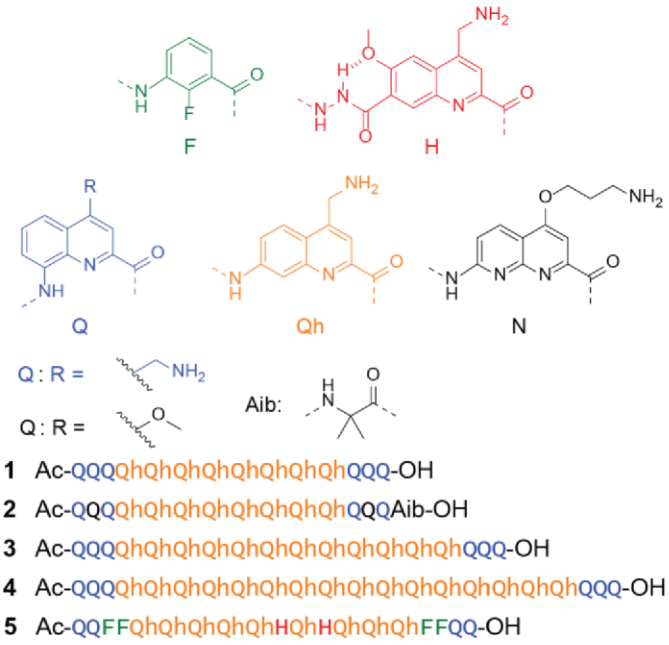


Fig.1 Structures of Q, Qh, F and H monomers and composition of sequences **1-5** investigated in this study. Different aromatic monomers are highlighted with colors. Q: blue (4-aminomethyl side chain) or black (4-methoxy side chain), Qh: orange, F: green, H: red.

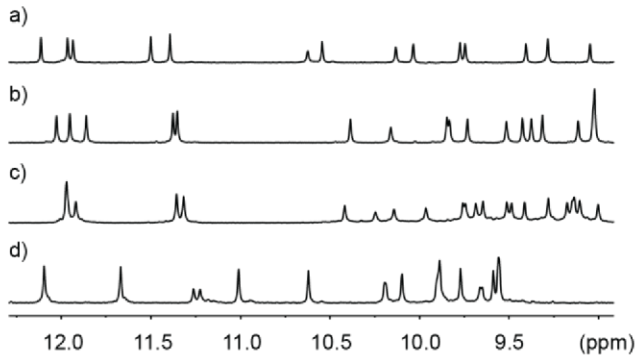


Fig.2 Excerpts from the 500 MHz ^{1}H NMR spectra of compounds **1** (a), **3** (b), **4** (c), and **5** (d), as ammonium trifluoroacetate salts, all recorded at 298 K in H_2O/D_2O 90:10 (v/v).

approximately 1.8, 2.7 or 3.6 helix turns, respectively. Sequence **2** is an analogue of **1** that could be synthesized in higher yield.

The cavities of **2** and **1** are identical and the two compound were no differentiated in binding studies. Energy minimized molecular models of **1**, **3** and **4** are shown in Fig. 3a-i. Computational analysis allowed for the calculation of the cavity dimensions. Since these sequences differ only in the length of their Qh segment, the diameter of the cavity is consistent throughout the series and only its length varies. The top views (down the helix axis) of molecular models illustrate the lower diameter of the Q_3 capping segments, but also that these imperfectly close the central cavity: the helix axes of Q_3 and Qh_n segments are not collinear as the result of the abrupt change of curvature at their junction. This observation led to the design of new sequence **5** (Fig. 1).

In sequence **5**, the Q_3 segments have been replaced by a Q_2F_2 and a F_2Q_2 segments at the N and C-termini, respectively. The contribution to helix diameter of monomer F is larger than

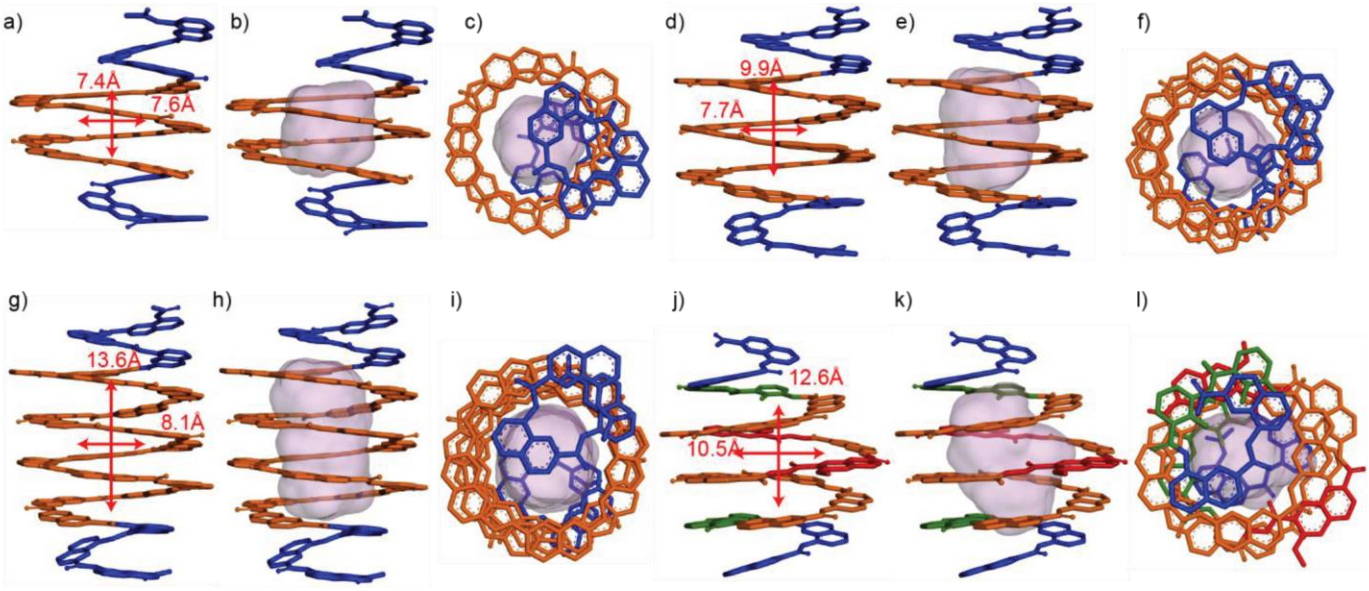


Fig.3 Energy-minimized capsule conformations (Maestro, MMFFs force field, water as implicit solvent)[50] of compound **1**, **3**, **4**, and **5**. Stick representations of a) **1**, d) **3**, g) **4** and j) **5**, showing their corresponding cavity length and diameter. Side views of b) **1**, e) **3**, h) **4** and k) **5** and top views of c) **1**, f) **3**, i) **4** and l) **5**, demonstrating calculated cavity volumes (254 Å^3, 352 Å^3, 550 Å^3 and 422 Å^3 for **1**, **3**, **4**, and **5** respectively). The volumes were calculated utilizing virtual rolling probes with inner probe radius of 1.4 Å and outer probe radius of 4 Å. Different monomer units are highlighted with the same colors as in Fig. 1 Q: blue, Qh: orange, F: green, H: red). Side chains have been omitted for clarity.

that of Qh but smaller than that of Q, leading to a better align of the caps with the central cavity (Fig. 3j-l). Another feature of **5** is the incorporation of two H units near the center of the sequence (Fig. 1). H units code for an even larger diameter than Qh and also introduce a good hydrogen bond acceptor – a hydrazide carbonyl group – in the cavity. As a result, the cavity of **5** is predicted to be wider than those of **1**-**4**, and thus potentially able to accommodate larger guests. Both H and F units have been used previously in the context of organic soluble aromatic oligoamide capsules.[33] For the purpose of this study, the H monomer was equipped with a new 4-aminomethyl side chain and both H and F were produced as Fmoc-acid precursors (with a Boc-amine side chain protection for H) ready for solid phase synthesis.[51] These monomer modifications as well as the synthesis of **5** are described in detail in the Electronic Supplementary Information (Schemes S1, S2).† Like for the other sequences, the ^{1}H NMR spectrum of **5** shows one set of sharp resonances consistent with its folding into a signal helical monomer (Fig. 2d). Of note, the solid phase preparation of **1**-**5** on low loading Wang or Protide resins and their subsequent purification by RP-HPLC are much quicker than earlier solution phase syntheses of organic soluble capsules, all the more so with the newly-developed automation of aromatic oligoamide foldamer synthesis.[52] This progress bodes well for the future implementation of iterative structure-based design of foldamer capsules as performed previously with organic-soluble hosts.[33,53]

Screening of binding properties

At the onset, it should be pointed that **1**-**5** were designed as water soluble helical capsules with sizeable cavities, but were not intended to binding any particular type of guest. Their binding properties were thus explored through a guest screening phase guided by structural considerations. Top views of Qh_4 and $(QhH)_2$ segments are shown in Fig. 4 and illustrate the arrangement of functional groups at the inner rims of the helices. The amide group all provide NH hydrogen bond donors. Endocyclic nitrogen atoms are hydrogen bond acceptors but they may not be available to guest molecules since they are already involved in intramolecular hydrogen bonding with the amide NHs. As mentioned above, the H units of **5** provide a carbonyl hydrogen acceptor and the aryl CH in position 8 of Qh units may constitute small lipophilic patches. NH groups have been used to binding to anions[25,54,55] and aryl CHs are somewhat polarized and can also serve that purpose.[56] Nevertheless, ^{1}H NMR spectroscopy showed no evidence of anion binding in **1**-**4**. Dianions such as sulfate generally caused precipitation, indicating their preferential interaction with the peripheral ammonium side chains.

Another speculative consideration about the binding properties of the water soluble aromatic foldamer capsules is a possible favorable role of entropy. We recently reported the crystal structure of an anionic analogue of **2** whose cavity was filled with fourteen clustered water molecules in the solid state (Fig. 5).[49] Even though the structure concerns a double helix with a cavity twice as large as that expected for **2**, the fact that the water molecules are crystallographically defined indicates the prevalence of one spatial arrangement and thus of one array of hydrogen bonds between the water molecules and between water and the inner wall of the capsule. X-ray data collection was performed at low temperature and room temperature solution behavior is more dynamic. Nevertheless, this observation points to possible gains of degrees of freedom, *i.e.* entropic benefits, upon replacing the water molecules by suitable guests.

Various guests were tested for binding with the four different capsules (Fig. 6). Guests were selected because their overall volumes were compatible with the corresponding expected capsule cavity size. They typically contained some hydrogen bond acceptors, donors and some hydrophobic patches. Most of them contained an aryl ring. Few of them were chiral or contained charges. But for a couple exceptions, guests were water soluble. Several molecules tested where introduced as variants of identified guests, e.g. benzyl alcohol and 4-aminomehtyl-pyridine as variants of 4-hydromethyl-pyridine.

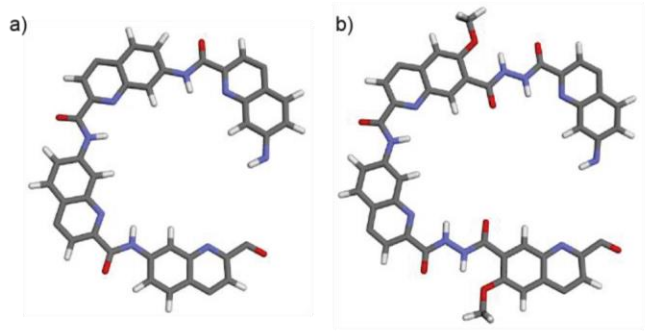


Fig. 4 Tube representation of molecular models of a) Qh_4, and b) $(QhH)_2$ segments highlighting the functional groups that converge towards the helix cavities of **1**-**5** and that will determine their molecular recognition properties. The models are part of the structures shown in Fig. 3. Side chains have been removed for clarity.

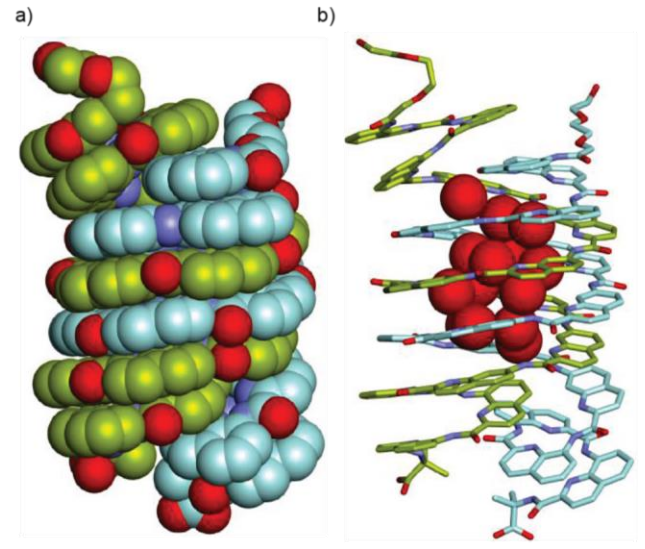


Fig. 5 Solid -state structure of the double helical dimer of an anionic analogue of **2** (details are shown in Fig.S1†).[49] a) Side view shown in CPK representation with color-coded strands. b) Same view in tube representation showing the oxygen atoms of fourteen crystallographically defined encapsulated water molecules. The hydrogen atoms of the water molecules were not visible in the electron density map. Side chains and CH and NH hydrogen atoms have been omitted for clarity.

Binding tests were performed by ^{1}H NMR titrations upon adding a solution of guest to a solution of the empty host in H_20/D_2O (9:10 *v*/*v*) (Fig. S2-16†). The absence of chemical shift variations were interpreted as an indication of an absence of binding in the concentration range used for NMR. This interpretation stemmed not only from the fact that host-guest interactions were expected to result in chemical shift variations but also from the fact that helical hosts undergo slight guest-induced changes of shape,[57] like a snake around its prey, and that these structural changes are also expected to result in chemical shift variations due to changes in the ring current effects associated with aromatic stacking in the helix. When binding was observed, titration data were fitted to a binding isotherm to determine association constants (K_a) as summarized in Table 1.

The titration of sequence **3** by 5-hydroxymethyl-2-methylpyridine is shown in Fig. 7 as a representative example. Characteristic changes of the ^{1}H NMR signals of the amide protons of the host and of the 2-methyl group of the guest during the titration indicate that binding is fast on the NMR time scale (Fig. 7a). Both upfield and downfield shifts of the NH resonances are observed and the maximal amplitude of the shifts at saturation vary greatly. The interpretation of these changes is not straightforward as they may result as much from host-guest interactions as from a change of shape of the host. For example, a downfield shift of an NH signal may result from stronger hydrogen bonding to the guest than to water, but also from a weaker intramolecular ring-current effect within the host due to a slight change of shape. One may nevertheless observe that the most deshielded protons at 11.5 ppm or above, which belong the terminal Q3 segments, are less altered by the presence of the guest, consistent with their location somewhat away from the binding site. Conversely, overlapping signals of the free hosts near 9.2 ppm are differentiated in presence of the guest. Also the upfield shift of the 2-methyl resonances of the guest upon binding indicate their exposure to a molar polar environment in the helix cavity than in water.

The sigmoidal shape of the titration curve observed for various protons (Fig. 7b, Fig. S7) cannot be accounted by a 1:1 binding isotherm. The data fitted well to a 1:2 binding model yielding. K_a values of 1870 M^{-1} and 800 M^{-1} for the first and second binding event, respectively, indicating a slight positive binding cooperativity. ESI-MS under mild (native) conditions,[58] revealed the presence of both 1:1 and 1:2 species in the gas phase (Fig. S17). A nuclear Overhauser enhancement spectroscopy (NOESY) measurement allowed for the observation of several distinct intermolecular correlations between host **3** and 5-hydroxymethyl-2-methylpyridine, suggesting close proximity in a structurally defined complex (Fig. 7d, blue circled). In addition, a correlation between the methyl and H6 protons of the guest (Fig. 7d purple circle) can hardly be accounted for by proximity between these protons with a guest molecule, all the more so that the equivalent methyl/H4 correlation is missing. Instead, the observed correlation is more likely intermolecular and would reflect proximity between two guest molecules within the1:2 complex. A plausible energy minimized model of the 1:2 complex could be build where the two guests were stacked face-to-face with an antiparallel orientation, thus placing their respective methyl groups at opposite ends of the cavity (Fig. 8). This arrangement brings the methyl of each guest in close proximity with the H6 proton of the other.

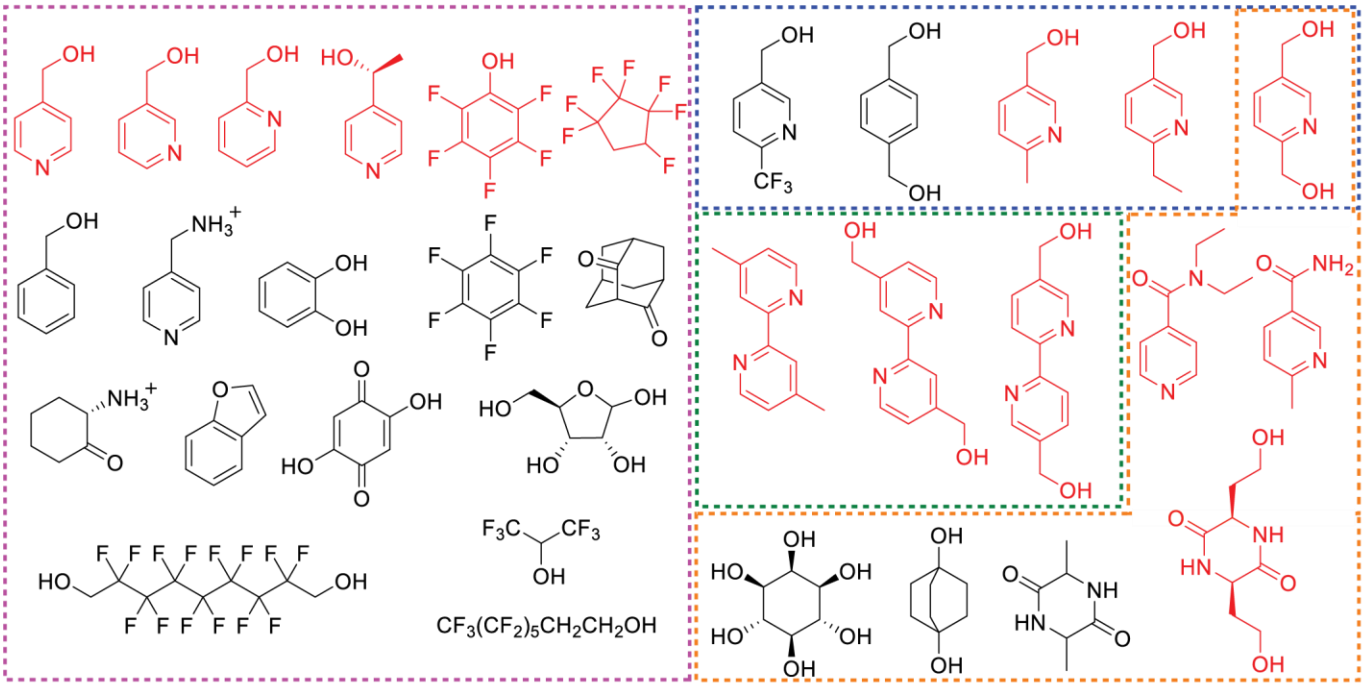


Fig. 6 Structural formulas of the guests tested for binding by **1**, **3**, **4**, and **5**. Guests for which binding was observed are shown in red. The dashed lines encircle groups of guest that were tested with sequence **1** (purple), **3** (blue), **4** (green) or 5 (orange). 3,6-Dimethylpiperazine-2,5-dione was used as a mixture of diastereomers.

Table 1 Octanol-water partition coefficient (P) of various guests and their association constants (K_a) with **1-5** in water at 25 °C as determined by ^{1}H NMR titration.

		1		**3**		**4**		**5**	
Guest	Log P[a]	K_{a1} (M^{-1})	K_{a2} (M^{-1})	K_{a1} (M^{-1})	K_{a2} (M^{-1})	K_{a1} (M^{-1})	K_{a2} (M^{-1})	K_{a1} (M^{-1})	K_{a2}(M^{-1})
4-pyridinemethanol	-0.06	45300 (39800)[b]	510 (1080)[b]						
3-pyridinemethanol	-0.17	1500[c]							
2-pyridinemethanol	-0.05	1080[c]							
(R)-(+)-α-Methyl-4-pyridinemethanol	0.13	1080[c]							
pentafluorophenol	1.57	350							
1,1,2,2,3,3,4-heptafluorocyclopentane	-[d]	-[e]							
5-hydroxymethyl-2-methylpyridine	0.29			1870	800				
(5-ethylpyridin-2-yl) methanol	0.97			2610	380				
pyridine-2,5-diyldimethanol	-0.76			3110	270			600	
4,4′-dimethyl-2,2′-dipyridyl	1.99					3720	1910		
[2,2'-bipyridine]-4,4'-diyldimethanol	0.30					20500	940		
[2,2'-bipyridine]-5,5'-diyldimethanol	0.57					2950 (11500)[b]	1350 (1100)[b]		
6-methylpyridine-3-carboxamide	-0.01							270	
N,*N*-diethylisonicotinamide	0.27							390	
cis-3,6-bis(2-hydroxyethyl) piperazine-2,5-dione	-2.01							26 (11)[b]	

[a]Measured by UV absorbance or HPLC. [b]$K_a$values in parenthesis were determined by ITC. [c]Titration was performed with **2** instead of **1**. [d]No UV absorbance could be detected for this guest. [e]Binding was observed but no K_a could be determined due to the low solubility in water.

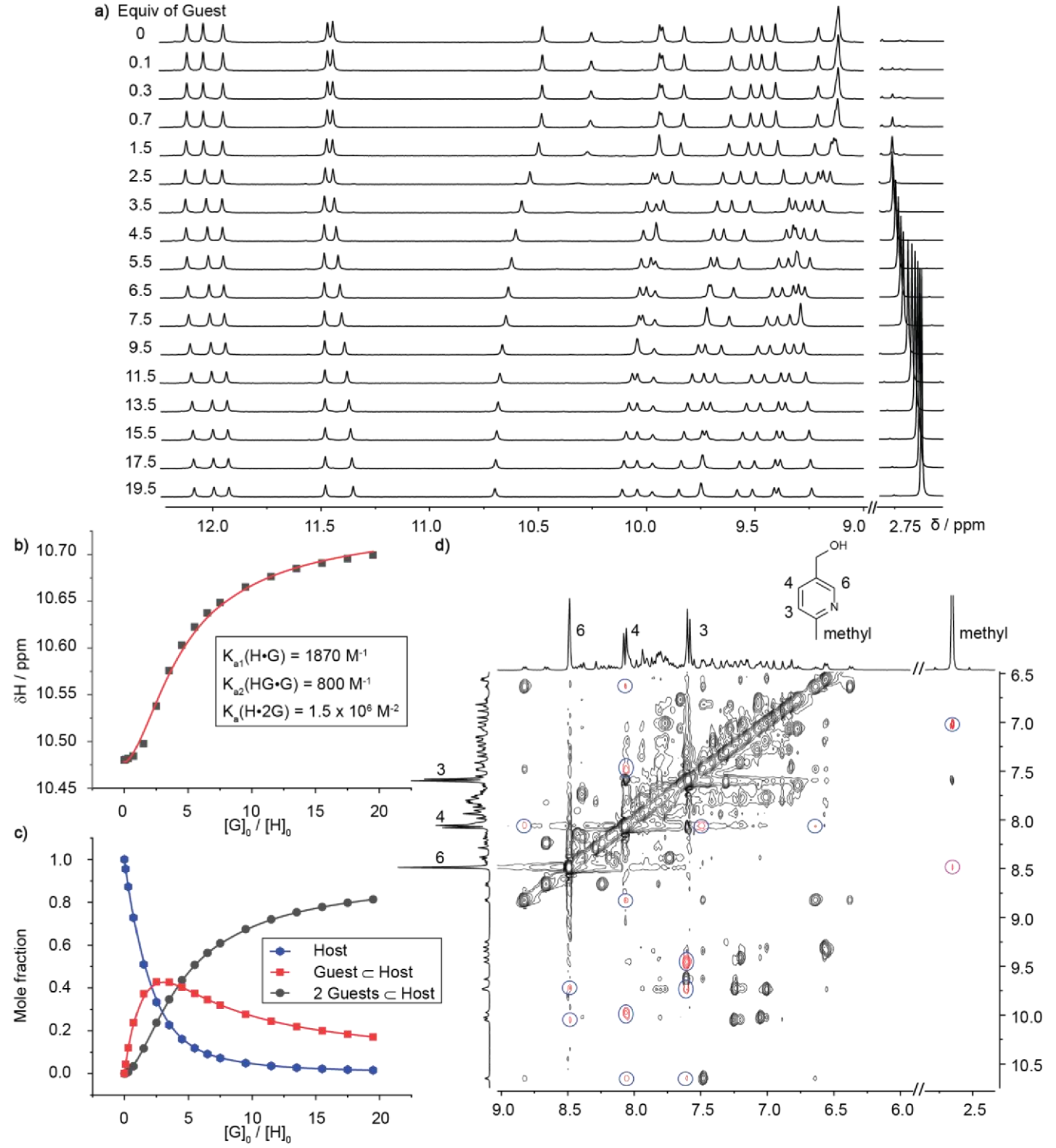


Fig. 7 a) ^{1}H NMR spectra (500 MHz) of **3** (0.45 mM in H_2O/D_2O 90:10 *v/v* at 25°C) titrated with 5-hydroxymethyl-2-methylpyridine. b) Plot of the chemical shift variation of a selected amide proton during the titration. The horizontal scale corresponds to the guest/host molar ratio. The red trace shows the curve fitting using a 1:2 host–guest binding isotherm. c) Calculated changes in mole fraction of compound **3** (blue trace), 1:1 complex (red trace) and 1:2 complex (black trace) as a function of the guest/host molar ratio. d) ^{1}H–^{1}H NOESY map of a sample containing compound **3** (1 mM) and 5-hydroxymethyl-2-methylpyridine (9 mM) in H_2O/D_2O (90:10 v/v). Some intermolecular NOE correlations are highlighted in red. Host-guest and guest-guest NOEs are pointed out by blue and pink circles, respectively. Guest peak labels are shown as numbers on the structural formula of the guest.

Binding selectivity, kinetics and thermodynamics

The fast exchange on the NMR time scale at room temperature between bound and unbound states highlighted for **3** and 5-hydroxymethyl-2-methylpyridine (Fig. 7a) is representative of the titration data recorded with **1**-**5**. The only exception concerns a hydrophobic guest, 1,1,2,2,3,3,3,3-heptafluorocyclopentane, whose association with **2** gives rises to a second set of ^{1}H NMR signals (Fig. 9a,b). Slow exchange kinetics can be reached with other guests upon cooling as

shown in Fig. 9c,d for **2** and 4-pyridinemethanol. Nevertheless, the fast exchange at room temperature contrast with the slow exchange generally observed for the binding of polar guests by organic soluble capsules, regardless of whether K_a values are large or not.[42] As mentioned in the section on design and synthesis, the folding of Qh is largely driven by solvophobic effects in contrast with most other units used previously that fold in any solvent. This may be conducive of greater conformational dynamics responsible for faster guest binding and release, but this hypothesis has not yet been tested.

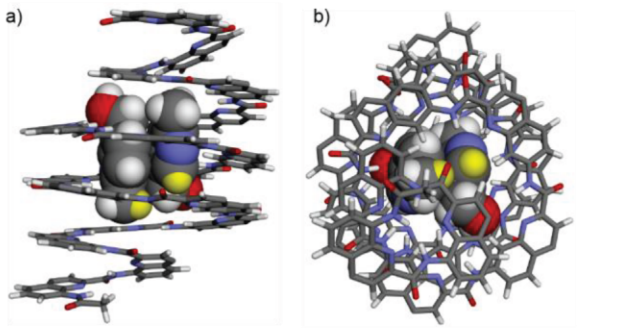


Fig.8 Computational modeling of the complex between compound **3** and 5-hydroxymethyl-2-methylpyridine. a) side view and b) bottom view of energy minimized models (MacroModel, MMFFs) of 1:2 host–guest binding model indicating position 6 proton and methyl proton proximity (target protons are highlighted in yellow). Side chains have been omitted for clarity.

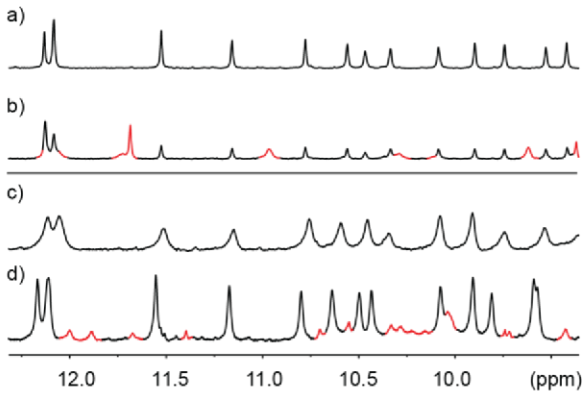


Fig.9 a) Excerpts of the ^{1}H NMR (500 MHz) at 298 K of compound **2** (0.5 mM) in H_2O/D_2O 90:10 (v/v). b) Excerpts of the ^{1}H NMR after compound **2** mixed with 1,1,2,2,3,3,4-Heptafluoro cyclopentane and then heated to 50°C for 30 mins. Excerpts of the ^{1}H NMR (500 MHz) at 298 K of compound **2** (1 mM) mixed with 4-pyridinemethanol (2 mM) in H_2O/D_2O 90:10 (v/v) at c) 25°C and d) 1°C. The signals of new species are highlighted in red.

A general feature of the guests for which binding was detected is that they are neither hydrophilic nor hydrophobic, *i.e.* their water/octanol partition coefficient (log *P*) is close to zero (Table 1). Exceptions are few and include two hydrophobic guests for **1**, namely pentafluorophenol‖ and 1,1,2,2,3,3,4-heptafluorocyclopentane, and one hydrophilic guest for **5**, namely cis-3,6-bis(2-hydroxyethyl) piperazine-2,5-dione. Heptafluorocyclopentane is actually so hydrophobic that no K_a could be determined because its concentration in the aqueous phase was unknown (it tends to phase separate). One may also note that hydrogen bonding to fluorine may contribute to the binding of the hydrophobic guests.[59] It is interesting to note that host-guest recognition in water has been largely documented for hydrophobic guests and, less frequently, for hydrophilic neutral organic guests,[60,61] but that guests that are neither hydrophilic nor hydrophobic have rarely been considered. Representative examples include 4-pyridinemethanol for **1** and 2,2'-bipyridine]-4,4'-diyldimethanol for **4**, as well as other hydroxymethyl-substituted pyridine guests (Table 1).

That pyridinemethanol-based guests bound better than more hydrophobic molecules is quite remarkable, and possibly a reflection of the fact that binding cavities are also not truly hydrophobic (Fig. 4). Natural receptor cavities are also often adorned by both polar and apolar patches. Binding was found to be sometimes quite selective to the guest structure. For example, **1** binds to 4-pyridinemethanol but not to benzyl alcohol and to pentafluorophenol‖ but not to hexafluorobenzene; **3** binds to 5-hydroxymethyl-2-methylpyridine but not to its trifluoromethyl analogue and also to 2,5-bis(hydroxymethyl)-pyridine but not to its benzenic analogue. This level of selectivity suggests that some attractive or repulsive directional interactions play an important role that may, in some cases, be rationalized based on models (Fig. S18-S20). Affinity cliffs were generally observed when trying to fit a guest too large for a given binding cavity. As highlighted in Table 1, many host-guest complex binding data were found to fit to a 1:2 binding isotherm with either slightly positive or slightly negative cooperativity. We assumed the binding mode involves the stacking of two guest molecules in the host cavity as proposed for complex between compound **3** and 5-hydroxymethyl-2-methylpyridine (Fig. 8). For chiral guests, *e.g.* *R*-(+)-α-methyl-4-pyridinemethanol for **1** and cis-3,6-bis(2-hydroxyethyl) piperazine-2,5-dione, we observed no induced CD as a reflection of an expectedly high barrier of helix handedness inversion in water.[62]

Binding constants ranged from 270 to 45300 M^{-1}, with the exception of the most polar guest cis-3,6-bis(2-hydroxyethyl) piperazine-2,5-dione which showed weaker binding. These values compare well with other host-guest systems involving neutral non-hydrophobic guests in water. To determine the thermodynamic parameters of binding, some titrations were repeated using ITC (isothermal titration calorimetry). Raw titrations data, *i.e.* before subtracting the blank, are shown in Fig. 10a,c and Figs S21 and S22. Thus, **1** was titrated with 4-pyridinemethanol and the data fitted well to a 1:2 binding model yielding K_a values of 39800 M^{-1} and 1080 M^{-1} that matched well with those calculated from the NMR titration (Fig. 10a,b). For the first binding event, a weak favorable binding enthalpy (ΔH_1) of -5.7 kJ.mol^{-1} was calculated. The binding entropy ($-T\Delta S_1$) was also favorable -20.5 kJ.mol^{-1} and appeared to represent the main driving force, presumably due to the release of ordered water from the host cavity (Fig. 5). The binding of the second guest was found to be propelled by an enormously advantageous change in entropy ($-T\Delta S_2$ = -94.6 kJ.mol^{-1}) that overcame a surprising enthalpic penalty of 77.4 kJ.mol^{-1}. The favorable entropy may again be assigned to the release of water molecules and is so strong that it allows for the formation of an inherently unstable complex.

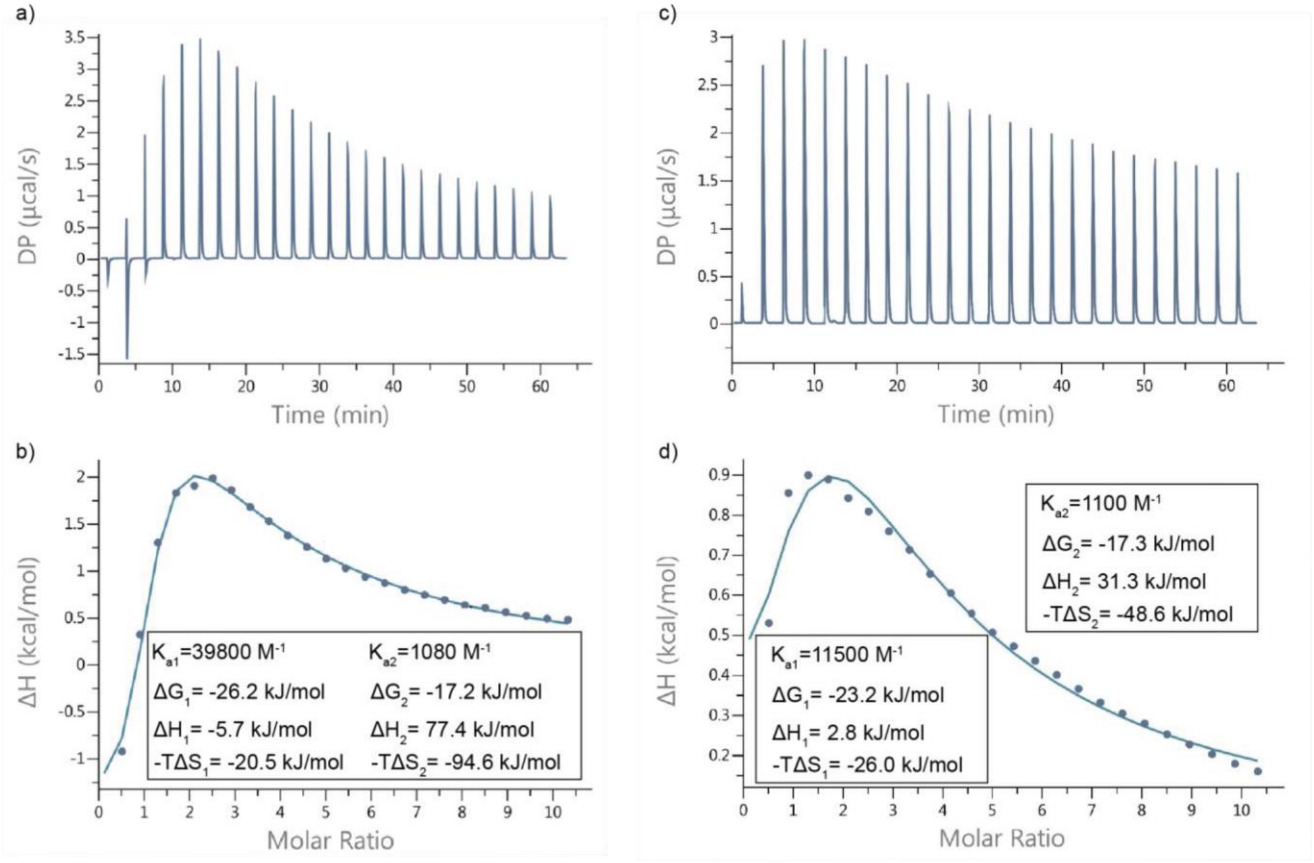


Fig. 10 ITC data for the titration of a, b) compound **1** (0.2 mM) with 4-pyridinemethanol and c, d) compound **4** (0.2 mM) with [2,2'-Bipyridine]-5,5'-diyldimethanol at 25 °C in H_2O. The solid blue line represents the best-fitting curve obtained assuming a 1:2 host–guest sequential binding model.

A titration of **4** with [2,2'-bipyridine]-5,5'-diyldimethanol revealed a similar trend (Fig. 10c,d). Both **4** with [2,2'-bipyridine]-5,5'-diyldimethanol are structurally related to, yet different from, **1** and 4-pyridinemethanol, respectively. *K*a values of 11500 M^{-1} and 1100 M^{-1} were calculated for the first and second binding events, respectively. This time the first binding event was slightly enthalpically disfavored (ΔH_1 = 2.8 kJ.mol^{-1}) and driven by entropy ($-T\Delta S_1$ = -26.0 kJ.mol^{-1}) while the second binding event was again strongly enthalpically disfavored (ΔH_2 = 31.3 kJ.mol^{-1}) and driven by entropy ($-T\Delta S_2$ = -48.6 kJ.mol^{-1}). An entropically driven association may in principle be further favored upon increasing temperature. We thus performed a second ITC titration of **1** with 4-pyridinemethanol at 40°C instead of 25°C (Fig. S21). However, instead of enhanced entropic effects, we observed a strong entropy-enthalpy compensation. In particular, the second binding event became even more enthalpically unfavorable (ΔH_2 = 99 kJ.mol^{-1} instead of 77.4 kJ.mol^{-1} at 25°C) leading to a slight drop in K_{a2} despite the larger entropic term. Finally, also performed an ITC titration of **5** with hydrophilic guest cis-3,6-bis(2-hydroxyethyl) piperazine-2,5-dione (Fig. S22). The data fitted with a 1:1 binding model yielding a K_a value of 11 M^{-1} of the same order of magnitude to that measured by NMR (26 M^{-1}). The enthalpic term was this time negative but both entropic and enthalpic terms were small and not easily amenable to an interpretation.

As pointed in the introduction, endothermic molecular recognition has rarely been reported.[1,24,25,27] It typically entails a entropic force that is so strong that it allows binding even when it sums up to an unfavorable balance between attractions and repulsions. The origin of the positive enthalpy in the two cases described above is difficult to identify. Are the interactions between the host and guest not sufficiently attractive in comparison to their interactions with water molecules? Does binding cause an unfavorable deformation of the host structure that went unnoticed in our molecular models? In any case, it is hard to invoke a high enthalpic cost of desolvation as was done in the case of anion binding. In contrast, the strong entropic drive can reasonably be assigned to the release of water molecules that we know are entrapped in the hosts' cavities (Fig. 5).

Conclusion

Helical aromatic foldamer capsules represent attractive molecular containers because they completely surround their guests, allowing for potential host-guest interactions to occur from all directions as well as seclusion from the solvent, and because of their inherently modular sequence structure that accelerates iterative structure-based design.[33,41,53] These properties had until now been exploited in organic solvents only. Here, we have introduced the first examples of water

soluble foldamer capsules displaying molecular recognition properties in water. We identified a peculiar privileged group of guests derived from hydroxymethyl-pyridines that are neither hydrophilic nor lipophilic. Binding in the low micromolar range was observed for the best system. Remarkably, binding was found to be endothermic in several instances. It thus relies on strongly favorable entropic terms assigned to the release of ordered water molecules from the binding cavity. The positive enthalpies of association also tell that there is much room from improving binding affinities upon introducing more favorable host-guest interactions. Molecular recognition in water may find applications in molecular separation,[63,64] and provide reactive or protective environments to guest molecules [65,66] The 1:2 associations that we observed also hint at the possible promotion of intermolecular reactions.[67] Overall, our results provide a foundation for such future developments using aromatic foldamer-based molecular recognition in water.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by the DFG (project HU 1766/6-1). B. Teng gratefully acknowledge the financial support from the China Scholarship Council.

Footnotes and References

‡ Another example is that of the recognition of adamantyl carboxylate by a hydrophobic receptor in presence of certain ions, but the enthalpy then reflected guest exchange rather than guest binding. See ref. [68].

§ Individual chemical functions that have this property have been termed “hydroneutral”. See ref. [69].

‖ For pentafluorophenol and only for this guest, the pH of the host solution was adjusted to 4 with acetic acid buffer, see Fig. S6.

1 S. Kubik, *ChemistryOpen*, 2022, **11**, e202200028.
2 N. K. Pinkin and M. L. Waters, *Org. Biomol. Chem.*, 2014, **12**, 7059–7067.
3 F. Guagnini, P. M. Antonik, M. L. Rennie, P. O’Byrne, A. R. Khan, R. Pinalli, E. Dalcanale and P. B. Crowley, *Angew. Chem. Int. Ed.*, 2018, **57**, 7126–7130.
4 A. Höing, A. Kirupakaran, C. Beuck, M. Pörschke, F. C. Niemeyer, T. Seiler, L. Hartmann, P. Bayer, T. Schrader and S. K. Knauer, *Biomacromolecules*, 2022, **23**, 4504–4518.
5 M. Vladimir and Y. Anatoly, Artificial Receptors for Chemical Sensors, WILEY-VCH, New Jersey, 2010.
6 K. Worm and F. P. Schmidtchen, *Angew. Chem. Int. Ed.*, 1995, **34**, 65–66.
7 A. Bianchi, M. Micheloni, P. Orioli, P. Paoletti and S. Mangani, *Inorganica Chim. Acta*, 1988, **146**, 153–159.
8 X. Lu, S. A. Zebaze Ndendjio, P. Y. Zavalij and L. Isaacs, *Org. Lett.*, 2020, **22**, 4833–4837.
9 W. Xue, P. Y. Zavalij and L. Isaacs, *Angew. Chem. Int. Ed.*, 2020, **59**, 13313–13319.
10 E. Jeamet, J. Septavaux, A. Héloin, M. Donnier-Maréchal, M. Dumartin, B. Ourri, P. Mandal, I. Huc, E. Bignon, E. Dumont, C. Morell, J.-P. Francoia, F. Perret, L. Vial and J. Leclaire, *Chem. Sci.*, 2018, **10**, 277–283.
11 R. M. Izatt, J. S. Bradshaw, S. A. Nielsen, J. D. Lamb, J. J. Christensen and D. Sen, *Chem. Rev.*, 1985, **85**, 271–339.
12 M. V. Rekharsky, T. Mori, C. Yang, Y. H. Ko, N. Selvapalam, H. Kim, D. Sobransingh, A. E. Kaifer, S. Liu, L. Isaacs, W. Chen, S. Moghaddam, M. K. Gilson, K. Kim and Y. Inoue, *Proc. Natl. Acad. Sci. U.S.A.*, 2007, **104**, 20737–20742.
13 D. A. Stauffer, R. E. Jr. Barrans and D. A. Dougherty, *J. Org. Chem.*, 1990, **55**, 2762–2767.
14 W. Liu, Y. Tan, L. O. Jones, B. Song, Q.-H. Guo, L. Zhang, Y. Qiu, Y. Feng, X.-Y. Chen, G. C. Schatz and J. F. Stoddart, *J. Am. Chem. Soc.*, 2021, **143**, 15688–15700.
15 M. Lisbjerg, B. M. Jessen, B. Rasmussen, B. E. Nielsen, A. Ø. Madsen and M. Pittelkow, *Chem. Sci.*, 2014, **5**, 2647–2650.
16 M. Lisbjerg, B. E. Nielsen, B. O. Milhøj, S. P. A. Sauer and M. Pittelkow, *Org. Biomol. Chem.*, 2014, **13**, 369–373.
17 G. Peñuelas-Haro and P. Ballester, *Chem. Sci.*, 2019, **10**, 2413–2423.
18 M. J. Langton, S. W. Robinson, I. Marques, V. Félix and P. D. Beer, *Nature Chem*, 2014, **6**, 1039–1043.
19 A. Borissov, I. Marques, J. Y. C. Lim, V. Félix, M. D. Smith and P. D. Beer, *J. Am. Chem. Soc.*, 2019, **141**, 4119–4129.
20 D. B. Smithrud, E. M. Sanford, I. Chao, S. B. Ferguson, D. R. Carcanague, J. D. Evanseck, K. N. Houk and F. Diederich, *Pure Appl. Chem.*, 1990, **62**, 2227–2236.
21 F. Biedermann, W. M. Nau and H.-J. Schneider, *Angew. Chem. Int. Ed.*, 2014, **53**, 11158–11171.
22 F. Biedermann, V. D. Uzunova, O. A. Scherman, W. M. Nau and A. De Simone, *J. Am. Chem. Soc.*, 2012, **134**, 15318–15323.
23 M. B. Hillyer and B. C. Gibb, *Annu Rev Phys Chem*, 2016, **67**, 307–329.
24 M. V. Rekharsky and Y. Inoue, *Chem. Rev.*, 1998, **98**, 1875–1918.
25 F. Sommer, Y. Marcus and S. Kubik, *ACS Omega*, 2017, **2**, 3669–3680.
26 Y. Liu, Y.-M. Zhang, S.-X. Sun, Y.-M. Li and R.-T. Chen, *J. Chem. Soc., Perkin Trans. 2*, 1997, 1609–1614.
27 M. Boiocchi, M. Bonizzoni, C. Ciarrocchi, L. Fabbrizzi, M. Invernici and M. Licchelli, *Chem. Eur. J.*, 2018, **24**, 5659–5666.
28 A. Tanatani, T. S. Hughes and J. S. Moore, *Angew. Chem. Int. Ed.*, 2002, **41**, 325–328.
29 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.
30 J. Y. Hwang, H.-G. Jeon, Y. R. Choi, J. Kim, P. Kang, S. Lee and K.-S. Jeong, *Org. Lett.*, 2017, **19**, 5625–5628.
31 J. Garric, J.-M. Léger and I. Huc, *Angew. Chem. Int. Ed.*, 2005, **44**, 1954–1958.
32 Q. Gan, Y. Ferrand, C. Bao, B. Kauffmann, A. Grélard, H. Jiang and I. Huc, *Science*, 2011, **331**, 1172–1175.
33 N. Chandramouli, Y. Ferrand, G. Lautrette, B. Kauffmann, C. D. Mackereth, M. Laguerre, D. Dubreuil and I. Huc, *Nature Chem*, 2015, **7**, 334–341.
34 Y. Ferrand, N. Chandramouli, A. M. Kendhale, C. Aube, B. Kauffmann, A. Grélard, M. Laguerre, D. Dubreuil and I. Huc, *J. Am. Chem. Soc.*, 2012, **134**, 11282–11288.
35 S. B. Seo, S. Lee, H.-G. Jeon and K.-S. Jeong, *Angew. Chem. Int. Ed.*, 2020, **59**, 10441–10445.

36 Y. Liu, F. C. Parks, W. Zhao and A. H. Flood, *J. Am. Chem. Soc.*, 2018, **140**, 15477–15486.
37 W. Wang, C. Zhang, S. Qi, X. Deng, B. Yang, J. Liu and Z. Dong, *J. Org. Chem.*, 2018, **83**, 1898–1902.
38 E. A. John, C. J. Massena and O. B. Berryman, *Chem. Rev.*, 2020, **120**, 2759–2782.
39 Y. Zhong, T. A. Sobiech, B. Kauffmann, B. Song, X. Li, Y. Ferrand, I. Huc and B. Gong, *Chem. Sci.*,2023, DOI:10.1039/D3SC00524K.
40 H. Juwarker and K.-S. Jeong, *Chem. Soc. Rev.*, 2010, **39**, 3664–3674.
41 Y. Ferrand and I. Huc, *Acc. Chem. Res.*, 2018, **51**, 970–977.
42 V. Koehler, A. Roy, I. Huc and Y. Ferrand, *Acc. Chem. Res.*, 2022, **55**, 1074–1085.
43 F. C. Parks, Y. Liu, S. Debnath, S. R. Stutsman, K. Raghavachari and A. H. Flood, *J. Am. Chem. Soc.*, 2018, **140**, 17711–17723.
44 M. T. Stone and J. S. Moore, *Org. Lett.*, 2004, **6**, 469–472.
45 N. Chandramouli, M. F. El-Behairy, G. Lautrette, Y. Ferrand and I. Huc, *Org. Biomol. Chem.*, 2016, **14**, 2466–2472.
46 Q. Wang, Y. Zhong, D. P. Miller, X. Lu, Q. Tang, Z.-L. Lu, E. Zurek, R. Liu and B. Gong, *J. Am. Chem. Soc.*, 2020, **142**, 2915–2924.
47 M. L. Singleton, G. Pirotte, B. Kauffmann, Y. Ferrand and I. Huc, *Angew. Chem. Int. Ed.*, 2014, **53**, 13140–13144.
48 B. Teng, J. Atcher, L. Allmendinger, C. Douat, Y. Ferrand and I. Huc, *Org. Biomol. Chem.*, 2023, **21**, 3525–3530.
49 B. Teng, P. K. Mandal, L. Allmendinger, C. Douat, Y. Ferrand and I. Huc, *Chem. Sci.*, submitted.
50 Maestro, Schrödinger, LLC, New York, NY, 2021.
51 B. Baptiste, C. Douat-Casassus, K. Laxmi-Reddy, F. Godde and I. Huc, *J. Org. Chem.*, 2010, **75**, 7175–7185.
52 V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat and I. Huc, *Chem. Eur. J.*, 2023, e202300898.
53 G. Lautrette, B. Wicher, B. Kauffmann, Y. Ferrand and I. Huc, *J. Am. Chem. Soc.*, 2016, **138**, 10314–10322.
54 S. Kubik, R. Goddard, R. Kirchner, D. Nolting and J. Seidel, *Angew. Chem. Int. Ed.*, 2001, **40**, 2648–2651.
55 S. Kubik and R. Goddard, *Proc. Natl. Acad. Sci. U.S.A.*, 2002, **99**, 5127–5132.
56 K. Choi and A. D. Hamilton, *J. Am. Chem. Soc.*, 2001, **123**, 2456–2457.
57 C. Bao, B. Kauffmann, Q. Gan, K. Srinivas, H. Jiang and I. Huc, *Angew. Chem. Int. Ed.*, 2008, **47**, 4153–4156.
58 A. J. R. Heck, *Nat Methods*, 2008, **5**, 927–933.
59 C. Dalvit, C. Invernizzi and A. Vulpetti, *Chem. Eur. J.*, 2014, **20**, 11058–11068.
60 G.-B. Huang, S.-H. Wang, H. Ke, L.-P. Yang and W. Jiang, *J. Am. Chem. Soc.*, 2016, **138**, 14550–14553.
61 Y. Ferrand, M. P. Crump and A. P. Davis, *Science*, 2007, **318**, 619–622.
62 T. Qi, V. Maurizot, H. Noguchi, T. Charoenraks, B. Kauffmann, M. Takafuji, H. Ihara and I. Huc, *Chem. Commun.*, 2012, **48**, 6337–6339.
63 D. Zhang, T. K. Ronson, Y.-Q. Zou and J. R. Nitschke, *Nat Rev Chem*, 2021, **5**, 168–182.
64 T. Ogoshi, K. Saito, R. Sueto, R. Kojima, Y. Hamada, S. Akine, A. M. P. Moeljadi, H. Hirao, T. Kakuta and T. Yamagishi, *Angew. Chem. Int. Ed.*, 2018, **57**, 1592–1595.
65 A. Das, I. Mandal, R. Venkatramani and J. Dasgupta, *Sci Adv*, 2019, **5**, eaav4806.
66 Z. Lin, J. Sun, B. Efremovska and R. Warmuth, *Chem. Eur. J.*, 2012, **18**, 12864–12872.
67 T. Murase and M. Fujita, *Chem. Rec.*, 2010, **10**, 342–347.
68 P. Sokkalingam, J. Shraberg, S. W. Rick and B. C. Gibb, *J. Am. Chem. Soc.*, 2016, **138**, 48–51.
69 N. Sagawa and T. Shikata, *Phys. Chem. Chem. Phys.*, 2014, **16**, 13262–13270.

Supplementary Information

Endothermic molecular recognition within helical aromatic oligoamide capsules in water

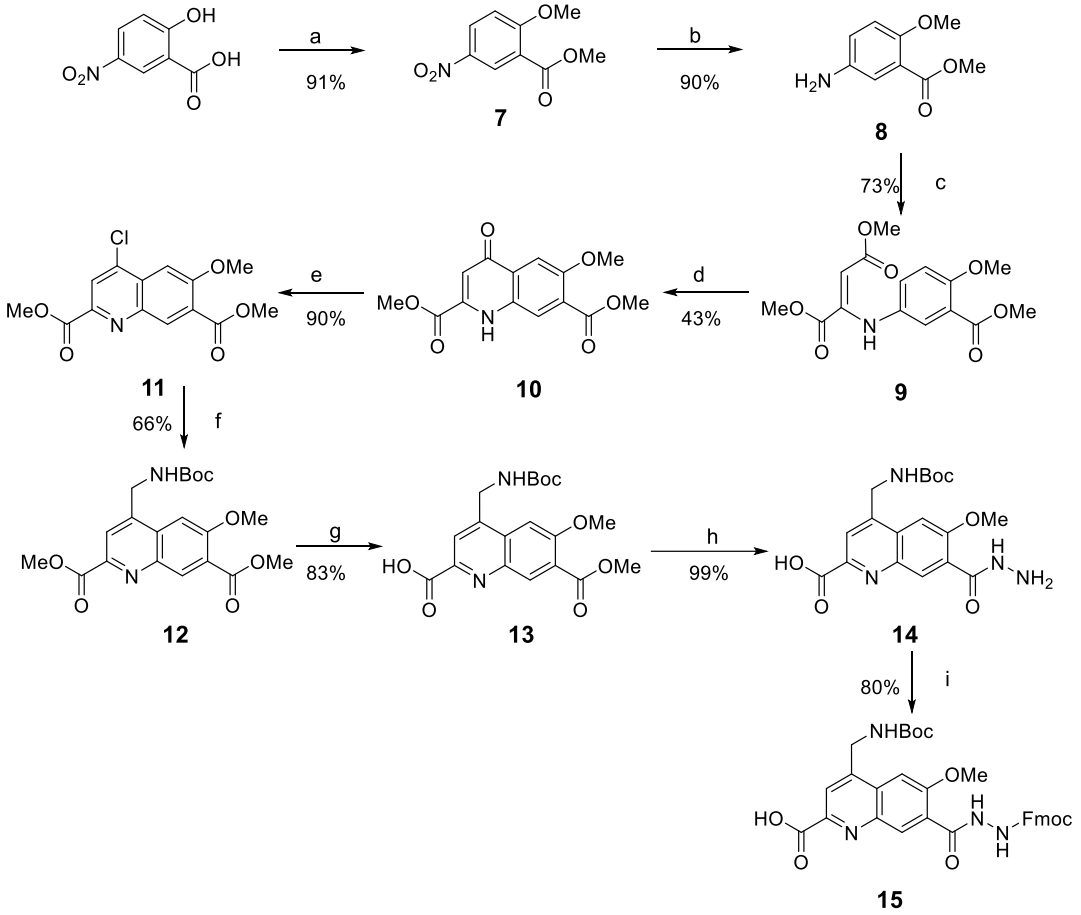
Binhao Teng, [a] Lars Allmendinger, [a] Céline Douat, [a] Yann Ferrand [b] and Ivan Huc* [a]

[a.] Department of Pharmacy, Ludwig-Maximilians-Universität, Butenandtstr. 5–13, 81377, München, Germany.

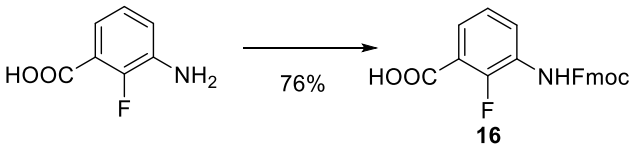
[b.] Univ. Bordeaux, CNRS, Bordeaux Institut National Polytechnique, CBMN UMR 5248, 2 rue Escarpit, 33600 Pessac, France.

Table of contents

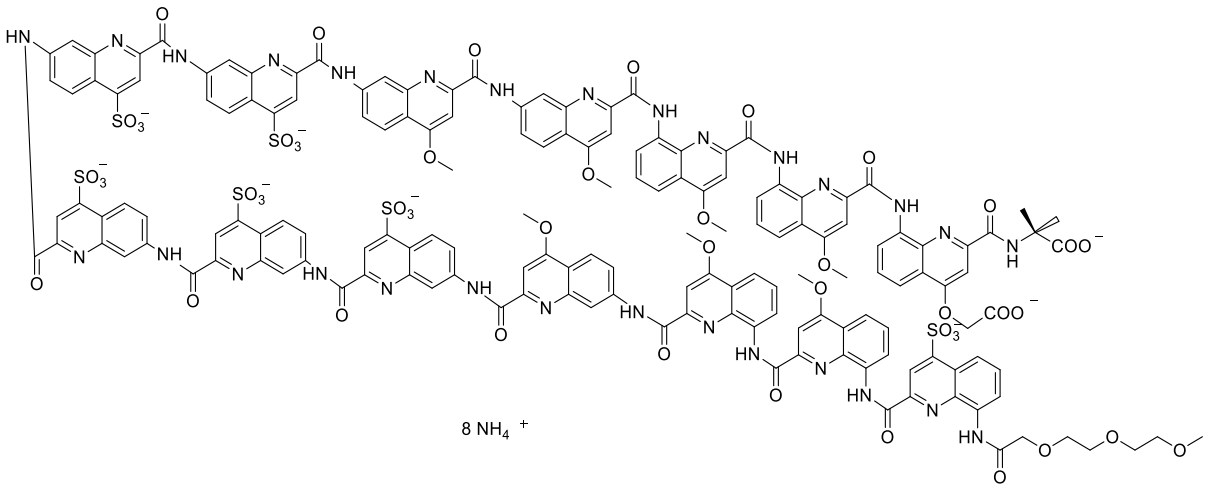
1 Supplementary figures



Scheme S1. Synthesis of H monomer **15**: a) K_2CO_3, MeI, DMF; b) Fe, CH_3COOH, MeOH, 65 °C; c) dimethyl acetylenedicarboxylate, MeOH, reflux; d) PhOPh, reflux; e) $POCl_3$, DMF; f) KF_3BCH_2NHBoc, $Pd(OAc)_2$, SPhos, K_2CO_3, toluene, H_2O, 85 °C g) NaOH, DCM, MeOH, 0 °C to rt.; h) NH_2NH_2, MeOH; i) Fmoc-Cl, dioxane, H_2O, 0 °C.



Scheme S2. Synthesis of F monomer **16**: Fmoc-Cl, $NaHCO_3$, dioxane, H_2O, 0 °C to r.t..



6

Figure S40. Chemical structural formula of double helix capsule crystal **6**.

NMR titration spectra and plots

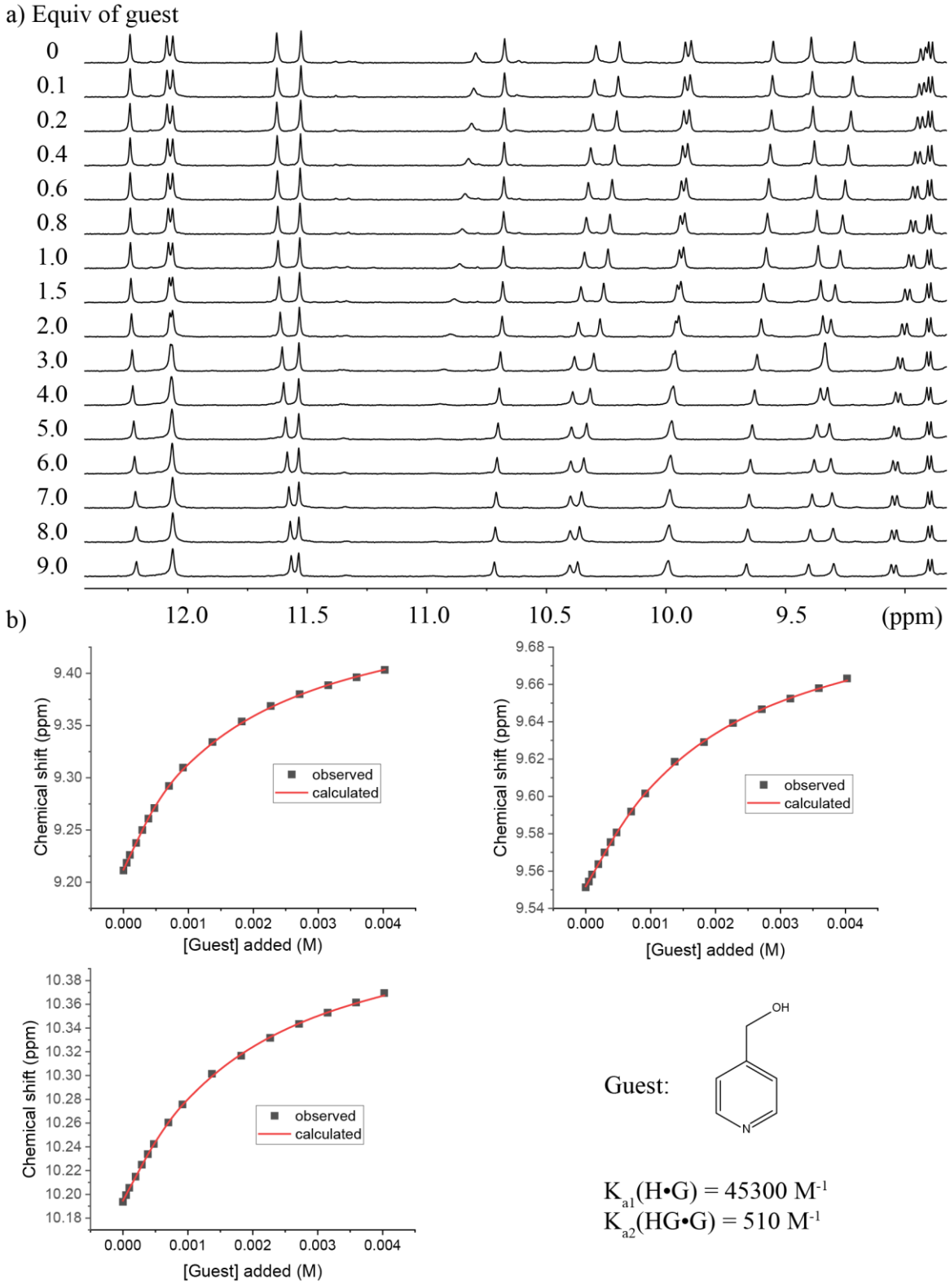


Figure S41. a) ^{1}H NMR spectra (500 MHz, H_2O/D_2O 90:10 (*v*/*v*), 25 °C) of oligomer **1** (0.5 mM) titrated with 4-pyridinemethanol. b) Plots showing chemical shift changes of 3 different amide protons upon addition of 4-pyridinemethanol. Red trace represents curve fitting using a 1:2 host−guest binding model. The limiting chemical shifts for the selected protons are 9.48 ppm, 9.71 ppm, and 10.44 ppm, respectively. The final constant value is an average of the three fit curve calculation.

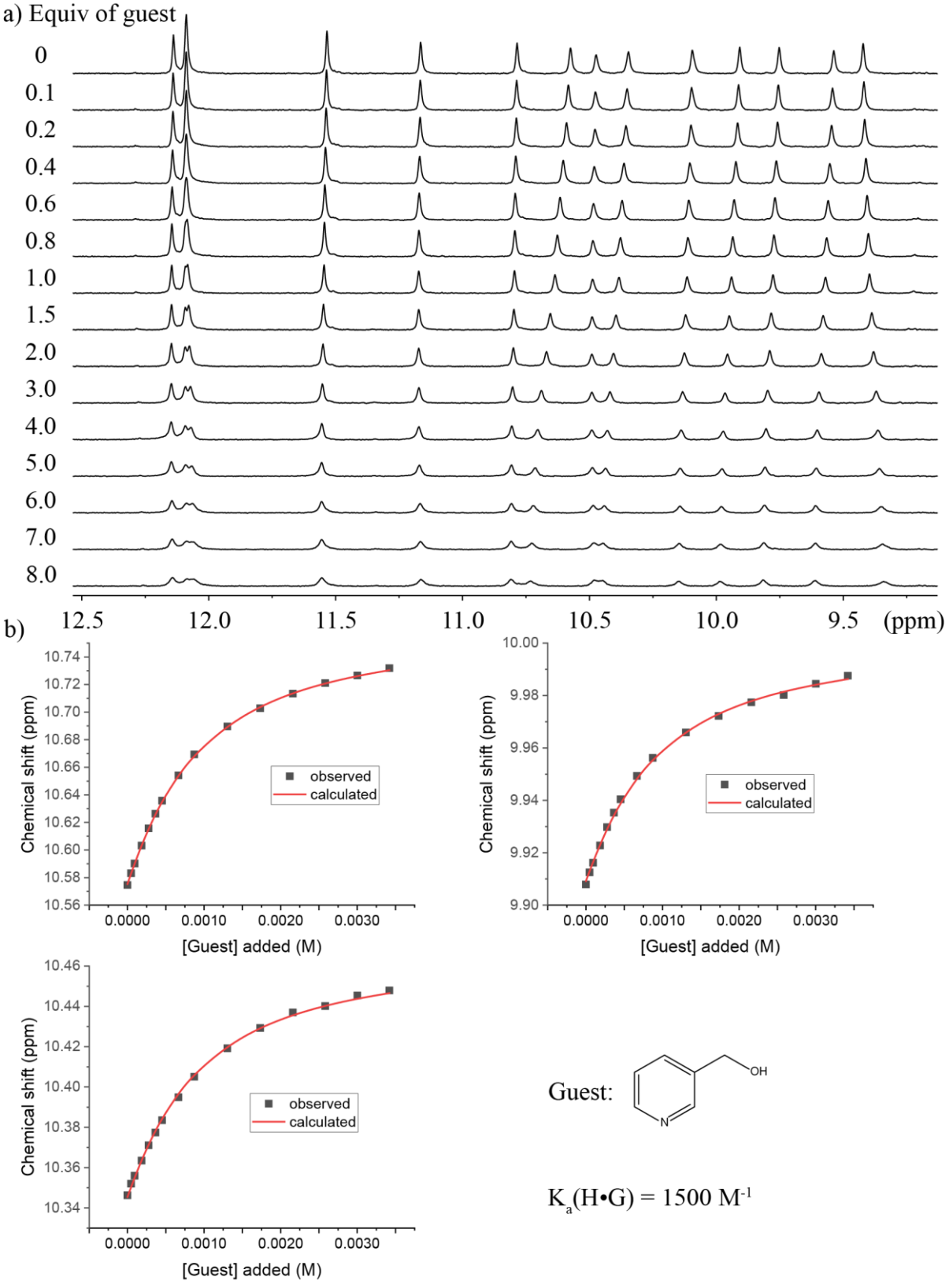


Figure S42. a) ^{1}H NMR spectra (500 MHz, H_2O/D_2O 90:10 (*v*/*v*), 25 °C) of oligomer **2** (0.5 mM) titrated with 3-pyridinemethanol. b) Plots showing chemical shift changes of 3 different amide protons upon addition of 3-pyridinemethanol. Red trace represents curve fitting using a 1:1 host−guest binding model. The limiting chemical shifts for the selected protons are 10.76 ppm, 10.00 ppm, and 10.47 ppm, respectively. The final constant value is an average of the three fit curve calculation.

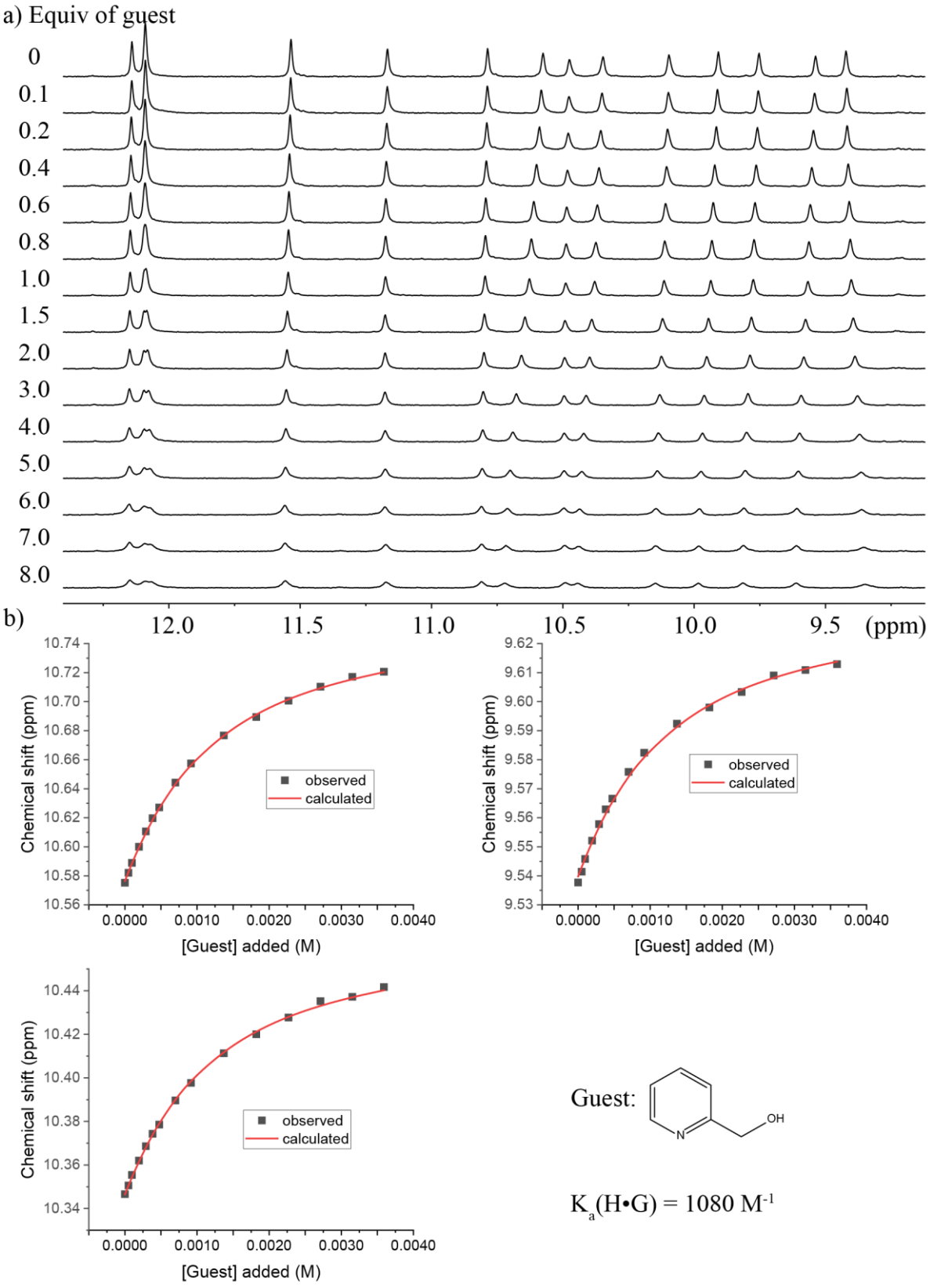


Figure S43. a) ^{1}H NMR spectra (500 MHz, H_2O/D_2O 90:10 (*v*/*v*), 25 °C) of oligomer **2** (0.5 mM) titrated with 2-pyridinemethanol. b) Plots showing chemical shift changes of 3 different amide protons upon addition of 2-pyridinemethanol. Red trace represents curve fitting using a 1:1 host−guest binding model. The limiting chemical shifts for the selected protons are 10.76 ppm, 9.63 ppm, and 10.47 ppm, respectively. The final constant value is an average of the three fit curve calculation.

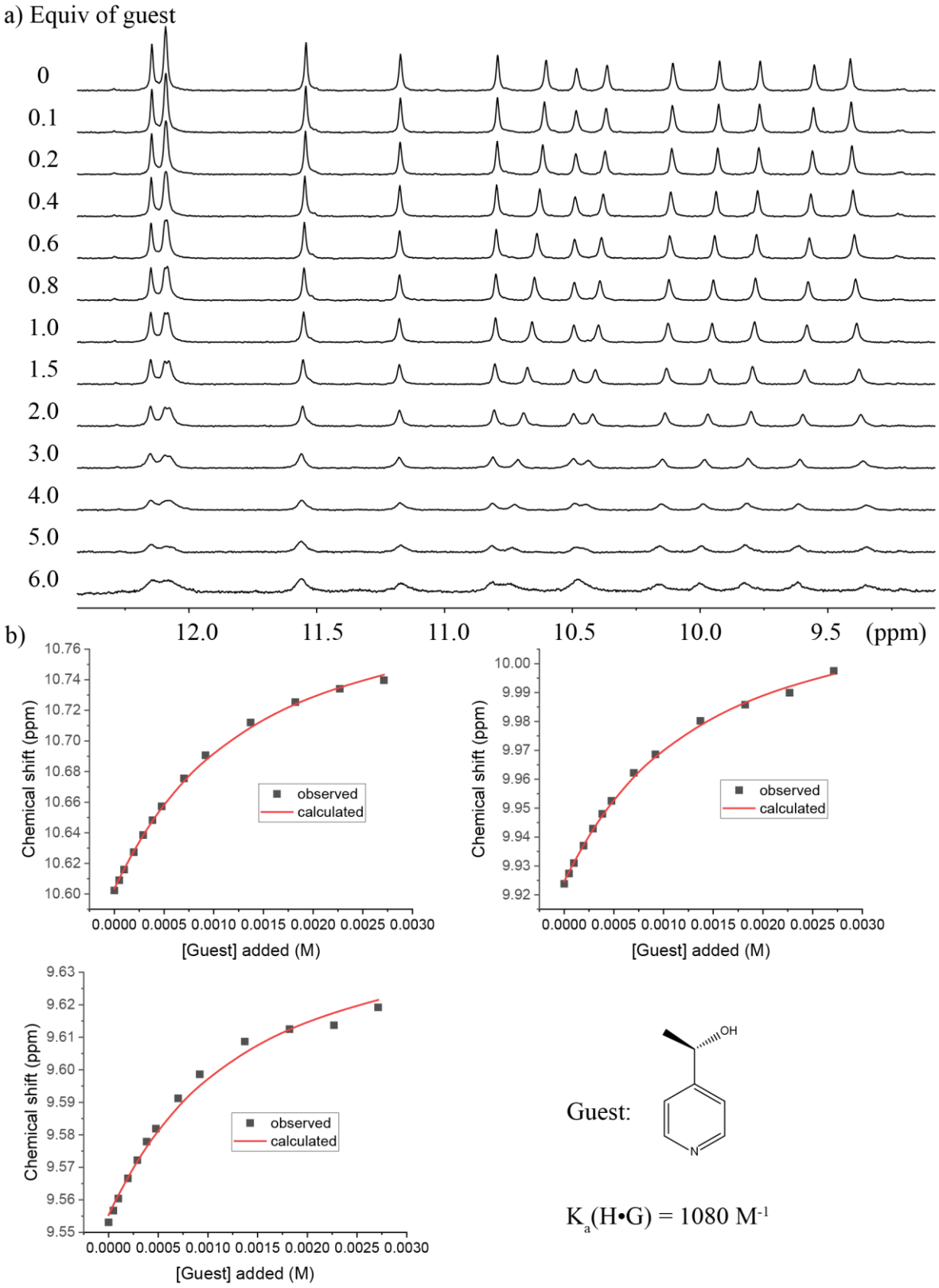


Figure S44. a) ^{1}H NMR spectra (500 MHz, H_2O/D_2O 90:10 (*v*/*v*), 25 °C) of oligomer **2** (0.5 mM) titrated with (*R*)-(+)-α-methyl-4-pyridinemethanol. b) Plots showing chemical shift changes of 3 different amide protons upon addition of (*R*)-(+)-α-methyl-4-pyridinemethanol. Red trace represents curve fitting using a 1:1 host−guest binding model. The limiting chemical shifts for the selected protons are 10.80 ppm, 10.02 ppm, and 9.65 ppm, respectively. The final constant value is an average of the three fit curve calculation.

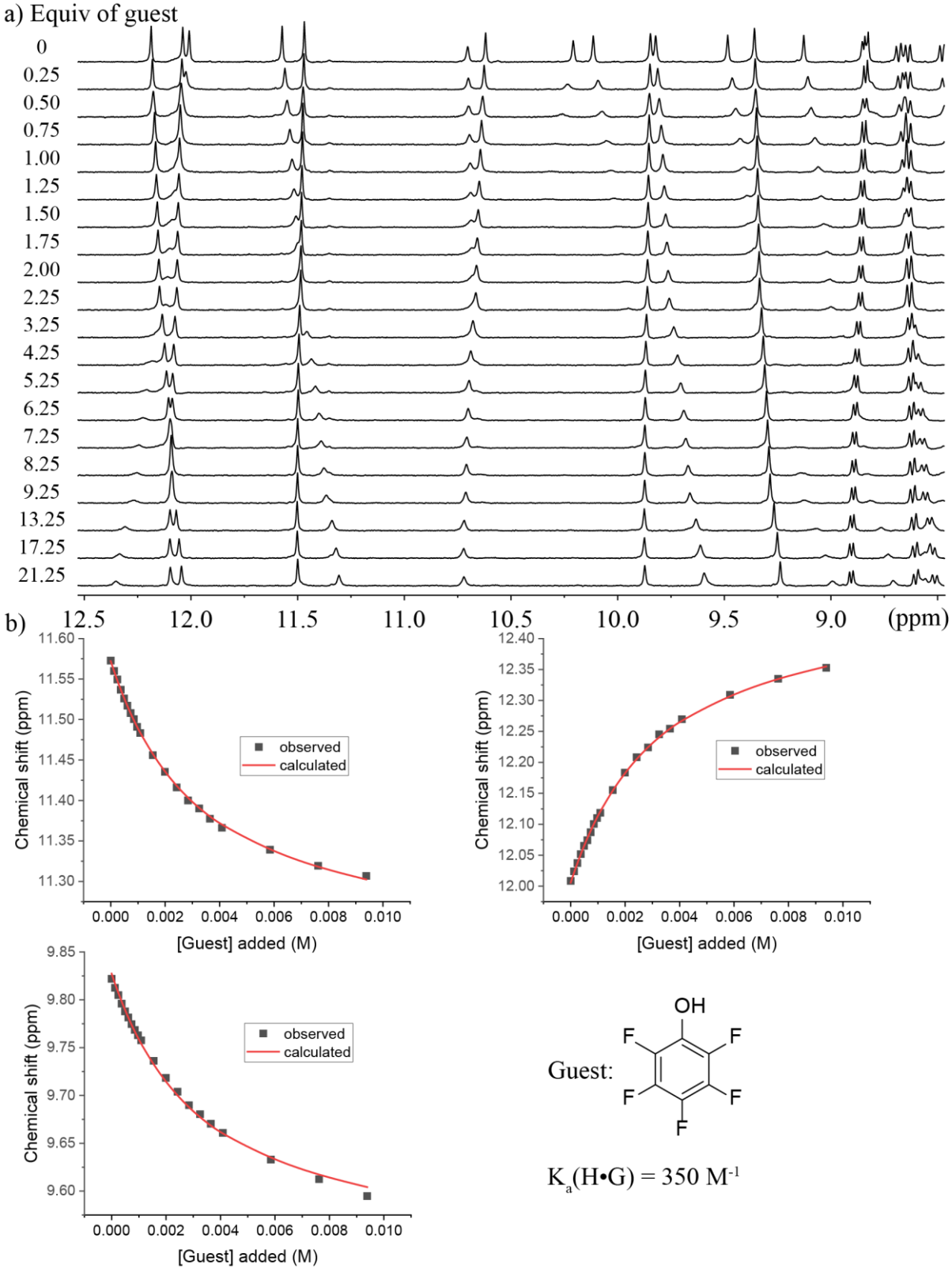


Figure S45. a) ^{1}H NMR spectra (500 MHz, 15 mM acetic acid buffer (pH = 4.0)/D_2O 90:10 (*v*/*v*), 25 °C) of oligomer **1** (0.5 mM) titrated with pentafluorophenol. b) Plots showing chemical shift changes of 3 different amide protons upon addition of pentafluorophenol. Red trace represents curve fitting using a 1:1 host−guest binding model. The limiting chemical shifts for the selected protons are 11.22 ppm, 12.47 ppm, and 9.53 ppm, respectively. The final constant value is an average of the three fit curve calculation.

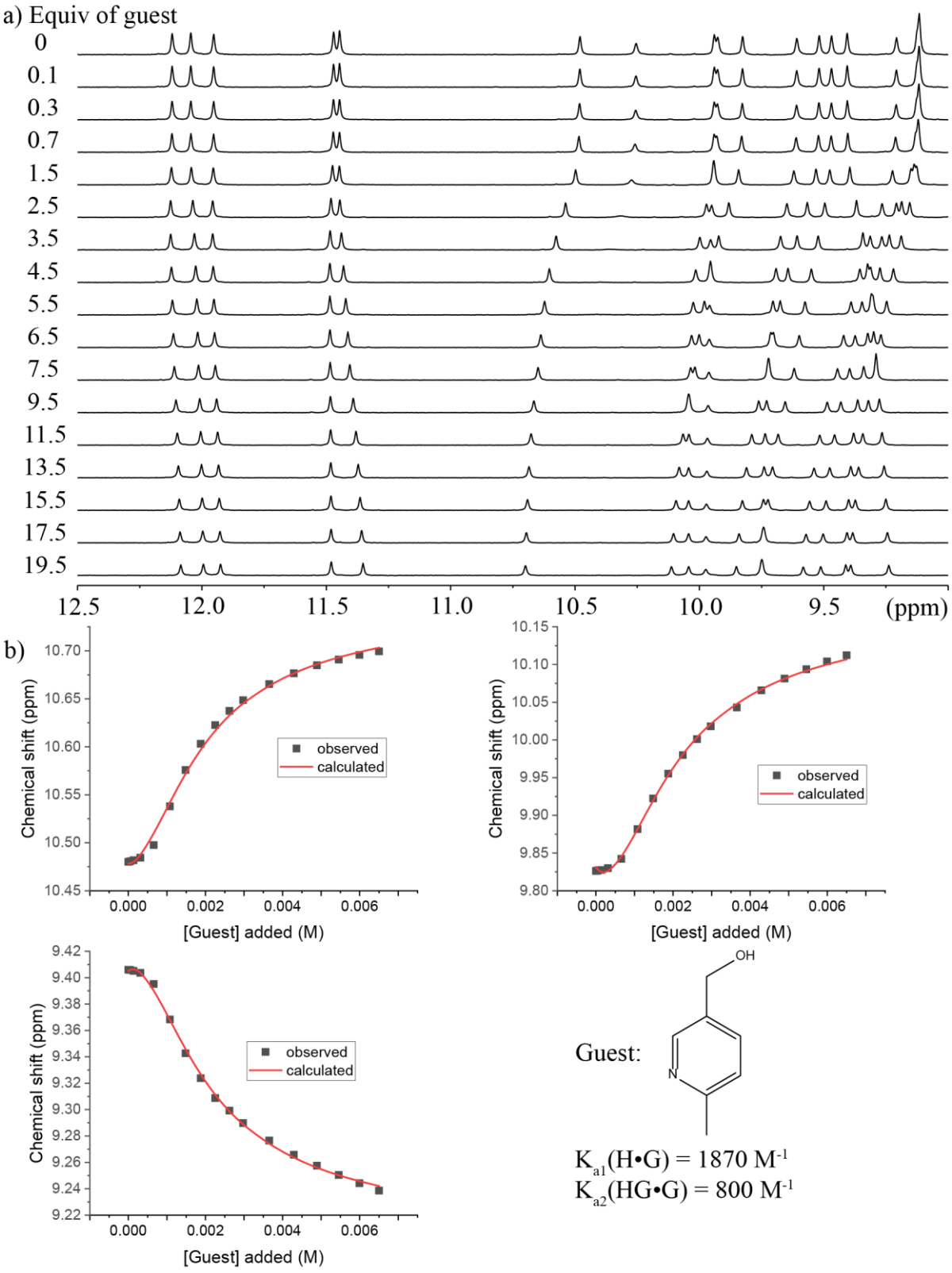


Figure S46. a) ^{1}H NMR spectra (500 MHz, H_2O/D_2O 90:10 (*v*/*v*), 25 °C) of oligomer **3** (0.45 mM) titrated with 5-hydroxymethyl-2-methylpyridine. b) Plots showing chemical shift changes of 3 different amide protons upon addition of 5-hydroxymethyl-2-methylpyridine. Red trace represents curve fitting using a 1:2 host−guest binding model. The limiting chemical shifts for the selected protons are 10.76 ppm, 10.19 ppm, and 9.20 ppm, respectively. The final constant value is an average of the three fit curve calculation.

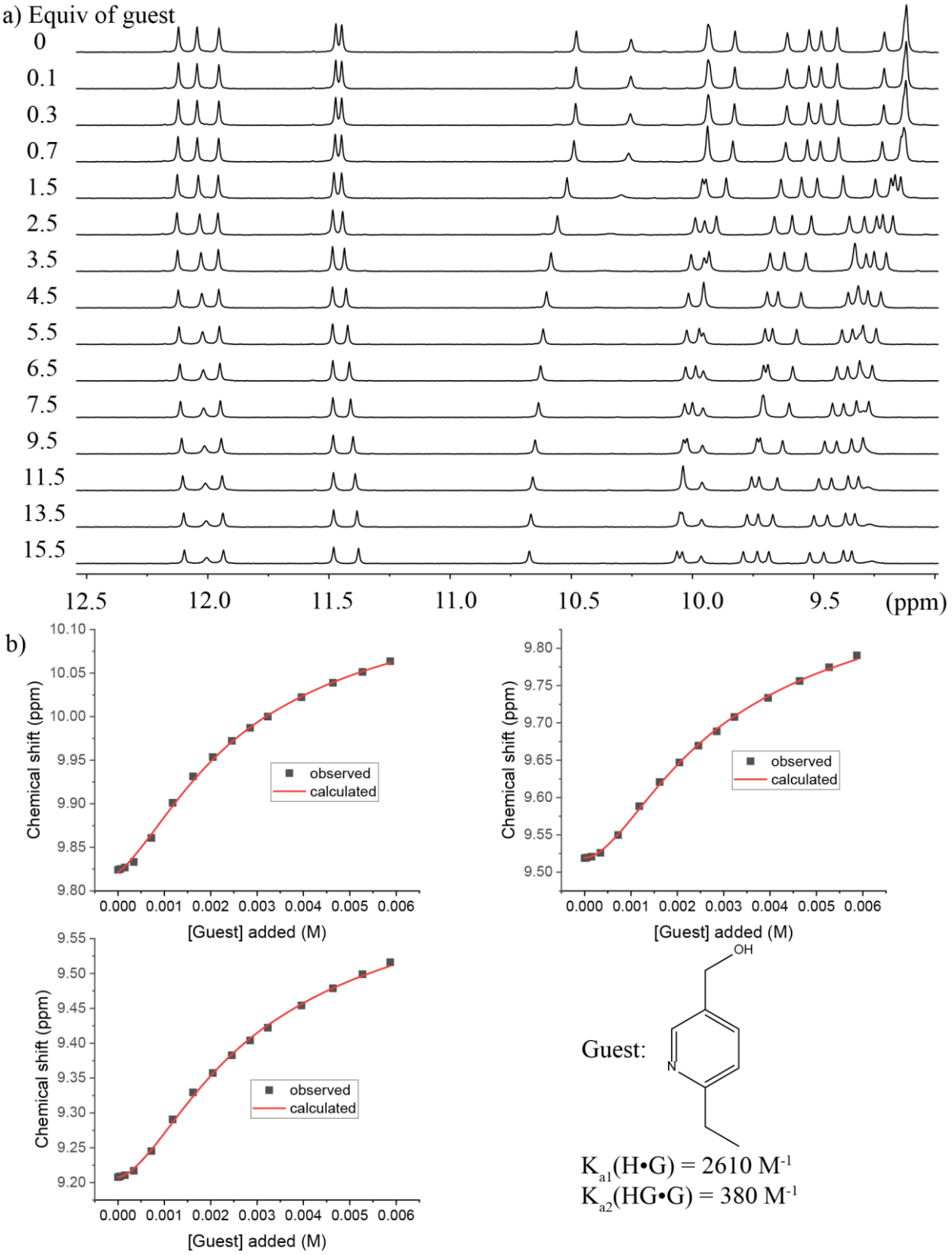


Figure S47. a) ^{1}H NMR spectra (500 MHz, H_2O/D_2O 90:10 (*v*/*v*), 25 °C) of oligomer **3** (0.5 mM) titrated with (5-ethylpyridin-2-yl) methanol. b) Plots showing chemical shift changes of 3 different amide protons upon addition of (5-ethylpyridin-2-yl) methanol. Red trace represents curve fitting using a 1:2 host−guest binding model. The limiting chemical shifts for the selected protons are 10.17 ppm, 9.93 ppm, and 9.67 ppm, respectively. The final constant value is an average of the three fit curve calculation.

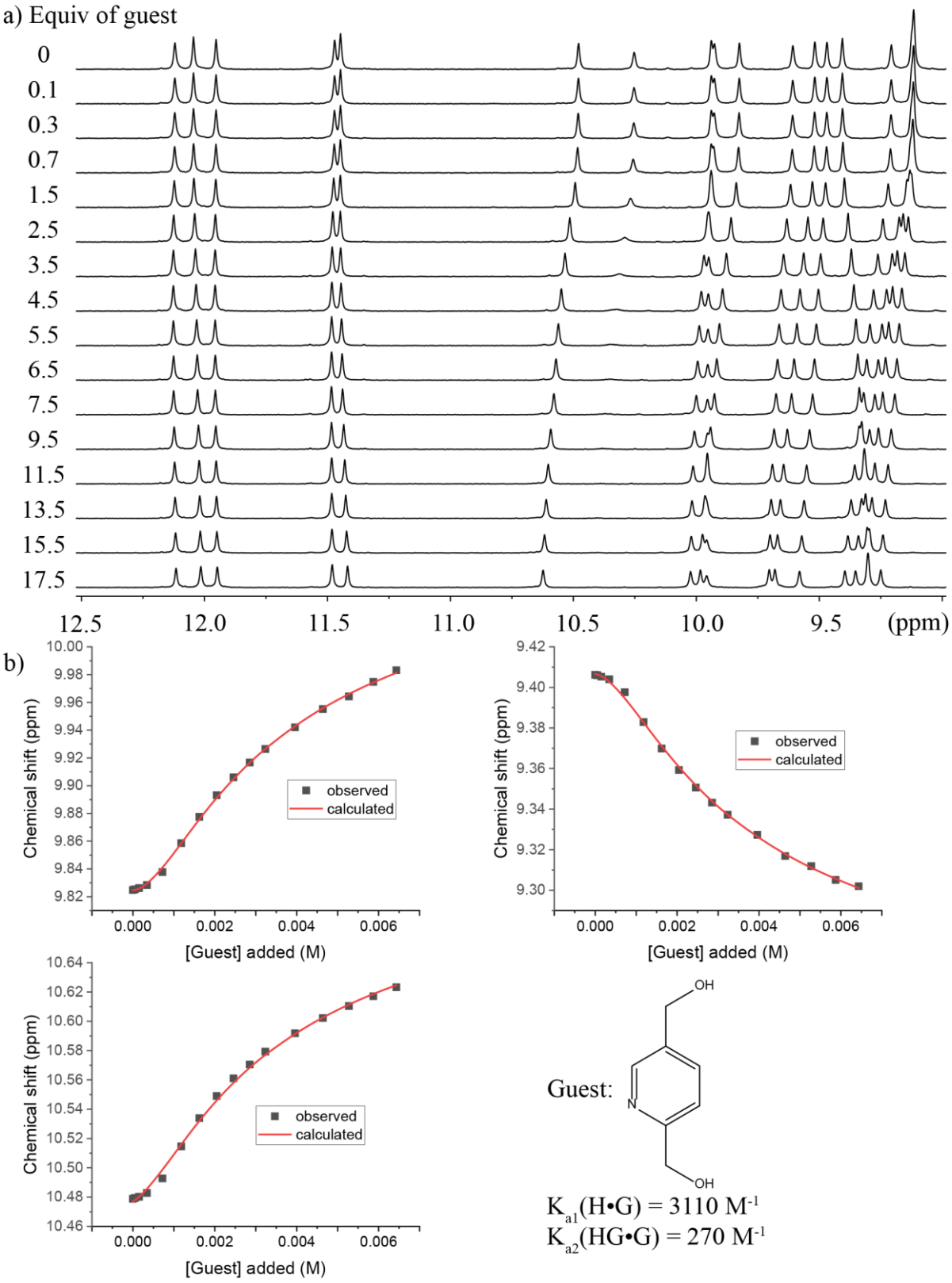


Figure S48. a) ^{1}H NMR spectra (500 MHz, H_2O/D_2O 90:10 (*v*/*v*), 25 °C) of oligomer **3** (0.5 mM) titrated with (5-ethylpyridin-2-yl) methanol. b) Plots showing chemical shift changes of 3 different amide protons upon addition of (5-ethylpyridin-2-yl) methanol. Red trace represents curve fitting using a 1:2 host−guest binding model. The limiting chemical shifts for the selected protons are 10.09 ppm, 9.23 ppm, and 10.71 ppm, respectively. The final constant value is an average of the three fit curve calculation.

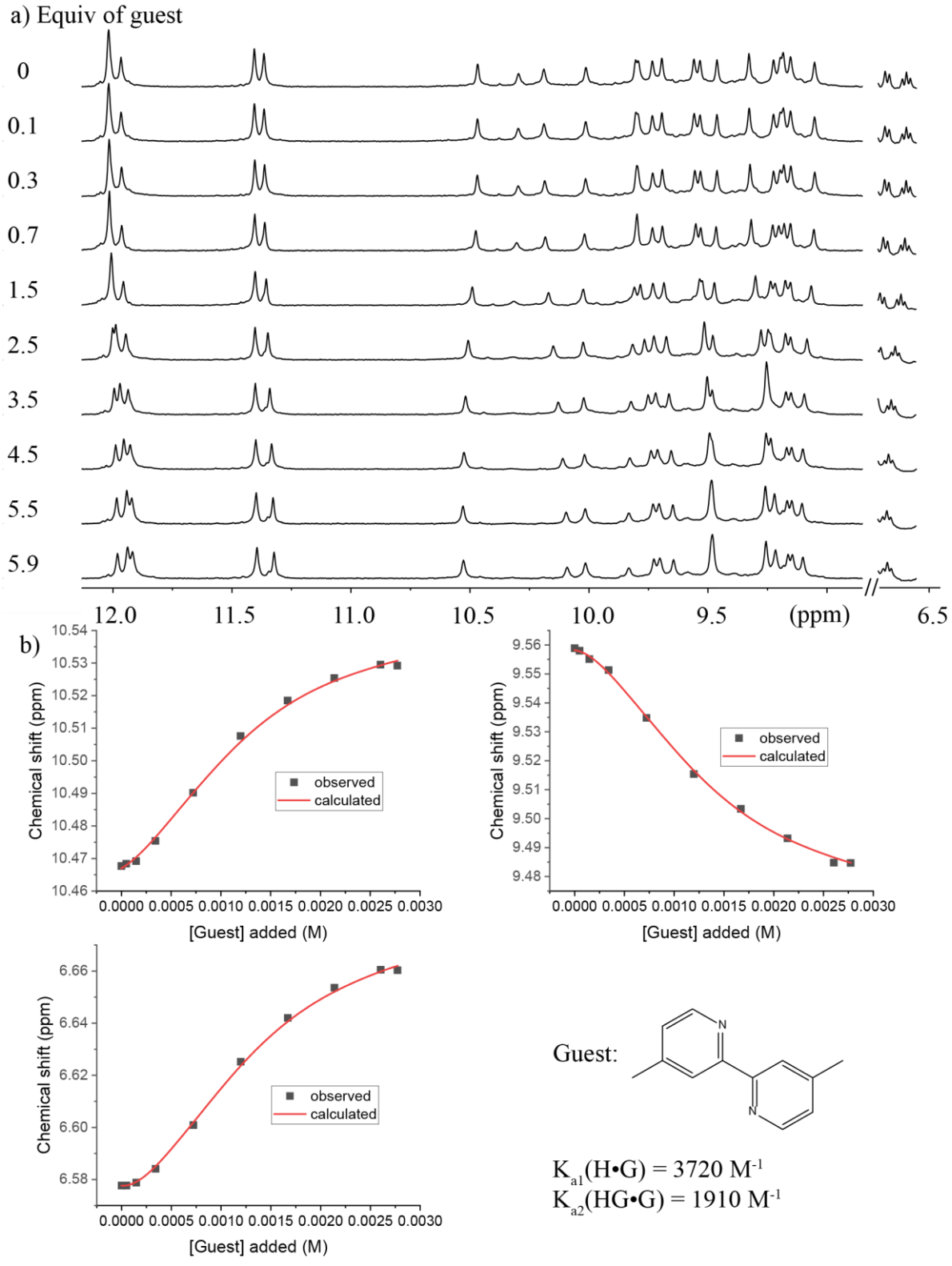


Figure S49. a) ^{1}H NMR spectra (500 MHz, H_2O/D_2O 90:10 (*v*/*v*), 25 °C) of oligomer **4** (0.5 mM) titrated with 4,4′-dimethyl-2,2′-dipyridyl. b) Plots showing chemical shift changes of 3 different amide protons upon addition of 4,4′-dimethyl-2,2′-dipyridyl. Red trace represents curve fitting using a 1:2 host−guest binding model. The limiting chemical shifts for the selected protons are 10.55 ppm, 9.46 ppm, and 6.69 ppm, respectively. The final constant value is an average of the three fit curve calculation.

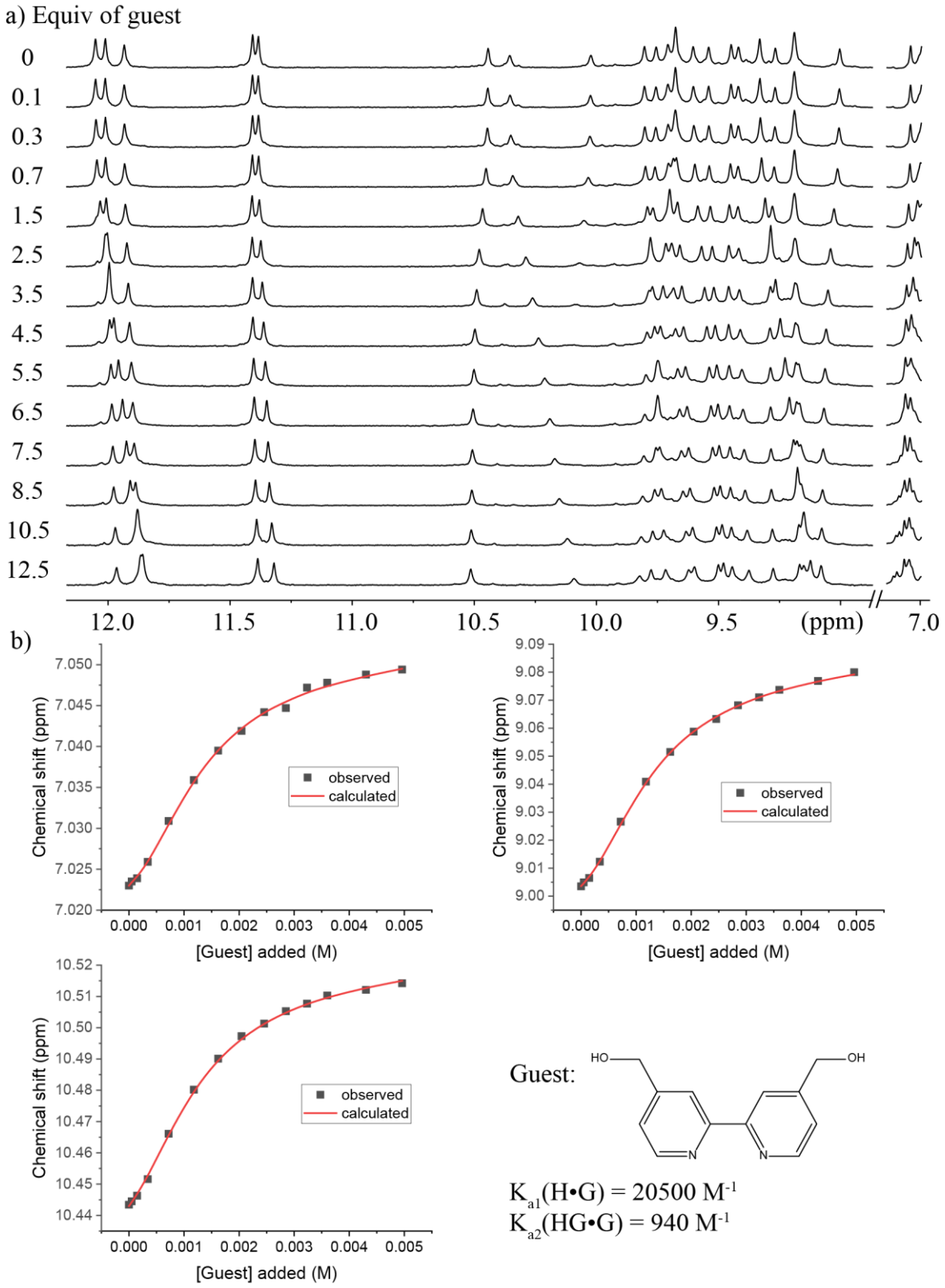


Figure S50. a) ^{1}H NMR spectra (500 MHz, H_2O/D_2O 90:10 (*v*/*v*), 25 °C) of oligomer **4** (0.5 mM) titrated with [2,2'-bipyridine]-4,4'-diyldimethanol. b) Plots showing chemical shift changes of 3 different amide protons upon addition of [2,2'-bipyridine]-4,4'-diyldimethanol. Red trace represents curve fitting using a 1:2 host−guest binding model. The limiting chemical shifts for the selected protons are 7.06 ppm, 9.10 ppm, and 10.53 ppm, respectively. The final constant value is an average of the three fit curve calculation.

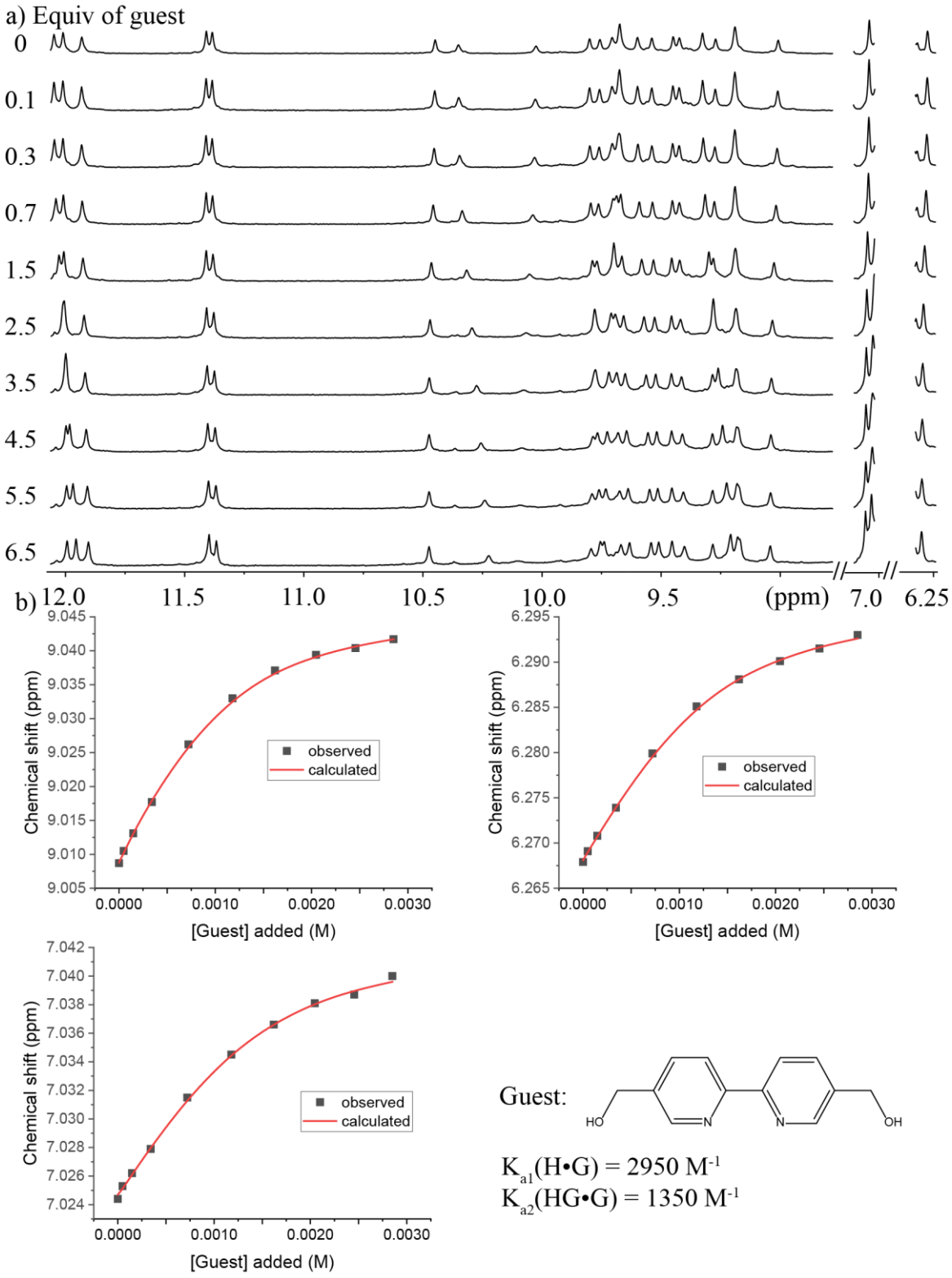


Figure S51. a) ^{1}H NMR spectra (500 MHz, H_2O/D_2O 90:10 (*v*/*v*), 25 °C) of oligomer **4** (0.5 mM) titrated with [2,2'-bipyridine]-5,5'-diyldimethanol. b) Plots showing chemical shift changes of 3 different amide protons upon addition of [2,2'-bipyridine]-5,5'-diyldimethanol. Red trace represents curve fitting using a 1:2 host−guest binding model. The limiting chemical shifts for the selected protons are 9.046 ppm, 6.297 ppm, and 7.043 ppm, respectively. The final constant value is an average of the three fit curve calculation.

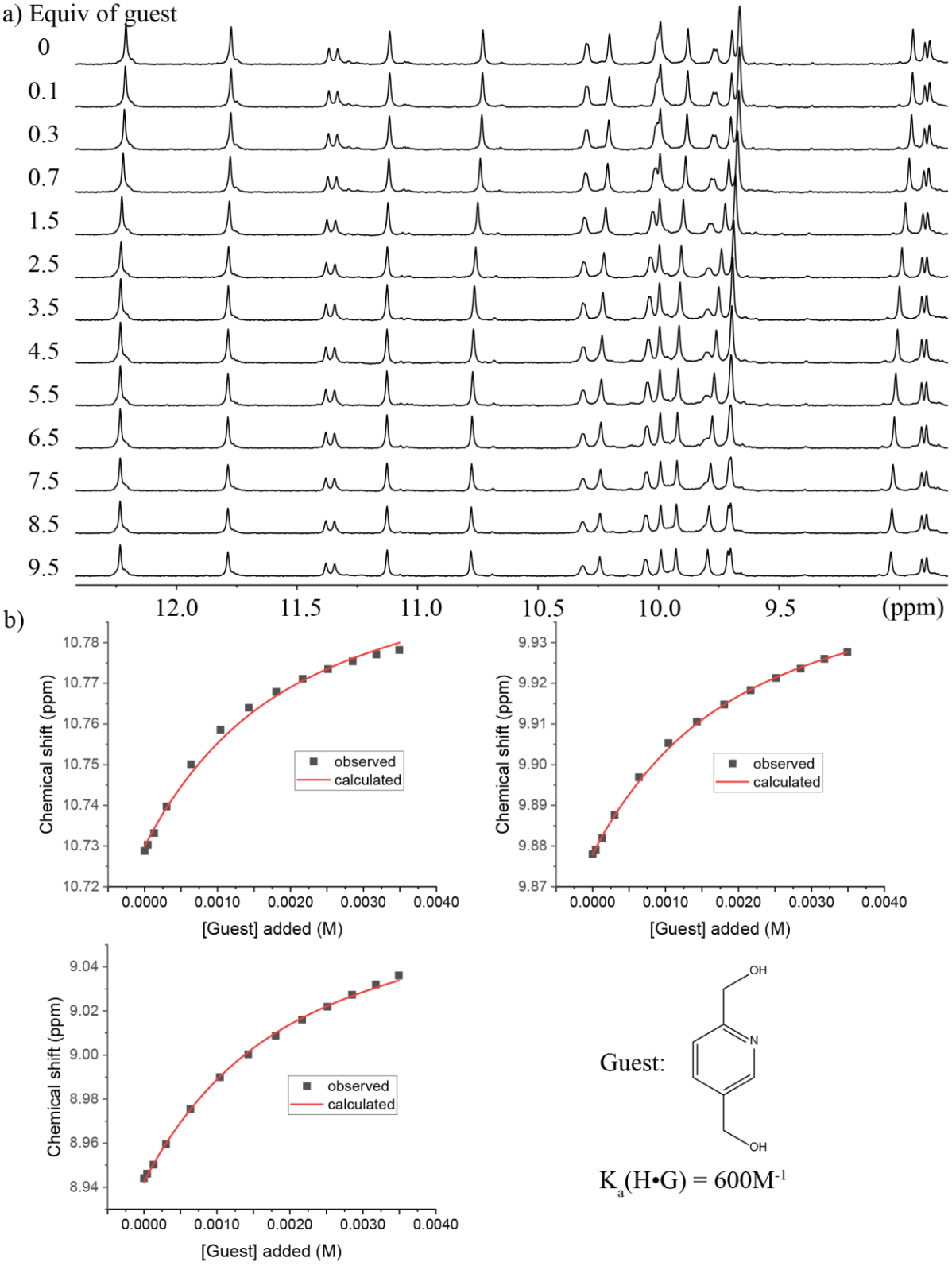


Figure S52. a) ^{1}H NMR spectra (500 MHz, H_2O/D_2O 90:10 (*v*/*v*), 25 °C) of oligomer **5** (0.44 mM) titrated with pyridine-2,5-diyldimethanol. b) Plots showing chemical shift changes of 3 different amide protons upon addition of pyridine-2,5-diyldimethanol. Red trace represents curve fitting using a 1:1 host−guest binding model. The limiting chemical shifts for the selected protons are 10.81 ppm, 9.95 ppm, and 9.08 ppm, respectively. The final constant value is an average of the three fit curve calculation.

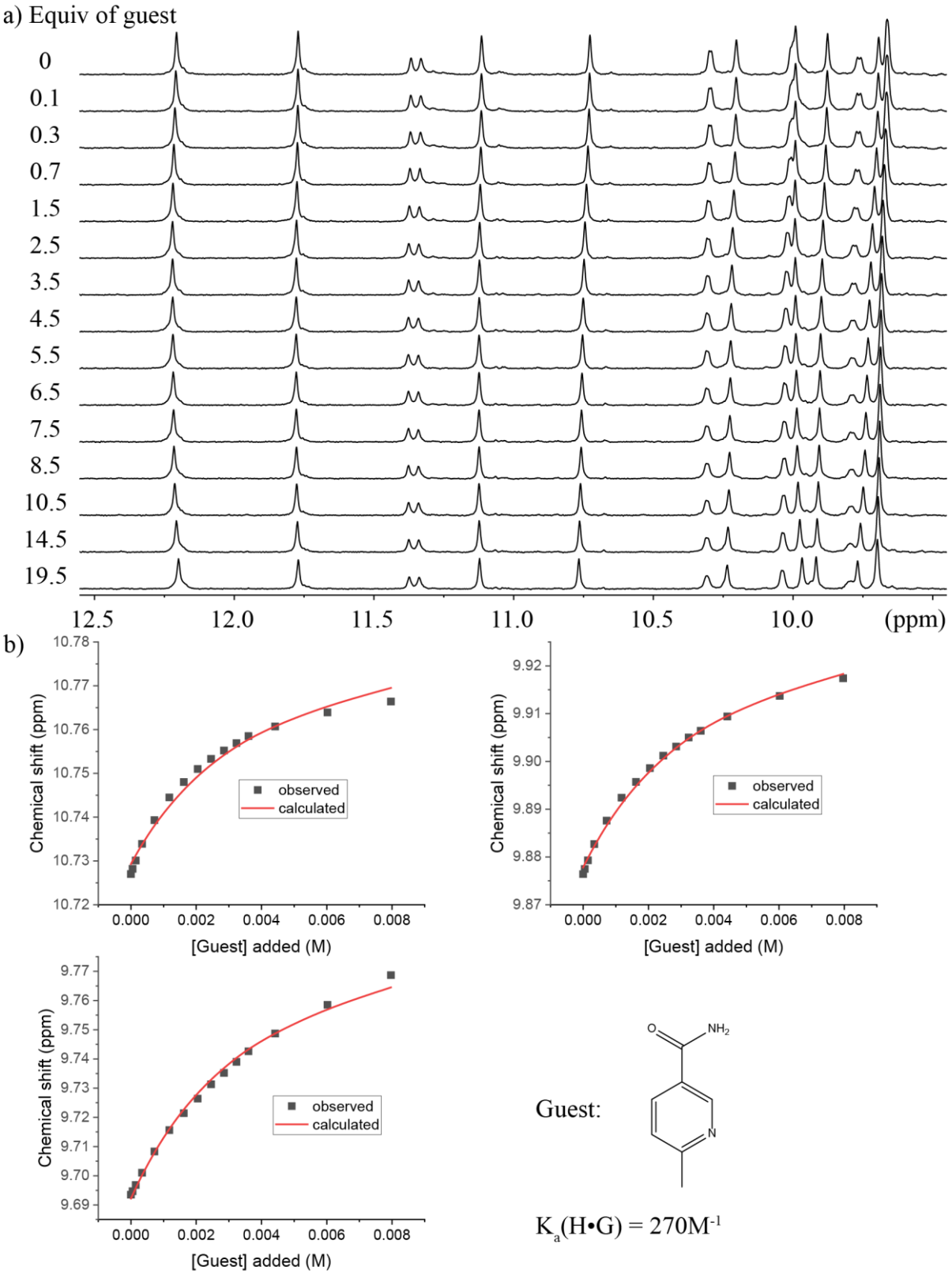


Figure S53. a) ^{1}H NMR spectra (500 MHz, H_2O/D_2O 90:10 (*v*/*v*), 25 °C) of oligomer **5** (0.5 mM) titrated with 6-methylpyridine-3-carboxamide. b) Plots showing chemical shift changes of 3 different amide protons upon addition of 6-methylpyridine-3-carboxamide. Red trace represents curve fitting using a 1:1 host−guest binding model. The limiting chemical shifts for the selected protons are 10.79 ppm, 9.94 ppm, and 9.80 ppm, respectively. The final constant value is an average of the three fit curve calculation.

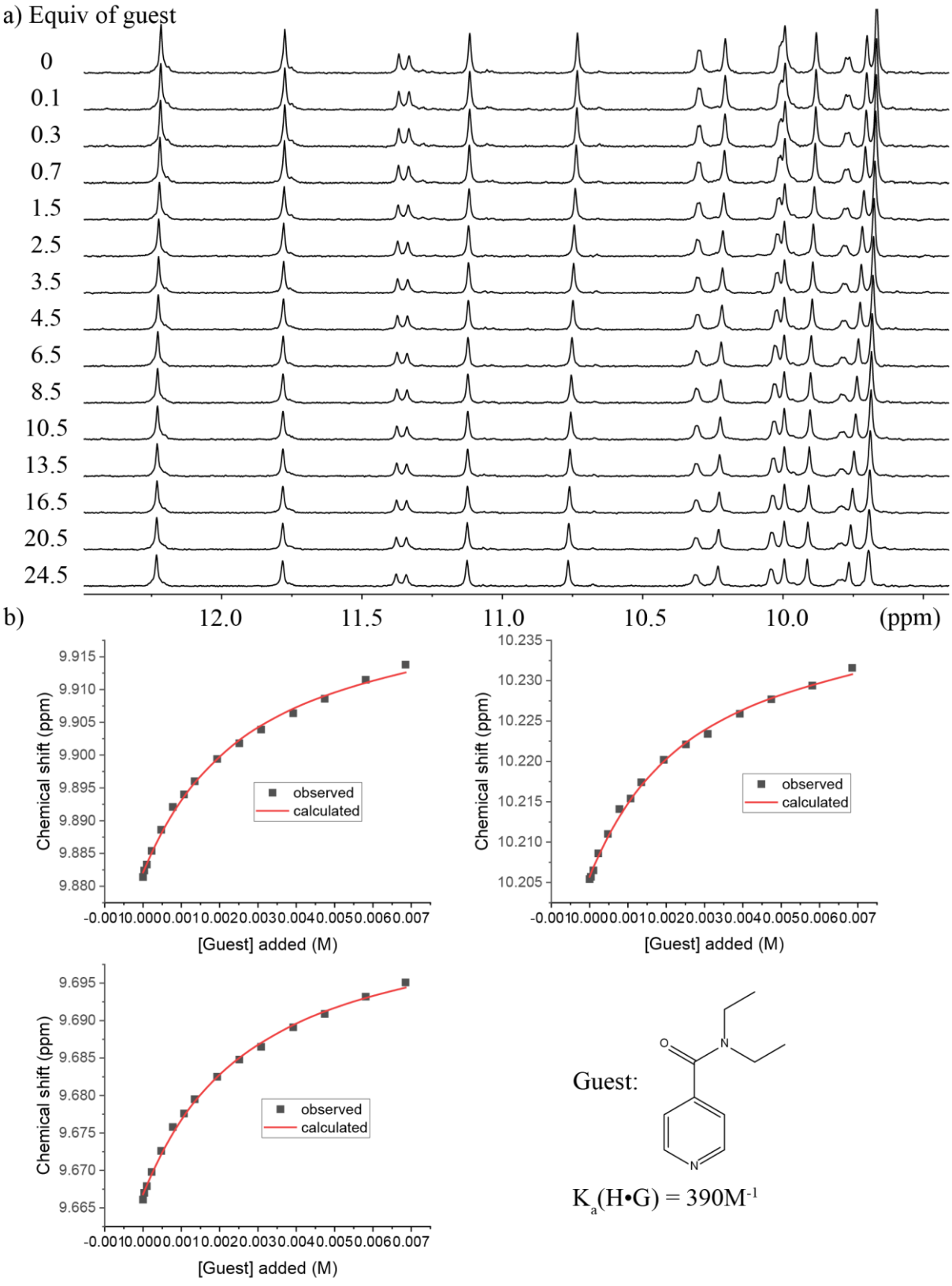


Figure S54. a) ^{1}H NMR spectra (500 MHz, H_2O/D_2O 90:10 (*v*/*v*), 25 °C) of oligomer **5** (0.33 mM) titrated with *N*,*N*-diethylisonicotinamide. b) Plots showing chemical shift changes of 3 different amide protons upon addition of *N*,*N*-diethylisonicotinamide. Red trace represents curve fitting using a 1:1 host−guest binding model. The limiting chemical shifts for the selected protons are 9.92 ppm, 10.24 ppm, and 9.71 ppm, respectively. The final constant value is an average of the three fit curve calculation.

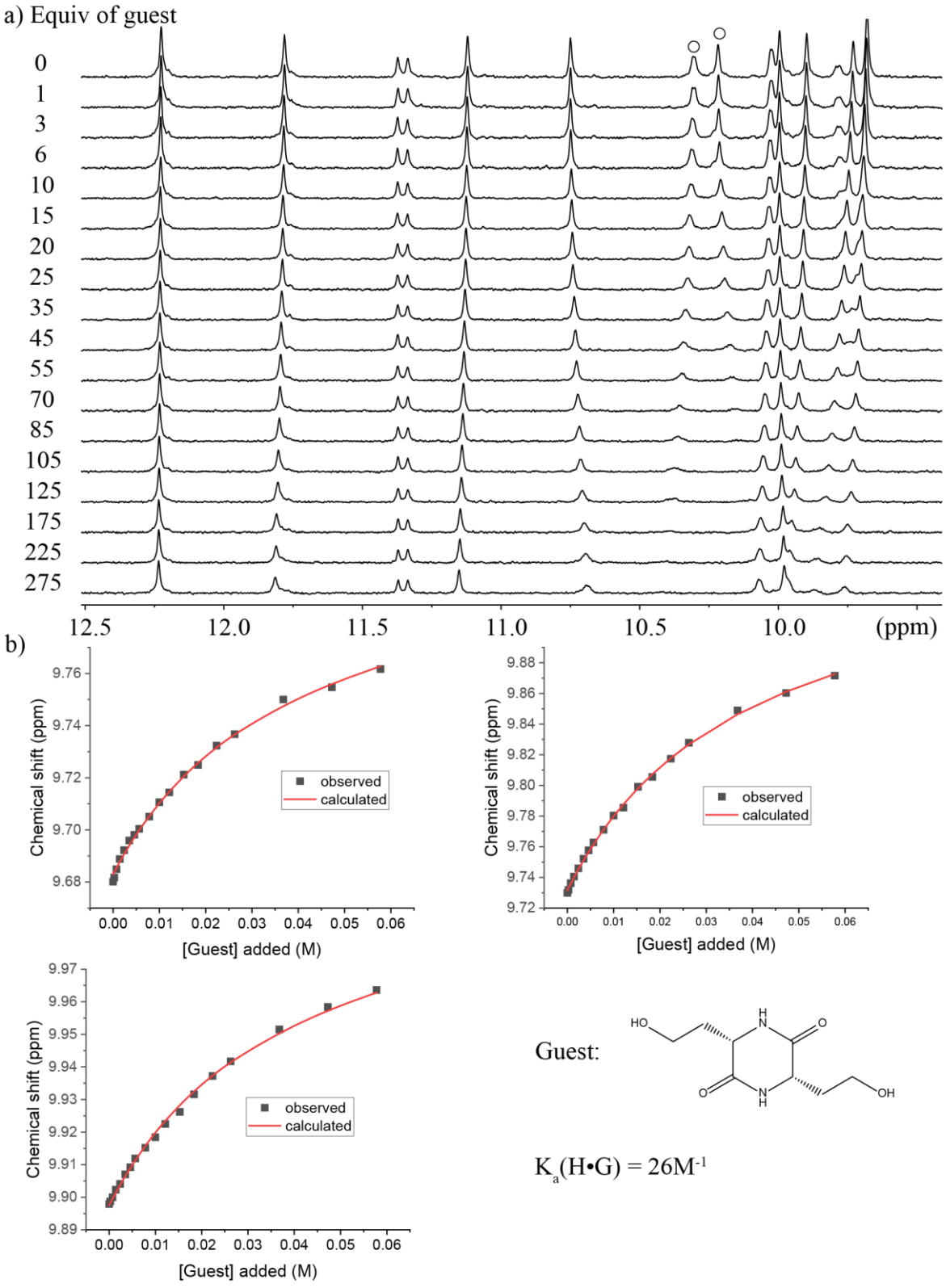


Figure S55. a) ^{1}H NMR spectra (500 MHz, H_2O/D_2O 90:10 (*v*/*v*), 25 °C) of oligomer **5** (0.25 mM) titrated with *cis*-3,6-bis(2-hydroxyethyl)piperazine-2,5-dione. The two proton signals that became broad are pointed out by black circle. b) Plots showing chemical shift changes of 3 different amide protons upon addition of *cis*-3,6-bis(2-hydroxyethyl)piperazine-2,5-dione. Red trace represents curve fitting using a 1:1 host−guest binding model. The limiting chemical shifts for the selected protons are 9.82 ppm, 9.97 ppm, and 10.00 ppm, respectively. The final constant value is an average of the three fit curve calculation.

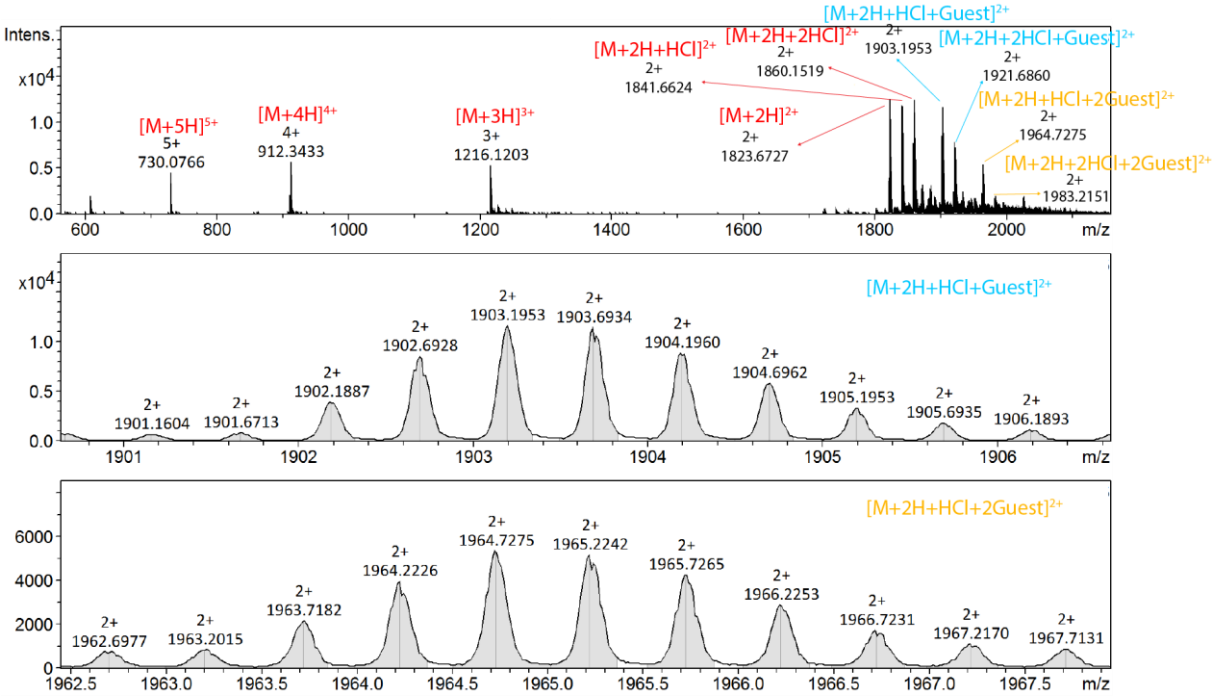


Figure S56. Host-guest complex in gas phase. ESI-MS spectrum of oligomer **3** and 5-hydroxymethyl-2-methylpyridine complex (direct infusion of an aqueous sample). Peaks corresponding to free **3**, guest ⊂ **3** and 2 guests ⊂ **3** are highlighted in red, blue and yellow, respectively.

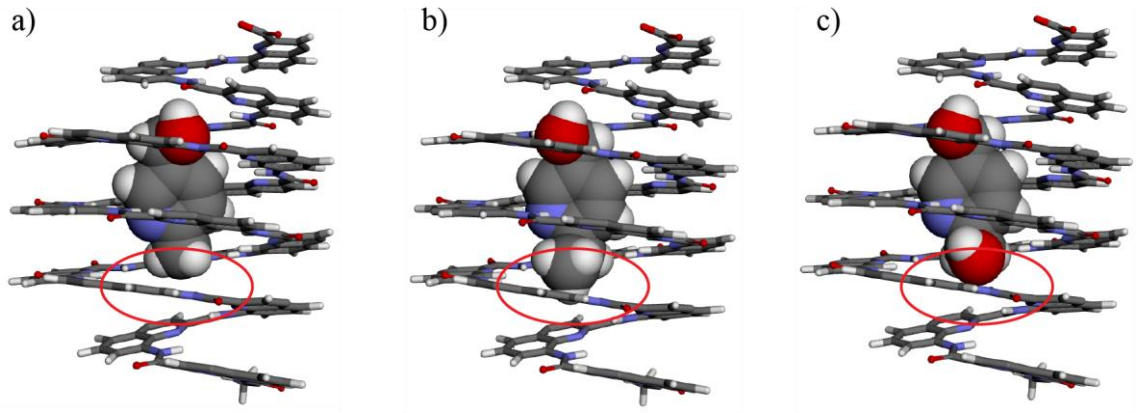


Figure S57. Computational modeling of the complex between oligomer 3 and different guests in 1:1 binding model. a) 5-hydroxymethyl-2-methylpyridine; b) (5-ethylpyridin-2-yl)methanol; c) pyridine-2,5-diyldimethanol, respectively. The energy minimized models (MacroModel, MMFFs) of 1:1 host−guest binding models indicate that (5-ethylpyridin-2-yl)methanol and pyridine-2,5-diyldimethanol can occupy more cavity space (red circle). Side chains have been omitted for clarity.

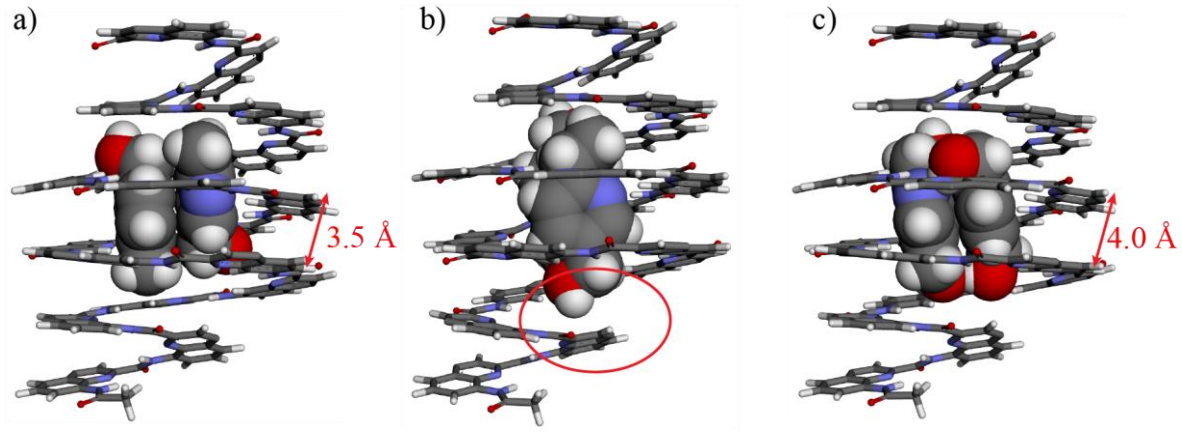


Figure S58. Computational modeling of the complex between oligomer 3 and different guests in 1:2 binding models. a) 5-hydroxymethyl-2-methylpyridine; b) (5-ethylpyridin-2-yl)methanol; c) pyridine-2,5-diyldimethanol, respectively. Two (5-ethylpyridin-2-yl)methanol push the cap segment and slightly open the capsule (red circle), according to the energy-minimized models (MacroModel, MMFFs) of 1:2 host-guest binding models. And binding two pyridine-2,5-diyldimethanol would slightly bend the central helical turn and increase distance (red arrow). Side chains have been omitted for clarity.

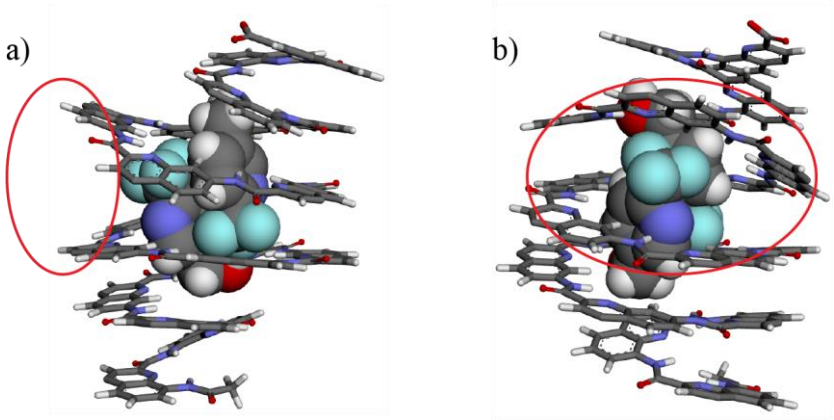


Figure S 59. Computational modeling of the complex between oligomer 3 and 6-(trifluoromethyl)pyridine-3-methanol. a) front view and b) side view of energy minimized models (Maestro MacroModel, MMFFs) of 1:2 host−guest binding model indicate two 6-(trifluoromethyl)pyridine-3-methanol can't fit inside the cavity and destroy the helical structure. Side chains have been omitted for clarity.

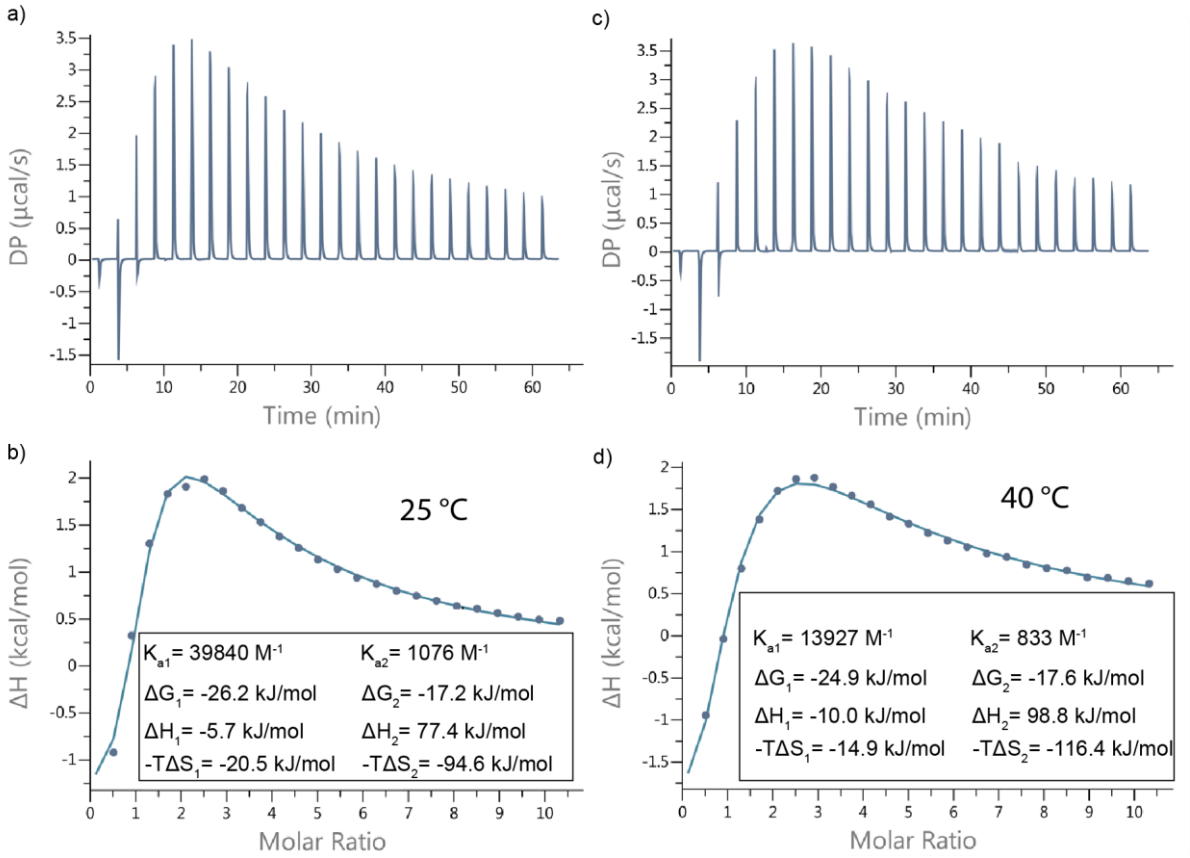


Figure S60. Different temperature isothermal microcalorimetric profile of the titration. ITC data for the titration of compound **1** with 4-pyridinemethanol at a) 25 °C and b) 40 °C in H_2O. The solid blue line represents the best-fitting curve obtained assuming a 1:2 host−guest sequential binding model.

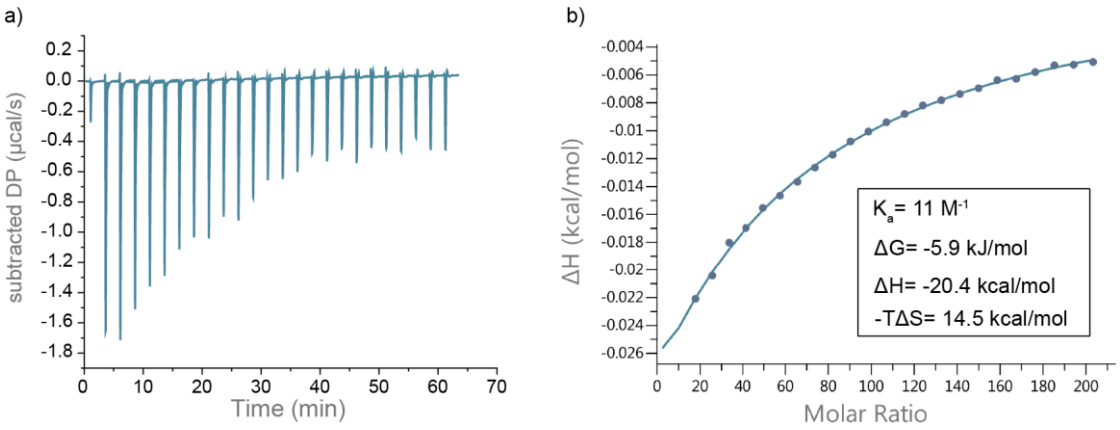


Figure S61. Isothermal microcalorimetric profile of the titration. Oligomer **5** (0.5 mM) titrated with *cis*-3,6-bis(2-hydroxyethyl)piperazine-2,5-dione at 25 °C. a) calorimetric trace subtracted by blank; b) The blue solid line represents the best-fitting curve obtained from a model assuming 1:1 stoichiometry.

2 Materials and Methods

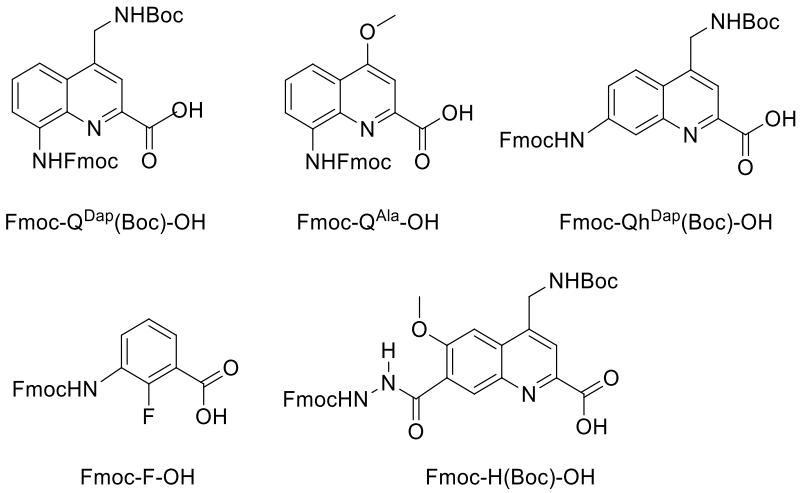


Figure S62. Fmoc-acid building blocks used in this study. Fmoc-Q^{Dap}(Boc)-OH[1], Fmoc-Q^{Ala}-OH[2] and Fmoc-Q^{H}(Boc)-OH[1] have been previously described. For a detailed procedure to Fmoc-H(Boc)-OH and Fmoc-F-OH, see section 2.3.

2.1 General

General. Commercial reagents (Suppliers: Abcr, Fisher Scientific, Merck, Sigma-Aldrich, TCI, BLDpharm or VWR) were used without further purification unless otherwise stated. Wang resin LL (100–200 mesh) was purchased from Sigma-Aldrich. Cl-MPA protide resin was purchased from CEM. Peptide grade *N*,*N*-dimethylformamide (DMF) was purchased from Carlo Erba. Anhydrous chloroform, triethylamine (TEA) and *N*,*N*-diisopropylethylamine (DIEA) were obtained via distillation over CaH_2 prior to use. Anhydrous tetrahydrofuran (THF) and dichloromethane (DCM) were obtained via an MBRAUN SPS-800 solvent purification system. Ultrapure water was obtained via a Sartorius arium® pro VF 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). Nuclear magnetic resonance (NMR) spectra were recorded on an Avance III HD 400 MHz Bruker BioSpin spectrometer or an Avance III HD 500 MHz Bruker BioSpin spectrometer equipped with a broad band observe 5-mm BB-H&FD CryoProbe™ Prodigy. Measurements were performed at 25 °C unless stated otherwise. Water suppression was performed with excitation sculpting method. Processing was done with MestReNova (v.12.0.0-20080) NMR processing software from Mestrelab Research. Chemical

shifts are reported in ppm and calibrated via residual solvent signals or 3-(trimethylsilyl)propionic-2,2,3,3-d_4 acid sodium salt (TMSP) when water suppression was applied. Signal multiplicities are abbreviated as s, singlet; d, doublet; t, triplet; q, quartet, and m, multiplet. Signals were assigned using ^{1}H-^{13}C HMQC and ^{1}H-^{13}C HMBC spectra. Electrospray ionization (ESI) mass spectra were recorded on Bruker microTOF II and Thermo Finnigan LTQ FT Ultra spectrometers. Electron ionization (EI) mass spectra were recorded on a Thermo Q Exactive GC Orbitrap or a Finnigan MAT 95 sector mass spectrometer. Analytical and semi-preparative reversed phase high performance liquid chromatography (RP-HPLC) was performed on a Thermo Fisher Scientific Ultimate 3000 HPLC System using Macherey-Nagel Nucleodur C18 Gravity columns (4 × 100 mm, 5 µm and 10 × 250 mm, 5 µm) and Macherey-Nagel Nucleodur C8 Gravity columns (4 × 50 mm, 5 µm and 10 × 100 mm, 5 µm) with different gradients of 0.1% TFA water and 0.1% TFA acetonitrile. All ultraviolet–visible (UV/Vis) absorbance measurements were done with a Jasco V-750 spectrophotometer instrument using a 1 cm quartz cuvette. Measurements were performed at 20 °C if not stated otherwise. Microwave-assisted solid phase foldamer synthesis (SPFS) was performed via a CEM® Discover Bio manual microwave peptide synthesizer. The temperature within the reactor vessel was monitored with an optical fiber probe. Automated SPFS was done on a PurePep® Chorus synthesizer (Gyros Protein Technologies) by applying an induction heating.

2.2 ^{1}H NMR investigations for host-guest interactions

^{1}H NMR Titrations in water were performed at 25 ˚C on a Bruker 500 MHz system with a cryoprobe. Aliquots from a stock solution containing the appropriate guests were added sequentially to an NMR tube containing the solution of corresponding oligomers. The ^{1}H NMR spectrum of the solution was acquired after each addition. The resulting NMR titration spectra were stacked by MestReNova software. Three protons' chemical shift upon guests binding and was used to determine the binding constant K_a. The resulting titration isotherms were fitted to a 1:1 or 1:2 host-guest binding model using HypNMR2008. The fitting data were plotted using OriginLab (v2018) software.

2.3 Isothermal titration calorimetry (ITC) titration

Isothermal titration calorimetry (ITC) titrations were performed by PEAQ-ITC microcalorimeter from Malvern at 25 ˚C. A Hastelloy cell was used with an active cell volume of 300 μL. The stirring speed was set at 750 rpm. Host and guest solutions were prepared in Milli-Q water using the Sartorius arium® pro VF ultrapure water system and allowed to equilibrate overnight if necessary. The hosts placed in the titration cell, and the guests were loaded into the titration syringe. In each titration experiment, 25 injections were performed with gradually decreased titration peaks. The heat of dilution was measured by titrating the guests into a blank solution. After subtracting the heat of dilution, the resulting data were analyzed with PEAQ-ITC software using a 1:1 host-guest binding model or 1:2 sequential binding model.

2.4 Octanol-water partition coefficient (P) determination

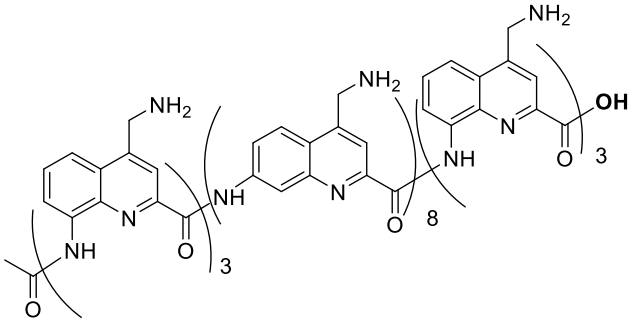
Octanol−water partition coefficient (log P) values were determined for guests that were used for host-guest interactions experiments. Each guest was dissolved in an octanol−water system and allowed to equilibrate (shaking overnight), and then both fractions were analyzed by UV spectroscopy or high-performance liquid chromatography. log P was calculated as log[ratio of the concentration in the octanol phase to the concentration in the aqueous phase].

2.5 Solid phase synthesis procedures

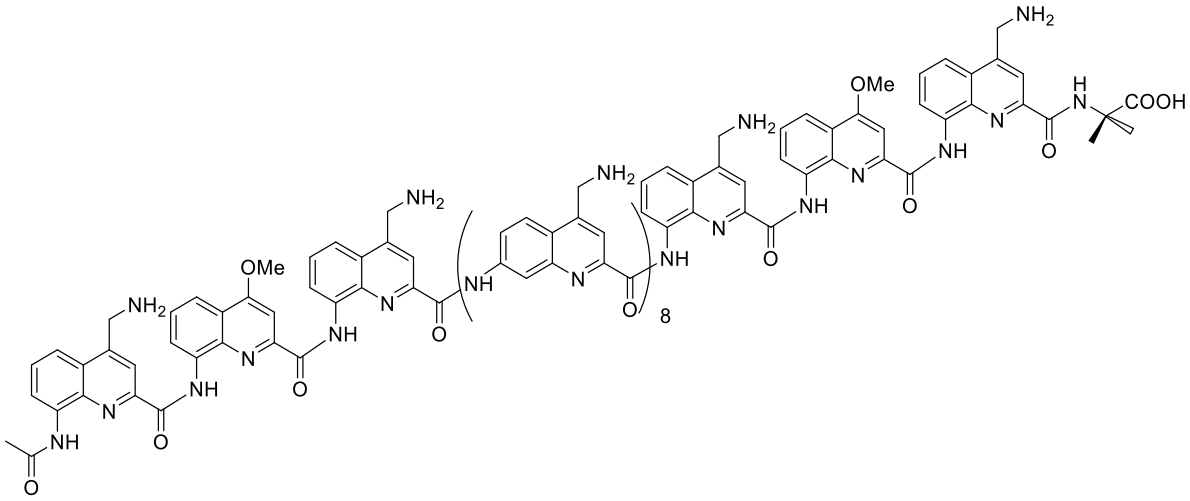
Oligomers were synthesized according to previously reported SPFS protocols,[3] hereafter referred to as manual synthesis method. The automated synthesis protocol has been reported in previous work.[4]

Final acetylation of the N-terminal aromatic amine: H_2N-$(Q)_n$-Wang resin and DIEA (6.0 equiv.) were suspended in anhydrous THF (1.25 mL) then acetyl chloride (4.0 equiv.) in anhydrous THF (1.25 mL) was added. The reaction vessel was then placed under microwave irradiation (25 W, ramp to 50 °C over 5 min, then hold at 50 °C for 15 min). The resin was filtered off and washed with anhydrous THF (2 x 3 mL). The coupling step was repeated twice using the same conditions. The resin was finally intensively washed with DCM, dried by passing N_2 flow through it before to be submitted to TFA cleavage

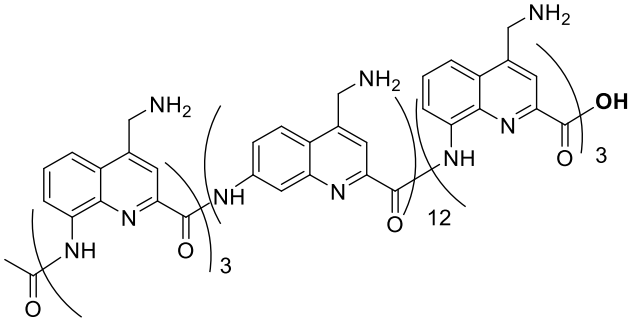
Resin cleavage: The resin-bound foldamer was placed in a syringe equipped with a filter and was then suspended in a solution of TFA/*i*Pr_3SiH/H_2O (95:2.5:2.5, *v*/*v*/*v*). The resin was next gently shaken for at least 2 h at room temperature before to be was filtered off and rinsed one time with TFA cleavage coktail. The foldamer was precipitated from the TFA cleavage solution by adding cold Et_2O and centrifugation to obtain the crude precipitate.



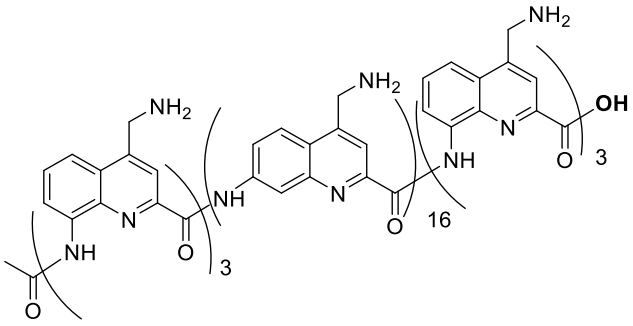
Oligomer 1: The synthesis of this oligomer has already been published previously.[1]



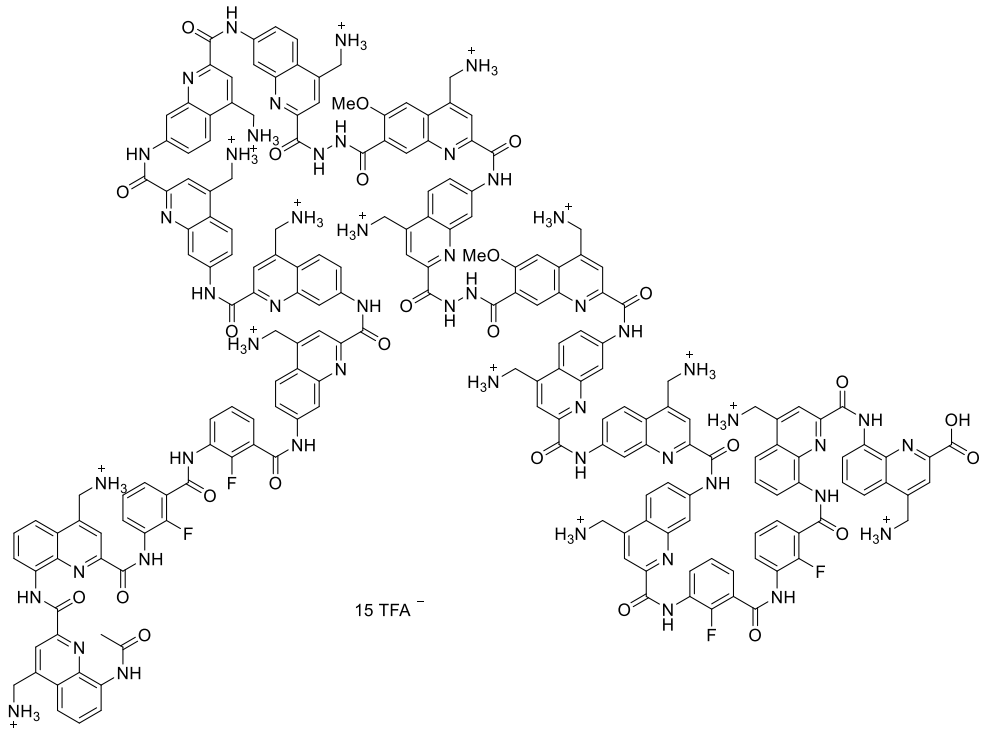
Oligomer 2: The synthesis of this oligomer has already been published previously.[5]



Oligomer 3: The synthesis of this oligomer has already been published previously.[1]



Oligomer 4: The synthesis of this oligomer has already been published previously.[1]



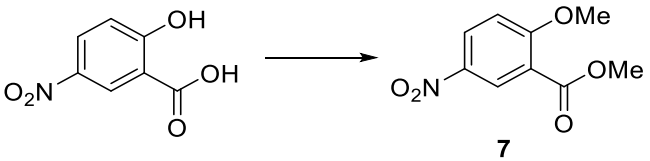
Oligomer 5: Oligomer **5** was synthesized starting from Cl-MPA-Protide resin (0.17 mmol g^{-1}, 24 µmol scale) according to the automated synthesis method. Loading of the first monomer: 0.12 mmol g^{-1} (71%). Final acetylation was carried out via the general acetylation method. After precipitation in cold Et_2O, the crude mixture was purified by semi prep RP-HPLC to give the **5** as a yellow solid (16 mg, 12%).

^{1}H NMR (500 MHz, H_2O/D_2O = (9:1, *v*/*v*)) δ 12.14 (s, 1H), 11.72 (s, 1H), 11.29 (d, *J* = 18.2 Hz, 1H), 11.06 (s, 1H), 10.67 (s, 1H), 10.24 (d, *J* = 4.8 Hz, 1H), 10.15 (s, 1H), 9.94 (d, *J* = 7.7 Hz, 2H), 9.82 (s, 1H), 9.70 (d, *J* = 6.7 Hz, 1H), 9.64 (s, 1H), 9.61 (d, *J* = 4.0 Hz, 1H), 8.88 (s, 1H), 8.83 (d, *J* = 10.0 Hz, 1H), 8.54 (d, *J* = 10.1 Hz, 1H), 8.47 (s, 1H), 8.36 (d, *J* = 9.6 Hz, 1H), 8.34 – 8.28 (m, 4H), 8.26 – 7.98 (m, 9H), 7.97 – 7.69 (m, 14H), 7.69 – 7.22 (m, 23H), 7.22 –

6.45 (m, 26H), 6.10 (t, *J* = 8.2 Hz, 1H), 5.95 (t, *J* = 8.1 Hz, 1H), 3.82 – 3.54 (m, 3H), 3.18 (dd, *J* = 23.1, 16.0 Hz, 2H).

HRMS: calcd. for $C_{199}F_4H_{163}N_{51}O_{25}$ $[M+2H]^{2+}$ 1871.1488; found 1871.1631.

2.6 Monomer synthesis procedures



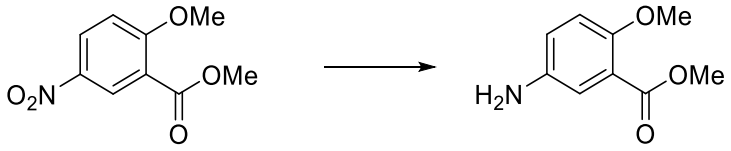
7

Compound 7. 5-Nitrosalicilic acid (20 g, 0.109 mol) and K_2CO_3 (60 g, 0.434 mol) was dissolved in DMF (200 mL) and methyl iodide (27.2 mL, 0.437 mol) was added. The reaction mixture was stirred for 48 h at r.t.. After TLC confirmed the reaction completion, the mixture was diluted with water and the aqueous layer was extracted with ethyl acetate three times. The organic layers were washed with 0.5 N HCl and brine, dried (Na_2SO_4) and concentrated under reduced pressure to give 21 g (91%) of **7** as a yellow solid.

^{1}H NMR (400 MHz, DMSO-d_6) δ 8.48 (d, *J* = 3.0 Hz, 1H), 8.43 (dd, *J* = 9.3, 3.0 Hz, 1H), 7.40 (d, *J* = 9.3 Hz, 1H), 3.98 (s, 3H), 3.84 (s, 3H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 164.3, 163.0, 140.0, 129.0, 126.5, 120.2, 113.5, 57.1, 52.6.

HRMS: calcd. for $C_9H_9NO_5$ $[M]^+$ 211.0475; found 211.0470.



7 **8**

Compound 8. Compound **7** (21 g, 0.099 mol) was dissolved in MeOH (200 mL) and CH_3COOH (200 mL). Then, iron powder (28 g, 0.501 mol) was added in portion. The reaction mixture was stirred 5 h. at 65 °C. After TLC confirmed the reaction completion, the mixture was filtered over celite and water was added; the mixture was then extracted with ethyl acetate three times. The organic layers were washed with saturated $NaHCO_3$ solution and brine, dried (Na_2SO_4) and then concentrated under reduced pressure to give 16.2g (90%) of **8** as a semi solid.

^{1}H NMR (500 MHz, DMSO-d_6) δ 6.90 (d, *J* = 2.9 Hz, 1H), 6.86 (d, *J* = 8.8 Hz, 1H), 6.74 (dd, *J* = 8.8, 2.9 Hz, 1H), 4.88 (s, 2H), 3.75 (s, 3H), 3.68 (s, 3H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 166.6, 149.6, 142.2, 120.5, 118.8, 115.7, 114.6, 56.5, 51.7.
HRMS: calcd. for $C_9H_{11}NO_3$ $[M]^+$ 181.0733; found 181.0732.

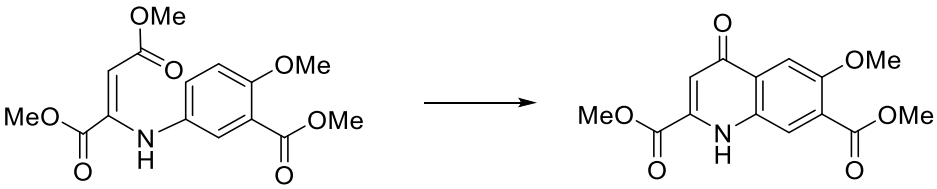
8 **9**

Compound 9. Compound **8** (8.50 g, 0.047 mol) was dissolved in methanol (135 mL) and dimethylacetylene dicarboxylate (7.77 g, 0.055 mol) was added. The resulting mixture was heated to reflux temperature for 12 h. After TLC confirmed the reaction completion, the reaction mixture was concentrated and purified by column chromatography to give 11 g (73%) of **9** as a yellow syrup.

^{1}H NMR (500 MHz, DMSO-d_6) δ 9.55 (s, 1H), 7.26 (d, J = 2.9 Hz, 1H), 7.16 (dd, J = 9.0, 2.9 Hz, 1H), 7.10 (d, J = 9.0 Hz, 1H), 5.24 (s, 1H), 3.79 (s, 3H), 3.78 (s, 3H), 3.65 (s, 3H), 3.65 (s, 3H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 168.3, 165.6, 164.3, 155.0, 147.5, 132.9, 126.1, 123.1, 120.0, 113.4, 92.3, 56.1, 53.0, 52.0, 51.0.

HRMS: calcd. for $C_{15}H_{16}NO_7$ $[M-H]^-$ 322.0921; found 322.0931.



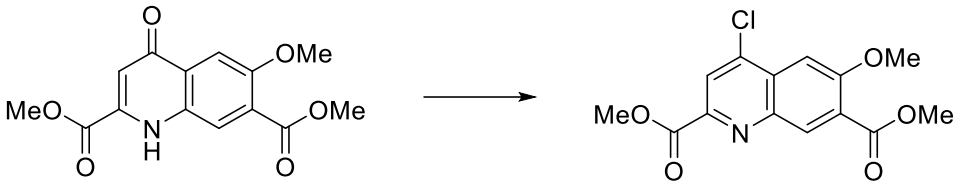
9 **10**

Compound 10. Diphenyl ether (350 mL) was placed in a 1 L round bottom flask and heated to its boiling point at 260 ˚C using a heating mantle. Then compound **9** (11 g, 0.034 mol) was added to the boiling solvent by means of a glass funnel. The mixture was kept boiling for 20 min; no condenser was used. The reaction mixture was left to cool down to r. t.. The resulting solid was filtered off and washed with cyclohexane several times, then dried under reduced pressure. The product was purified by triturating in hot MeOH to give 4.3 g (43%) of **10** as a light yellow solid.

^{1}H NMR (500 MHz, DMSO-d_6) δ 12.26 (s, 1H), 8.29 (s, 1H), 7.60 (s, 1H), 6.65 (s, 1H), 3.98 (s, 3H), 3.91 (s, 3H), 3.88 (s, 3H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 176.5, 165.6, 162.5, 153.8, 137.6, 133.5, 128.7, 126.6, 122.7, 109.2, 104.9, 56.1, 53.6, 52.6.

HRMS: calcd. for $C_{14}H_{13}NO_6$ $[M]^+$ 291.0737; found 291.0733.



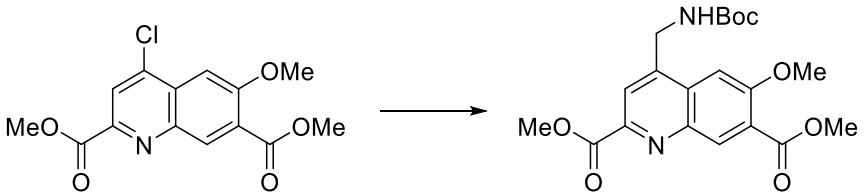
10 **11**

Compound 11. Compound **10** (2.30 g, 0.008 mol) was dissolved in dry DMF (23 mL) and $POCl_3$ (0.8 mL, 0.009 mmol) was added slowly. The reaction mixture was stirred for 15 min at 80 °C. After TLC confirmed the reaction completion, the mixture was poured onto ice and neutralised with saturated $NaHCO_3$ solution. Then the mixture was extracted with DCM. The combined organic layers were washed with subsequently water and brine, dried (Na_2SO_4) and then concentrated under reduced pressure to give 2.20 g (90%) of **11** as a white solid.

^{1}H NMR (500 MHz, DMSO-d_6) δ 8.43 (s, 1H), 8.29 (s, 1H), 7.60 (s, 1H), 4.06 (s, 3H), 3.97 (s, 3H), 3.91 (s, 3H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 165.1, 164.2, 157.3, 146.1, 142.2, 140.7, 132.9, 129.3, 128.1, 122.5, 102.6, 56.6, 52.9, 52.8.

HRMS: calcd. for $C_{14}H_{12}ClNO_5$ $[M]^+$ 309.0399; found 309.0391.



11 **12**

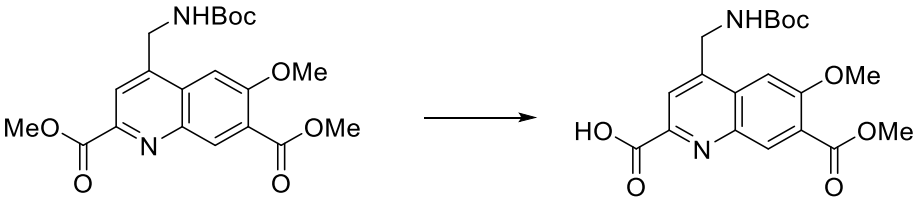
Compound 12. Compound **11** (0.92 g, 0.0030 mol), KF_3BCH_2NHBoc (0.62 g, 0.0031 mol), $Pd(OAc)_2$ (0.036 g, 0.00013 mol), SPhos (0.123 g, 0.0003 mol) and K_2CO_3 (1.04 g, 0.009 mol) were added in a sealed tube and the mixture was purged 3 times with argon. Then toluene (9.6 mL) and water (2.4 mL) were added and the reaction mixture was heated to 85 °C for 30 h. After TLC confirmed the reaction completion, phosphate buffer pH = 7 was added and the mixture extracted with DCM. The combined organic layers were subsequently washed with water and brine, dried (Na_2SO_4) and then concentrated under reduced pressure. The resulting residue was purified by column chromatography on silica. The recovered quinoline derivative

11 was used to repeat the reaction and the product that was obtained after column was precipitated in DCM/diethyl ether to give 0.8 g (66%) of **12** as white solid.

^{1}H NMR (500 MHz, DMSO-d_6) δ 8.36 (s, 1H), 8.01 (s, 1H), 7.70 (t, J = 6.1 Hz, 1H), 7.60 (s, 1H), 4.70 (d, J = 6.1 Hz, 2H), 4.02 (s, 3H), 3.95 (s, 3H), 3.89 (s, 3H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 165.5, 165.3, 156.2, 155.9, 145.9, 145.7, 141.6, 132.8, 129.8, 126.7, 119.8, 103.0, 78.5, 56.5, 52.7, 52.6, 40.8, 28.2.

HRMS: calcd. for $C_{20}H_{25}N_2O_7$ $[M+H]^+$ 405.1656; found 405.1659.

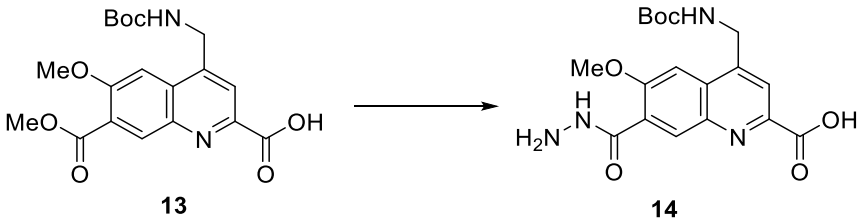


Compound 13. Compound **12** (2.63 g, 0.0065 mol) was dissolved in DCM (105 mL) and NaOH (0.281 g, 0.007mol) in MeOH (25 mL) was added slowly at 0 °C. The mixture stirred for 3 h at 0 °C and allowed to warm to r.t. overnight. After TLC confirmed the reaction completion, solvent was evaporated then water was added. The aqueous layer was washed with DCM and neutralized by 0.1M HCl to allow precipitate formation, the mixture was filtered and dried under reduced pressure to give 2.1 g (83%) of **13** as white solid.

^{1}H NMR (500 MHz, DMSO-d_6) δ 8.35 (s, 1H), 8.01 (s, 1H), 7.69 (t, J = 6.0 Hz, 1H), 7.60 (s, 1H), 4.69 (d, J = 6.0 Hz, 2H), 4.02 (s, 3H), 3.89 (s, 3H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 166.4, 165.6, 156.0, 155.9, 145.4, 141.4, 132.8, 129.6, 126.3, 120.0, 103.0, 78.4, 56.4, 52.6, 40.9, 28.2.

HRMS: calcd. for $C_{19}H_{23}N_2O_7$ $[M+H]^+$ 391.1500; found 391.1502.

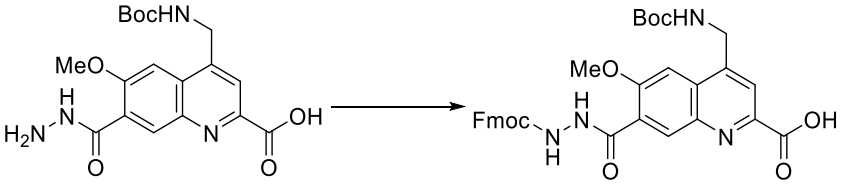


Compound 14. Compound **13** (1.69 g, 0.0043mol) was dissolved in methanol (17 mL) and cooled down to 0 °C. Hydrazine (2.02 mL, 0.042mol) was added slowly and the mixture was stirred for 3 h at r. t.. After HPLC confirmed the reaction completion, the solvent was evaporated and the residues was dried under reduced pressure over night to obtain 1.69 g (99%) **14** as white solid.

^{1}H NMR (500 MHz, DMSO-d_6) δ 9.50 (s, 1H), 8.19 (s, 1H), 7.94 (s, 1H), 7.61 (t, J = 6.1 Hz, 1H), 7.47 (s, 1H), 4.61 (d, J = 6.2 Hz, 2H), 3.97 (s, 3H), 1.43 (s, 9H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 166.9, 164.3, 155.9, 155.6, 144.7, 144.1, 141.9, 131.7, 129.1, 128.5, 119.7, 102.2, 78.3, 56.2, 40.9, 28.3.

HRMS: calcd. for $C_{18}H_{23}N_4O_6$ $[M+H]^+$ 391.1612; found 391.1614.

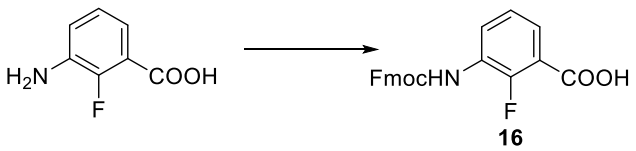


Compound 15. Compound **14** (1.73g, 0.0044 mol) was dissolved in dioxane (137 mL) and water (103.7 mL). The mixture was cooled to 0 °C by ice-bath and a solution of Fmoc-Cl (1.49 g, 0.006 mol) in dioxane (17.3 mL) was added slowly. The resulting mixture is stirred at 0 °C for 2 h. After TLC confirmed the reaction completion, the mixture was frozen in liquid nitrogen and lyophilized. The obtained powder was precipitated in acetonitrile then precipitated in DCM/diethyl ether to give 2.1 g (80%) of **15** as a white powder.

^{1}H NMR (400 MHz, DMSO-d_6) δ 10.21 (s, 1H), 9.61 (s, 1H), 8.24 (s, 1H), 8.01 (s, 1H), 7.92 (d, J = 7.5 Hz, 2H), 7.78 (d, J = 7.3 Hz, 2H), 7.70 (t, J = 6.1 Hz, 1H), 7.59 (s, 1H), 7.44 (t, J = 7.5 Hz, 2H), 7.36 (t, J = 7.3 Hz, 2H), 4.70 (d, J = 6.1 Hz, 2H), 4.40 (d, J = 7.0 Hz, 2H), 4.31 (t, J = 7.0 Hz, 1H), 4.02 (s, 3H), 1.43 (s, 9H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 175.4, 166.7, 165.2, 156.5, 156.4, 156.4, 150.5, 147.1, 145.9, 144.2, 142.1, 141.2, 132.5, 129.7, 129.2, 128.2, 127.6, 125.8, 120.7, 120.0, 103.1, 78.9, 66.8, 56.8, 47.0, 41.3, 28.7.

HRMS: calcd. for $C_{33}H_{33}N_4O_8$ $[M+H]^+$ 613.2293; found 613.2381.



Compound 16. 3-Amino-2-fluorobenzoic acid (5 g, 0.032 mol) was dissolved in dioxane (500 mL) and water (350 mL) and cooled to 0 °C by ice-bath. A solution of Fmoc-Cl (10.84 g, 0.042 mol) in dioxane (100 mL) was added slowly. The resulting mixture was stirred at 0 °C for 1 h and allowed to warm to r.t. overnight. After TLC confirmed the reaction completion, the solution was acidified to pH = 4 using a saturated citric acid solution. Then water was added

and the aqueous layers were extracted with DCM three times. The combined organic layers were washed with water and brine, dried (Na_2SO_4) and then concentrated under reduced pressure. The residue was precipitated in acetonitrile then precipitated in MeOH to give 9.2 g (76%) of **16** as a white powder.

^{1}H NMR (500 MHz, DMSO-d_6) δ 13.28 (s, 1H), 9.60 (s, 1H), 7.91 (dt, *J* = 7.6, 1.0 Hz, 2H), 7.75 (d, *J* = 7.6 Hz, 2H), 7.58 (ddd, *J* = 8.1, 6.5, 1.8 Hz, 1H), 7.43 (tt, *J* = 7.5, 0.8 Hz, 2H), 7.34 (td, *J* = 7.5, 1.2 Hz, 2H), 7.21 (t, *J* = 7.9 Hz, 1H), 4.45 (d, *J* = 7.0 Hz, 2H), 4.31 (t, *J* = 7.0 Hz, 1H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 164.9, 164.9, 153.9, 143.7, 140.8, 127.7, 127.2, 127.1, 127.1, 126.8, 125.3, 123.9, 123.8, 120.2, 119.9, 119.8, 66.3, 46.5.

HRMS: calcd. for $C_{22}H_{16}FNO_4Na$ $[M+Na]^+$ 400.0956; found 400.0963.

3 Spectra and chromatograms of new compounds

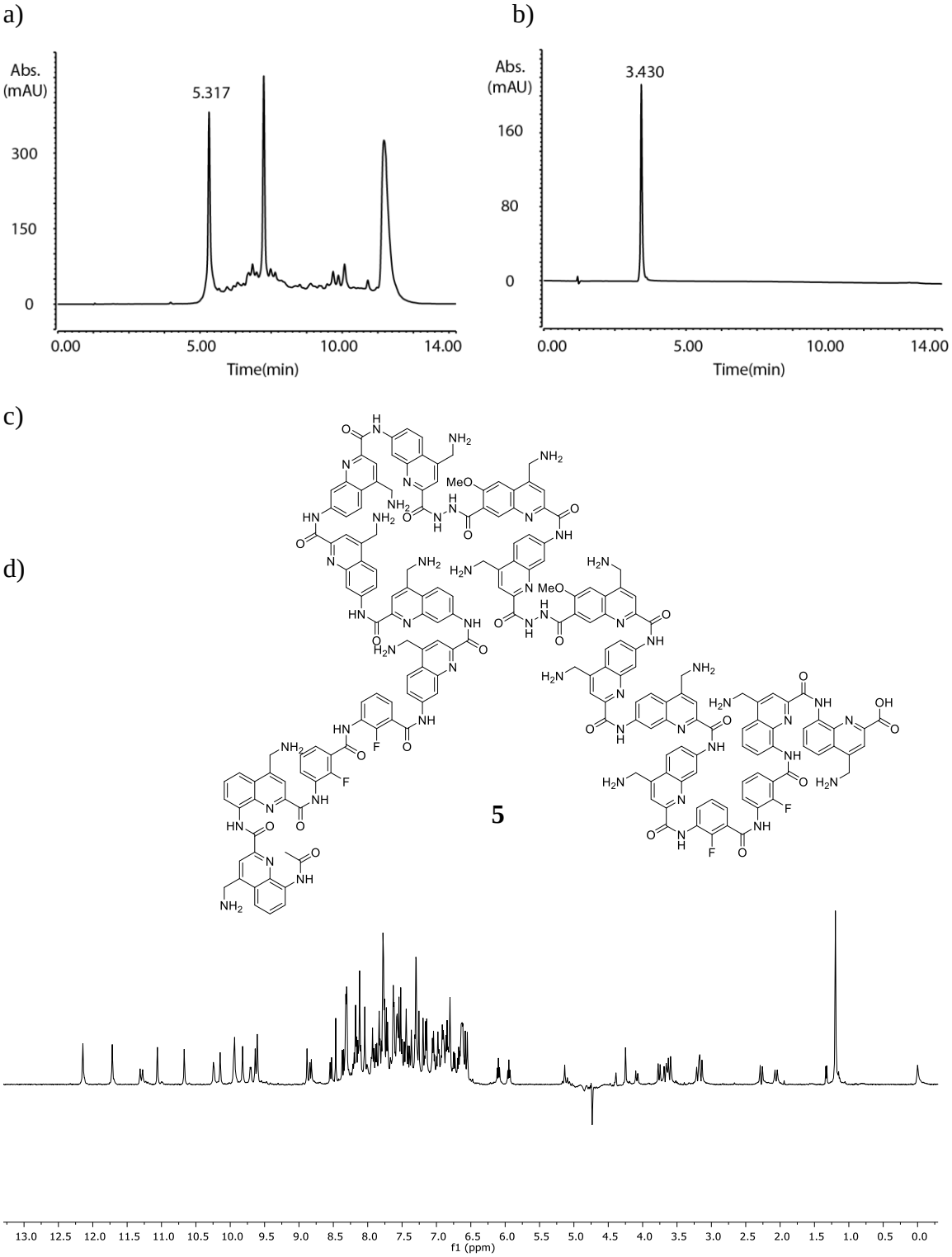


Figure S63. Analytical data of oligomer **5**. HPLC chromatograms (a) after cleavage from the resin (C8, 5 to 30 B% over 10 min, λ = 300 nm) and (b) after purification (C8, 5 to 60 B% over 10 min, λ = 300 nm); A: 0.1% TFA water, B: 0.1% TFA acetonitrile. (c) Chemical structure of oligomer **5**. (d) ^{1}H NMR spectrum with water suppression (500 MHz, H_2O/D_2O = (9:1, *v*/*v*), 25 °C).

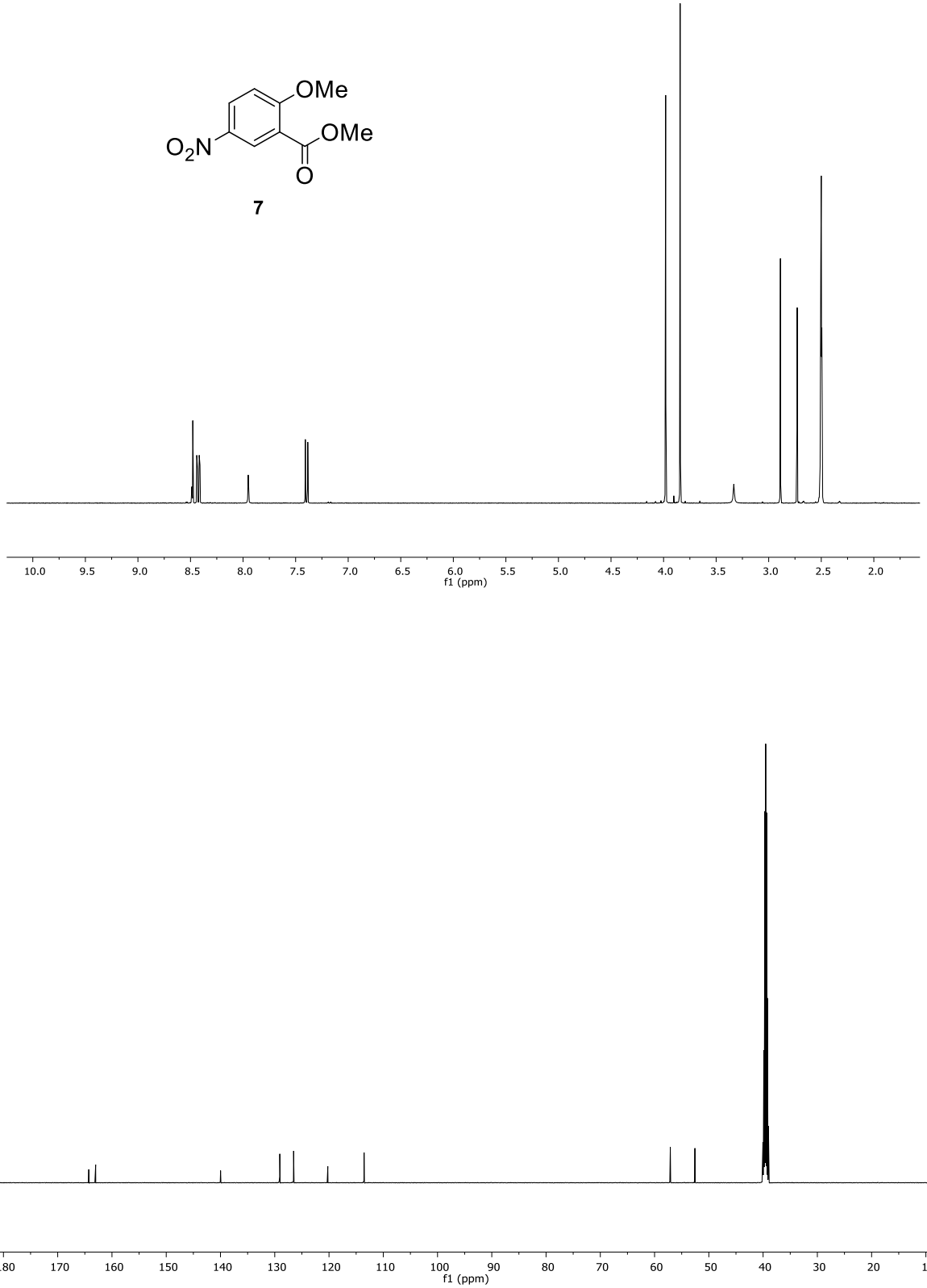


Figure S64. NMR spectra of compound **7** : ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

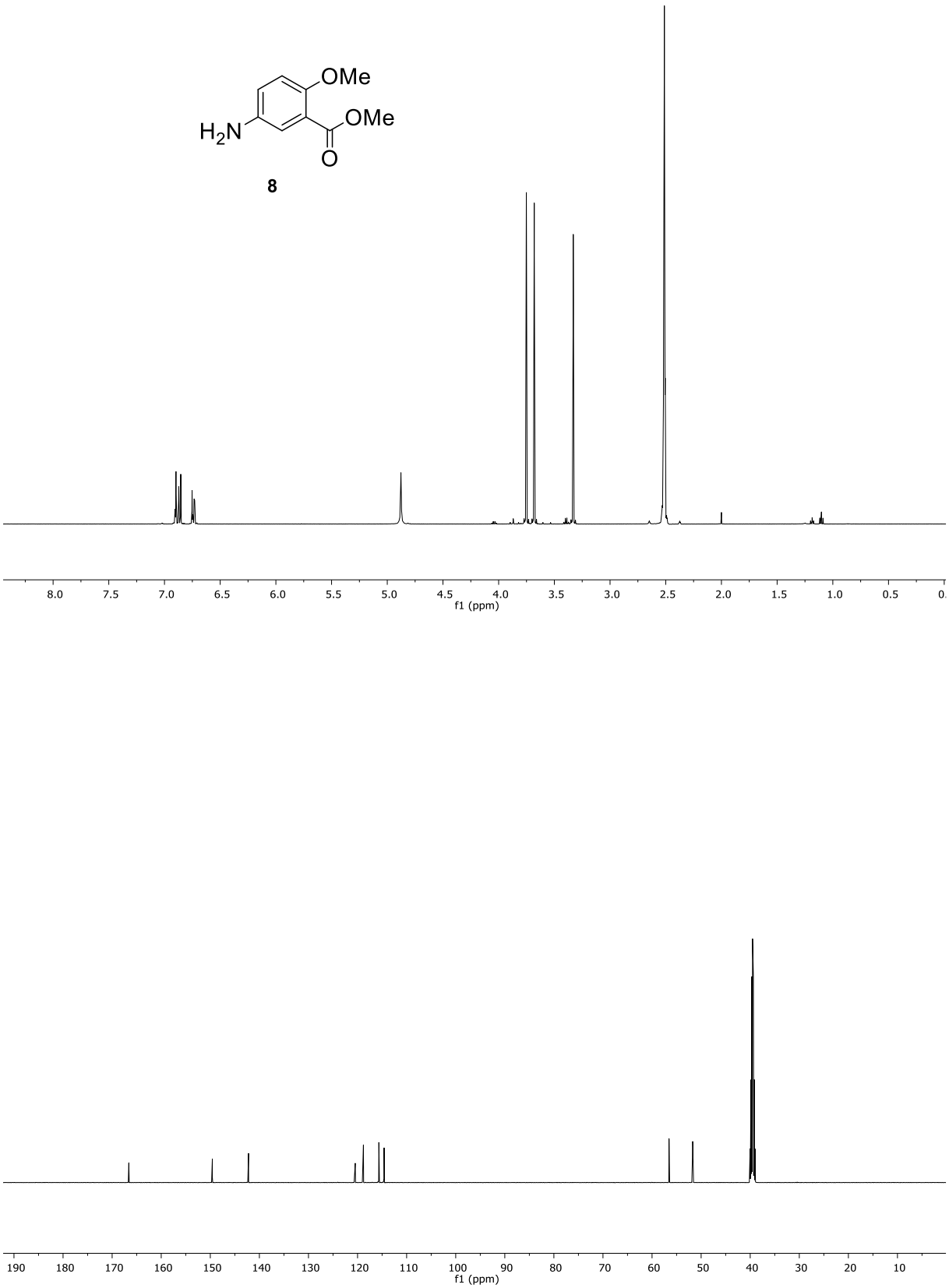


Figure S65. NMR spectra of compound **8** : ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

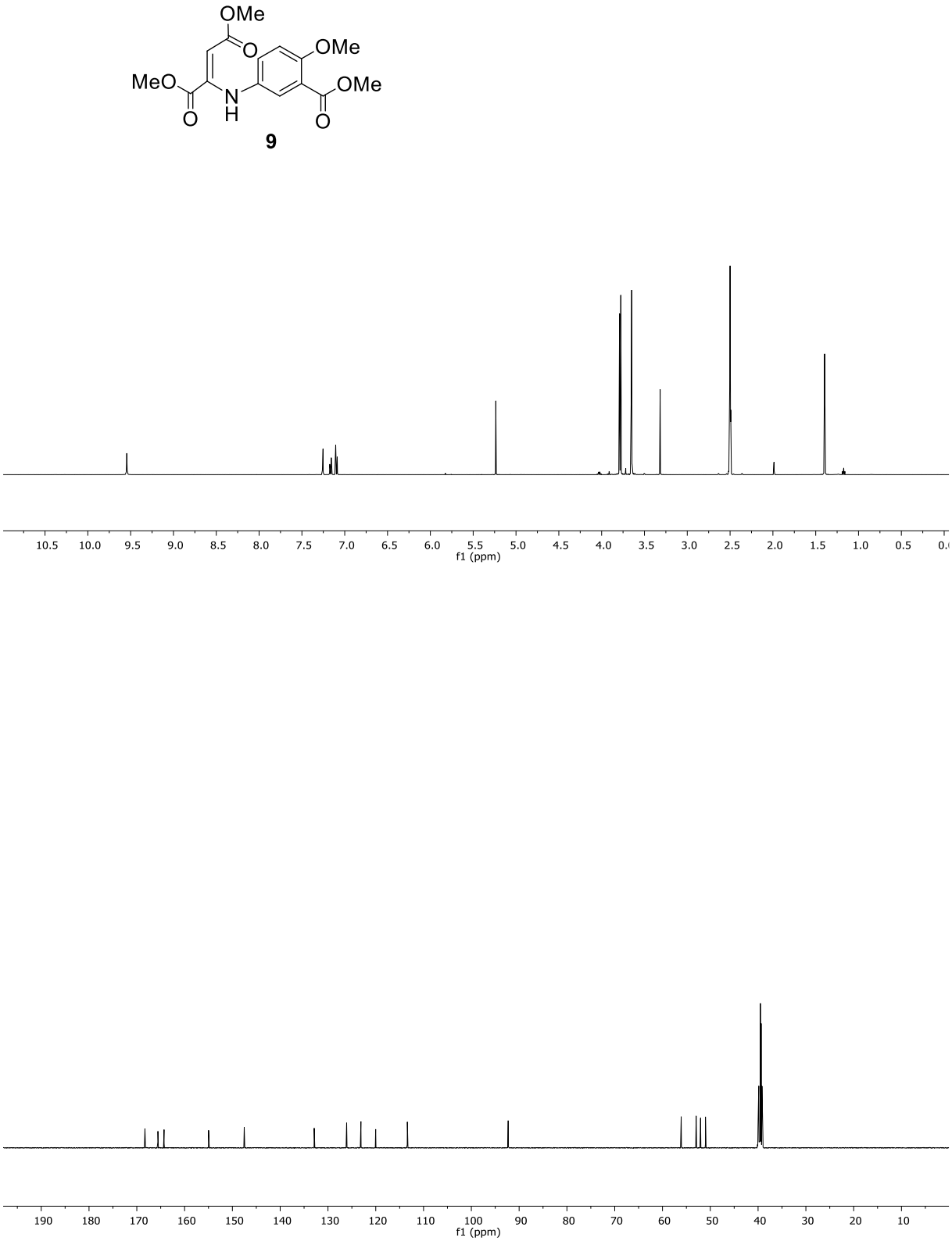


Figure S66. NMR spectra of compound **9** : ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

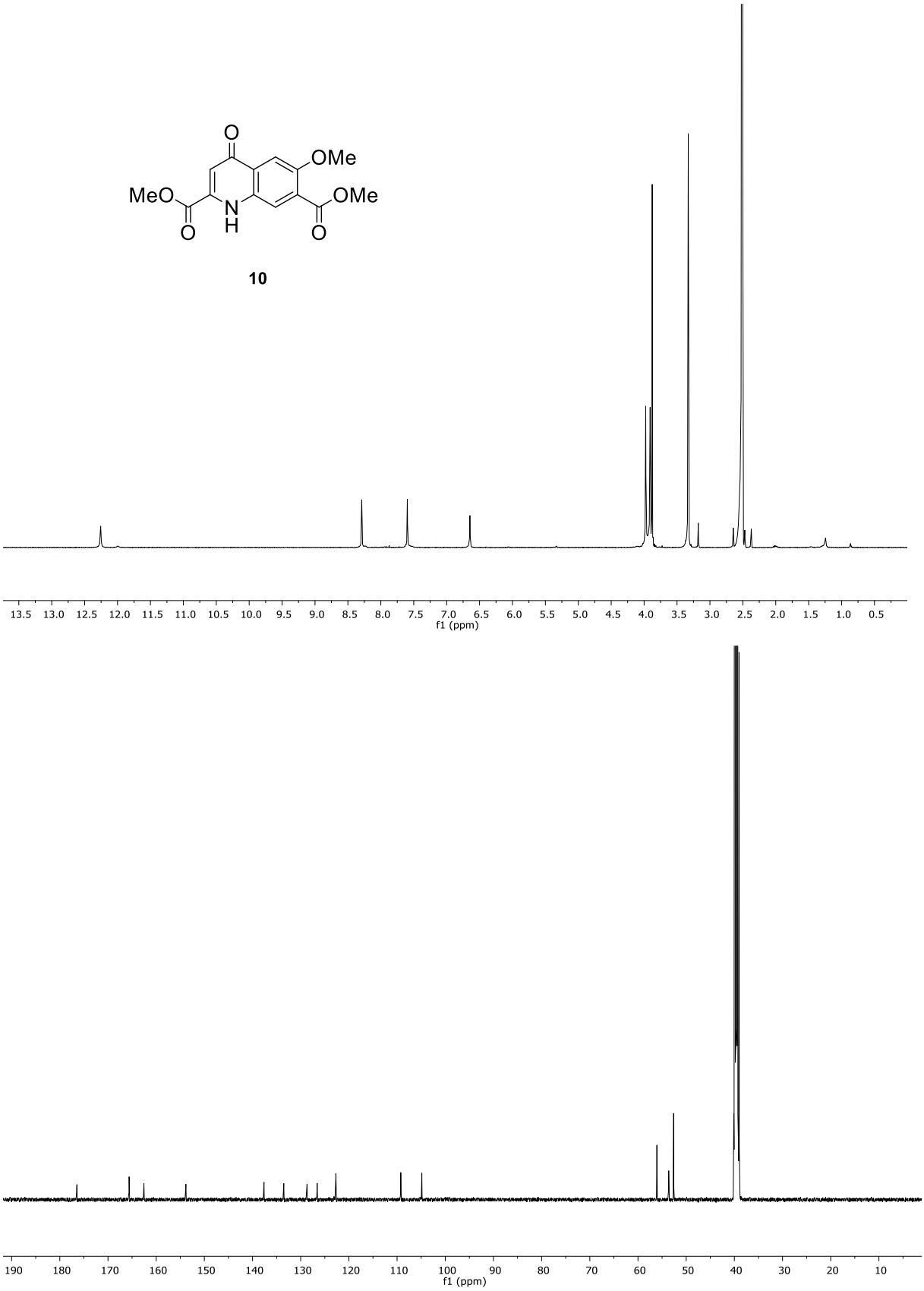


Figure S67. NMR spectra of compound **10** : ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

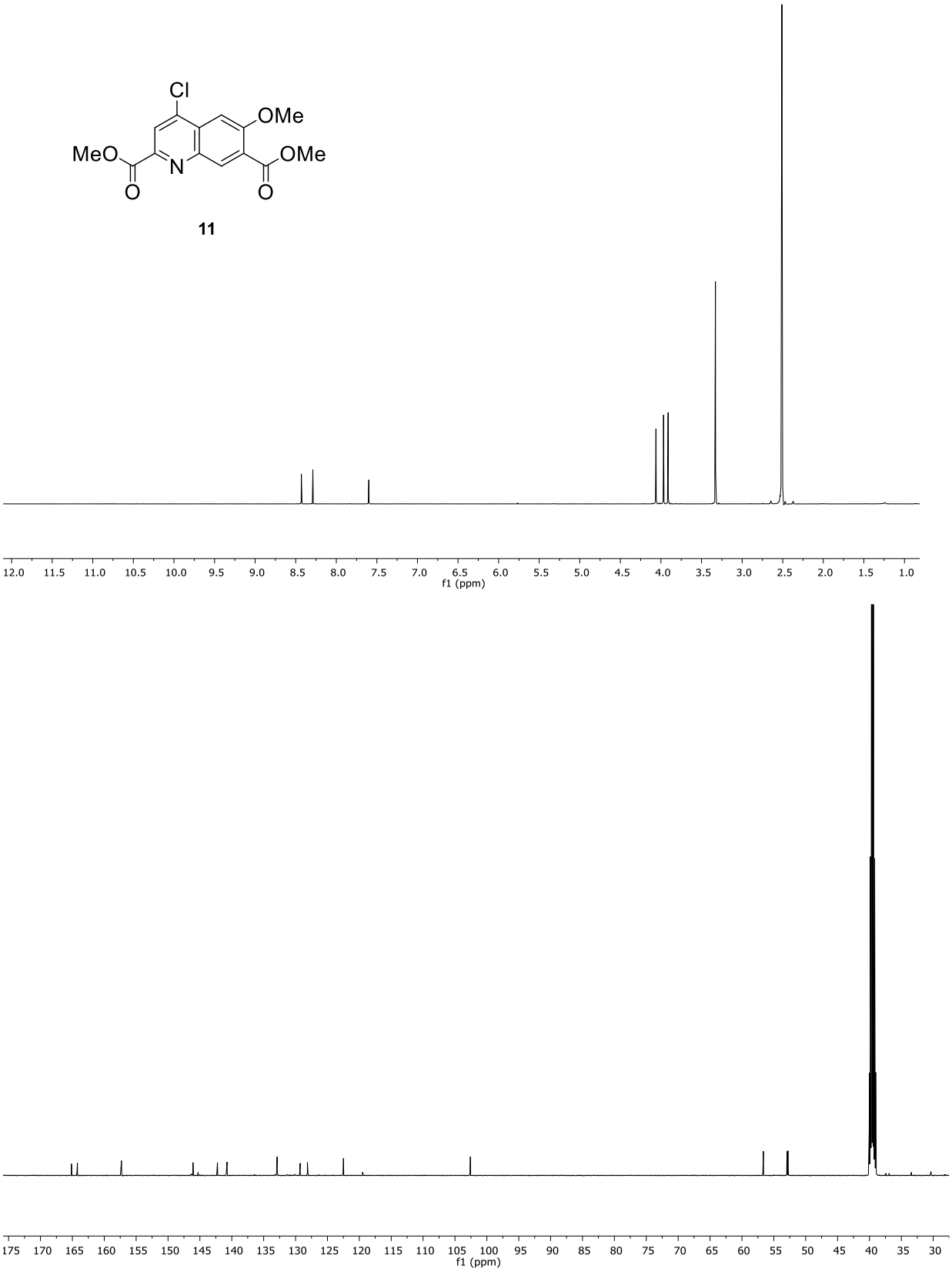


Figure S68. NMR spectra of compound **11** : ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

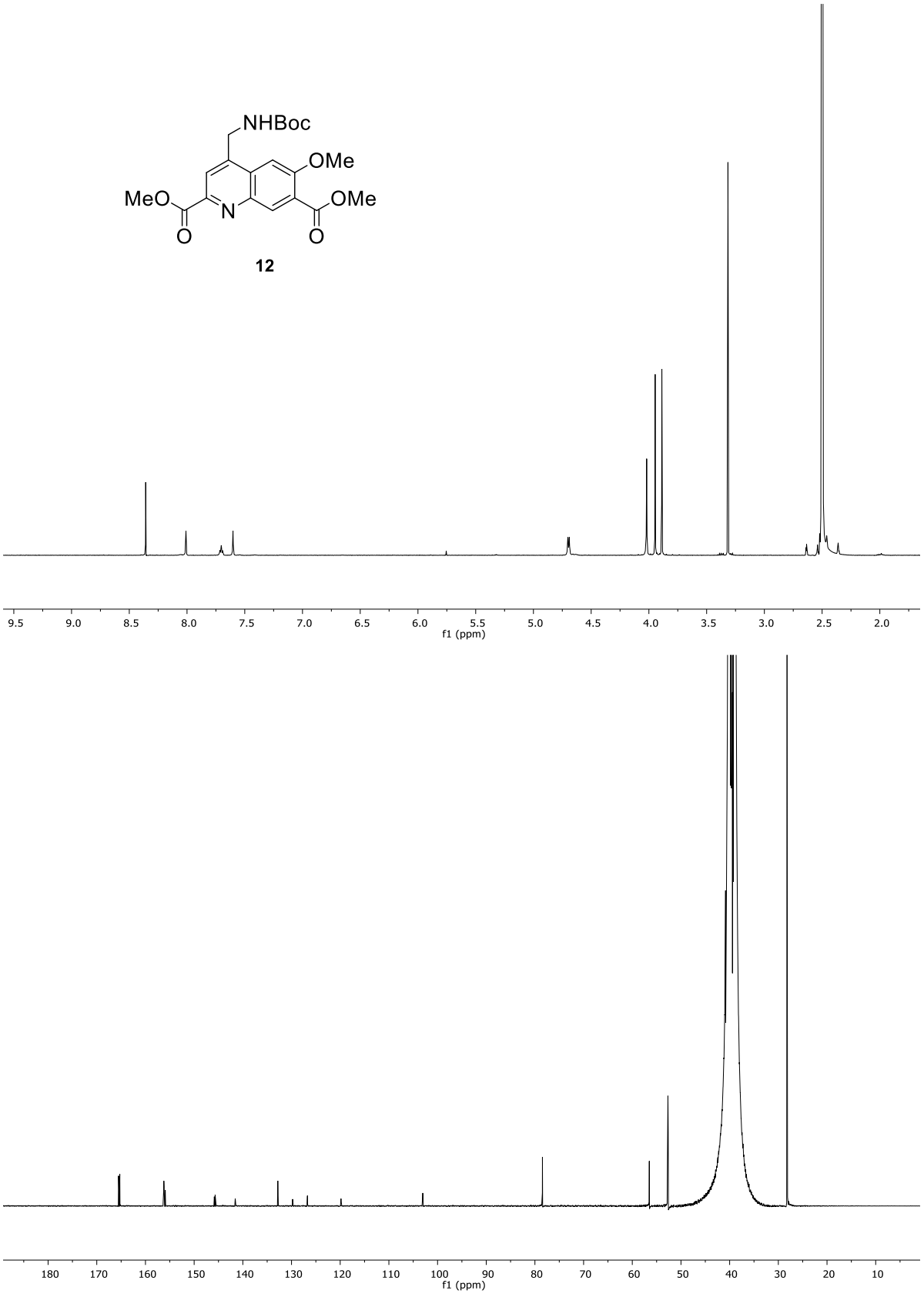


Figure S69. NMR spectra of compound **12** : ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

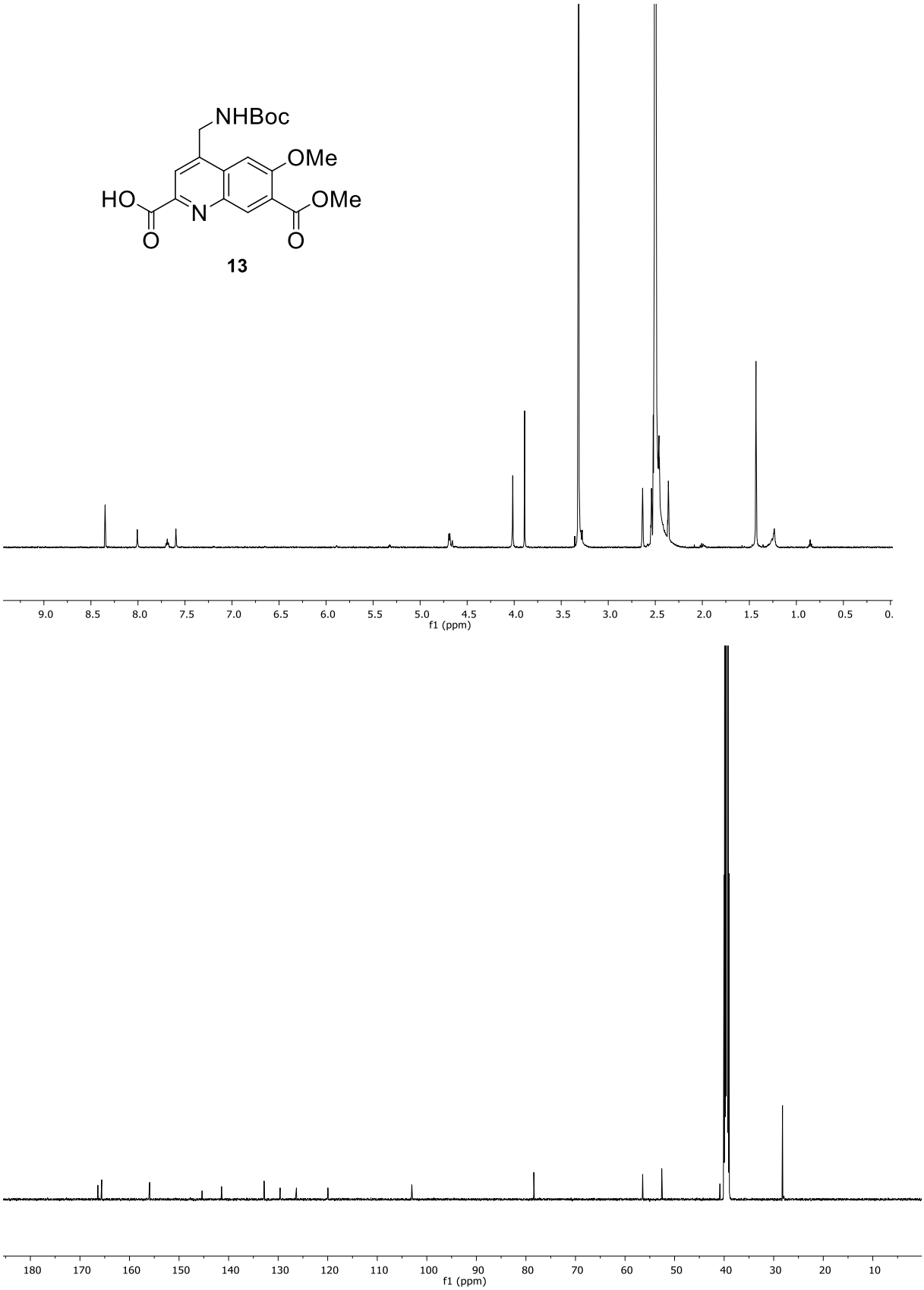


Figure S70. NMR spectra of compound **13** : ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

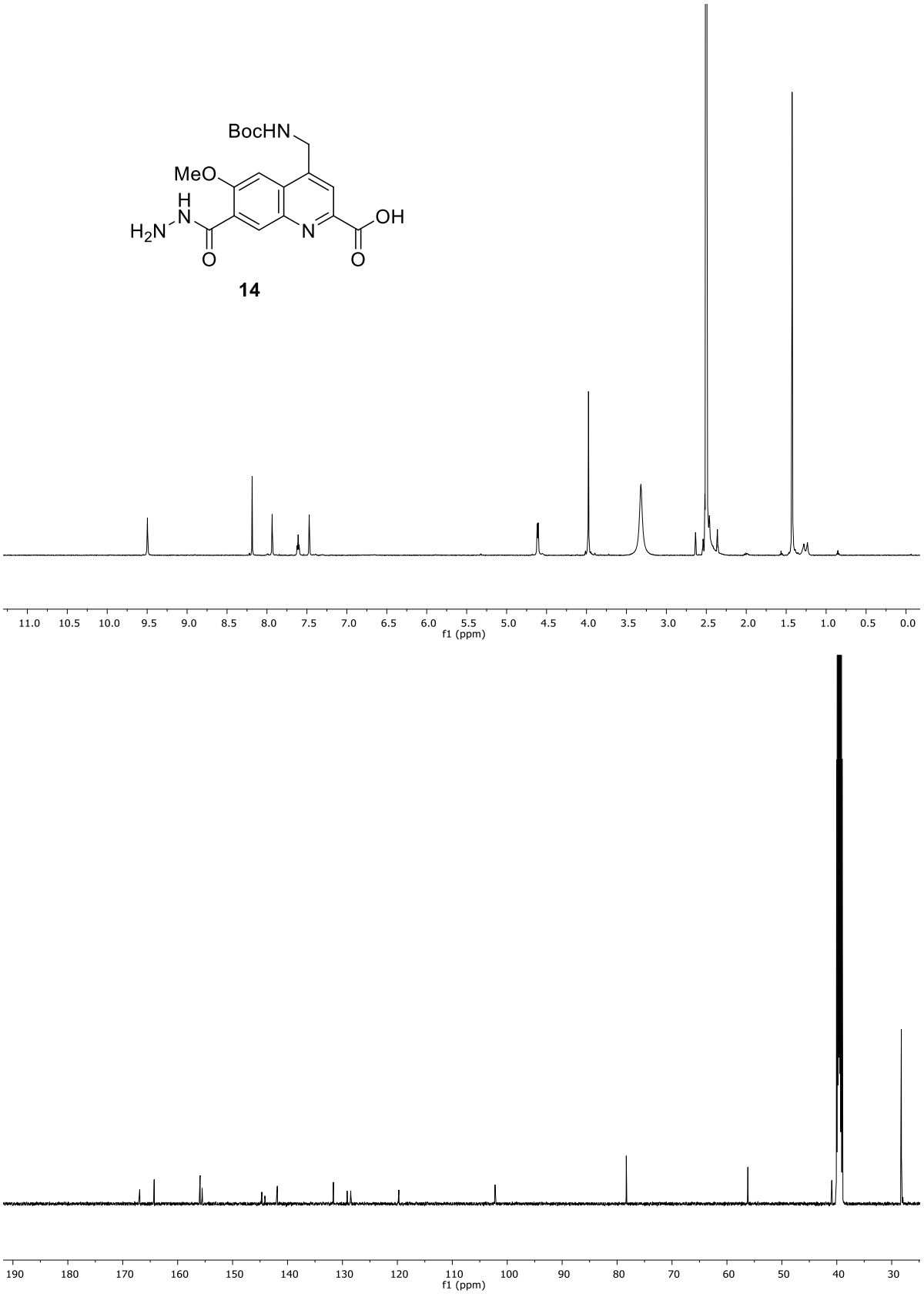


Figure S71. NMR spectra of compound **14** : ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

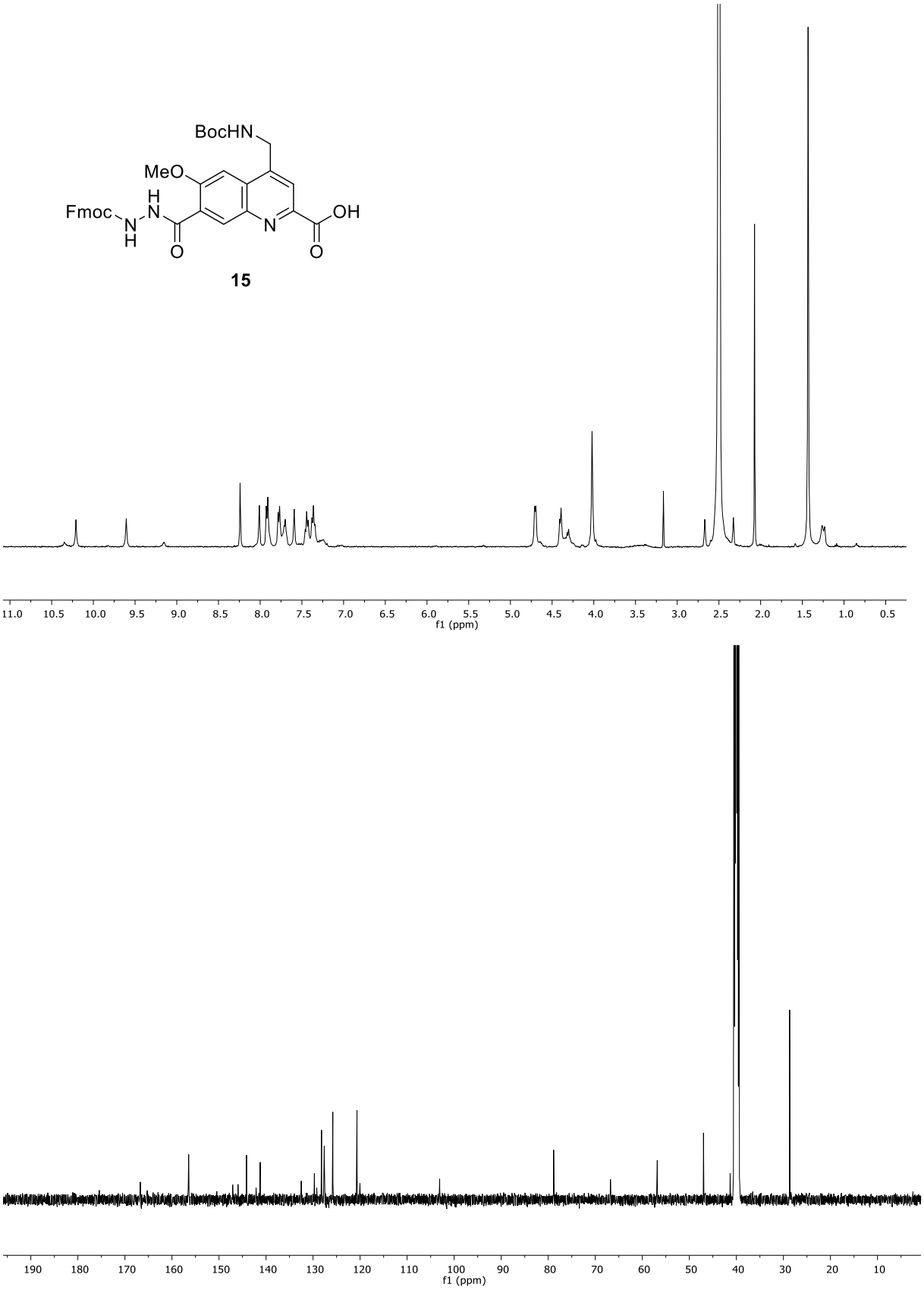


Figure S72. NMR spectra of compound **15** : ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

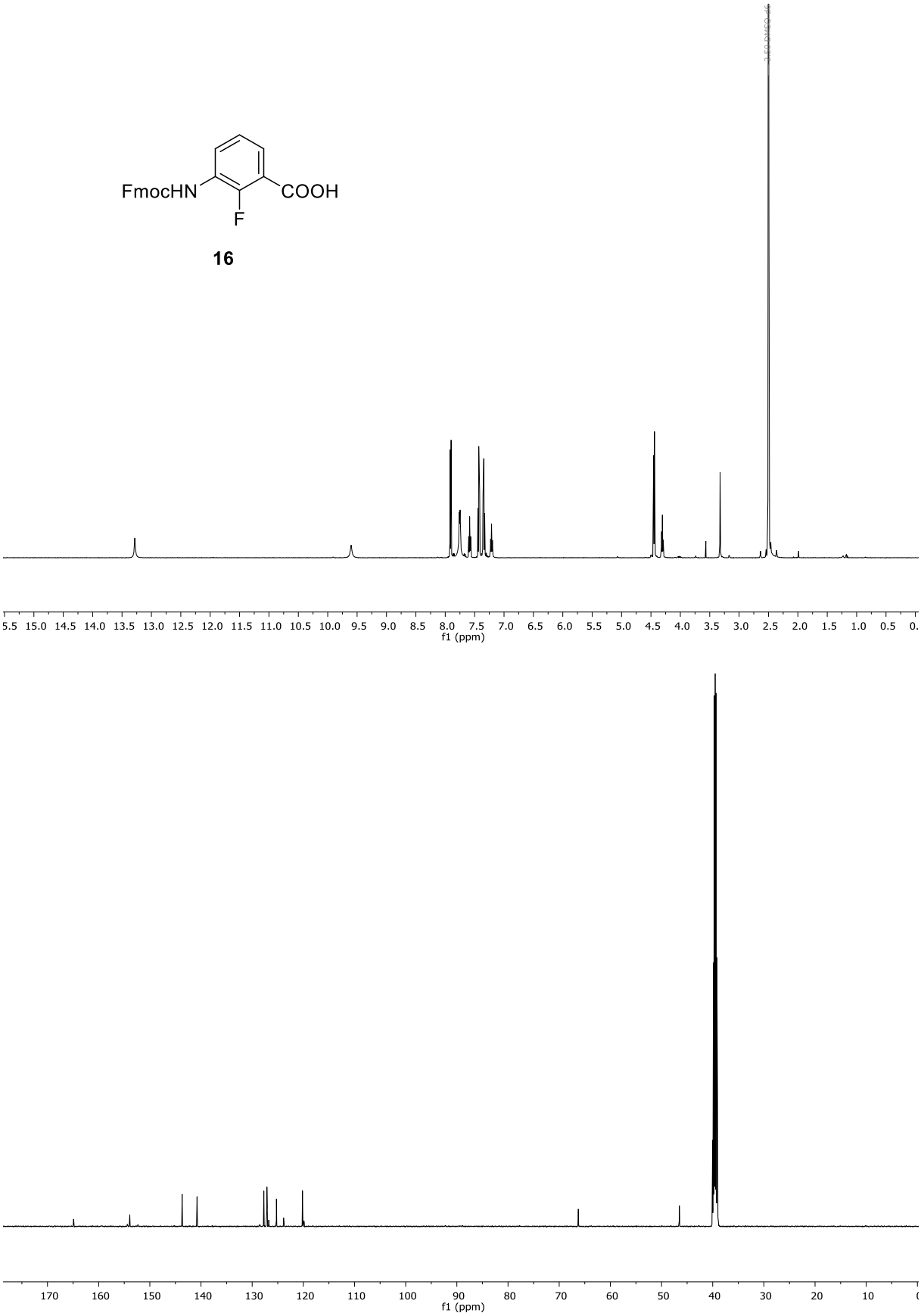


Figure S73. NMR spectra of compound **16** : ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

4 Reference

1 B. Teng, J. Atcher, L. Allmendinger, C. Douat, Y. Ferrand and I. Huc, *Org. Biomol. Chem.*, 2023, DOI: 10.1039/D3OB00473B.

2 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.

3 S. Dengler, P. K. Mandal, L. Allmendinger, C. Douat and I. Huc, *Chem. Sci*., 2021, **12**, 11004–11012.

4 V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat, and I. Huc, *Chem. Eur. J*, 2023, DOI: 10.1002/chem.202301120.

5 B. Teng, P. K. Mandal, L. Allmendinger, C. Douat, Y. Ferrand and I. Huc, *Chem. Sci*., submitedd.

6 Controlling aromatic helix dimerization in water by tuning charge repulsions

The secondary and tertiary structures of biomacromolecules are essential to their biological activity. As a crucial aggregation pattern for constructing functional structures, the dimerization method has been employed in a variety of ways. The halogen bonds could be utilized to connect two bowl-shaped segments and accomplish dimerization (Figure 15a).[8] Another strategy is increasing the helical diameter by altering backbone units (Figure 15b) which decreases the stereo hindrance between two different strands, the π-π stacking and hydrophobic effect will facilitate double helix formation automatically.[80] In addition, using a type of force to bind two strands together is an effective approach; salt bridges[81] and metal ion chelation[82] can function as molecular glue to facilitate dimerization (Figure 15c, d). Although many investigations have been done on dimerization, side-chain optimization studies that facilitate dimer formation is still rare. To obtain helical capsule crystals in water, the side chain variation sequences were synthesized, and surprisingly, multiple unique dimerization patterns were observed without any backbone modification.

Although the investigation of oligoamide helical capsules in the solution state was conducted effectively in water, solid-state research was not possible due to the extremely water-soluble aminomethyl side chain. Thus, a new synthetic route needs to be developed for solubility optimization. The established synthetic protocol, which is employed to introduce sulfonic and methoxy side chains, was used to synthesize negatively charged and neutral Q^H monomers. The whole sequence synthesis was performed by solid phase synthesis with in situ Appel condition. The first aromatic oligoamide foldamers capsule crystal that grew from water was obtained when negative side chain was performed.

As described in the following submitted in *Chemical Science*, these dimerization patterns were facilitated by electron repulsion and hydrophobic effect. A head-to-head dimer was produced by substituting methoxy side chains at both ends, which decreased the charge at the C terminal. After introducing a stereo hindrance to prevent head-to-head dimerization, the other two methoxy side chains at the central Q^H segment resulted in the formation of parallel double helices. This aggregation pattern had never previously been observed. The hemi-capsule sequences with flexible backbone can also form double helices in water. By variation of side chains, a new aggregation species appeared. Using an energy-minimizing model, the tetrameric

helix model was developed in relation to the crystal structure of the parallel double helix. In conclusion, our study highlighted the distinct aggregation behaviors in water for flexible backbone foldamers, which will guide the design of oligoamide helical capsules in the future.

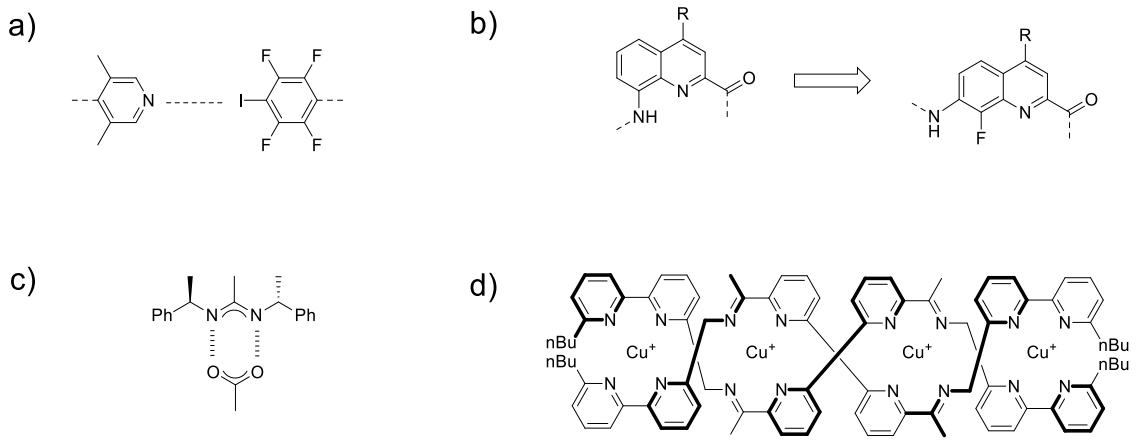


Figure 15. Overview of different strategies to carry out dimerization. a) Link two segments by halogen bonding. b) Increase the helical diameter to promote double helix formation. c) Construct double helix by introducing salt bridges between two strands. d) Promote dimerization by utilizing guests as a molecular glue.

Contribution: I conceived the ideas for this project in collaboration with I. Huc. Synthetic routes and conditions for sulfonic and methoxy side chains Q^H monomers were developed by me. Oligomer synthesis and analysis were done by me. Energy minimized model was done by me. L. Allmendinger was involved in the design and execution of NMR experiments. Crystallization experiments were conducted in collaboration with P. K. Mandal. X-ray measurements were done with assistance of Dr G. Leonard, and the structure was refined by P. K. Mandal. The automated solid phase synthesis was done in collaboration with C. Douat. The manuscript was written by me and in collaboration with I. Huc and P. K. Mandal.

ARTICLE

Controlling aromatic helix dimerization in water by tuning charge repulsions

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Binhao Teng,[a] Pradeep K. Mandal,[a] Lars Allmendinger,[a] Céline Douat,[a] Yann Ferrand[b] and Ivan Huc*[a]

Several helically folded aromatic oligoamide foldamer sequences were designed and synthesized. The sequences were all water-soluble thanks to the charged side chains borne by the monomers. Replacing a few, sometimes only two, charged side chains by neutral methoxy groups was shown to trigger various types of aggregation processes leading to the formation of head-to-head stacked dimers of single helices, double helical duplexes and a quadruplex, none of which would form in organic solvent with organic soluble analogues. The nature and structure of the aggregates were supported by concentration and solvent dependent NMR studies, ^{1}H DOSY experiments, mass spectrometry, and X-ray crystallography or energy-minimized models. The hydrophobic effect appears to be the main driving force for aggregation but it can be finely modulated by the presence or absence of a small number of charges to an extent that had no precedent in aromatic foldamer architectures. These results will serve as a benchmark for future foldamer design in water.

Introduction

Molecular aggregation in water often rests on a balance between hydrophobic effects that act as a driving force, and hydrophilic interactions mediated by polar groups without which molecules may not be water-soluble. When hydrophilic groups consist of ions, it is generally beneficial that a net total charge exists. Zwitterions may cancel each other's hydrophilic interactions and tend to be poorly water-soluble. Illustrations of this hydrophobic-hydrophilic balance can be found in objects as diverse as a simple charged surfactant and a polyanion such as DNA. In both cases, one may say that aggregation – micelle formation for the former and double helix association for the latter – occurs 'despite' electrostatic repulsions between charges. Tuning these repulsions may in turn result in changes in aggregation properties. For instance, upon screening electrostatic interactions, spherical micelles may transit to cylindrical micelles or bilayer phases.[1] Conversely, DNA forms more stable duplexes with its neutral analogue PNA than with itself.[2] Here, we show that minor changes in the hydrophobic-hydrophilic balance of helical aromatic foldamers result in great variations of their propensity to form different types of discrete aggregates in water.

Aromatic helices are an important class of foldamers with aryl groups in their main chain. They have attracted attention for their ability to serve as scaffolds for molecular recognition either in their interior[3–9] or at their surface,[10–12] in ion[13–15] and water[16–18] transport through membranes, in molecular machinery,[19,20] and for their charge transport[21,22] and chiroptical properties.[23,24] Aromatic helix folding may be driven by solvophobic effects,[25–31] the presence of neutral or ionic guests,[32–36] steric congestion combined with aryl-aryl interactions,[37,38] or local conformational preferences.[39–43] Aromatic helices may undergo different types of aggregation: stacking into continuous[44] or discrete[45] tubular objects, side-by-side bundling,[46] or intertwining in doubly,[9,47–54] triply,[36,55,56] or quadruply[57,58] stranded structure. Helix aggregation may contribute to the properties mentioned above. For example, double helices may also act as molecular containers.[9,19,59,60] Double helix formation has been shown to depend on solvent[52,53,61] and on the presence of guests.[62,63] A few reports have also highlighted the role of interactions between side chains.[50,51] However, aggregation has been investigated mostly in organic solvents. Reports on aggregation in water are few[45,52–54] and have not made use of the hydrophobic-hydrophilic charge balance. We now report that removing as few as two or four out of fourteen positive charges from a helical oligoamide can drive its association into discrete stacked dimers or double helical dimers, respectively. In a shorter sequence, removing four out of seven charges allows for the formation of a tetramer.

Results and discussion

Design and synthesis

Q^R and Qf^R monomers bearing various R solubilizing groups in position 4 have been widely investigated as building blocks of helically folded oligoamides (Fig. 1a).[4,39,45,46,53,58,64] In these sequences, folding is driven by local electrostatic attractions and repulsions between the amide NH or CO and neighboring quinoline endocyclic nitrogen atoms or exocyclic fluorine atoms that create a conformational bias at aryl-amide rotatable bonds, resulting in a well-defined curvature of the aromatic strand and eventually promoting a helical shape.[41,42] These helices are stable in all kinds of solvents and are particularly stable in water due to hydrophobic

[a] Department of Pharmacy, Ludwig-Maximilians-Universität, Butenandtstr. 5–13, 81377, München, Germany. E-mail: ivan.huc@cup.lmu.de
[b] Univ. Bordeaux, CNRS, Bordeaux Institut National Polytechnique, CBMN UMR 5248, 2 rue Escarpit, 33600 Pessac, France.

†Electronic Supplementary Information (ESI) available: Synthetic protocols, crystallographic studies and characterisation of new compounds. CCDC 2216788. For ESI and crystallographic data in CIF or other electronic format See DOI: 10.1039/x0xx00000x

ARTICLE

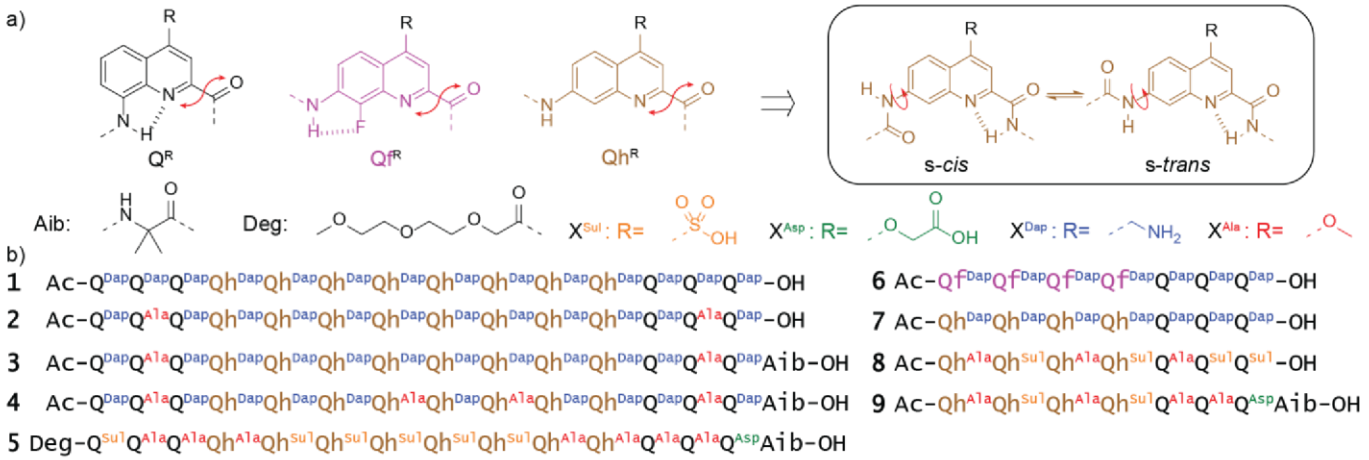


Fig. 1 a) Structure of the building blocks Q^R, Qf^R and Qh^R bearing R side chains in position 4 and composition of sequences **1**-**9** investigated in this study. Different quinoline monomers and side chains are highlighted with different colors. Monomers: Q: black, Qh: brown, Qf: purple; side chains: Dap: blue, Ala: red, Sul: gold, Asp: green). X^R refers to any Q^R, Qh^R or Qh^R quinoline unit bearing an R side-chain. b) Sequences prepared for this study.

effects associated with aromatic stacking.[65] We recently introduced Qh^{Dap} as an analogue of Qf^{Dap} lacking the conformational control associated with the fluorine atom of Qf, which gives rise to a conformational equilibrium at the 7-carbonylamino-quinoline bond of Qh.[66] Using sequence **1**, we showed that this additional conformational degree of freedom hampers folding in DMSO but that hydrophobic effects still promote helix folding in water, thus favoring the s-*trans* conformation of each Qh unit. Sequence **1** comprises a central $(Qh^{Dap})_8$ segment flanked with $(Q^{Dap})_3$ segments. Qh, like Qf, was expected to generate a wider helix diameter possibly conducive of dimerization into multistranded helices through intertwining of the strands.[53,58] In contrast, Q_n oligomers possess a narrower diameter and do not form multiple helices. Also, they have been shown to prevent double helix formation of wider helical segments placed between them.[4,67] Indeed, such a double helix would have two Q_n segments at each end (one at each end of both strands), which would expectedly create steric hindrance and perturb folding.[67] As expected, sequence **1** was shown to be a single helix in water.[66] It may possess a cavity potentially capable of encapsulating guest molecules, as in designs that we and others have presented previously.[4,7,30] Its molecular recognition properties will be described elsewhere. The structure of **1** in solution was elucidated using ^{1}H NMR spectroscopy, but attempts to produce crystals suitable for X-ray diffraction analysis were unsuccessful. We hypothesized that the high charge density at the surface of single helical **1** was, at least in part, responsible for the difficulty to produce crystals. In the context of this work, new sequences **2**-**5** were thus synthesized as analogues of **1** in which two, four or seven charged side chains were replaced by neutral methoxy groups in the hope to obtain crystals. In the case of **5**, the seven remaining charged side chains were all made anionic. As shown in the following, these variations gave rise to various types of unanticipated aggregation behavior (and also to one crystal structure).

In the meantime, sequences **6**-**9** were prepared to purposely investigate possible double helix formation. Sequence **6** possesses Qf monomers that encode helicity regardless of the solvent. Analogues of **6** have been shown to form double helical dimers, including in water.[53] Indeed, this sequence contains only one Q_3 segment and may thus accommodate two of them in a head-to-tail double helical dimer. Because of their additional degree of conformational freedom, it was unknown whether Qh monomers would also undergo double helix formation. Sequence **7**, was thus also prepared, that amounts to the seven C-terminal residues of **1**. Sequences **8** and **9** are analogues of **7**, bearing anionic residues and three or four methoxy side chains.

The preparation of these new sequences entailed the synthesis of two new monomers ready for oligoamide synthesis on solid phase, Fmoc-Qh^{Ala}-OH and Fmoc-Qh^{Sul}-OH. Details of their preparation is reported in the supporting information (Schemes S1,S2†). All sequences were synthesized on low loading Wang or Cl-MPA Protide resins using either manual or automated protocols.[68,69]

Head-to-head dimerization of single helices

In order to reduce water solubility and promote crystal growth in water, oligomer **2** was produced as an analogue of **1** in which two charged aminomethyl side chains were replaced with neutral and isosteric methoxy groups near the ends. Crystals suitable for X-ray diffraction analysis could not be obtained. In contrast with **1**, the ^{1}H NMR spectrum of **2** in water showed two sets of sharp signals, the proportion of which varies with concentration (Fig. 2a), therefore

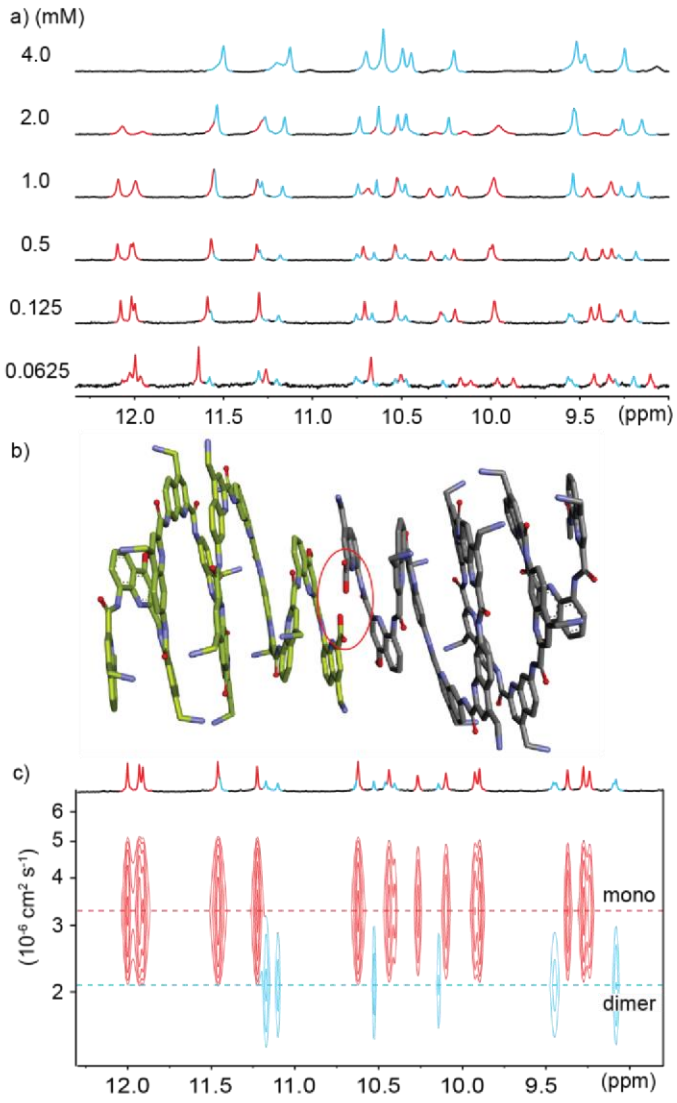


Fig. 2 a) Excerpts of the ^{1}H NMR (500 MHz) spectra of compound **2** at different concentrations in 45 mM sodium acetate aqueous buffer pH 4.0 (D_2O/H_2O, 10:90 v/v). Selected amide proton signals belonging to either the monomeric or dimeric form are highlighted in red and blue, respectively. b) Energy minimized model (Maestro, MMFFs force field, water as implicit solvent)[71] of a dimer of **2** formed by stacking two C-terminal cross-sections. The two distinct molecules engaged in binding are colored in light green and gray, respectively. The carboxylic acid and carboxylate involved in intermolecular hydrogen bonding at the C-termini are circled in red. c) Amide region of the ^{1}H DOSY NMR (500 MHz) spectrum of compound **2** (0.49 mM) in the same solvent as in a). The two different hydrodynamic radii related to different levels of the signals in the ^{1}H DOSY spectrum are indicated by dashed lines. Signals corresponding to the monomeric or dimeric form are highlighted in red and blue, respectively.

demonstrating a process of reversible aggregation in slow exchange on the NMR timescale. The different sizes of the two species were highlighted by a ^{1}H DOSY experiment clearly showing different hydrodynamic radii (Fig. 2c). The number of resonances indicated that the aggregate is symmetrical, *i.e.*, its helical subcomponents are in the same environment. The signals of the aggregate are upfield-shifted, suggesting that aggregation involved aromatic stacking. The dimeric species $[2M+4H]^{4+}$ and $[2M+5H]^{5+}$ were detected in the gas phase by ESI-MS (Fig. S1†) indicating that the aggregation is probably a dimerization. Upon integrating the NMR signals of the two species, a dimerization constant of 970 M^{-1} was calculated.‡

The above observations are in principle compatible with either the formation of a symmetrical double helix,[51,53,70] or with head-to-head stacking of two single helices as reported for Q_n oligomers bearing anionic residues.[45] Evidence in favor of the latter came from the behavior of **3**, an analogue of **2** bearing a C-terminal Aib. The ^{1}H NMR spectrum of **3**, shows a single set of signals at chemical shift values very similar to those of the monomer of **2** (Fig. S2) suggesting that the Aib has suppressed dimerization. This C-terminal extension would not be expected to much alter double helix formation, but it is not compatible with the stacking of the C-terminal cross-sections of two helices which normally requires direct hydrogen bonding between a carboxylate and a carboxylic acid of the two strands involved (Fig. S3).[45] Furthermore, the ratio of monomer and dimer of **2** was shown to reach equilibrium within minutes after a dilution (Fig. S4), whereas kinetics would be expected to be much slower – of the order of hours or days – for a long double helix.[42] Based on these experiments, a plausible energy-minimized model of a stacked dimer of **2** was constructed (Fig. 2b). It is quite remarkable that the proximity between two charges at these positions is sufficient to completely suppress dimerization. These findings thus extend the formation of stacked Q_n helix dimers to cationic sequences.

We also found that dimerization was enhanced by increasing buffer concentration, which we interpreted as the result of an enhanced polarity of the solvent that promotes aggregation via a stronger hydrophobic effects and better screening of charge repulsions. In an attempt to decrease charge repulsion by neutralizing some cationic residues, the pH was raised to 5.6 (sodium acetate buffer, 45 mM), but this led to precipitation and no NMR signals could be recorded.

Double helices may accommodate four Q_3 segments

Using a C-terminal Aib extension to avoid head-to-head stacking of single helices, sequence **4** was designed to further reduce its water solubility and favor crystal growth. Thus, two additional positively charged side chains were replaced by two neutral methoxy groups. Unexpectedly, this change led to another kind of aggregation. Despite the Aib termination, the ^{1}H NMR spectrum of compound **4** in water again revealed two sets of sharp signals whose proportions vary with concentration (Fig. 3a). ^{1}H DOSY NMR confirmed that the aggregate – the species more abundant at high concentration – has a larger hydrodynamic radius (Fig. 3b). Furthermore, dimeric species $[2M+4H]^{4+}$ and $[2M+5H]^{5+}$ were detected in the gas phase by ESI-MS (Fig. S5) hinting at a dimerization process again. Integration of the ^{1}H NMR signals allowed for the calculation of a dimerization constant of 32500 M^{-1} at pH 4 in 13.5 mM sodium acetate buffer.‡ The dimerization of **4** was enhanced upon increasing buffer concentration more significantly than the dimerization of **2**, $(\mathbf{4})_2$ being formed quantitatively at 0.5 mM in 45 mM sodium acetate while it represents only 76% at the same concentration in 13.5 mM sodium acetate (Fig. 3a).

The exchange between monomeric and dimeric **4** was found to be much slower than between monomeric and dimeric **2**. After a dilution, several hours of incubation were required to reach equilibrium, *i.e.* until proportions stop evolving (Fig. S6). This indicates a larger kinetic barrier between the two species, that is, a change in molecular structure more significant than just stacking two single helices. We tried to obtain crystals of **4** to elucidate its

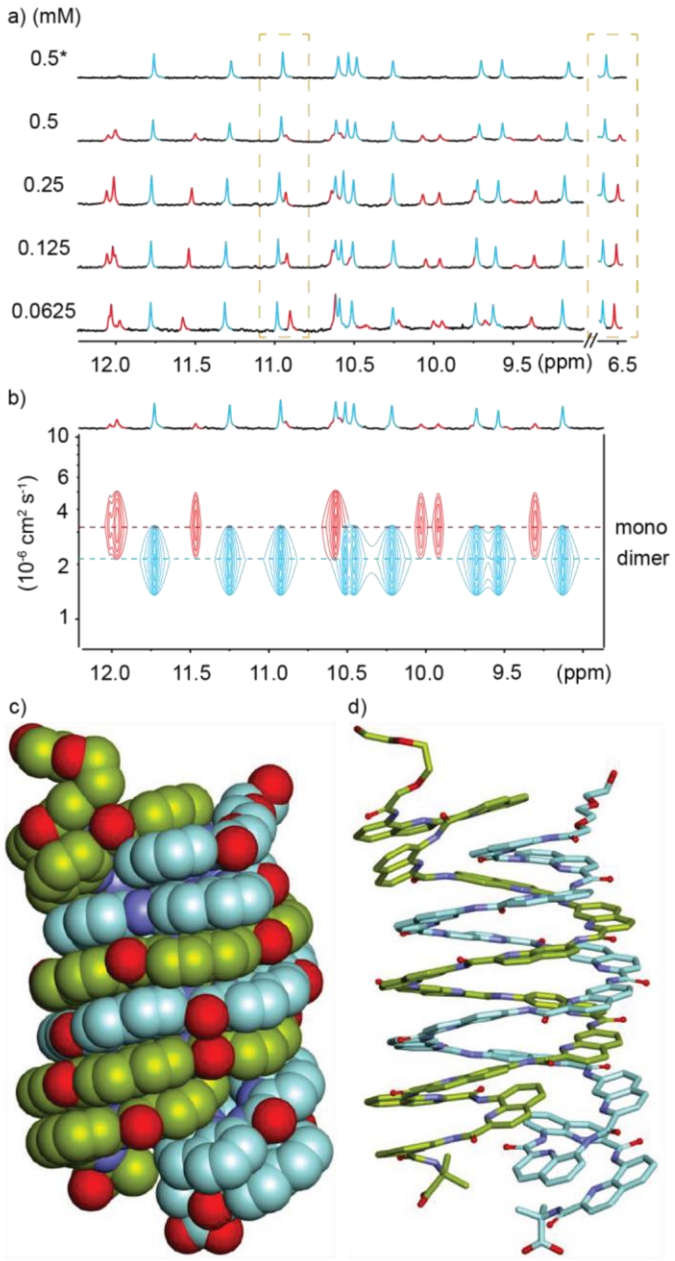


Fig. 3 a) Excerpts of the ^{1}H NMR (500 MHz) spectra of compound **4** at different concentrations in 13.5 mM (* means in 45 mM) sodium acetate aqueous buffer pH 4.0 (D_2O/H_2O, 10:90 v/v). Selected amide proton signals belonging to either the monomeric or dimeric form are highlighted in red and blue, respectively. b) Amide-region of the ^{1}H DOSY NMR (500 MHz) spectrum of compound **4** (0.5 mM) in 13.5 mM sodium acetate aqueous buffer pH 4.0 (D_2O/H_2O, 10:90 v/v). The two different levels of signals related to different molecularities in the ^{1}H DOSY spectrum are indicated by dashed lines. Signals corresponding to the monomeric or dimeric form are highlighted in red and blue. Solid-state structure of $(\mathbf{5})_2$: Side view shown in c) CPK and d) stick representation for different strands. In both representations, two distinct strands are colored in light green or cyan, respectively. Side chains have been omitted for clarity.

structure and aggregation mode but without success, as for all cationic oligomers in our hands so far. We therefore turned to anionic derivatives that have crystallized more readily in other systems[68] and prepared oligomer **5**. This sequence was not conceived as a direct analogue of **4**, as it contains more and differently located neutral methoxy side chains. The solution behavior of **5** is also more complex and its NMR spectrum shows broader peaks (Fig. S7). Nevertheless, single crystals of **5** that diffracted at atomic resolution could be obtained, revealing a parallel double helix that forms despite the presence of Q_3 segments at each end of the two strands (Fig. 3c, Fig. S8), and despite the greater flexibility of Qh-containing sequences.

The double helix of $(\mathbf{5})_2$ only involves the central Qh_8 segments and spans two full turns. The two strands are offset by half a turn and engage in extensive intermolecular aromatic stacking. The four Q_3 segments are all single helical and the steric crowd that they create is somehow accommodated. For that, two outer Q_3 segments are slightly tilted away from the duplex and overlap with two inner Q_3 segments. This arrangement suggests that segments having more than three Q units may not be accommodated as easily, but this hypothesis was not tested. The structure also shows that terminal Aib units play no role in aggregation. A molecular model of the single helical monomer of **5** was built (Fig. S9) and its solvent accessible surface was calculated to be 1790 Å^2, compared to 2690 Å^2 for the duplex solid state structure. Thus, the duplex surface amounts to about 75% of the surface of two monomers, giving a measure of how much is hidden from water upon dimerization. Of note, the double helix $(\mathbf{5})_2$ is parallel and lacks any symmetry. Its two strands are inequivalent. In contrast, the number of ^{1}H NMR resonances of $(\mathbf{4})_2$ suggests a symmetrical structure (Fig. 3a). Dimer $(\mathbf{4})_2$ may instead be a symmetrical anti-parallel dimer, or perhaps its NMR spectrum represents an average between two degenerate parallel conformations. In any case, we think it reasonable to propose that the crystal structure of $(\mathbf{5})_2$ informs about the aggregation mode of **4**. We also note that the structure of $(\mathbf{5})_2$ shares a key design feature with **1**-**4** and with other helical capsules:[4] it possesses a sizeable cavity surrounded by the helix backbone (350 Å^3 in this case, see Fig. S8). The solid state structure of $(\mathbf{5})_2$ represents the first example of a water-soluble helical capsule crystal structure.

To conclude this section, it is again remarkable that differences as small as those that exist between **3** and **4** – minus two charged residues – can trigger such a drastic change of aggregation behavior. We also infer that the monomeric nature of **1** and **3** is not just the result of the bulky terminal Q_3 segments, but also of charge repulsions within double helical dimers.

Hemi-capsule dimerization

To explore further the modes of dimerization of helically folded aromatic oligoamides in water, several hemicapsule sequences were synthesized. A sequence differing from **6** only by the length of the cationic side chains has already been shown to form a double helical dimer in water.[53] The behavior of **6** was consistent. Upon adding water to a DMSO-d_6 solution of **6**, a species assigned to a single helical monomer in DMSO is replaced by another species assigned to a double helical dimer (Fig. 4a). In the gas phase, ESI-MS showed a substantial population of $[2M+2H]^{2+}$ and $[2M+3H]^{3+}$ dimeric species (Fig. S10). The dimerization constant of **6** was too high to observe monomeric species by ^{1}H NMR upon dilution a solution in water (Fig. S11). Also, given the multiplicity of NMR signals, one can infer that $(\mathbf{6})_2$ has a symmetrical structure and propose a molecular model as shown in Fig. 5a.

Similar experiments were carried out using analogous sequence **7** in which Qf units were replaced by Qh units that should provide

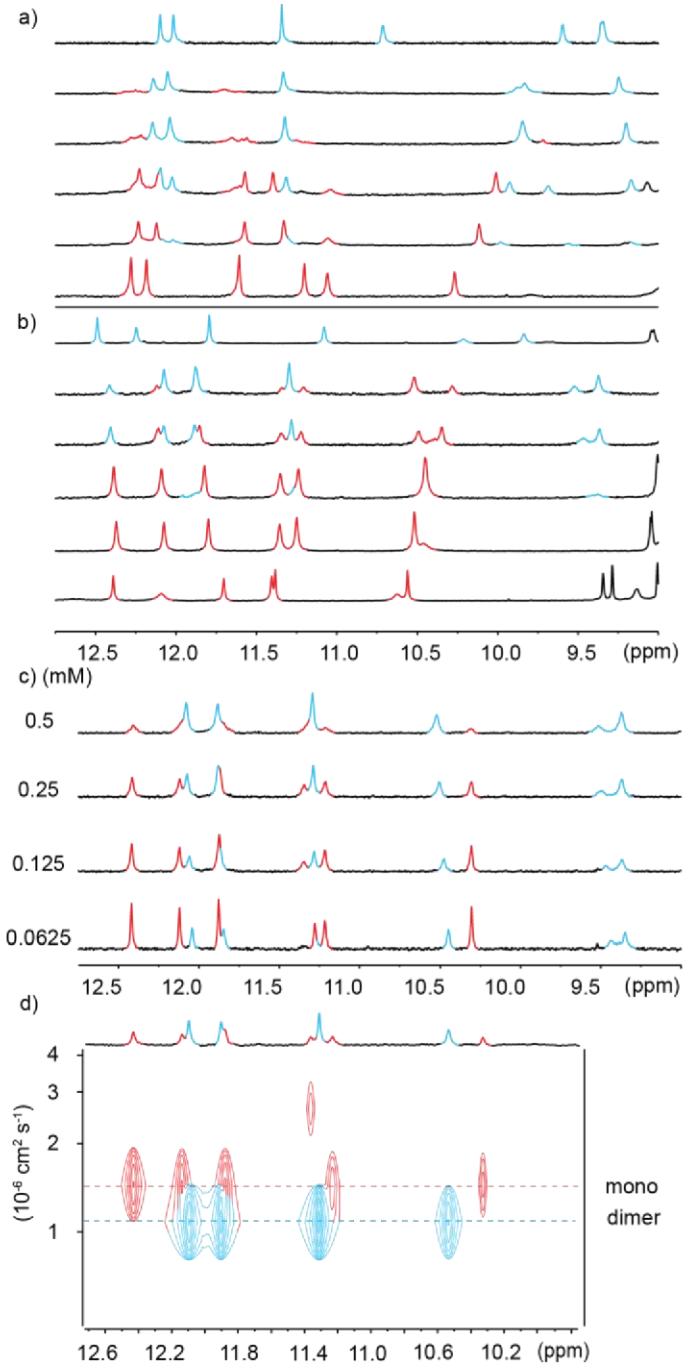


Fig. 4 Excerpts of the ^{1}H NMR (500 MHz) spectra at 298 K of **6** (a) and **7** (b) in DMSO-d_6/H_2O mixtures. The percentage of DMSO is from top to bottom: a) 0, 50, 66, 70, 76 and 100; and b) 0, 33, 40, 50, 57 and 100. Note: 0% DMSO means the spectrum was recorded in H_2O/D_2O (90:10 v/v). Red and blue amide proton signals stand for either the monomeric or dimeric form, respectively. c) Excerpts of the ^{1}H NMR spectra of **7** at different concentrations in DMSO-d_6/H_2O (40:60 v/v). Selected amide proton signals belonging to either the monomeric or dimeric form are highlighted in red and blue, respectively. d) Amide NH region of the ^{1}H DOSY NMR (500 MHz) spectrum of **7** (0.32 mM) in DMSO-d_6/H_2O (40:60 v/v). The two different levels of signals related to different hydrodynamic radii are indicated by dashed lines. Signals corresponding to the monomeric or dimeric form are highlighted in red and blue.

more degrees of conformational freedom. A similar transition as for **6** was observed upon adding water to a DMSO-d_6 solution of **7** (Fig. 4b). That the two species correspond to an aggregation phenomenon was ascertained by ^{1}H DOSY NMR experiment and from the

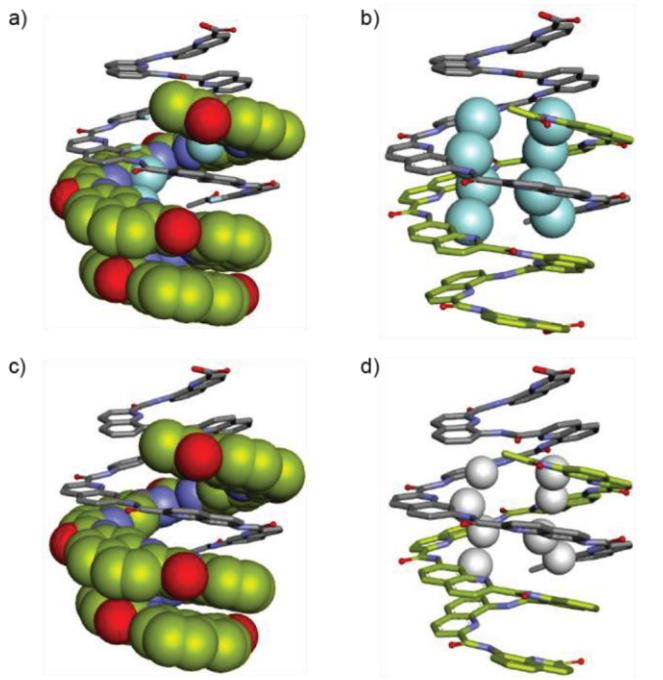


Fig. 5 Energy minimized models (Maestro, MMFFs force field, water as implicit solvent)[71] of (**6**)$_2$ (a,b) and (**7**)$_2$ (c,d). In a) and c), one strand is shown in space filling representation and the other in stick representation. In b) and d), both strands are shown in stick representation but the fluorine atoms in position 8 of Qf units and the hydrogen atoms in position 8 of Qh units are shown in space filling representations and colored in cyan and white, respectively. In all representations two distinct strands are colored in light green or gray, respectively. Side chains and most hydrogen atoms have been omitted for clarity.

concentration dependence of the proportion of the two species (Fig. 4c,d) in DMSO-d_6/H_2O (40:60 v/v). A dimerization constant of 10^4 M^{-1} in this solvent could be calculated. ESI-MS confirmed the presence of $[2M+2H]^{2+}$ and $[2M+3H]^{3+}$ dimeric species in the gas phase (Fig. S12). Again, no dissociation was observed upon diluting a solution of **7** in water (Fig. S13). Nevertheless, less DMSO is required to disrupt (**7**)$_2$ than to disrupt (**6**)$_2$ (Fig. 4a,b). In short, Qh units also promote double helix formation, albeit slightly less efficiently than Qf monomers, and the model shown in Fig. 5b can be proposed for (**7**)$_2$. The dimeric segment involves eight Qh units, allowing some cooperative solvophobic folding to occur. In contrast, the four units of monomeric **7** cannot in principle undergo cooperative folding because they do not span a full helix turn. Estimations of the solvent-accessible surface of (**7**)$_2$ and of different conformations of monomeric **7** suggest that the surface of a dimer amounts to 61 to 71% of the surface of two monomers (Fig. S14), shedding light on the significant hydrophobic effects observed. Interestingly when mixing **6** and **7** in water, only homodimers (**6**)$_2$ and (**7**)$_2$ were observed (Fig. S15, S16), indicating a sort of narcissistic self-sorting that contrasts with heterodimerization observed in other systems.[67,72]

Double helices may also dimerize

In view of the success at getting crystals of anionic sequences, oligomer **8** was introduced as an analogue of **7** having both methoxy and sulfonic acid side chains. At this stage, the neutral methoxy groups did not alter aggregation properties and **8** behaved like **7** in pure water (Fig. S17) and in water/DMSO mixtures (Fig. S18). $[2M-3H]^{3-}$ and $[2M-2H]^{2-}$ dimeric species could be observed by ESI-MS in

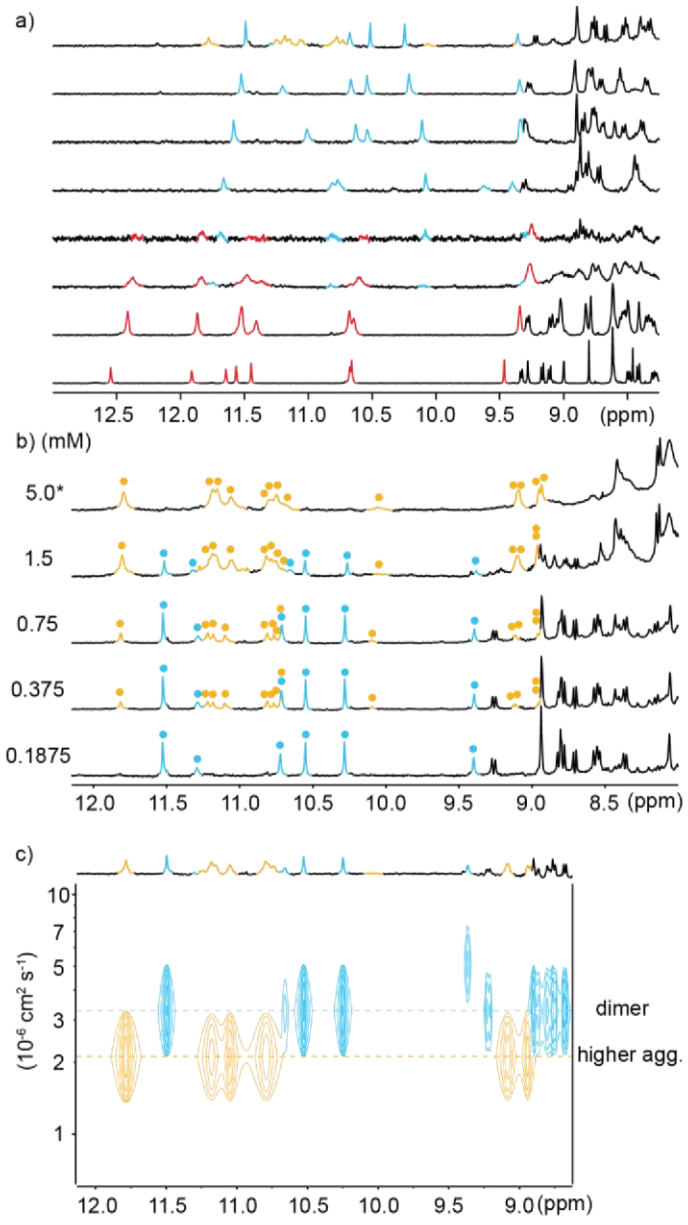


Fig. 6 Excerpts of the ^{1}H NMR (500 MHz) titration at 298 K of **8** (0.72 mM) in DMSO-d_6/buffer mixtures in the following DMSO percentage (top to bottom): (a) 0, 16.7, 50, 66.6, 72.5, 76.9, 83.3 and 100. Note: buffer means 13 mM ammonium acetate aqueous buffer pH 8.5. 0% DMSO means ^{1}H NMR (500 MHz) spectrum of **8** in buffer (buffer/D_2O 90:10 v/v). Red, blue and yellow amide proton signals stand for the monomeric, dimeric and tetrameric form, respectively. b) Excerpts of the ^{1}H NMR (500 MHz) spectra of compound **8** at different concentrations in 11.7 mM (* means in 47.7 mM) ammonium acetate aqueous buffer pH 8.5. Selected amide proton signals belonging to either the dimeric or tetrameric form are highlighted in blue and yellow, respectively. d) Amide-region of the ^{1}H DOSY NMR (500 MHz) spectrum of compound **8** (1 mM) in 11.7 mM ammonium acetate aqueous buffer pH 8.5. The two different levels of signals related to different molecularities in the ^{1}H DOSY spectrum are indicated by dashed lines. Signals corresponding to the dimeric or tetrameric form are highlighted in blue and yellow.

the gas phase (Fig. S19). However, sequence **9** in which one more sulfonic acid has been replaced by a methoxy group exhibited a new behavior.§

Unlike **6**, **7** and **8**, the ^{1}H NMR spectrum of **9** in water showed two sets of signals (Fig. 6a). Upon adding water to a DMSO-d_6 solution of **9**, a first transition occurred near 30% of water, which we assigned

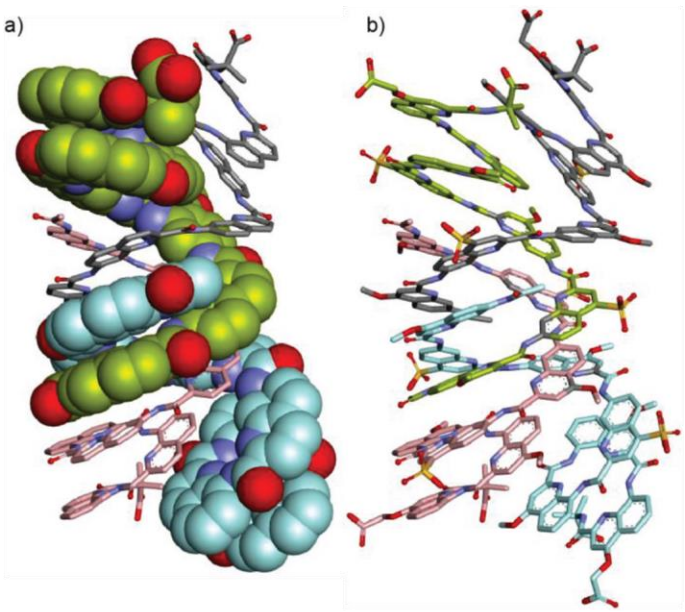


Fig. 7 Energy minimized model (Maestro, MMFFs force field, water as implicit solvent)[71] of $(\mathbf{9})_4$: a) Side view with two strands in space-filling representation and two strands in stick representation. Side chains and hydrogen atoms have been omitted for clarity. b) Same side view including the side chains with all four strands in stick representation. In all representations, four distinct strands are colored in light green, gray, cyan and pink, respectively. The symmetry of the structure is such that gray and cyan strands on one hand and light green and pink strands on the other hand, are equivalent.

to the expected monomer-to-double-helical-dimer transition, and a second transition occurred when reaching 100% water (Fig. 6a). All three species are in slow exchange on the NMR times scale. The proportion of the two species present in water was shown to depend on concentration: the relative intensity of the signals assigned to the double helical dimer decreases when concentration increases (Fig. 6b), indicating that the new species is an aggregate of higher molecularity than a dimer. The conversion is quantitative at 5 mM. ^{1}H DOSY NMR confirmed that the higher aggregate has a larger hydrodynamic radius (Fig. 6c). Native ESI-MS analysis could only be performed at concentrations too low for the higher aggregate to be present in significant amount and only the $[2M-2H]^{2-}$ dimeric species was detectable in the gas phase (Fig. S20). However, the molecularity of the higher aggregate may be inferred from its number of ^{1}H NMR resonances. This number is double the number of resonances of $(\mathbf{9})_2$, meaning the aggregate possesses two inequivalent strands in equal proportions, ruling out a possible trimer and making a tetramer the most likely possibility.

The time required to form $(\mathbf{9})_4$ after dissolving **9** in water was found to be of the order of hours (Fig. S21), indicating that significant structural changes are required. A plausible molecular model for $(\mathbf{9})_4$ was proposed based on the structure of $(\mathbf{5})_2$ (Fig. 7). Taking the C-terminal half of $(\mathbf{5})_2$ (bottom part of Fig. 3c), a parallel duplex $(\mathbf{9})_2$ can be proposed distinct from the anti-parallel duplex that prevails at lower concentration. Next, two of these parallel duplexes can be intertwined in an antiparallel fashion through their N-terminal Qh_4 segments to produce the C_2-symmetric quadruplex shown in Fig. S7 having two pairs of inequivalent strands. This model is of course speculative. It nevertheless allows pointing to some factors that could play a role: its solvent-accessible surface is smaller than that of

two anti-parallel dimers (Fig. S22); removing one charge from a monomer means removing four charges from the tetramer; two MeO of **9** that replaced sulfonic acids of **8** are found to be buried between two hydrophobic aromatic surfaces pointing to the destabilization of such a structure in the case of **8**. Besides electrostatic effects between charges, the fact that an electron withdrawing group in **8** has been replaced by a donor in **9** might also influence interactions between aryl rings but this is hard to estimate. Single crystals of **9** have been obtained that diffracted up to 1.5 Å but no structure could be solved so far.

Conclusions

We have shown that the formation of at least three different types of discrete aggregates of helical aromatic foldamers can be triggered in water by minor changes in side chain properties. These changes always consisted in replacing a small number of charged residues, sometimes only one, by a neutral methoxy group. Specifically, a head-to-tail stacked dimer of two single helices, a double helical duplex with four terminal Q_3 segments, and a quadruplex were shown to form whereas related sequences containing a few more charged residues did not. The last two aggregates are new. All aggregates are kinetically stable. They are in slow exchange with their subcomponents on the NMR timescale. In the case of the double helix and the quadruplex, kinetic stability is high and equilibrium may take hours to be reached. Hydrophobic effects appear to be a major driving force for aggregation, but these effects are counterbalanced by the side chain charges. Earlier observations in organic solvents had shown that Q_3 segments prevent the formation of some types of multistranded helices. We now have shown that this effect can be overcome in water unless sufficient charge repulsions are also at play. Overall, these results provide a foundation for future aromatic foldamer structure design in water. In addition, the monomeric and multimeric species presented here possess cavities in which guest molecules may be recognized, and secluded from the surrounding aqueous environment. These cavities, in particular those generated by more than one strand, may undergo rearrangements triggered by external stimuli and exhibit controlled release or guest swapping properties. Progress in this direction is being made and will be reported in due course.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by the DFG (project HU 1766/6-1). B. Teng gratefully acknowledges financial support from the China Scholarship Council. We thank D. Gill and E. Merlet for contributing some synthetic precursors and G. Leonard (ID30B, ESRF) for assistance during data collection at the Synchrotron beamline.

Footnotes and References

‡ Care should be taken during integration as the intensity of the ^{1}H NMR signals of exchangeable NH protons may be affected by the water suppression. This effect is more pronounced at lower concentration and results in lower-than-expected intensities for protons that exchange most. Typically, these results in different intensities for signals belonging to the same species that would be expected to have the same integration. The NH resonances belonging to Q units found at highest chemical shit values typically exchange least and may be preferred for accurate integration.

§ Compound **9** was not produced as an analogue of **8** but as another candidate for crystal growth, hence the other differences between the two. Crystal structures, including that of $(\mathbf{5})_2$, suggest that C-terminal Aib units, as in **9**, do not influence aggregation. There is also no ground to invoke a role of Q^{Asp} as the seventh residue in **9** instead of Q^{Sul} in **8**. We thus assign the different aggregation behavior of these compounds to the $Q^{Sul}6Q^{Ala}$ mutation.

1 D. Lombardo, M. A. Kiselev, S. Magazù and P. Calandra, *Adv. Condens. Matter Phys.*, 2015, **2015**, e151683.
2 P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Science*, 1991, **254**, 1497–1500.
3 R. B. Prince, S. A. Barnes and J. S. Moore, *J. Am. Chem. Soc.*, 2000, **122**, 2758–2762.
4 Y. Ferrand and I. Huc, *Acc. Chem. Res.*, 2018, **51**, 970–977.
5 H. Juwarker and K.-S. Jeong, *Chem. Soc. Rev.*, 2010, **39**, 3664–3674.
6 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.
7 W. Wang, C. Zhang, S. Qi, X. Deng, B. Yang, J. Liu and Z. Dong, *J. Org. Chem.*, 2018, **83**, 1898–1902.
8 Y. Liu, F. C. Parks, E. G. Sheetz, C.-H. Chen and A. H. Flood, *J. Am. Chem. Soc.*, 2021, **143**, 3191–3204.
9 Y. Zhong, T. A. Sobiech, B. Kauffmann, B. Song, X. Li, Y. Ferrand, I. Huc and B. Gong, *Chem. Sci.*, 2023, DOI:10.1039/D3SC00524K.
10 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.
11 K. Ziach, C. Chollet, V. Parissi, P. Prabhakaran, M. Marchivie, V. Corvaglia, P. P. Bose, K. Laxmi-Reddy, F. Godde, J.-M. Schmitter, S. Chaignepain, P. Pourquier and I. Huc, *Nature Chem*, 2018, **10**, 511–518.
12 S. Kumar, A. Henning-Knechtel, I. Chehade, M. Magzoub and A. D. Hamilton, *J. Am. Chem. Soc.*, 2017, **139**, 17098–17108.
13 J. Shen, R. Ye, Z. Liu and H. Zeng, *Angew. Chem. Int. Ed.*, 2022, **61**, e202200259.
14 L. Zhang, C. Zhang, X. Dong and Z. Dong, *Angew. Chem. Int. Ed.*, 2023, **135**, e202214194.
15 X. Wei, G. Zhang, Y. Shen, Y. Zhong, R. Liu, N. Yang, F. Y. Al-mkhaizim, M. A. Kline, L. He, M. Li, Z.-L. Lu, Z. Shao and B. Gong, *J. Am. Chem. Soc.*, 2016, **138**, 2749–2754.
16 Y. Shen, F. Fei, Y. Zhong, C. Fan, J. Sun, J. Hu, B. Gong, D. M. Czajkowsky and Z. Shao, *ACS Cent. Sci.*, 2021, **7**, 2092–2098.
17 A. Roy, J. Shen, H. Joshi, W. Song, Y.-M. Tu, R. Chowdhury, R. Ye, N. Li, C. Ren, M. Kumar, A. Aksimentiev and H. Zeng, *Nat. Nanotechnol.*, 2021, **16**, 911–917.
18 D.-D. Su and M. Barboiu, *CCS Chem.*, 2022, **5**, 279–291.
19 X. Wang, B. Wicher, Y. Ferrand and I. Huc, *J. Am. Chem. Soc.*, 2017, **139**, 9350–9358.
20 M. Gauthier, V. Koehler, C. Clavel, B. Kauffmann, I. Huc, Y. Ferrand and F. Coutrot, *Angew. Chem. Int. Ed.*, 2021, **60**, 8380–8384.
21 A. Méndez-Ardoy, N. Markandeya, X. Li, Y.-T. Tsai, G. Pecastaings, T. Buffeteau, V. Maurizot, L. Muccioli, F. Castet, I. Huc and D. M. Bassani, *Chem. Sci.*, 2017, **8**, 7251–7257.

22 X. Li, N. Markandeya, G. Jonusauskas, N. D. McClenaghan, V. Maurizot, S. A. Denisov and I. Huc, *J. Am. Chem. Soc.*, 2016, **138**, 13568–13578.
23 K. Moreno, E. Merlet, N. McClenaghan, T. Buffeteau, Y. Ferrand and C. Olivier, *ChemPlusChem*, 2021, **86**, 496–503.
24 D. Zheng, L. Zheng, C. Yu, Y. Zhan, Y. Wang and H. Jiang, *Org. Lett.*, 2019, **21**, 2555–2559.
25 X. Hu, A. Schulz, J. O. Lindner, M. Grüne, D. Bialas and F. Würthner, *Chem. Sci.*, 2021, **12**, 8342–8352.
26 J. C. Nelson, J. G. Saven, J. S. Moore and P. G. Wolynes, *Science*, 1997, **277**, 1793–1796.
27 H. Goto, H. Katagiri, Y. Furusho and E. Yashima, *J. Am. Chem. Soc.*, 2006, **128**, 7176–7178.
28 K. P. de Carvasal, N. Aissaoui, G. Vergoten, G. Bellot, J.-J. Vasseur, M. Smietana and F. Morvan, *Chem. Commun.*, 2021, **57**, 4130–4133.
29 Y. Zhao and J. S. Moore, in *Foldamers*, John Wiley & Sons, Ltd, 2007, pp. 75–108.
30 Y. Hua, Y. Liu, C.-H. Chen and A. H. Flood, *J. Am. Chem. Soc.*, 2013, **135**, 14401–14412.
31 K. P. Divya, S. Sreejith, C. H. Suresh, D. S. Philips and A. Ajayaghosh, *Chem Asian J*, 2013, **8**, 1579–1586.
32 K.-J. Chang, B.-N. Kang, M.-H. Lee and K.-S. Jeong, *J. Am. Chem. Soc.*, 2005, **127**, 12214–12215.
33 H. Juwarker and K.-S. Jeong, *Chem. Soc. Rev.*, 2010, **39**, 3664–3674.
34 V. R. Naidu, M. C. Kim, J. Suk, H.-J. Kim, M. Lee, E. Sim and K.-S. Jeong, *Org. Lett.*, 2008, **10**, 5373–5376.
35 E. A. John, C. J. Massena and O. B. Berryman, *Chem. Rev.*, 2020, **120**, 2759–2782.
36 C. J. Massena, D. A. Decato and O. B. Berryman, *Angew. Chem. Int. Ed.*, 2018, **57**, 16109–16113.
37 C. S. Hartley, *Acc. Chem. Res.*, 2016, **49**, 646–654.
38 S. Peddi, M. C. Bookout, G. N. Vemuri and C. S. Hartley, *J. Org. Chem.*, 2022, **87**, 3686–3690.
39 H. Jiang, J.-M. Léger and I. Huc, *J. Am. Chem. Soc.*, 2003, **125**, 3448–3449.
40 X. Yang, A. L. Brown, M. Furukawa, S. Li, W. E. Gardinier, E. J. Bukowski, F. V. Bright, C. Zheng, X. C. Zeng and B. Gong, *Chem. Commun.*, 2003, 56–57.
41 A. M. Abramyan, Z. Liu and V. Pophristic, *Chem. Commun.*, 2015, **52**, 669–672.
42 Z. Liu, A. M. Abramyan and V. Pophristic, *New J. Chem.*, 2015, **39**, 3229–3240.
43 L. Yuan, H. Zeng, K. Yamato, A. R. Sanford, W. Feng, H. S. Atreya, D. K. Sukumaran, T. Szyperski and B. Gong, *J. Am. Chem. Soc.*, 2004, **126**, 16528–16537.
44 V. Maurizot, J. Duhamel, I. Huc, B. Wicher, K. Lulic, J. Wang and A. Galan, *Chem. Commun.*, 2023, DOI:10.1039/D3CC00866E.
45 D. Bindl, P. K. Mandal, L. Allmendinger and I. Huc, *Angew. Chem. Int. Ed.*, 2022, **61**, e202116509.
46 S. De, B. Chi, T. Granier, T. Qi, V. Maurizot and I. Huc, *Nature Chem*, 2018, **10**, 51–57.
47 F. C. Parks, Y. Liu, S. Debnath, S. R. Stutsman, K. Raghavachari and A. H. Flood, *J. Am. Chem. Soc.*, 2018, **140**, 17711–17723.
48 N. Ousaka, M. Itakura, A. Nagasaka, M. Ito, T. Hattori, D. Taura, T. Ikai and E. Yashima, *J. Am. Chem. Soc.*, 2021, **143**, 4346–4358.
49 A. Urushima, D. Taura, M. Tanaka, N. Horimoto, J. Tanabe, N. Ousaka, T. Mori and E. Yashima, *Angew. Chem. Int. Ed.*, 2020, **59**, 7478–7486.
50 N. Saito, R. Terakawa, M. Shigeno, R. Amemiya and M. Yamaguchi, *J. Org. Chem.*, 2011, **76**, 4841–4858.
51 D. Haldar, H. Jiang, J.-M. Léger and I. Huc, *Angew. Chem. Int. Ed.*, 2006, **45**, 5483–5486.
52 H. Goto, Y. Furusho, K. Miwa and E. Yashima, *J. Am. Chem. Soc.*, 2009, **131**, 4710–4719.
53 J. Shang, Q. Gan, S. J. Dawson, F. Rosu, H. Jiang, Y. Ferrand and I. Huc, *Org. Lett.*, 2014, **16**, 4992–4995.
54 Y. Wang, Y. He, Z. Yu, J. Gao, S. ten Brinck, C. Slebodnick, G. B. Fahs, C. J. Zanelotti, M. Hegde, R. B. Moore, B. Ensing, T. J. Dingemans, R. Qiao and L. A. Madsen, *Nat Commun*, 2019, **10**, 801.
55 Y. Ferrand, A. M. Kendhale, J. Garric, B. Kauffmann and I. Huc, *Angew. Chem. Int. Ed.*, 2010, **49**, 1778–1781.
56 C. J. Massena, N. B. Wageling, D. A. Decato, E. Martin Rodriguez, A. M. Rose and O. B. Berryman, *Angew. Chem. Int. Ed.*, 2016, **55**, 12398–12402.
57 K. Zhang, C. Ma, N. Li, C. Lu, D. Li, S. Fu and Q. Gan, *Chem. Commun.*, 2019, **55**, 10968–10971.
58 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.
59 Y. Liu, F. C. Parks, W. Zhao and A. H. Flood, *J. Am. Chem. Soc.*, 2018, **140**, 15477–15486.
60 E. A. John, C. J. Massena and O. B. Berryman, *Chem. Rev.*, 2020, **120**, 2759–2782.
61 H. Sugiura, Y. Nigorikawa, Y. Saiki, K. Nakamura and M. Yamaguchi, *J. Am. Chem. Soc.*, 2004, **126**, 14858–14864.
62 H. Li, L. Kou, L. Liang, B. Li, W. Zhao, X.-J. Yang and B. Wu, *Chem. Sci.*, 2022, **13**, 4915–4921.
63 Q. Gan, Y. Ferrand, N. Chandramouli, B. Kauffmann, C. Aube, D. Dubreuil and I. Huc, *J. Am. Chem. Soc.*, 2012, **134**, 15656–15659.
64 C. Dolain, A. Grélard, M. Laguerre, H. Jiang, V. Maurizot and I. Huc, *Chem. Eur. J.*, 2005, **11**, 6135–6144.
65 T. Qi, V. Maurizot, H. Noguchi, T. Charoenraks, B. Kauffmann, M. Takafuji, H. Ihara and I. Huc, *Chem. Commun.*, 2012, **48**, 6337–6339.
66 B. Teng, J. Atcher, L. Allmendinger, C. Douat, Y. Ferrand and I. Huc, *Org. Biomol. Chem.*, 2023, DOI:10.1039/D3OB00473B.
67 M. L. Singleton, G. Pirotte, B. Kauffmann, Y. Ferrand and I. Huc, *Angew. Chem. Int. Ed.*, 2014, **53**, 13140–13144.
68 X. Hu, S. J. Dawson, P. K. Mandal, X. de Hatten, B. Baptiste and I. Huc, *Chem. Sci.*, 2017, **8**, 3741–3749.
69 V. Corvaglia, F. Sanchez, F. Menke, C. Douat and I. Huc, *Chem. Eur. J.*, 2023, DOI: 10.1002/chem.202301120.
70 B. Baptiste, J. Zhu, D. Haldar, B. Kauffmann, J.-M. Léger and I. Huc, *Chem Asian J*, 2010, **5**, 1364–1375.
71 Maestro, Schrödinger, LLC, New York, NY, 2021.
72 C. Zhan, J.-M. Léger and I. Huc, *Angew. Chem. Int. Ed.*, 2006, **45**, 4625–4628.

Supplementary Information

Controlling aromatic helix dimerization in water by tuning charge repulsions

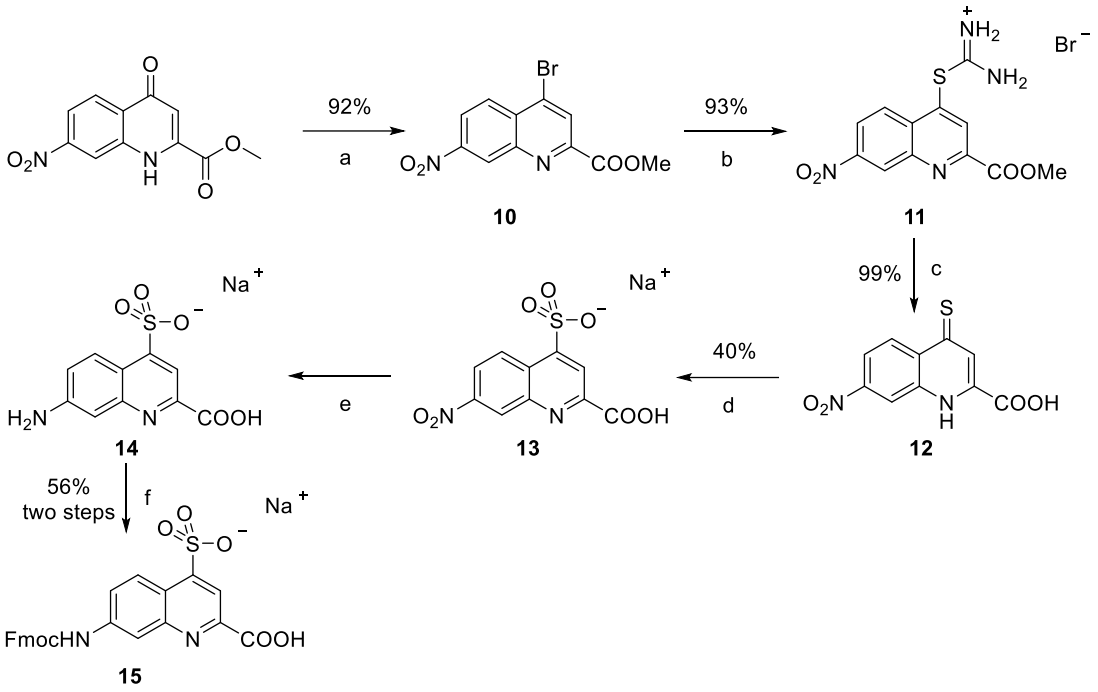
Binhao Teng, [a] Pradeep K. Mandal, [a] Lars Allmendinger, [a] Céline Douat,[a] Yann Ferrand [b] and Ivan Huc* [a]

[a.] Department of Pharmacy, Ludwig-Maximilians-Universität, Butenandtstr. 5–13, 81377, München, Germany.

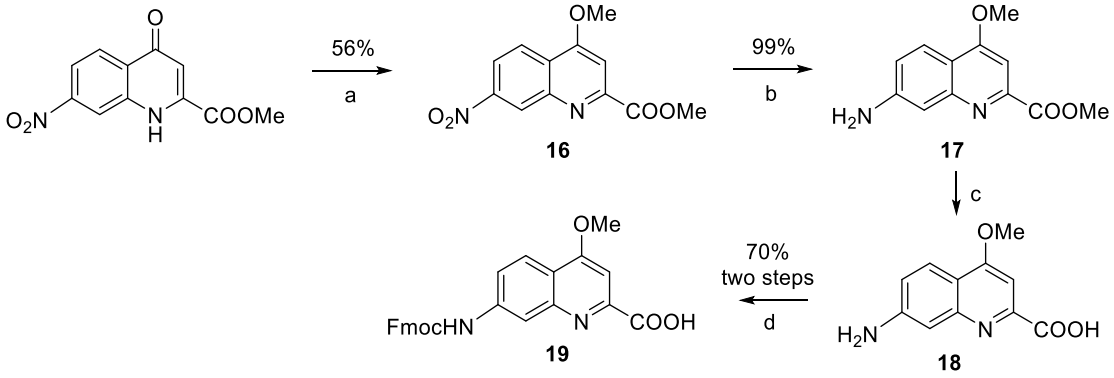
[b.] Univ. Bordeaux, CNRS, Bordeaux Institut National Polytechnique, CBMN UMR 5248, 2 rue Escarpit, 33600 Pessac, France.

Table of contents

1 Supplementary Schemes and Figures



Scheme S1. Synthesis of **Qh^{Sul}** monomer: a) $POBr_3$, DMF; b) thiourea, acetone, reflux; c) NaOH, MeOH/H_2O, r.t.; d) 30% H_2O_2, formic acid, 48 °C; e) Pd/C, H_2, MeOH; f) Fmoc-Cl, $NaHCO_3$, dioxane, H_2O, 0 ^oC to r.t..



Scheme S2. Synthesis of **Qh^{Ala}** monomer: a) PPh_3, MeOH, DIAD, THF; b) Pd/C, H_2, MeOH/THF; c) LiOH, THF, H_2O, 0 °C; d) Fmoc-Cl, $NaHCO_3$, dioxane, H_2O, 0 ^oC to r.t..

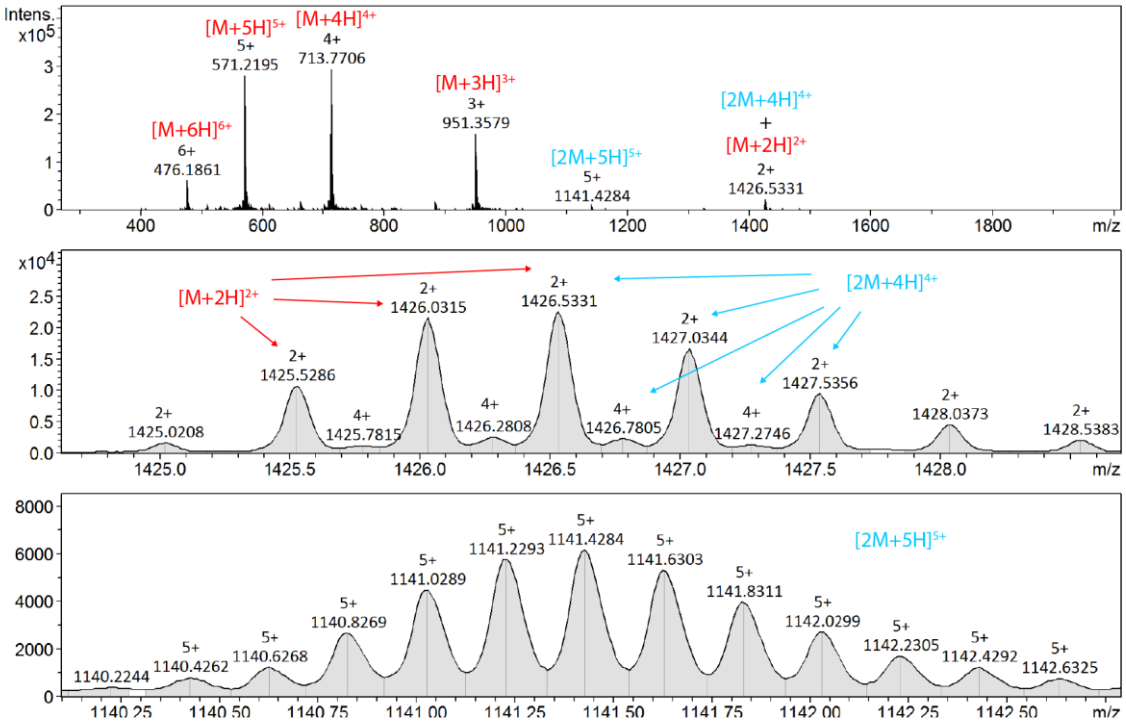


Figure S74. ESI-MS spectrum of oligomer 2 (direct infusion of an aqueous sample). Peaks corresponding to monomeric and dimeric species are highlighted in red and blue, respectively. The $[M+2H]^{2+}$ peak is superimposed by the $[2M+4H]^{4+}$ peak.

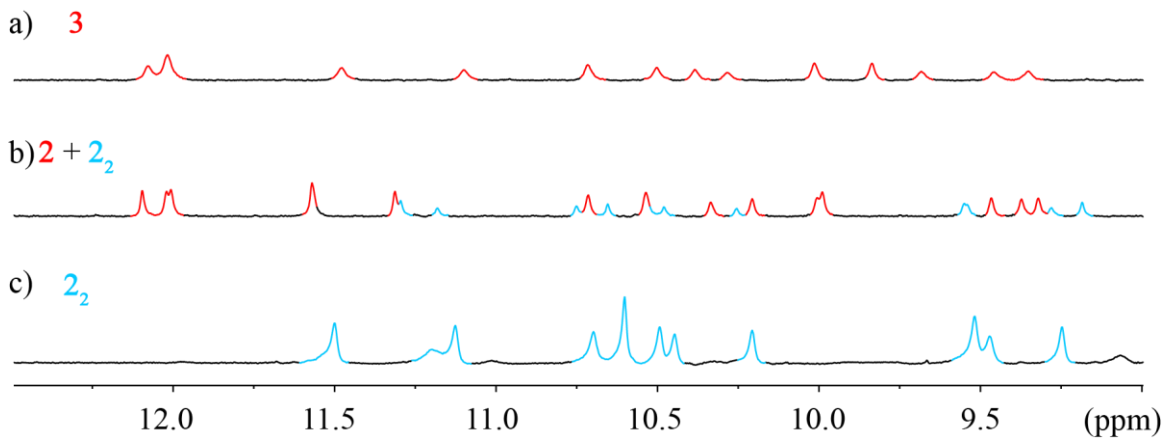


Figure S75. The comparison between oligomer 3 and monomeric oligomer 2. Excerpts of ^{1}H NMR (500 MHz, water suppression) at 25 °C: a) **3** at 0.5 mM and **2** at b) 0.5 mM and c) 4 mM in 45 mM sodium acetate aqueous buffer pH 4.0. Signals corresponding to monomeric and dimeric species are highlighted in red and blue, respectively.

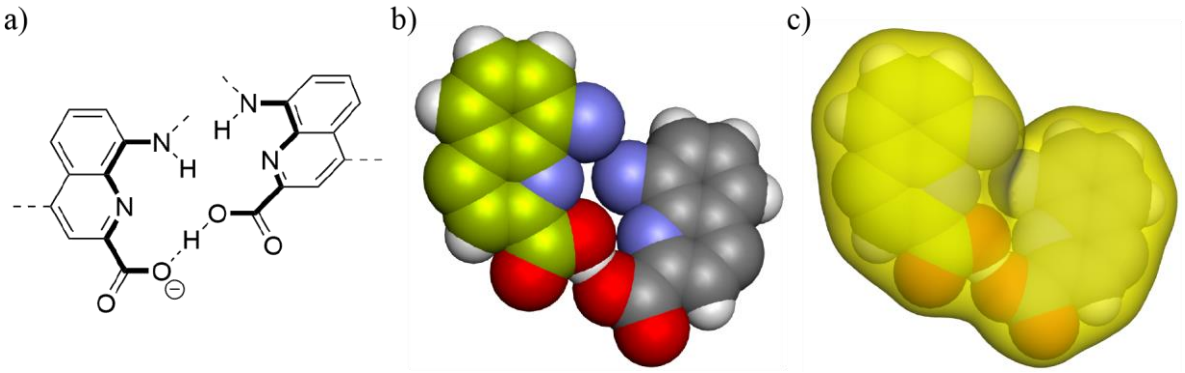


Figure S76. Computational analysis of the fragment of the helix–helix binding interface. The models showed the C-terminal monomers in space-filling representation in a) chemical structural formula, b) energy minimized model in CPK style and c) the calculated solvent accessible surface. Side chains are omitted for clarity. Solvent-accessible surface area was calculated using discovery studio software with a radius of 1.4 Å. The assignment of the carboxylic acid and carboxylate function is tentative, based on orientation.

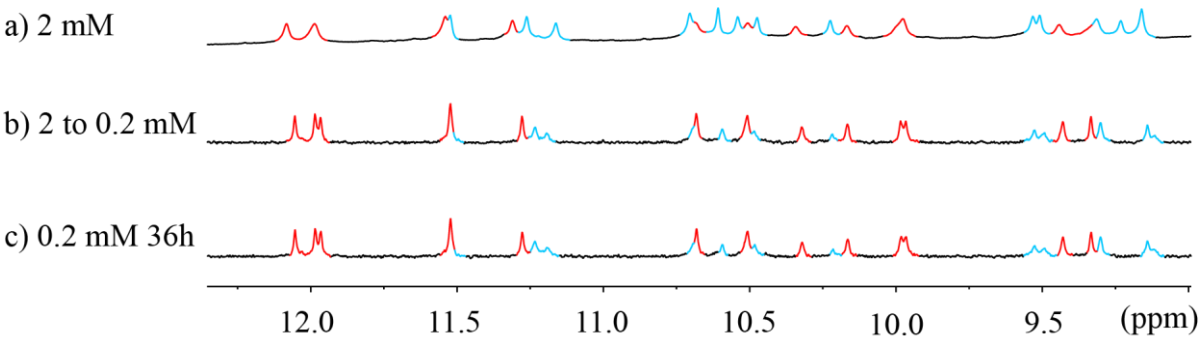


Figure S77. Dilution experiments of oligomer 2. Excerpts of ^{1}H NMR (500 MHz, water suppression) at 25 °C of oligomer **2** in H_2O/D_2O (9:1, *v*/*v*). a) no difference in NMR of **2** at 2 mM was observed in variable incubation time (0.5 h, 1 h and 4 h, only 4 h is shown). b) The NMR was immediately measured after diluting 2 mM of **2** to 0.2 mM and showed dimer to monomer equilibrium. c) A diluted solution of **2** was incubated for 36 h and showed no proportion change between the two species. This means that **2** reached monomer-dimer equilibrium immediately upon being dissolved . Peaks corresponding to monomeric and dimeric species are highlighted in red and blue, respectively.

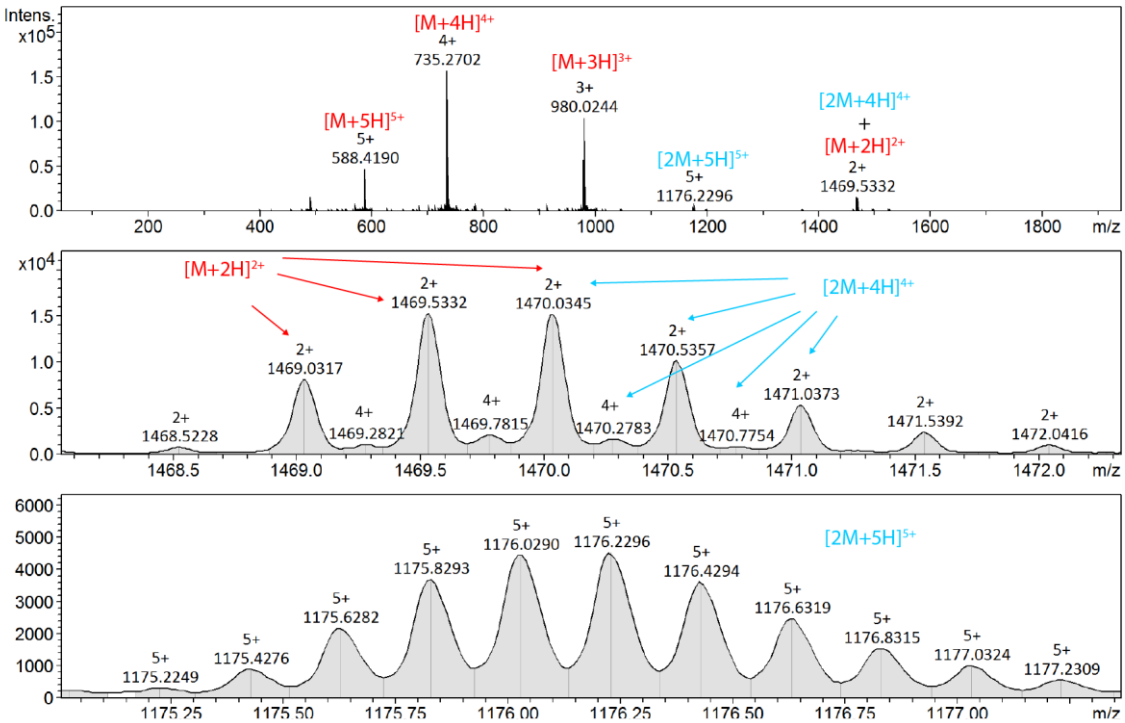


Figure S78. ESI-MS spectrum of oligomer 4 (direct infusion of an aqueous sample). Peaks corresponding to monomeric and dimeric species are highlighted in red and blue, respectively. The $[M+2H]^{2+}$ peak is superimposed by the $[2M+4H]^{4+}$ peak.

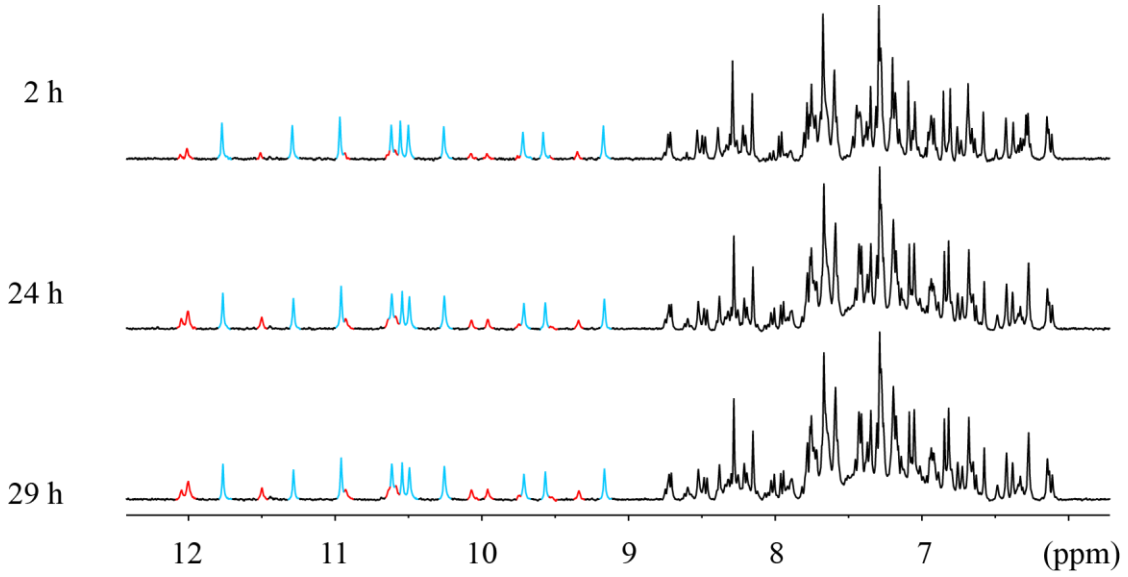


Figure S79. Variable incubation time ^{1}H NMR spectra of 4. Excerpts of ^{1}H NMR (500 MHz, water suppression) at 25 °C of oligomer **4** at 0.5 mM in 13.5 mM sodium acetate aqueous buffer pH 4.0 obtained after varying incubation times. Signals corresponding to monomeric and dimeric species are highlighted in red and blue, respectively.

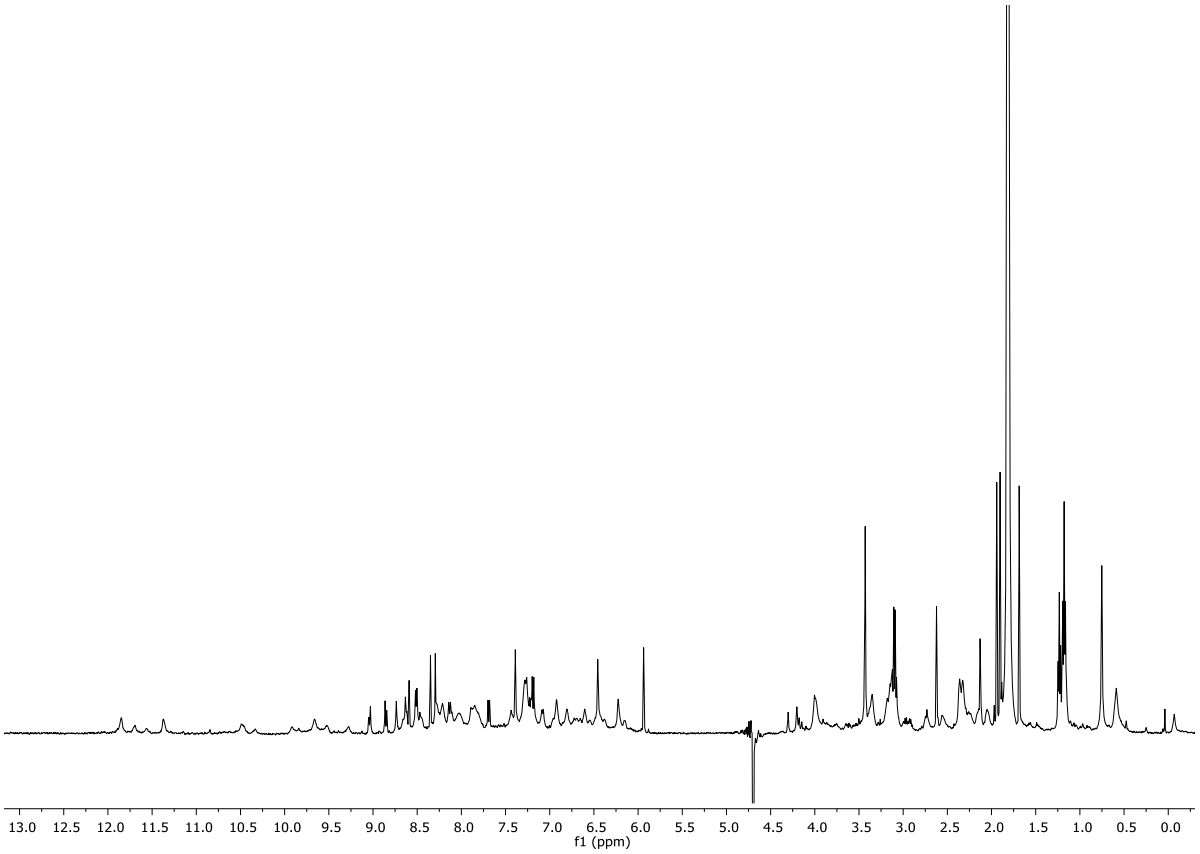


Figure S80. ^{1}H NMR spectra of oligomer 5 in buffered water. ^{1}H NMR spectrum of oligomer **5** at 0.125 mM with water suppression (500 MHz, 13 mM ammonium acetate aqueous buffer pH 8.5/D_2O (9:1, *v*/*v*), 25 °C).

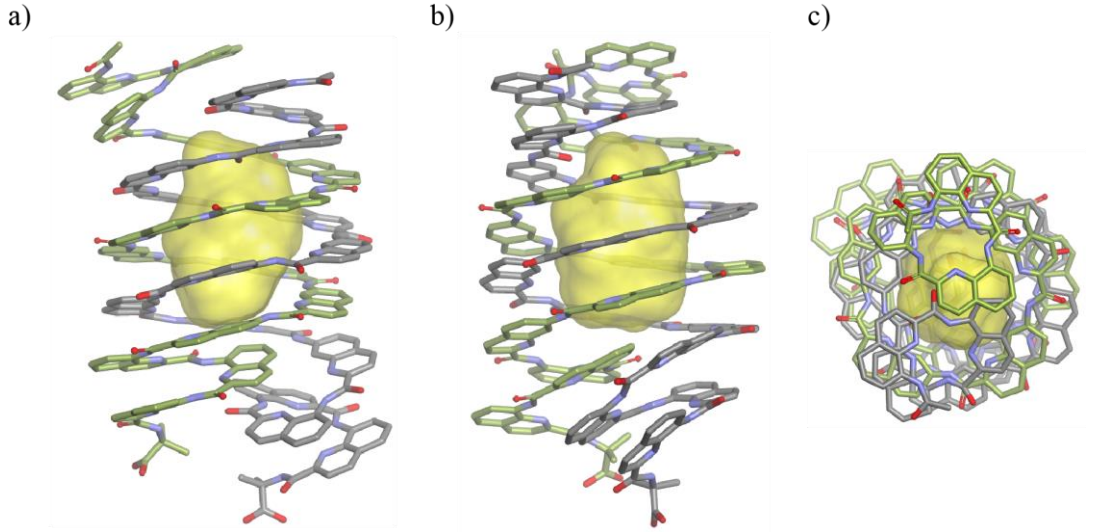


Figure S81. Cavity volume calculation of (5)$_2$. The front view, side view, and top view of (**5**)$_2$ are shown in tube representation where the cavity volume of 346 Å^3 is shown as a transparent yellow isosurface (a, b, c). In all representations, each strand is colored in green and grey, respectively. Side chains and N-terminal Deg tail were omitted for clarity.

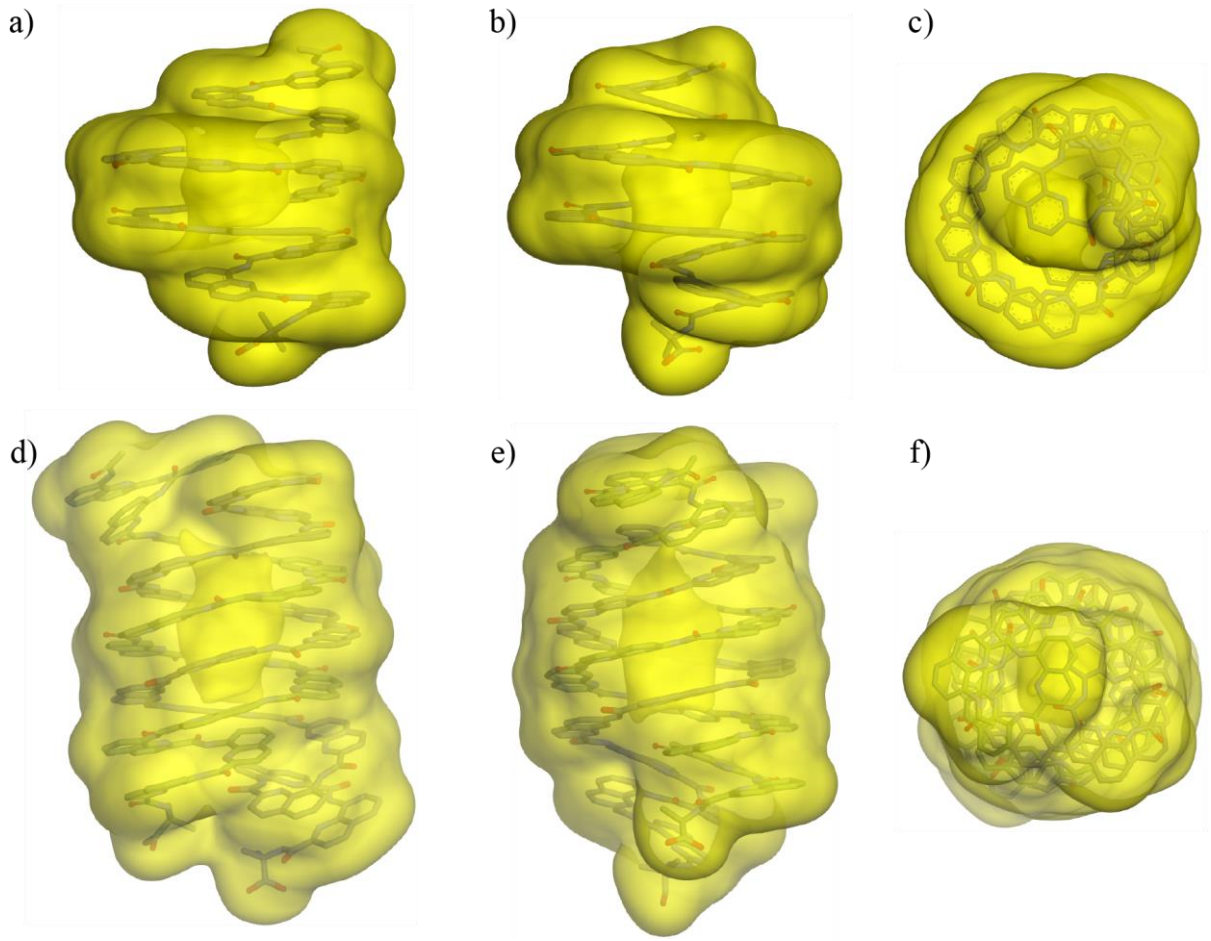


Figure S82. Solvent accessible surface of two conformations of 5: (a) front view, b) side view and c) top view of single helix **5**. Solvent accessible surface of: (d) front view, e) side view and f) top view of duplex $(\mathbf{5})_2$. All solvent accessible surfaces are shown as transparent yellow isosurfaces. Solvent-accessible surface area was calculated by discovery studio software rolling ball method with a radius of 1.4 Å. The solvent accessible surface was predicted to be 2690 $Å^2$ for the double helix state and 1790 $Å^2$ for the single helix. Side chains have been removed for the surface estimation.

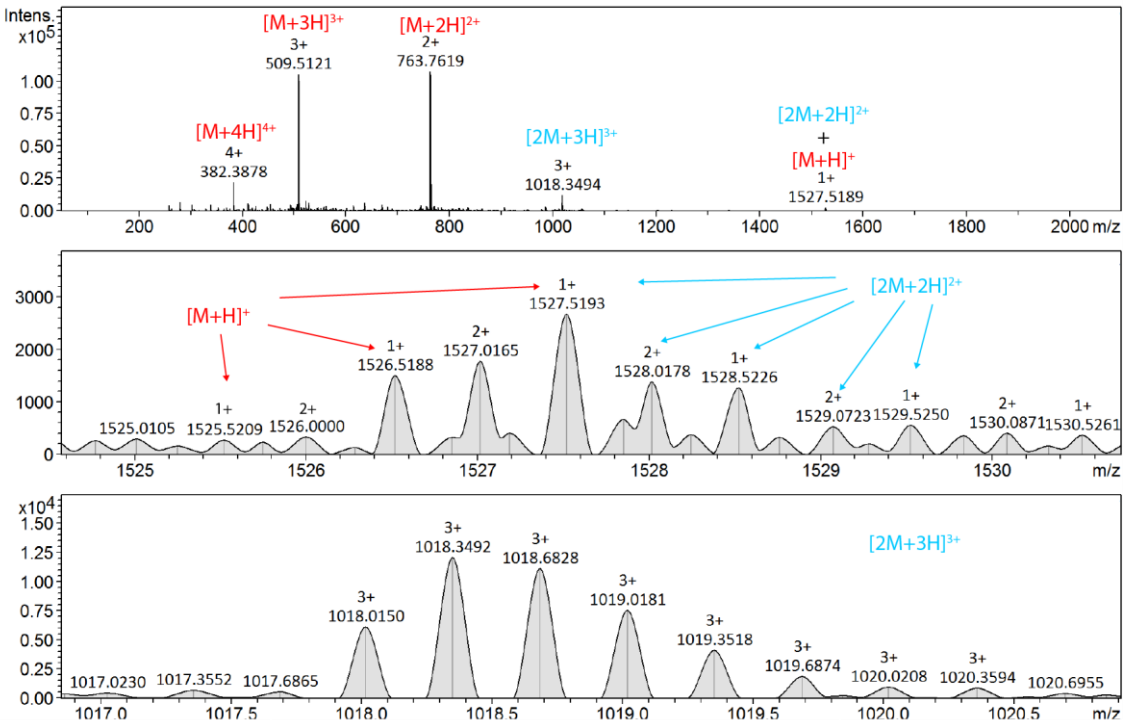


Figure S83. ESI-MS spectrum of oligomer 6 (direct infusion of an aqueous sample). Peaks corresponding to monomeric and dimeric species are highlighted in red and blue, respectively. The $[M+H]^+$ peak is superimposed by the $[2M+2H]^{2+}$ peak, while the mass envelope suggests a major population of the $[2M+2H]^{2+}$ species.

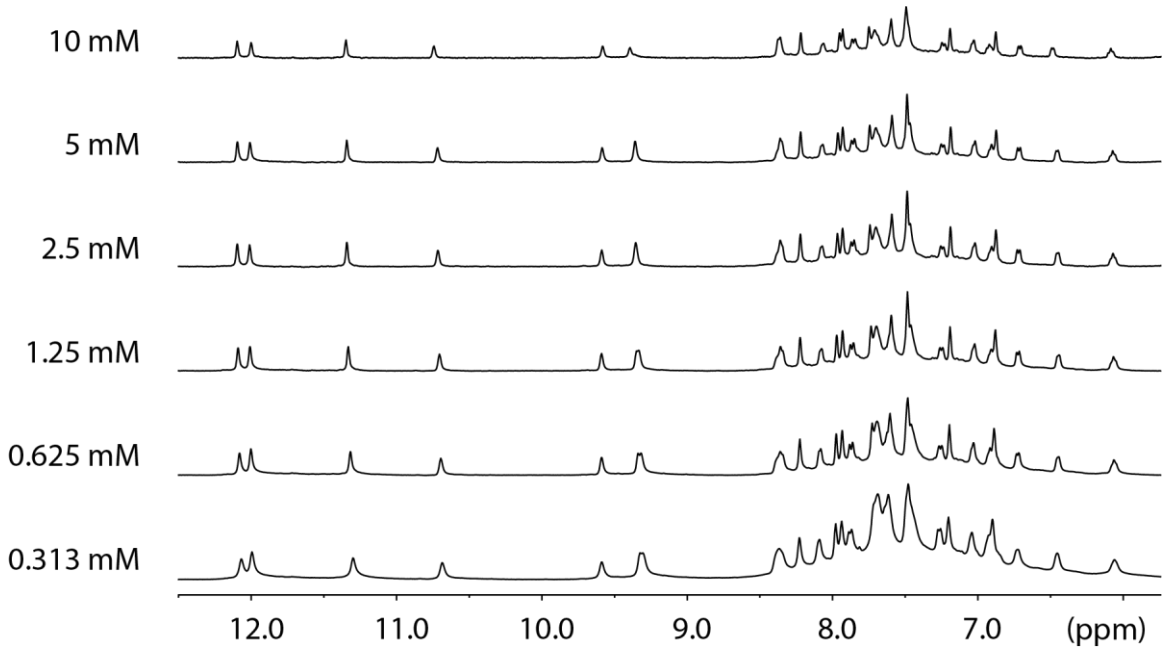


Figure S84. Variable concentration ^{1}H NMR spectra of oligomer 6. The spectra were recorded with water suppression (500 MHz, H_2O/D_2O (9:1, *v*/*v*), 25 °C).

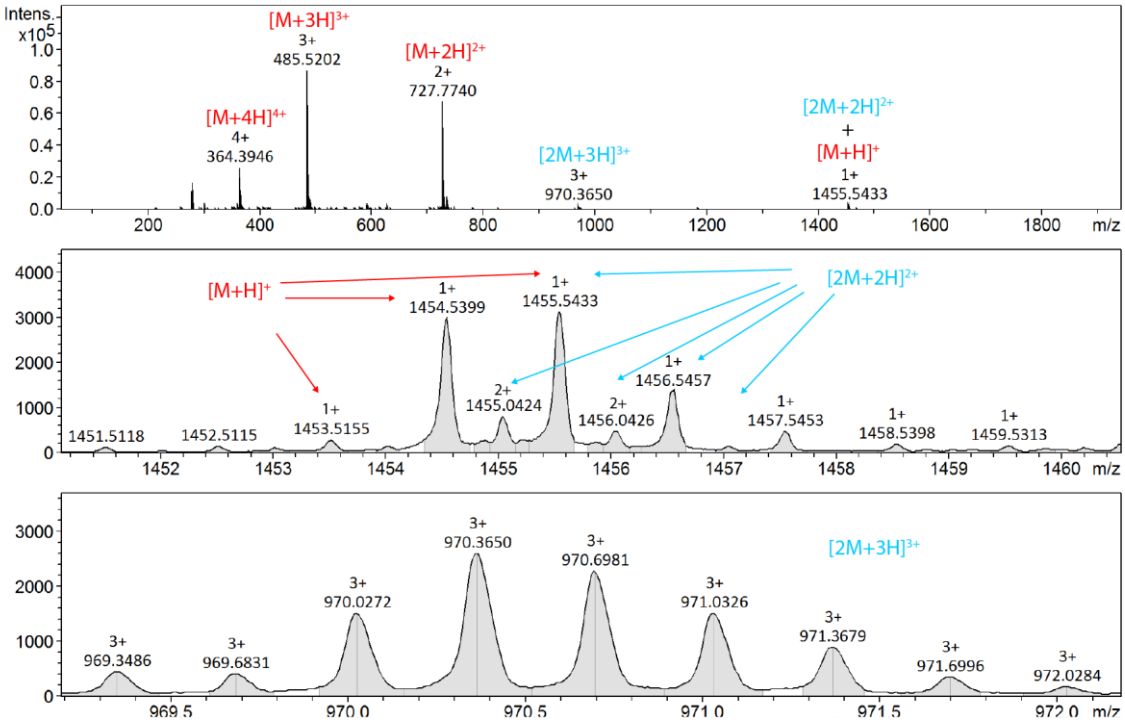


Figure S85. ESI-MS spectrum of oligomer 7 (direct infusion of an aqueous sample). Peaks corresponding to monomeric and dimeric species are highlighted in red and blue, respectively. The $[M+H]^+$ peak is superimposed by the $[2M+2H]^{2+}$ peak.

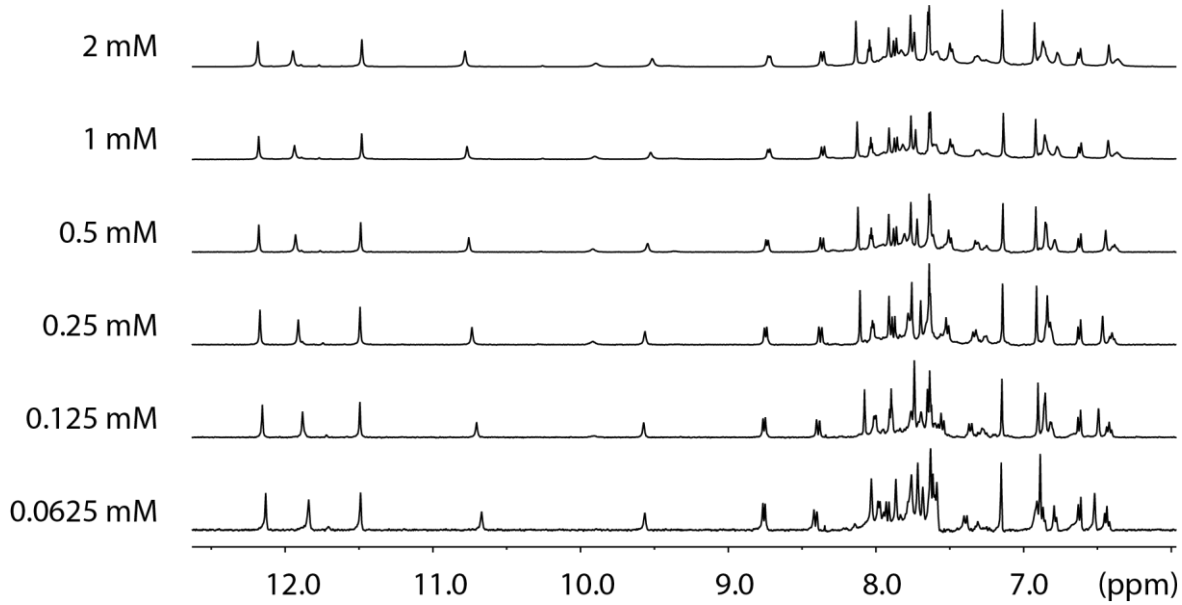


Figure S86. Variable concentration ^{1}H NMR spectra of oligomer 7. The spectra were recorded with water suppression (500 MHz, H_2O/D_2O (9:1, *v*/*v*), 25 °C).

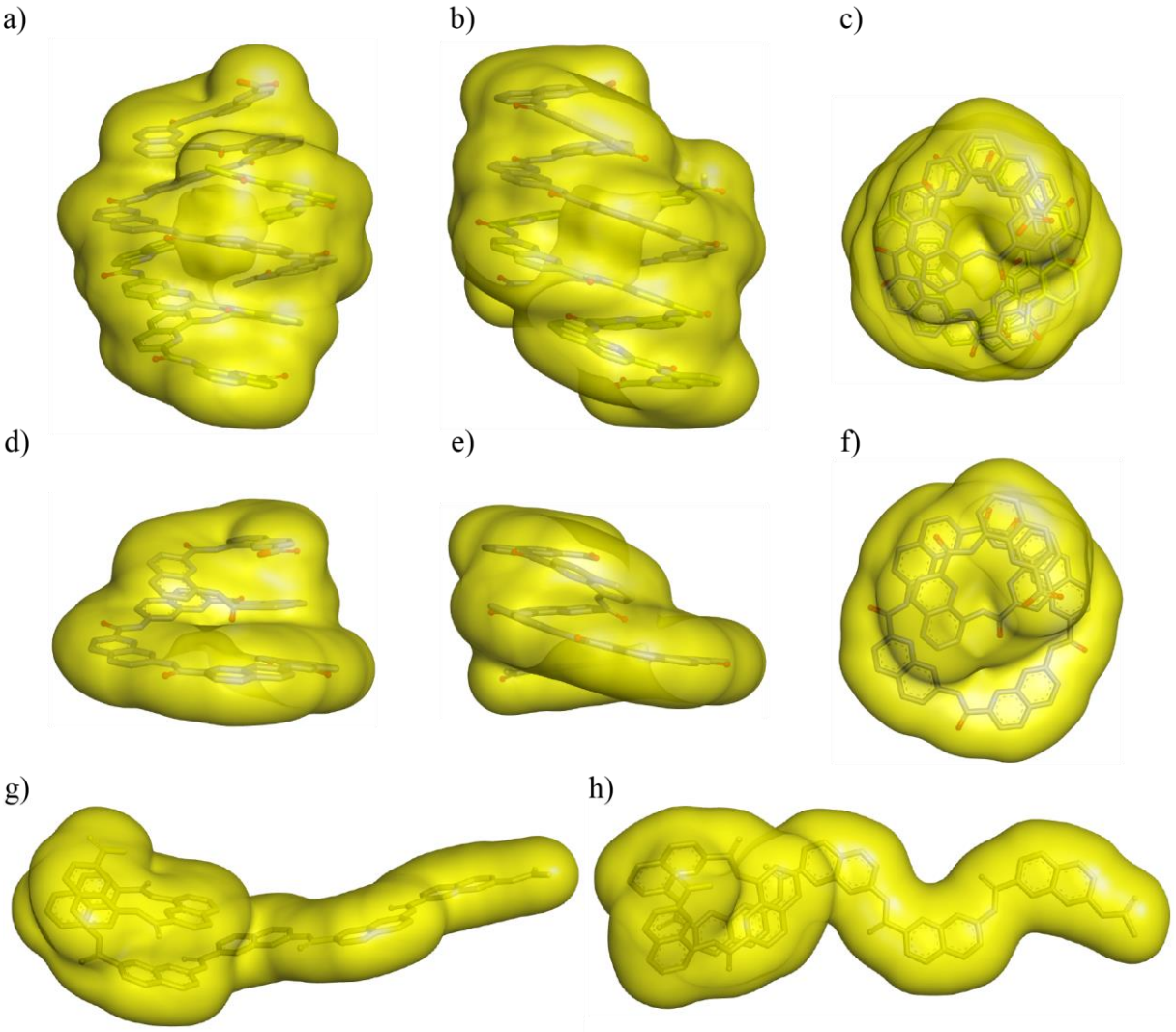


Figure S87. Solvent accessible surface of three conformations of oligomer 7: (a) front view, b) side view and c) top view of duplex $(\mathbf{7})_2$. Solvent accessible surface of: d) front view, e) side view and f) top view of single helix **7**. Solvent accessible surface of: g) front view and h) top view of extended state **7**. All solvent accessible surfaces are shown as transparent yellow isosurfaces. Solvent-accessible surface area was calculated by discovery studio software rolling ball method with a radius of 1.4 Å. The solvent accessible surface was predicted to be 1680 $Å^2$ for the double helix state, 1380 $Å^2$ for the extended monomer and 1160 $Å^2$ for the single helix. Thus, it can be estimated that the double helix form exposes to the solvent only around 61% or 71% of the area exposed by two individual monomeric species. Side chains have been removed for the surface estimation.

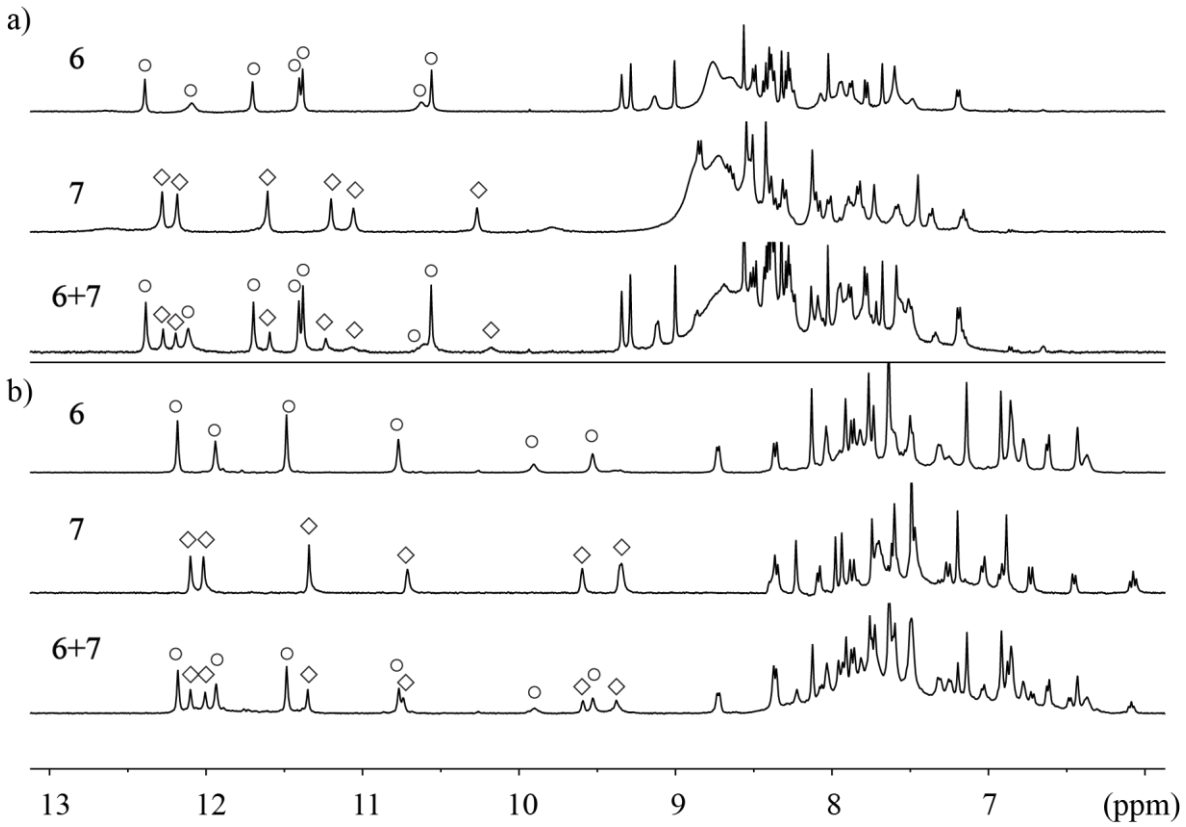


Figure S88. Hybridization experiments of 6 and 7. Excerpts of ^{1}H NMR (500 MHz, water suppression) at 25 °C of oligomer **6**, oligomer **7**, and the mixture of two oligomers at 0.5 mM in a) DMSO-d_6 and b) H_2O/D_2O (9:1, *v*/*v*). Signals corresponding to oligomer **6** and **7** are marked in circles and diamonds, respectively.

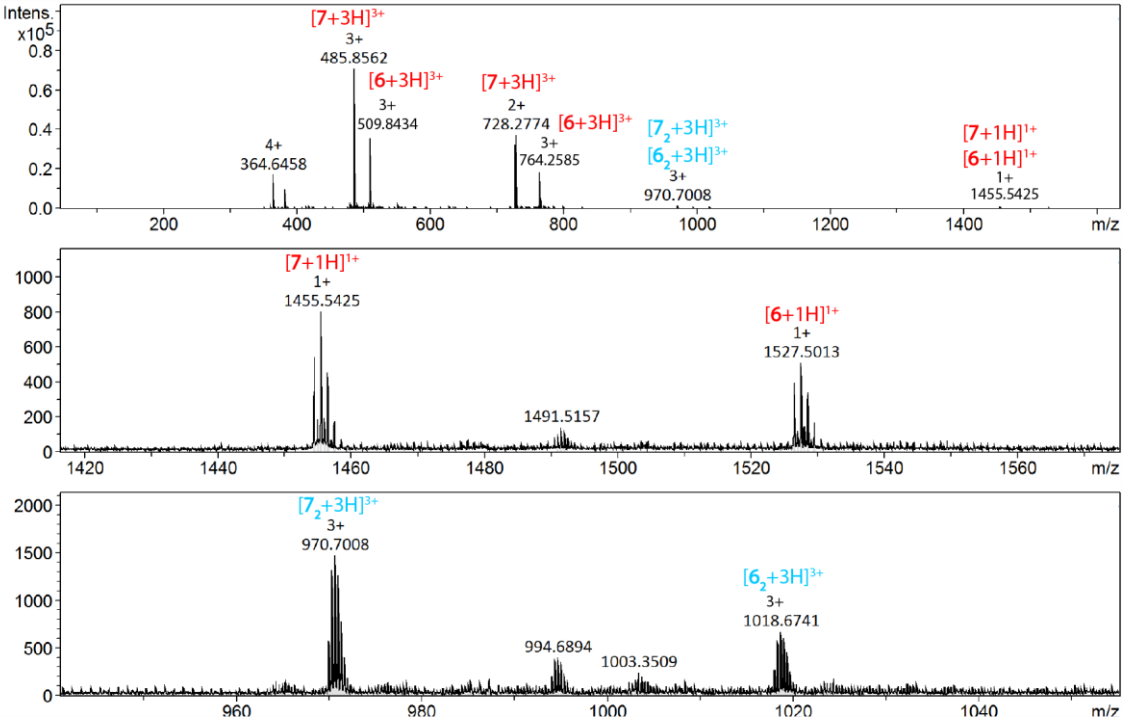


Figure S89. ESI-MS spectrum of oligomer 6 and 7 mixture (direct infusion of an aqueous sample). Peaks corresponding to monomeric and dimeric species are highlighted in red and blue, respectively.

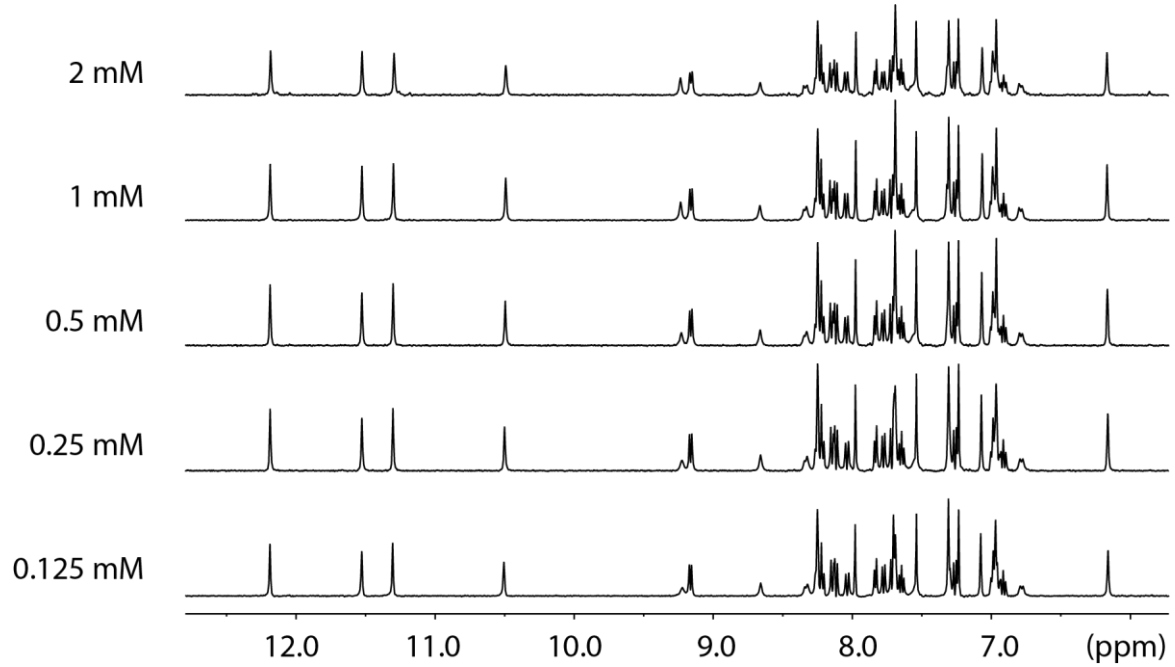


Figure S90. Variable concentration ^{1}H NMR spectra of oligomer 8. The spectra were recorded with water suppression (500 MHz) in 12 mM ammonium acetate aqueous buffer pH 8.5 at 25 °C.

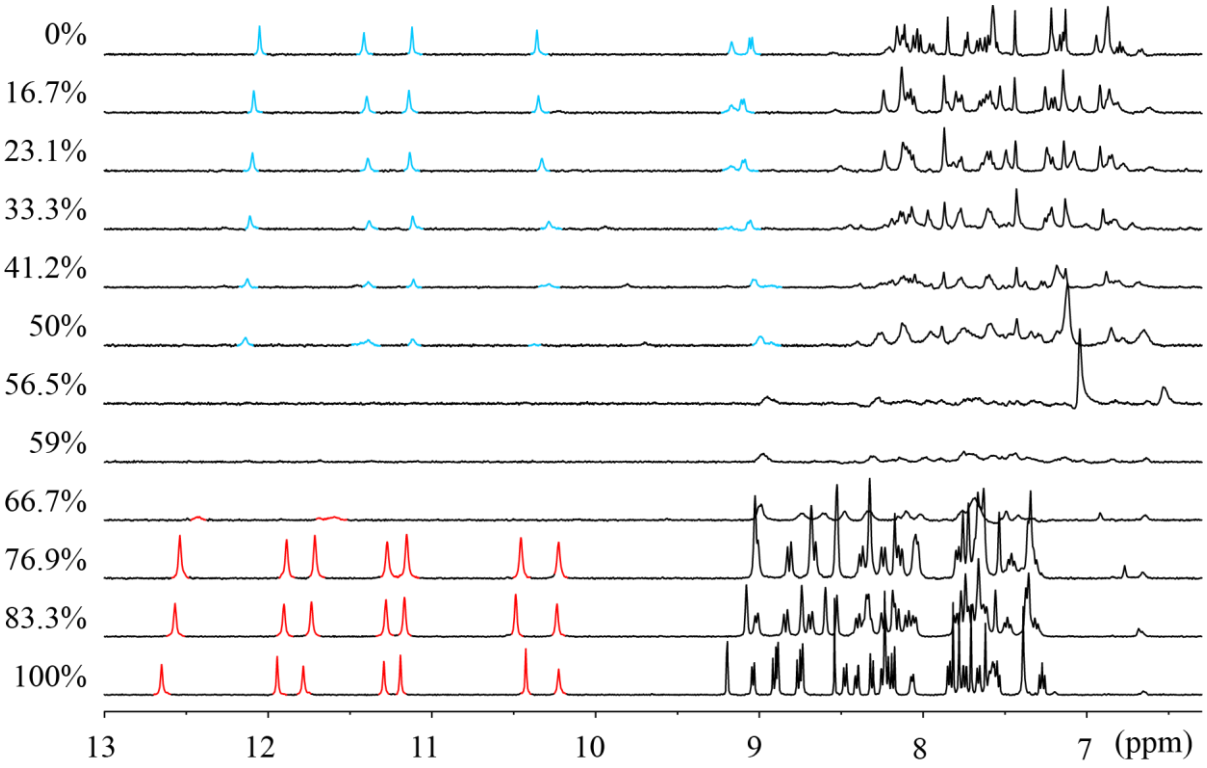


Figure S91. DMSO-d_6/H_2O variation ^{1}H NMR spectra of 8. ^{1}H NMR (500 MHz) spectra at 25 °C of oligomer **8** in DMSO-d_6/H_2O mixtures with different DMSO contents. Signals corresponding to monomeric and dimeric species are highlighted in red and blue, respectively. Note: 0% DMSO means ^{1}H NMR (500 MHz) spectra of oligomer **8** in H_2O/D_2O (9:1, *v*/*v*).

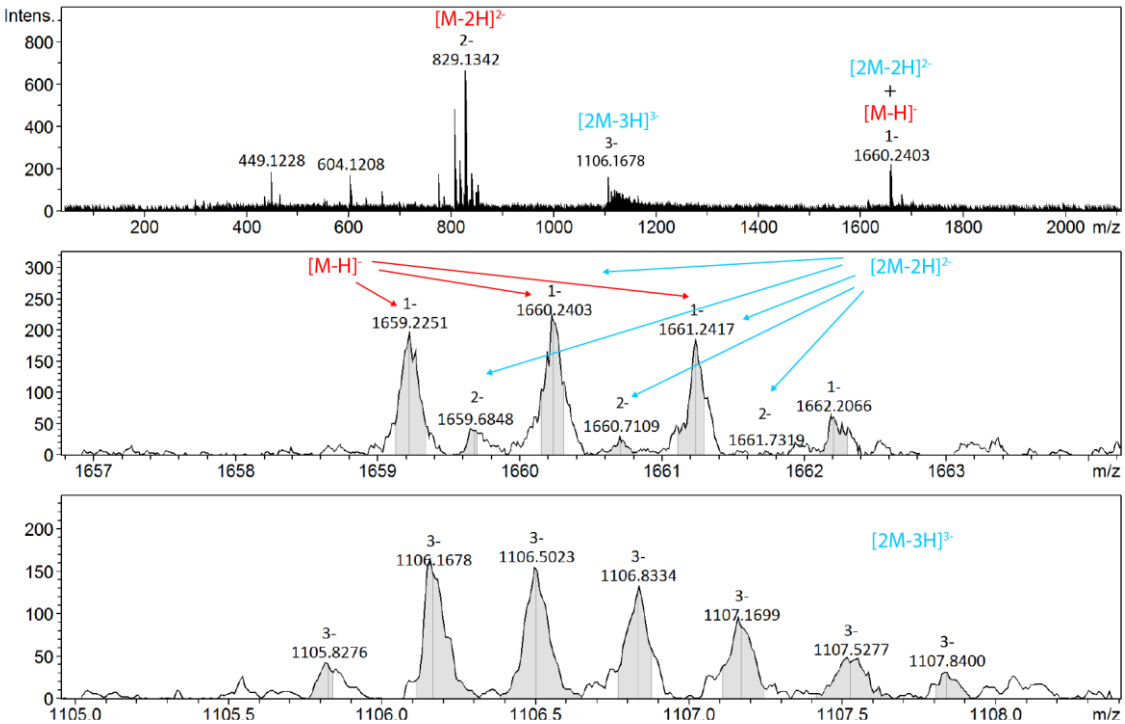


Figure S92. ESI-MS spectrum of oligomer 8 (direct infusion of an aqueous sample). Peaks corresponding to monomeric and dimeric species are highlighted in red and blue, respectively. The $[M-H]^-$ peak is superimposed by the $[2M-2H]^{2-}$ peak.

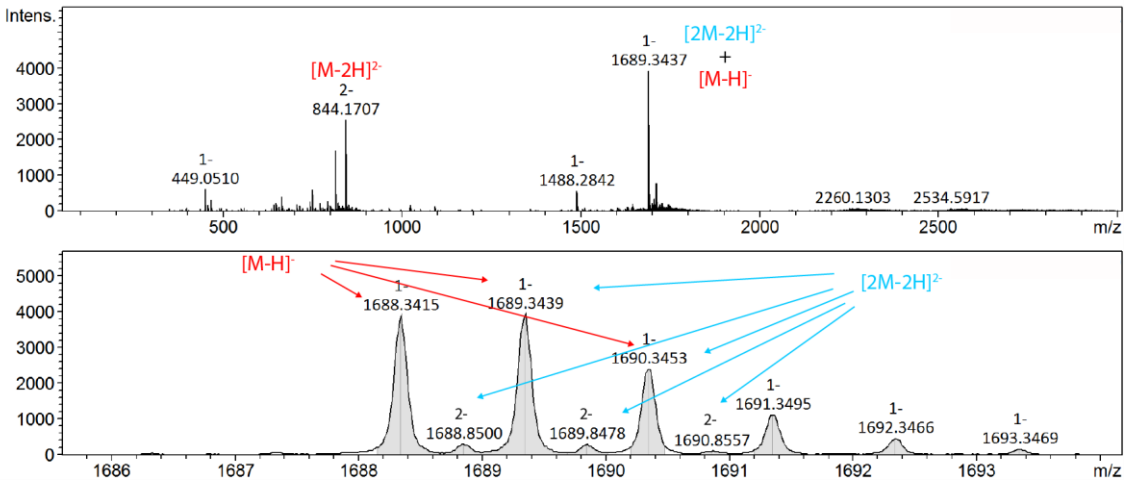


Figure S93. ESI-MS spectrum of oligomer 9 (direct infusion of an aqueous sample). Peaks corresponding to monomeric and dimeric species are highlighted in red and blue, respectively. The $[M-H]^-$ peak is superimposed by the $[2M-2H]^{2-}$ peak.

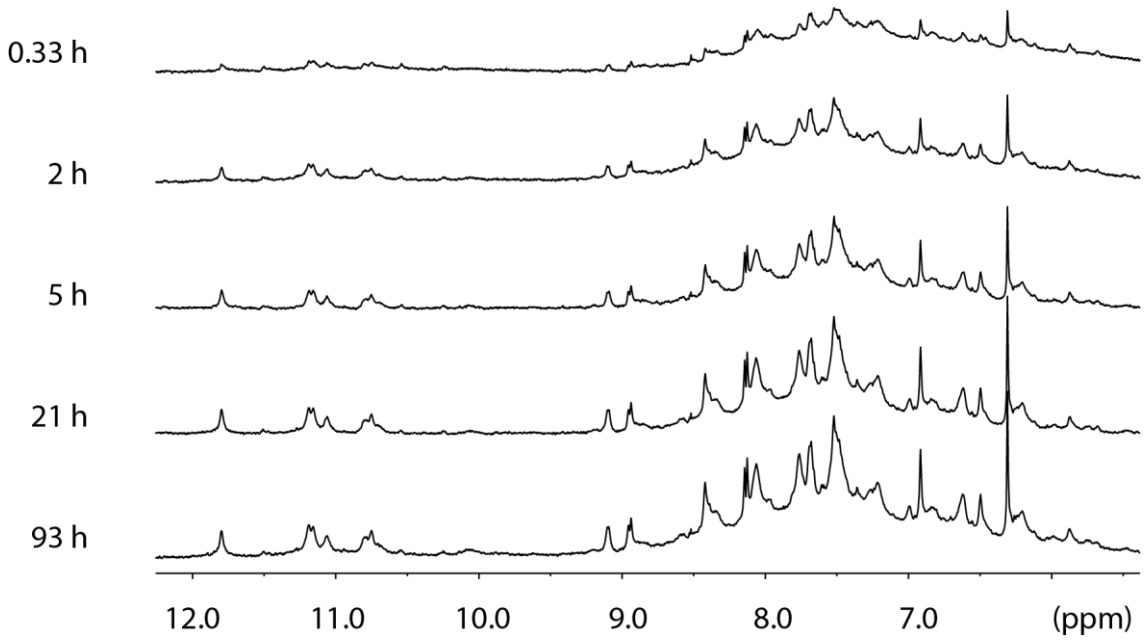


Figure S94. Variable incubation time ^{1}H NMR spectra of 9 in buffer solution. Excerpts of ^{1}H NMR (500 MHz, water suppression) at 25 °C of oligomer **9** at 5 mM in 48 mM ammonium acetate aqueous buffer pH 8.5 in the different incubation times.

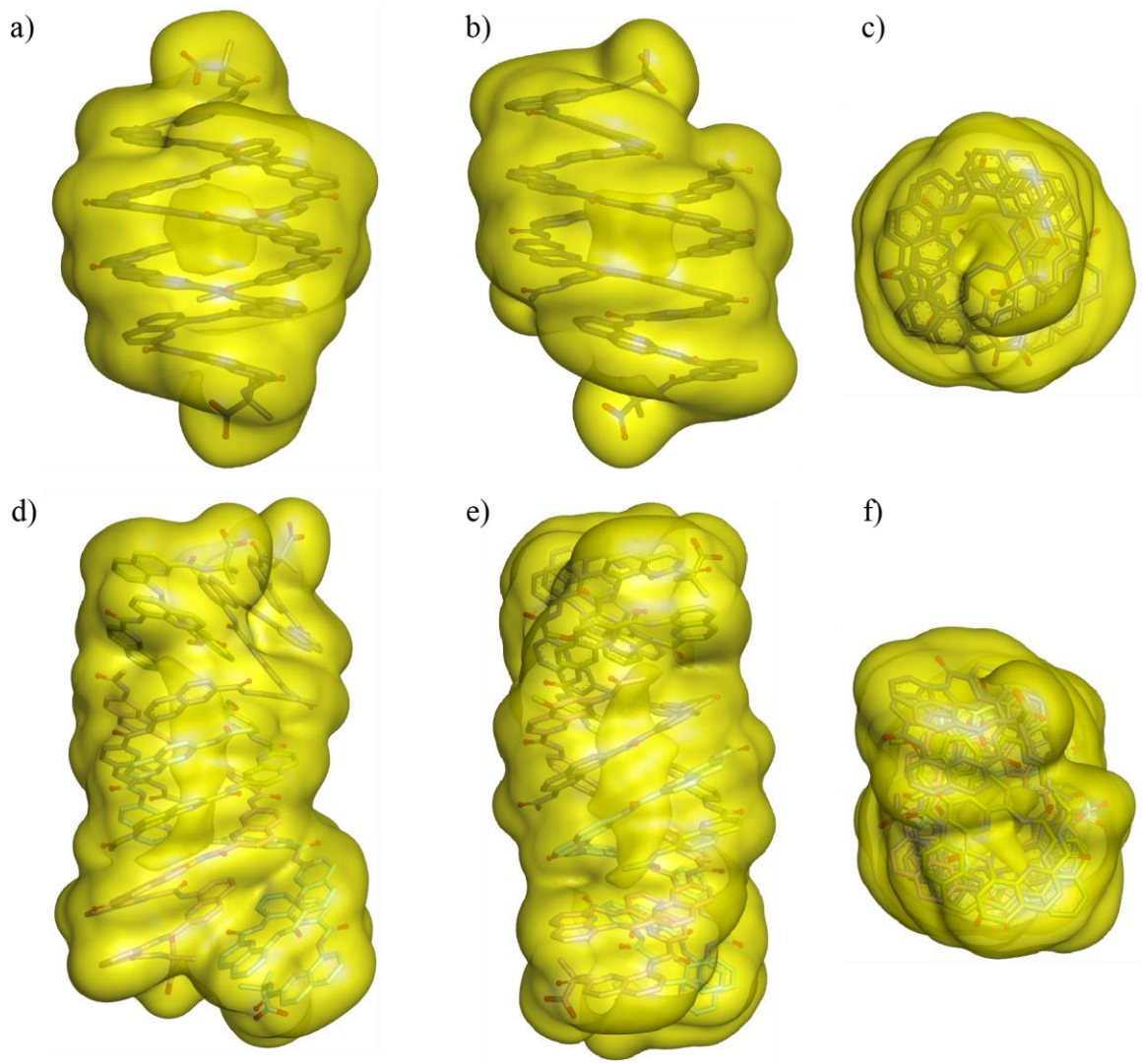


Figure S95. Solvent accessible surface of two conformations of 9: (a) front view, b) side view and c) top view of duplex (**9**)$_2$. Solvent accessible surface of: d) front view, e) side view and f) top view of tetrameric helix (**9**)$_4$. All solvent accessible surfaces are shown as transparent yellow isosurfaces. Solvent-accessible surface area was calculated by discovery studio software rolling ball method with a radius of 1.4 Å. The solvent accessible surface was predicted to be 3210 Å^2 for the tetra helix state and 1810 Å^2 for the double helix state. Thus, it can be estimated that the tetra helix form exposes to the solvent only around 89% of the area exposed by two individual double species.

2 Materials and Methods

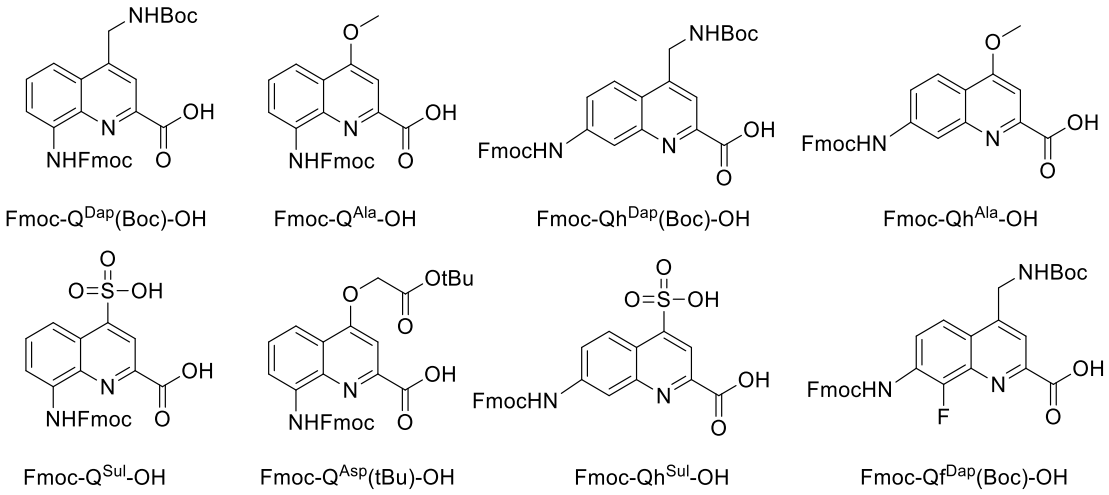


Figure S96. Protected Fmoc-acid building blocks used in this study. Fmoc-Q^{Dap}(Boc)-OH[1], Fmoc-Q^{Ala}-OH[2], Fmoc-Q^{Sul}-OH[3], Fmoc-Q^{Asp}(*t*Bu)-OH[4], Fmoc-Qh^{Dap}(Boc)-OH[1] and Fmoc-Qf^{Dap}(Boc)-OH[1] have been described previously. For detailed synthetic procedures to prepare Fmoc-Qh^{Ala}-OH and Fmoc-Qh^{Sul}-OH, see section 2.3.

2.1 General

General. Commercial reagents (Suppliers: Abcr, Fisher Scientific, Merck, Sigma-Aldrich, TCI, BLDpharm or VWR) were used without further purification unless otherwise stated. LL Wang resin (100–200 mesh) was purchased from Sigma-Aldrich. Cl-MPA protide resin was purchased from CEM. Peptide grade *N*,*N*-dimethylformamide (DMF) was purchased from Carlo Erba. Anhydrous chloroform, triethylamine (TEA) and *N*,*N*-diisopropylethylamine (DIEA) were obtained via distillation over CaH_2 prior to use. Anhydrous tetrahydrofuran (THF) and dichloromethane (DCM) were obtained via an MBRAUN SPS-800 solvent purification system. Ultrapure water was obtained via a Sartorius arium® pro VF 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). Nuclear magnetic resonance (NMR) spectra were recorded on an Avance III HD 400 MHz Bruker BioSpin spectrometer or an Avance III HD 500 MHz Bruker BioSpin spectrometer equipped with a broad band observe 5-mm BB-H&FD CryoProbe™ Prodigy. Measurements were performed at 25 °C unless stated otherwise. Water suppression was performed with excitation sculpting method. Processing was done with

MestReNova (v.12.0.0-20080) NMR processing software from Mestrelab Research. Chemical shifts are reported in ppm and calibrated via residual solvent signals or 3-(trimethylsilyl)propionic-2,2,3,3-d_4 acid sodium salt (TMSP) when water suppression was applied. Signal multiplicities are abbreviated as s, singlet; d, doublet; t, triplet; q, quartet, and m, multiplet. Signals were assigned using ^{1}H-^{13}C HMQC and ^{1}H-^{13}C HMBC spectra. Electrospray ionization (ESI) mass spectra were recorded on Bruker microTOF II and Thermo Finnigan LTQ FT Ultra spectrometers. Electron ionization (EI) mass spectra were recorded on a Thermo Q Exactive GC Orbitrap or a Finnigan MAT 95 sector mass spectrometer. Analytical and semi-preparative reversed phase (RP) high performance liquid chromatography (HPLC) were performed on a Thermo Fisher Scientific Ultimate 3000 HPLC System using Macherey-Nagel Nucleodur C18 Gravity columns (4 × 100 mm, 5 µm and 10 × 250 mm, 5 µm) or Macherey-Nagel Nucleodur C8 Gravity columns (4 × 50 mm, 5 µm and 10 × 100 mm, 5 µm) with different gradients of 0.1% TFA water and 0.1% TFA acetonitrile. All ultraviolet–visible (UV/Vis) absorbance measurements were done with a Jasco V-750 spectrophotometer instrument using a 1 cm quartz cuvette. Measurements were performed at 20 °C if not stated otherwise. Microwave-assisted solid phase foldamer synthesis (SPFS) was performed with a CEM® Discover Bio manual microwave apparatus. The temperature within the reactor vessel was monitored with an optical fiber probe. Automated SPFS was done on a PurePep® Chorus synthesizer (Gyros Protein Technologies) by applying induction heating. The DOSY spectra were recorded on a Avance III HD 500 MHz Bruker BioSpin spectrometer equipped with a broad band observe 5-mm BB-H&FD CryProbe™ Prodigy with a pulse sequence with stimulated echo using bipolar gradient pulses for diffusion and a 3-9-19 watergate solvent suppression pulse sequence from the Bruker pulse program library (stebpgp1s19). The diffusion delay Δ (big delta) was set to 175 ms and the diffusion gradient pulse length δ (little delta) was set to 1.1 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, 128 transients of 32k complex data points were acquired. DOSY processing was performed with the DOSY processing tool from MestReNova (v.12.0.0-20080).

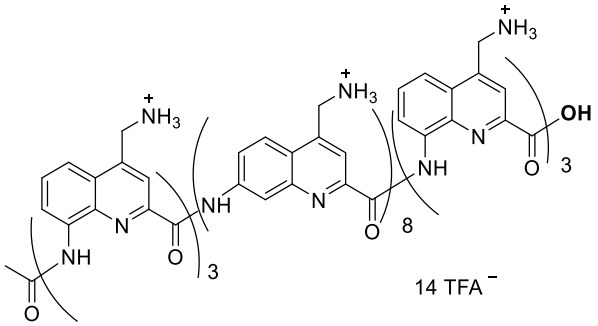
2.2 Solid phase synthesis procedures

Some oligomers were synthesized according to previously reported SPFS protocols,[5] hereafter referred to as manual synthesis method. Of note, in the case of Q^{Sul} and Qh^{Sul}-containing sequences, 2% DBU in NMP was employed for Fmoc deprotection, thus avoiding the additional

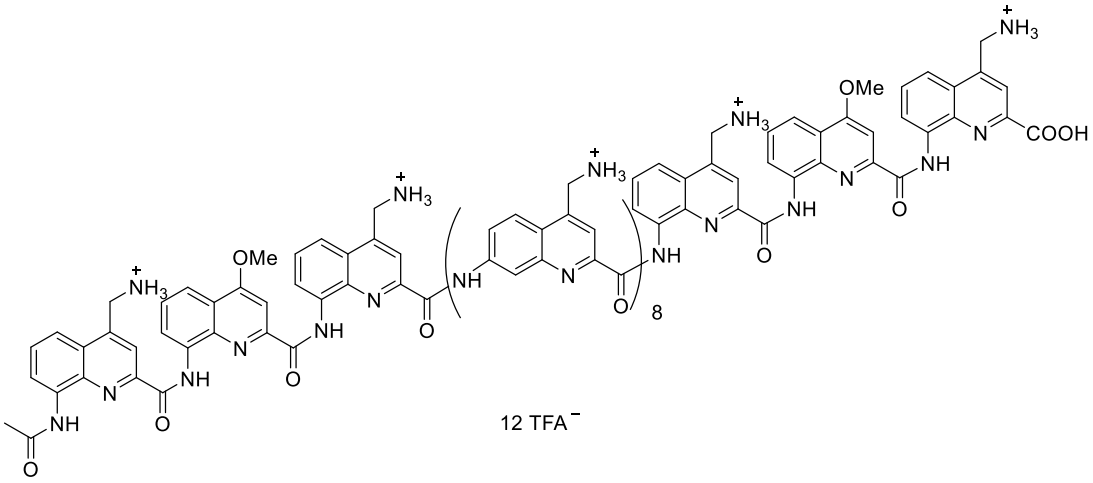
washing steps with 20% DIEA in NMP like previously reported.[5] Other oligomers were prepared using the automated foldamer solid phase synthesis protocols recently reported.[6]

Final acetylation of the N-terminal aromatic amine: H_2N-$(Q)_n$-Wang resin and DIEA (6.0 equiv.) were suspended in anhydrous THF (1.25 mL) then acetyl chloride (4.0 equiv.) in anhydrous THF (1.25 mL) was added. The reaction vessel was then placed under microwave irradiation (25 W, ramp to 50 °C over 5 min, then hold at 50 °C for 15 min). The resin was filtered off and washed with anhydrous THF (2 × 3 mL). This acetylation step was repeated once using the same conditions.

Resin cleavage: The resin-bound foldamer was placed in a syringe equipped with a filter, washed with DMF (3 × 3 mL) and DCM (3 × 3 mL), and dried by passing N_2 flow through it. It was then suspended in a solution of TFA/*i*Pr_3SiH/H_2O (95:2.5:2.5, *v*/*v*/*v*). The resin was next gently shaken for at least 2 h at room temperature, then filtered off and rinsed one time with TFA. The foldamer was precipitated from the TFA cleavage solution by adding cold Et_2O and centrifugation to obtain a crude precipitate.

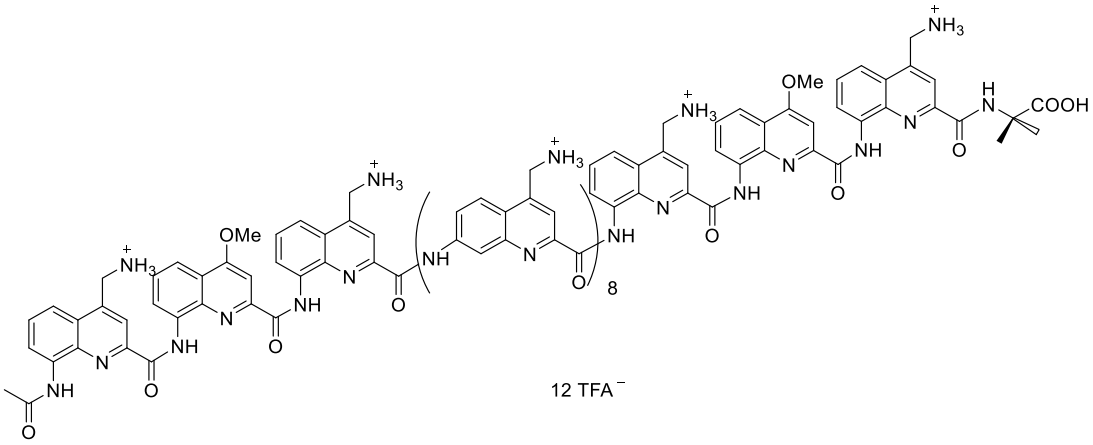


Oligomer 1: The synthesis of this oligomer has already been published.[1]



Oligomer 2: Oligomer **2** was synthesized starting from LL-Wang resin (0.37 mmol g^{-1}, 25 µmol scale) according to the manual synthesis method. Loading of the first monomer: 0.25 mmol g^{-1} (70%). Final acetylation was carried out via the general acetylation method. After precipitation in cold Et_2O, the crude mixture was purified by semi prep RP-HPLC to give the **2** as a yellow solid (48 mg, 45%).

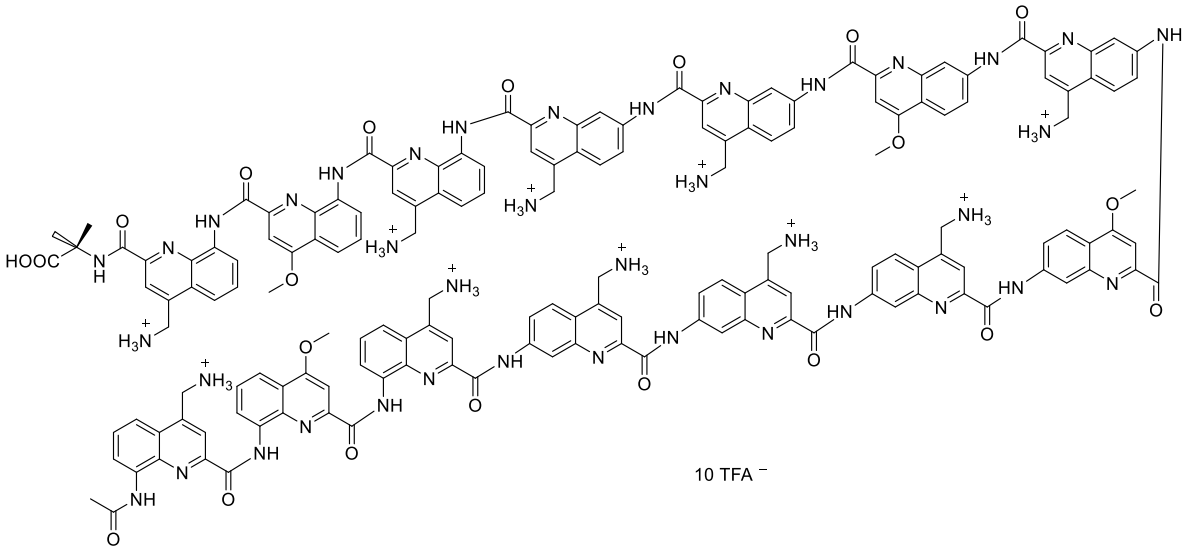
^{1}H NMR (500 MHz, H_2O/D_2O (9:1, *v*/*v*)) δ 12.11 (s, 1H), 12.03 (s, 1H), 12.01 (s, 1H), 11.56 (s, 1H), 11.33 (s, 1H), 10.72 (s, 1H), 10.54 (s, 1H), 10.37 (s, 1H), 10.20 (s, 1H), 10.03 (s, 1H), 10.00 (s, 1H), 9.54 (s, 1H), 9.47 (s, 1H), 9.38 (s, 1H), 9.34 (s, 1H), 8.82 – 8.71 (m, 2H), 8.61 (d, *J* = 10.7 Hz, 1H), 8.52 (d, *J* = 11.9 Hz, 2H), 8.39 – 7.99 (m, 18H), 7.99 – 7.71 (m, 13H), 7.71 – 7.52 (m, 12H), 7.52 – 7.07 (m, 26H), 7.07 – 6.85 (m, 12H), 6.84 – 6.57 (m, 6H), 6.56 – 6.27 (m, 6H), 3.53 (s, 3H), 3.21 (s, 3H).
HRMS: calcd. for $C_{156}H_{129}N_{40}O_{18}$ $[M+H]^+$ 2850.0403; found 2849.9474.



Oligomer 3: Oligomer **3** was synthesized starting from LL-Wang resin (0.37 mmol g^{-1}, 23 µmol scale) according to the automated synthesis method. Loading of the first monomer: 0.26 mmol g^{-1} (70%). Final acetylation was carried out via the general acetylation method. After precipitation in cold Et_2O, the crude mixture was purified by semi prep RP-HPLC to give the **3** as a yellow solid (56 mg, 56%).

^{1}H NMR (500 MHz, H_2O/D_2O (9:1, *v*/*v*)) δ 12.07 (s, 1H), 12.03 (s, 2H), 11.47 (s, 1H), 11.10 (s, 1H), 10.72 (s, 1H), 10.50 (s, 1H), 10.41 (s, 1H), 10.28 (s, 1H), 10.03 (s, 1H), 9.84 (s, 1H), 9.68 (s, 1H), 9.47 (s, 1H), 9.36 (s, 1H), 8.77 (d, *J* = 10.7 Hz, 1H), 8.74 – 8.63 (m, 2H), 8.53 (d, *J* = 10.6 Hz, 1H), 8.46 (d, *J* = 10.1 Hz, 1H), 8.37 (s, 1H), 8.34 – 8.14 (m, 8H), 8.10 – 7.99 (m, 5H), 7.98 – 7.58 (m, 16H), 7.58 – 6.85 (m, 27H), 6.74 – 6.44 (m, 7H), 6.33 (d, *J* = 9.0 Hz, 1H), 3.49 (s, 3H), 3.22 (s, 3H), 1.32 (s, 3H), 0.72 (s, 3H), 0.45 (s, 3H).

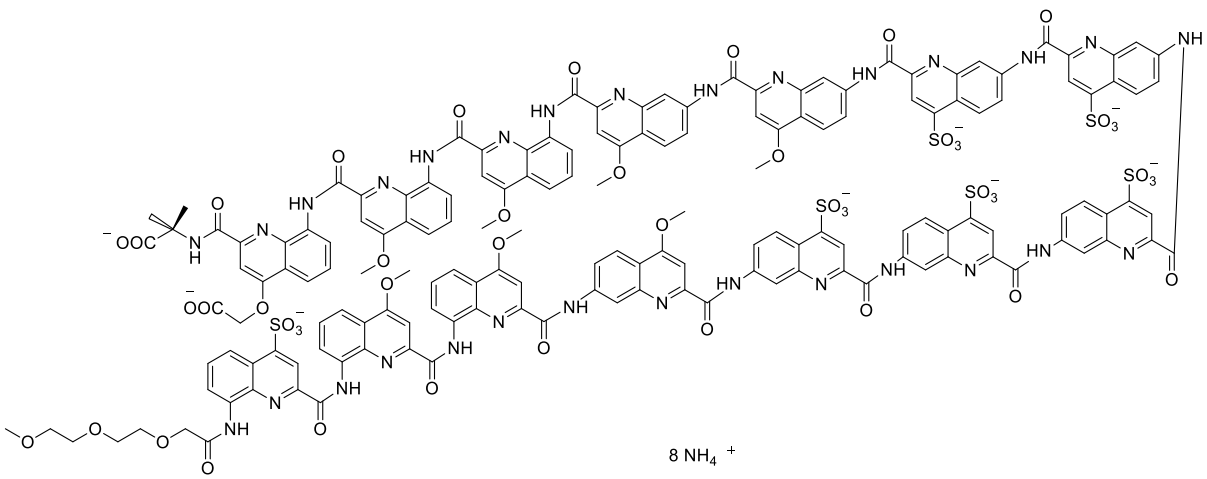
HRMS: calcd. for $C_{160}H_{137}N_{41}O_{19}$ $[M+2H]^{2+}$ 1468.0502; found 1468.1223.



Oligomer 4: Oligomer **4** was synthesized starting from LL-Wang resin (0.37 mmol g^{-1}, 29 µmol scale) according to the manual synthesis method. Loading of the first monomer: 0.28 mmol g^{-1} (76%). Final acetylation was carried out via the general acetylation method. After precipitation in cold Et_2O, the crude mixture was purified by semi prep RP-HPLC to give the **4** as a yellow solid (42 mg, 36%).

^{1}H NMR (500 MHz, H_2O/D_2O (9:1, *v*/*v*)) δ 11.70 (s, 1H), 11.22 (s, 1H), 10.89 (s, 1H), 10.55 (s, 1H), 10.48 (s, 1H), 10.43 (s, 1H), 10.19 (s, 1H), 9.65 (s, 1H), 9.50 (s, 1H), 9.10 (s, 1H), 8.65 (d, *J* = 7.8 Hz, 1H), 8.46 (s, 1H), 8.41 (d, *J* = 10.1 Hz, 1H), 8.31 (s, 1H), 8.29 – 8.06 (m, 7H), 7.89 (d, *J* = 9.7 Hz, 1H), 7.76 – 7.47 (m, 21H), 7.46 – 7.17 (m, 22H), 7.17 – 6.81 (m, 20H), 6.81 – 6.46 (m, 12H), 6.45 – 6.17 (m, 7H), 6.15 – 6.00 (m, 2H), 3.90 (s, 3H), 3.22 (s, 3H), 3.04 (s, 3H), 1.82 (s, 3H).

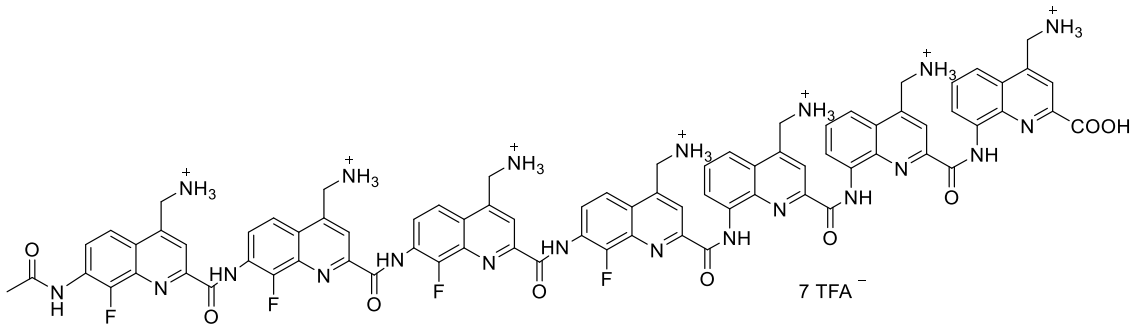
HRMS: calcd. for $C_{160}H_{135}N_{39}O_{21}$ $[M+2H]^{2+}$ 1469.0342; found 1469.0595.



Oligomer 5: Oligomer **5** was synthesized starting from Cl-MPA-Protide resin (0.15 mmol g^{-1}, 20 µmol scale) according to the manual synthesis method. Loading of the first monomer: 0.13 mmol g^{-1} (86%). After precipitation in cold Et_2O, the crude mixture was purified by semi prep RP-HPLC to give the **5** as a yellow solid (15 mg, 22%).

^{1}H NMR (500 MHz, ammonium acetate aqueous buffer pH 8.5/D_2O (9:1, *v*/*v*)) δ 11.85 (s, 1H), 11.70 (s, 1H), 11.56 (s, 1H), 11.37 (s, 1H), 10.48 (s, 1H), 10.33 (s, 1H), 9.92 (s, 1H), 9.66 (s, 1H), 9.52 (s, 1H), 9.28 (s, 1H), 9.04 (d, *J* = 9.7 Hz, 1H), 8.85 (d, *J* = 9.4 Hz, 1H), 8.74 (s, 1H), 8.70 – 7.75 (m, 29H), 7.69 (d, *J* = 8.7 Hz, 2H), 7.50 – 7.03 (m, 19H), 7.01 – 6.11 (m, 21H), 5.94 (s, 2H), 4.00 (s, 4H), 3.43 (s, 5H), 3.10 (m, 13H), 2.62 (s, 3H), 2.34 (d, *J* = 18.6 Hz, 8H), 1.21 (m, 16H), 0.75 (s, 6H).

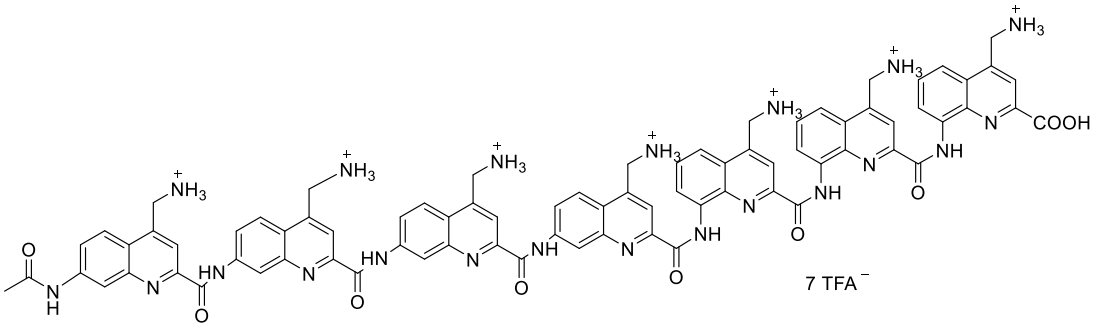
HRMS: calcd. for $C_{160}H_{119}N_{29}O_{48}S_6$ $[M-2H]^{2-}$ 1702.8038; found 1702.7972.



Oligomer 6: Oligomer **6** was synthesized starting from LL-Wang resin (0.44 mmol g^{-1}, 26 µmol scale) according to the manual synthesis method. Loading of the first monomer: 0.32 mmol g^{-1} (74%). Final acetylation was carried out via the general acetylation method. After precipitation in cold Et_2O to give the **6** as a grey solid (47 mg, 78%).

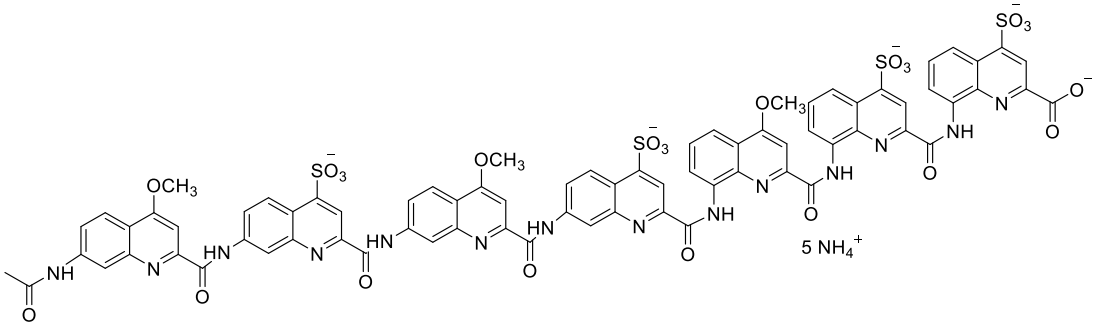
^{1}H NMR (400 MHz, H_2O/D_2O (9:1, *v*/*v*)) δ 12.10 (s, 1H), 12.02 (s, 1H), 11.34 (s, 1H), 10.71 (s, 1H), 9.60 (s, 1H), 9.35 (d, *J* = 5.5 Hz, 2H), 8.43 – 8.31 (m, 2H), 8.23 (s, 1H), 8.09 (d, *J* = 7.3 Hz, 1H), 7.98 (s, 1H), 7.94 (s, 1H), 7.89 (s, 1H), 7.86 (s, 1H), 7.77 – 7.63 (m, 6H), 7.63 – 7.54 (m, 3H), 7.53 – 7.40 (m, 7H), 7.26 (d, *J* = 9.8 Hz, 3H), 7.20 (s, 1H), 7.04 (d, *J* = 8.5 Hz, 2H), 6.96 – 6.80 (m, 3H), 6.73 (d, *J* = 9.0 Hz, 1H), 6.45 (d, *J* = 8.2 Hz, 1H), 6.08 (t, *J* = 8.3 Hz, 1H), 4.05 (m, 1H), 3.72 (d, *J* = 14.8 Hz, 1H), 3.47 (q, *J* = 7.3 Hz, 1H), 2.74 (d, *J* = 16.4 Hz, 1H), 1.75 (s, 3H).

HRMS: calcd. for $C_{79}H_{64}F_4N_{21}O_9$ $[M+H]^+$ 1526.5127; found 1526.5634.



Oligomer 7: Oligomer **7** was synthesized starting from LL-Wang resin (0.44 mmol g^{-1}, 23 µmol scale) according to the manual synthesis method. Loading of the first monomer: 0.31 mmol g^{-1} (70%). Final acetylation was carried out via the general acetylation method. After precipitation in cold Et_2O to give the **7** as a grey solid (30 mg, 88%).

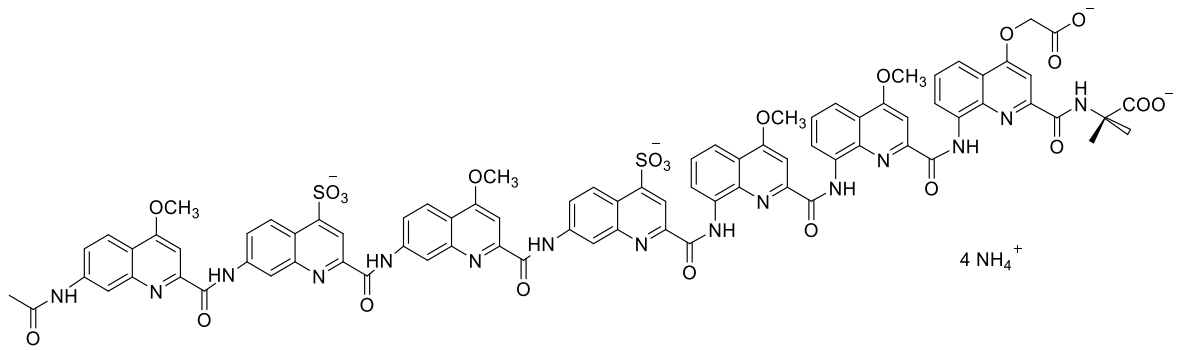
^{1}H NMR (500 MHz, H_2O/D_2O (9:1, *v*/*v*)) δ 12.18 (s, 1H), 11.94 (s, 1H), 11.49 (s, 1H), 10.77 (s, 1H), 9.91 (s, 1H), 9.53 (s, 1H), 8.73 (d, *J* = 8.1 Hz, 1H), 8.36 (d, *J* = 10.3 Hz, 1H), 8.13 (s, 2H), 8.04 (t, *J* = 4.5 Hz, 2H), 7.98 – 7.70 (m, 10H), 7.69 – 7.42 (m, 8H), 7.40 – 7.19 (m, 3H), 7.14 (s, 2H), 6.85 (m, 6H), 6.62 (d, *J* = 8.9 Hz, 2H), 6.43 (s, 1H), 6.37 (s, 1H).
HRMS: calcd. for $C_{79}H_{68}N_{21}O_9$ $[M+H]^+$ 1454.5503; found 1454.5802.



Oligomer 8: Oligomer **8** was synthesized starting from Cl-MPA-Protide resin (0.15 mmol g^{-1}, 20 µmol scale) according to the manual synthesis method. Loading of the first monomer: 0.14 mmol g^{-1} (90%). Final acetylation was carried out via the general acetylation method. After precipitation in cold Et_2O, the crude mixture was purified by semi prep RP-HPLC to give the **8** as a yellow solid (5 mg, 15%).

^{1}H NMR (500 MHz, ammonium acetate aqueous buffer pH 8.5/D_2O (9:1, *v*/*v*)) δ 12.19 (s, 1H), 11.53 (s, 1H), 11.30 (s, 1H), 10.49 (s, 1H), 9.23 (s, 1H), 9.16 (d, *J* = 8.5 Hz, 1H), 8.66 (s, 1H), 8.34 (d, *J* = 9.9 Hz, 1H), 8.29 – 8.02 (m, 6H), 7.97 (s, 1H), 7.87 – 7.61 (m, 6H), 7.54 (s, 1H), 7.34 – 7.22 (m, 4H), 7.07 (s, 1H), 7.03 – 6.87 (m, 3H), 6.79 (d, *J* = 10.9 Hz, 1H), 6.17 (s, 1H), 4.14 (s, 1H), 4.11 (s, 1H), 3.15 (s, 3H), 1.38 (s, 3H), 1.18 – 1.12 (m, 1H).

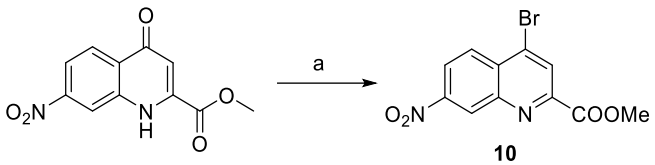
HRMS: calcd. for $C_{75}H_{51}N_{14}O_{24}S_4$ $[M-H]^-$ 1659.2078; found 1659.2261.



Oligomer 9: Oligomer **9** was synthesized starting from Cl-MPA-Protide resin (0.15 mmol g^{-1}, 18 µmol scale) according to the manual synthesis method. Loading of the first monomer: 0.11 mmol g^{-1} (75%). Final acetylation was carried out via the general acetylation method. After precipitation in cold Et_2O, the crude mixture was purified by semi prep RP-HPLC to give the **9** as a yellow solid (10 mg, 33%).

^{1}H NMR (500 MHz, ammonium acetate aqueous buffer pH 8.5/D_2O (9:1, *v*/*v*)) δ 11.53 (s, 1H), 10.72 (s, 1H), 10.55 (s, 1H), 10.28 (s, 1H), 9.40 (s, 1H), 9.27 (s, 1H), 9.25 (s, 1H), 8.94 (s, 2H), 8.80 (t, *J* = 12.3 Hz, 2H), 8.71 (d, *J* = 10.0 Hz, 1H), 8.60 – 8.51 (m, 2H), 8.37 (d, *J* = 10.8 Hz, 1H), 8.06 (s, 1H), 7.70 (s, 1H), 7.61 (t, *J* = 9.0 Hz, 2H), 7.54 (s, 1H), 7.39 (s, 1H), 7.31 – 7.17 (m, 3H), 6.90 (d, *J* = 8.3 Hz, 1H), 6.87 – 6.78 (m, 3H), 6.49 (s, 1H), 6.18 (d, *J* = 10.5 Hz, 1H), 6.13 (s, 1H), 5.90 (d, *J* = 7.3 Hz, 2H), 5.68 (s, 1H), 4.07 (s, 3H), 3.61 (s, 3H), 3.45 (d, *J* = 1.9 Hz, 3H), 3.12 (s, 3H), 2.32 (d, *J* = 2.0 Hz, 3H).
HRMS: calcd. for $C_{82}H_{62}N_{15}O_{23}S_2$ $[M-H]^-$ 1688.3579; found 1688.3416.

2.3 Monomer synthesis procedures



10

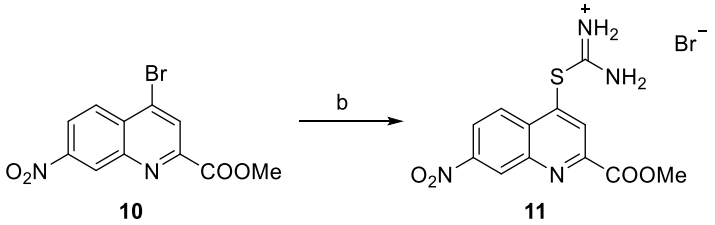
Compound 10. Methyl 7-nitro-4-oxo-1,4-dihydroquinoline-2-carboxylate (4.20 g, 0.017 mol) was dissolved in dry DMF (42 mL) under N_2. Then $POBr_3$ (6.45 g, 0.022 mol) was added. The reaction mixture was stirred for 5 h at 80 °C. After TLC confirmed the reaction completion, the mixture was poured into ice and neutralised with saturated $NaHCO_3$ solution. Then the mixture was extracted with DCM. The combined organic layers were washed with water and brine,

dried (Na_2SO_4) and then concentrated under reduced pressure to give 4.8 g (92%) of **10** as a solid.

^{1}H NMR (500 MHz, DMSO-d_6) δ 8.98 (d, J = 2.2 Hz, 1H), 8.59 (s, 1H), 8.58 (dd, J = 9.1, 2.2 Hz, 1H), 8.47 (d, J = 9.1 Hz, 1H), 4.00 (s, 3H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 163.6, 149.8, 149.1, 146.3, 134.7, 131.1, 129.0, 127.3, 126.1, 123.5, 53.2.

HRMS: calcd. for $C_{11}H_7N_2O_4Br$ $[M]^-$ 309.9584; found 309.9594.

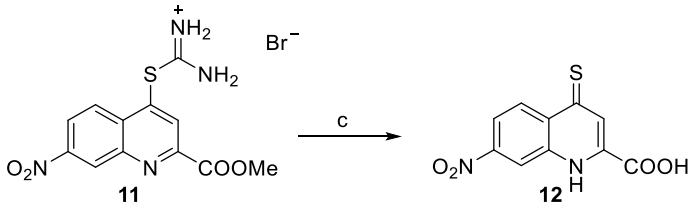


Compound 11. Compound **10** (4.76 g, 0.015 mol) and thiourea (2.33 g, 0.031 mol) were dissolved in acetone (75 mL). The reaction mixture was refluxed under N_2 for 3 h. After TLC confirmed the reaction completion, the reaction mixture was left to cool down to r.t.. The precipitate was filtered and washed with acetone and then dried under reduced pressure to give 5.47 g (93%) of **11** as a white powder.

^{1}H NMR (500 MHz, DMSO-d_6) δ 9.26 (s, 2H), 9.14 (s, 2H), 9.08 (d, J = 2.3 Hz, 1H), 8.67 (s, 1H), 8.63 (dd, J = 9.2, 2.3 Hz, 1H), 8.51 (d, J = 9.2 Hz, 1H), 4.03 (s, 3H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 167.0, 164.0, 149.9, 149.0, 146.8, 134.9, 132.6, 132.3, 127.1, 126.3, 123.7, 53.3.

HRMS: calcd. for $C_{12}H_{11}N_4O_4S$ $[M+H]^+$ 307.0496; found 307.0496.

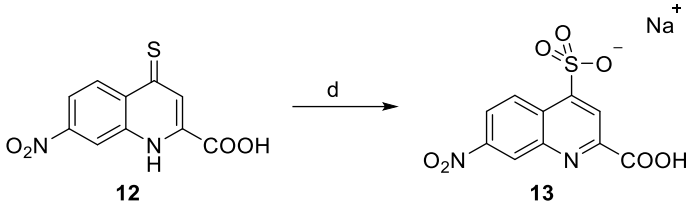


Compound 12. Compound **11** (5.47 g, 0.014 mol) was dissolved in MeOH (55 mL), then NaOH (2.83 g, 0.071 mol) in H_2O (17.6 mL) was added slowly. The mixture was stirred for 3 h at r.t.. After NMR confirmed the reaction completion, 1 M HCl was then added to acidify the mixture to approximately pH 4 to allow for full precipitation. The precipitate was filtered and washed with a small amount of water and then dried under reduced pressure to give 3.5 g (99%) of **12** as a dark red solid.

^{1}H NMR (500 MHz, DMSO-d_6) δ 9.00 (d, J = 2.3 Hz, 1H), 8.75 (d, J = 9.2 Hz, 1H), 8.19 (dd, J = 9.2, 2.4 Hz, 1H), 7.81 (s, 1H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 195.2, 163.1, 149.7, 136.0, 135.5, 133.3, 130.3, 126.2, 119.3, 117.4.

HRMS: calcd. for $C_{10}H_5N_2O_4S$ $[M-H]^-$ 248.9965; found 248.9978.

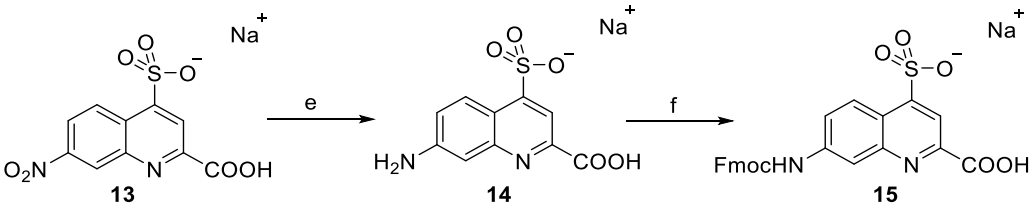


Compound 13. A solution of 30% H_2O_2 (4 mL, 0.04 mol) in formic acid (29 mL) was stirred at room temperature under N_2 for 1 h. Then a suspension of compound **12** (2.00 g, 0.008 mol) in formic acid (18 mL) was added carefully. The reaction mixture was stirred for 5 h at 48 °C under N_2. After HPLC confirmed the reaction completion, the reaction mixture was cooled to 0 °C by ice-bath and then NaCl solid was added slowly until a white precipitate formed. The precipitate was filtered while it was still cold. The solid was washed with a small amount of cold water and then dried under reduced pressure to give 1.00 g (40%) of **13** as a white powder.

^{1}H NMR (500 MHz, DMSO-d_6) δ 9.06 (d, J = 9.3 Hz, 1H), 8.67 (s, 1H), 8.92 (d, J = 2.4 Hz, 1H), 8.49 (dd, J = 9.3, 2.4 Hz, 1H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 165.5, 153.5, 151.0, 148.0, 146.7, 129.7, 128.0, 125.4, 121.4, 120.1.

HRMS: calcd. for $C_{10}H_5N_2O_7S$ $[M-H]^-$ 296.9812; found 296.9824.



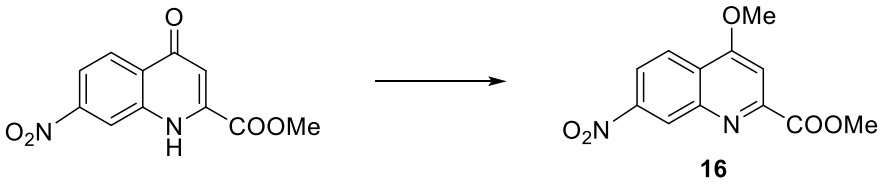
Compound 15. Compound **13** (0.96 g, 0.003 mol) was dissolved in MeOH (221 mL) under N_2. Then 10 wt.% Pd/C (0.34 g, 35% w/w) was added and N_2 was replaced by H_2. The mixture was stirred for 4 h at r.t. and then another 10 wt.% Pd/C (0.34 g, 35% w/w) was added and the mixture was stirred for another 4 h. The resulting mixture was filtered through celite and solvents were evaporated under reduced pressure to give **14** as a bright yellow powder which was used directly in the next step without further purification. **14** was dissolved in 10% (w/w) $NaHCO_3$ solution (58 mL) and the mixture was cooled down to 0 °C. Fmoc-Cl (1086 mg, 0.004

mol) in dioxane (34 mL) was added dropwise at 0 °C to the resulting slurry over 1 h. The mixture was stirred at 0 °C for another hour, and allowed to warm to room temperature overnight. After TLC confirmed the reaction completion, the solution was acidified to pH = 4 using 1 M HCl solution. The aqueous mixture was washed several times with Et_2O until precipitation formed. The solid was filtered and washed with Et_2O and a small amount of H_2O. The resulting solid was dried under reduced pressure to give 860 mg (56%) of **15** as a yellow powder. Note: Additional experiments indicated the product contains some sodium salt; nevertheless, the salt doesn't affect the use of SPS.

^{1}H NMR (500 MHz, DMSO-d_6) δ 10.21 (s, 1H), 8.73 (d, J = 9.3 Hz, 1H), 8.34 (s, 1H), 8.23 (s, 1H), 7.93 (dt, J = 7.6, 1.0 Hz, 2H), 7.81 (dd, J = 7.5, 1.1 Hz, 2H), 7.77 (d, J = 9.4 Hz, 1H), 7.45 (td, J = 7.7, 1.1 Hz, 2H), 7.38 (td, J = 7.5, 1.2 Hz, 2H), 4.58 (d, J = 6.6 Hz, 2H), 4.37 (t, J = 6.6 Hz, 1H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 166.5, 153.9, 153.5, 148.9, 148.9, 144.2, 141.3, 141.0, 128.3, 128.2, 127.6, 125.7, 121.9, 121.1, 120.7, 116.7, 115.2, 66.4, 47.1.

HRMS: calcd. for $C_{25}H_{17}N_2O_7S$ $[M-H]^-$ 489.0751; found 489.0763.



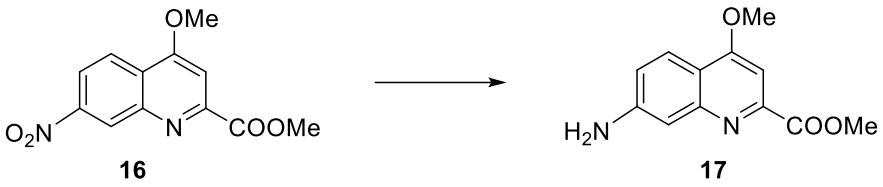
16

Compound 16. Methyl 7-nitro-4-oxo-1,4-dihydroquinoline-2-carboxylate (1 g, 0.040 mmol), PPh_3 (1.59 g, 0.061 mmol) and MeOH (245 μL, 0.060 mmol) were dissolved in THF (30 mL) under N_2. Then diisopropyl azodicarboxylate (1.2 mL, 0.060 mmol) was added dropwise. The mixture was stirred for 14 h at r.t.. After TLC confirmed the reaction completion, the mixture was filtered and concentrated. The residue was precipitated in MeOH to give 600 mg (57%) of **16** as a white solid.

^{1}H NMR (500 MHz, DMSO-d_6) δ 8.86 (d, J = 2.1 Hz, 1H), 8.43 (d, J = 9.2 Hz, 1H), 8.39 (dd, J = 9.2, 2.2 Hz, 1H), 7.72 (s, 1H), 4.20 (s, 3H), 3.99 (s, 3H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 164.9, 162.9, 151.4, 148.7, 146.8, 125.1, 124.7, 124.1, 120.9, 102.9, 57.0, 53.0.

HRMS: calcd. for $C_{12}H_{11}N_2O_5$ $[M+H]^+$ 263.0662; found 263.0769.

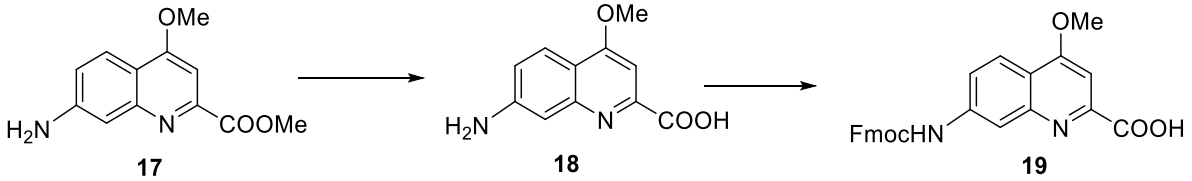


Compound 17. Compound **16** (3.97 g, 0.015 mmol) was dissolved in THF (266 mL) and MeOH (400 mL) under N_2. Then 10 wt.% Pd/C (1 g, 35% *w*/*w*) was added and N_2 was replaced by H_2. The mixture was stirred for 3 h at room temperature. After TLC confirmed the reaction completion, the mixture was filtered and concentrated to give 3.49 g (99%) of **17** as a yellow powder.

^{1}H NMR (500 MHz, DMSO-d_6) δ 7.85 (d, *J* = 8.9 Hz, 1H), 7.20 (s, 1H), 7.03 (dd, *J* = 8.9, 2.2 Hz, 1H), 6.95 (d, *J* = 2.2 Hz, 1H), 5.89 (s, 2H), 4.02 (s, 3H), 3.90 (s, 3H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 166.0, 162.4, 151.1, 150.4, 148.7, 122.1, 119.6, 112.9, 106.8, 96.8, 55.9, 52.4.

HRMS: calcd. for $C_{12}H_{12}N_2O_3Na$ $[M+Na]^+$ 255.0740; found 255.0739.



Compound 19. Compound **17** (3.81 g, 0.016 mol) was dissolved in dioxane (324 mL) then $LiOH{\cdot}H_2O$ (1.03 g, 0.024 mol) in water (102 mL) was added slowly at 0 °C. The mixture was stirred for 1 h at 0 °C, then neutralized by 0.1 mol·L^{-1} HCl. Then water (120 mL) and $NaHCO_3$ (4.13 g, 0.049 mol) were added and cooled down to 0 °C by ice-bath. Fmoc-Cl (5.52 g, 0.02 mol) is added slowly as a solution in dioxane. The resulting mixture is stirred at 0 °C for 1 h and allowed to warm to room temperature overnight. After TLC confirmed the reaction completion, the solution was acidified to pH = 4 using a saturated citric acid solution. Then water was added and the aqueous layers were extracted with DCM. The combined organic layers were washed with water and brine, dried (Na_2SO_4) and then concentrated under reduced pressure. The residue was triturated and sonicated in DCM, acetonitrile and water. The final slurry was lyophilized to give 5.06 g (70%) of **19** as a yellow powder.

^{1}H NMR (500 MHz, DMSO-d_6) δ 10.23 (s, 1H), 8.30 (s, 1H), 8.09 (d, *J* = 9.0 Hz, 1H), 7.92 (d, *J* = 7.5 Hz, 2H), 7.78 (d, *J* = 7.4 Hz, 2H), 7.73 (d, *J* = 8.2 Hz, 1H), 7.45 (s, 1H), 7.43 (t, *J* = 7.5 Hz, 2H), 7.36 (td, *J* = 7.4, 1.2 Hz, 2H), 4.59 (d, *J* = 6.5 Hz, 2H), 4.36 (t, *J* = 6.5 Hz, 1H), 4.10 (s, 3H).

^{13}C NMR (126 MHz, DMSO-d_6) δ 165.9, 163.1, 153.4, 150.9, 147.8, 143.7, 141.4, 140.8, 127.7, 127.2, 125.1, 122.3, 120.4, 120.2, 116.8, 114.2, 99.1, 65.8, 56.4, 46.6.

HRMS: calcd. for $C_{26}H_{21}N_2O_5$ $[M+H]^+$ 441.1445; found 441.1447.

3 X-ray crystallographic analysis of the structure of $(5)_2$

Aqueous solution of $(\mathbf{5})_2$ was prepared by dissolving the lyophilized powder in water to a final concentration of 5mM. Crystallization trials were carried with commercial sparse matrix screens using standard sitting drop vapor diffusion method at 293 K. X-ray quality crystals were obtained after three weeks by the addition of 0.8 µl of $(\mathbf{1})_2$ and 0.8 µl of 15% w/v polyethylene glycol 8000, 100 mM MES buffer at pH 6.5 and 200 mM calcium acetate in the reservoir. For low temperature diffraction measurement, a crystal was fished using a microloop, cryo-protected using 30% v/v polyethylene glycol 400 and plunged into liquid nitrogen.

X-ray diffraction data was collected at the ID30B[7] beamline in European Synchrotron Radiation Facility (ESRF), Grenoble. Diffraction data was measured at T = 100 K, λ = 0.9184 Å. The crystal was exposed for 0.02 s and 0.2° oscillation per frame and a rotation pass of 360° was measured using a EIGER2 X 9M 450µm Si sensor (Dectris) detector. Diffraction data was processed using the program *XDS*[8]. The crystal belonged to the space group *P*-1 with unit cell parameters: a = 25.288 (3) Å, b = 27.896 (4) Å, c = 41.288 (2) Å, α = 90.710 (4)°, β = 99.819 (1)°, γ = 116.180 (1)°; V = 25629 (6) Å^3 and two helices per asymmetric unit (Z = 4, Z' = 2). The structure was solved with *SHELXT*[9] structure solution program using Intrinsic Phasing and refined by full-matrix least-squares method on F^2 with *SHELXL*-2014[10] within *Olex2*[11]. After each refinement step, visual inspection of the model and the electron-density maps were carried out using *Olex2*[11] and *Coot*[12] using $2F_o - F_c$ and difference Fourier ($F_o - F_c$) maps. The initial structure revealed all main-chain atoms of a double-helical capsule. The CH_3 group of diethylene glycol tail was severely disordered in both helices and omitted. AFIX, SADI, DFIX and FLAT instructions were used to improve the geometry of molecules. Restraints on anisotropic displacement parameters were implemented with DELU, ISOR and EADP instructions. All non-H atoms were refined with anisotropic displacement parameters. After several attempts to model the disordered side chains and diethylene glycol tail, the SQUEEZE[13] procedure was used to flatten the electron density map. Very disordered solvent molecules were removed. Calculated total potential solvent accessible void volume and electron count per cell are 7403.1 Å^3 and 2220 respectively. Hydrogen atoms for $(\mathbf{5})_2$ were placed at idealized positions.

Statistics of data collection and refinement of $(\mathbf{5})_2$ are described in Table S1. The final cif file of $(\mathbf{5})_2$ was examined in IUCr's *checkCIF* algorithm. Due to the large volume fractions of disordered solvent molecules, weak diffraction intensity and poor resolution, a number of A- and B- level alerts remain in the *checkCIF* file. These alerts are inherent to the data set and

refinement procedures. They are listed below and were divided into two groups. The first group demonstrates weak quality of the data and refinement statistics when compared to those expected for small molecule structures from highly diffracting crystals. The second group is concerned to decisions made during refinement and explained below. Atomic coordinates and structure factors of (**5**)$_2$ was deposited in the Cambridge Crystallographic Data Centre (CCDC) with accession code 2216788. The data is available free of charge upon request (www.ccdc.cam.ac.uk/).

CheckCIF validation of (**1**)$_2$:

Group 1 alerts (these illustrate weak quality of data and refinement statistics if compared to small molecule structures from highly diffracting crystals):

THETM01_ALERT_3_A The value of sine(theta_max)/wavelength is less than 0.550
Calculated sin(theta_max)/wavelength = 0.4762
PLAT029_ALERT_3_A _diffrn_measured_fraction_theta_full value Low . 0.917
PLAT084_ALERT_3_B High wR2 Value (i.e. > 0.25) ..... 0.44 Report
PLAT241_ALERT_2_B High 'MainMol' Ueq as Compared to Neighbors of Check
PLAT242_ALERT_2_B Low 'MainMol' Ueq as Compared to Neighbors of Check
PLAT260_ALERT_2_B Large Average Ueq of Residue Including O0F5 0.353 Check
PLAT260_ALERT_2_B Large Average Ueq of Residue Including O0F9 0.310 Check
PLAT340_ALERT_3_B Low Bond Precision on C-C Bonds 0.01725 Ang.
PLAT414_ALERT_2_B Short Intra XH3 .. XHn H30C ..H302 1.75 Ang.
PLAT430_ALERT_2_A Short Inter D...A Contact O016 ..O277 2.76 Ang.

Group 2 alert (is connected with decision made during refinement and explained below):

PLAT196_ALERT_1_B No TEMP record and _measurement_temperature .NE. 293 Degree
Measurement temperature was 100 K.
PLAT306_ALERT_2_B Isolated Oxygen Atom (H-atoms Missing ?) Check
Dummy O atom was introduced into refinement.
PLAT780_ALERT_1_B Coordinates do not Form a Properly Connected Set
The connected set is through individual helices in the double helical capsule.

Table S1. Crystallographic data and refinement details for $(\mathbf{5})_2$.

Identification code	$(\mathbf{1})_2$
Emperical formula	$C_{159}\,H_{108.5}\,Ca\,N_{29}\,O_{62.5}\,S_6$
Formula weight	3657.68
Temperature	100 K
Wavelength	0.9184 Å
Crystal system	Triclinic
Space group	*P*-1
Unit cell dimensions	a = 25.288 (3) Å b = 27.896 (4) Å c = 41.288 (2) Å α = 90.710 (4)° β = 99.819 (1)° γ = 116.180 (1)°
Volume	25629 (5) $Å^3$
Z	4
Density (calculated)	0.948 g/cm^3
Absorption coefficient	0.266 μ/mm^{-1}
Colour and shape	Pale yellow, prism
Crystal size	0.200 x 0.100 x 0.050 mm
Index ranges	$-23 \leq h \leq 23$ $-26 \leq k \leq 26$ $-39 \leq l \leq 39$
Reflections collected	137893
R_{int}	0.0663
Data/restraints/parameters	42538/223/4438
Goodness-of-fit on F^2	1.909
Final R indexes [$I > 2\sigma(I)$]	R_1 = 0.1361 wR_2 = 0.4112
Final R indexes [all data]	R_1 = 0.1527 wR_2 = 0.4398
Largest diff. peak and hole	1.16/-0.65 e $Å^{-3}$
CCDC #	2216788

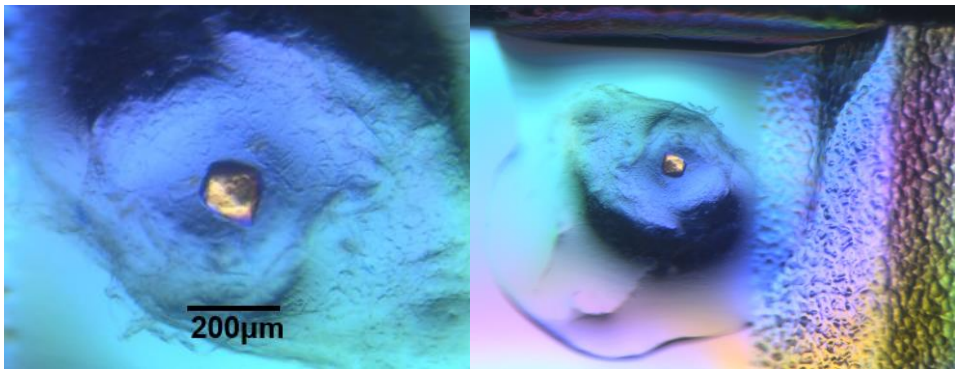


Figure S97.: Single crystal of $(\mathbf{5})_2$ observed under cross-polarizing microscope

4 NMR spectra and RP-HPLC chromatograms of new compounds

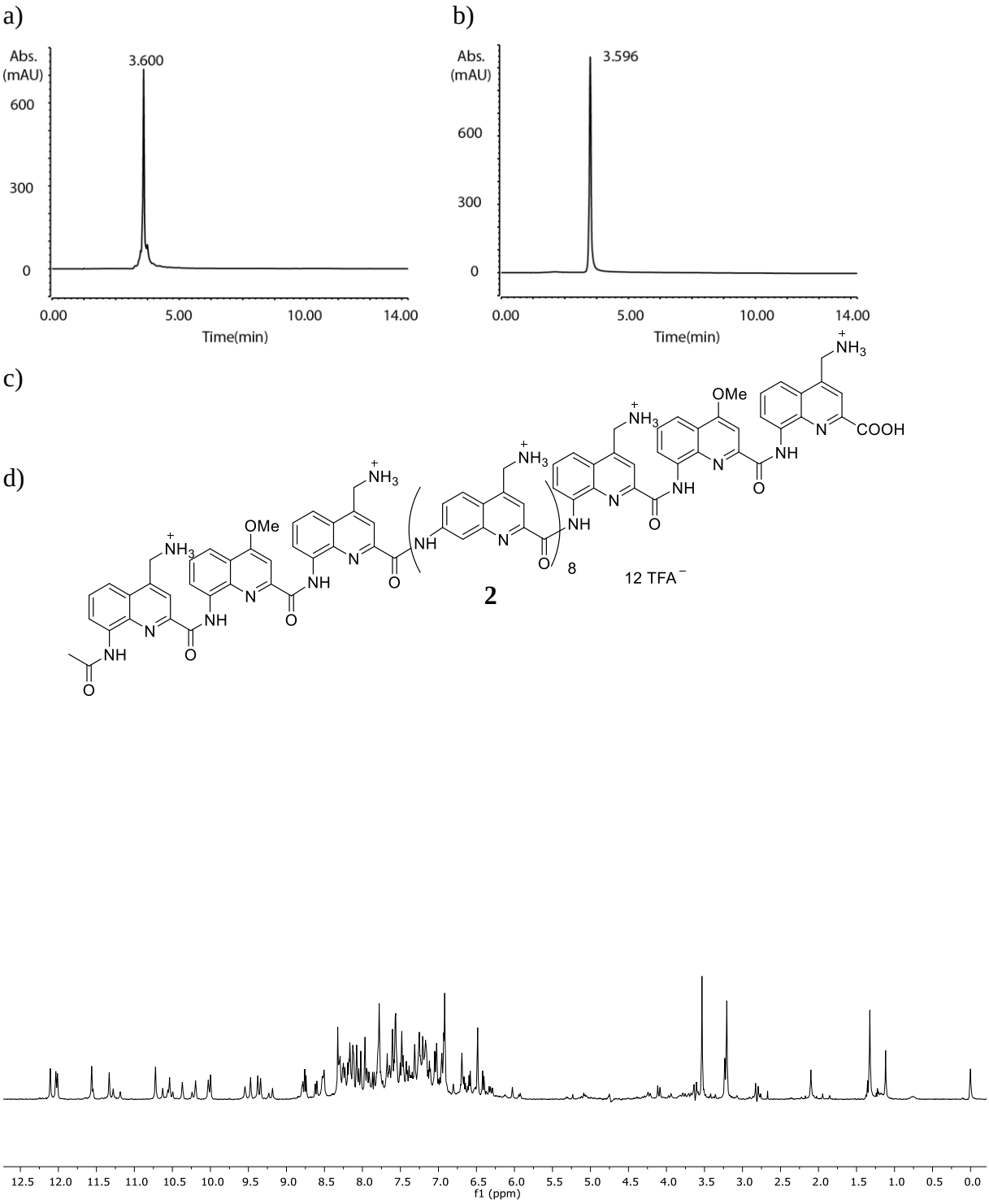


Figure S98. Analytical data of compound **2**. RP-HPLC chromatograms (a) recovered crude from TFA cleavage (C18, 5 to 100 B% over 10 min, λ = 300 nm) and (b) after purification (C18, 5 to 100 B% over 10 min, λ = 300 nm); A: 0.1% TFA water, B: 0.1% TFA acetonitrile. (c) Chemical structure of compound **2**. (d) ^{1}H NMR spectrum with water suppression (500 MHz, H_2O/D_2O = (9:1, *v*/*v*), 25 °C).

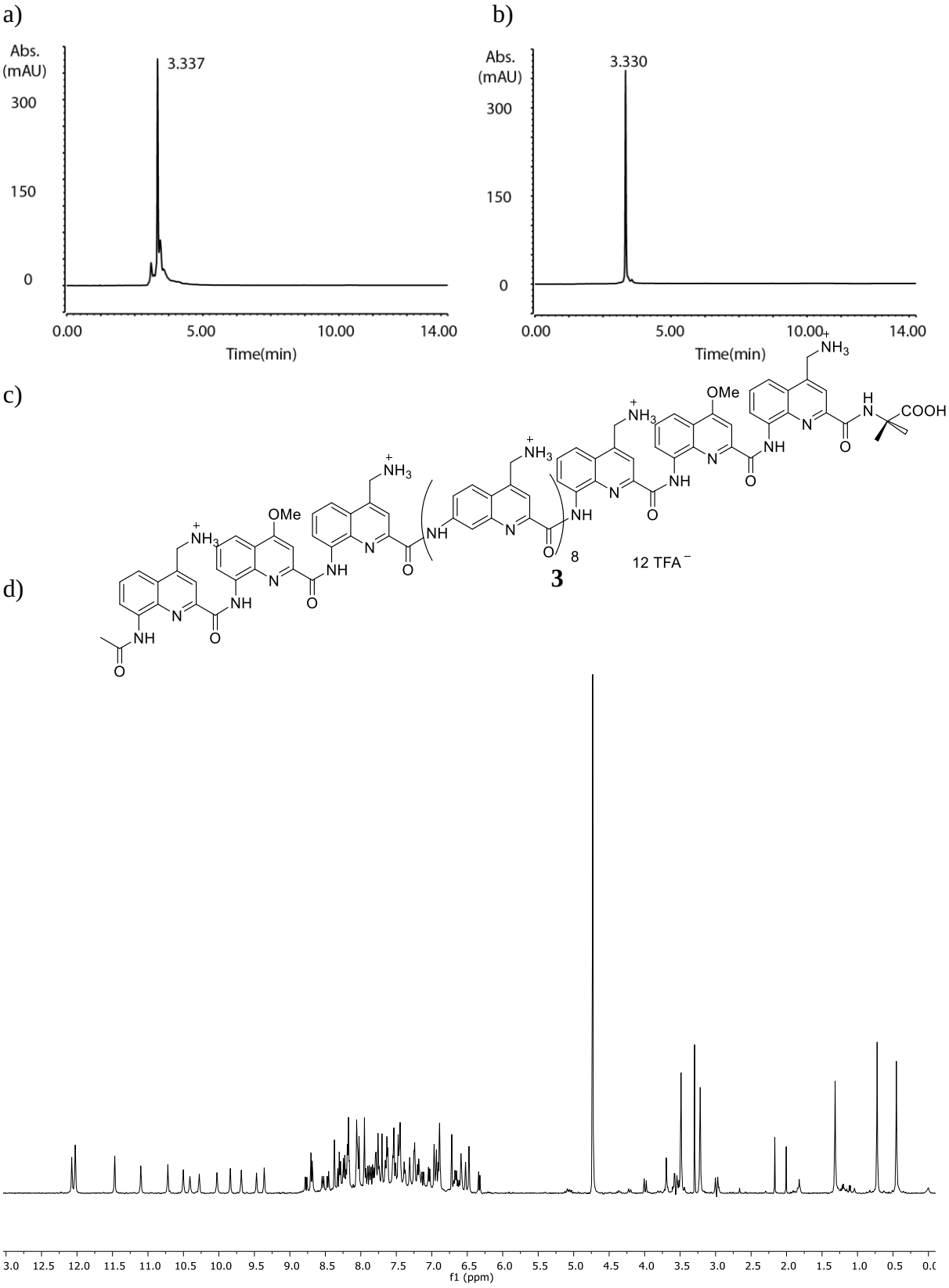


Figure S99. Analytical data of compound **3**. RP-HPLC chromatograms (a) recovered crude from TFA cleavage (C18, 5 to 100 B% over 10 min, λ = 300 nm) and (b) after purification (C18, 5 to 100 B% over 10 min, λ = 300 nm); A: 0.1% TFA water, B: 0.1% TFA acetonitrile. (c) Chemical structure of compound **3**. (d) ^{1}H NMR spectrum with water suppression (500 MHz, H_2O/D_2O (9:1, *v*/*v*), 25 °C).

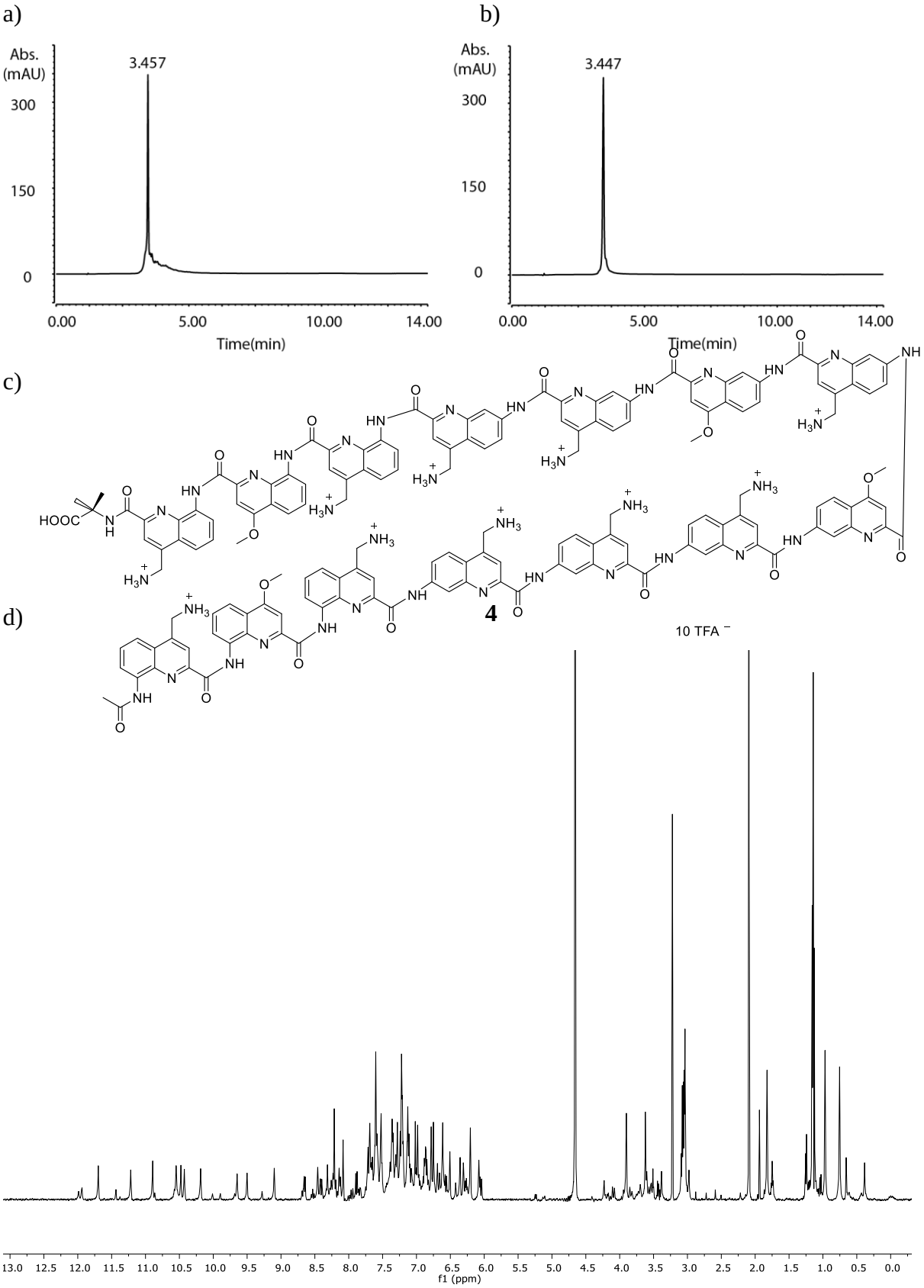


Figure S100. Analytical data of compound **4**. RP-HPLC chromatograms (a) recovered crude from TFA cleavage (C18, 5 to 100 B% over 10 min, λ = 300 nm) and (b) after purification (C18, 5 to 100 B% over 10 min, λ = 300 nm); A: 0.1% TFA water, B: 0.1% TFA acetonitrile. (c) Chemical structure of compound **4**. (d) ^{1}H NMR spectrum with water suppression (500 MHz, H_2O/D_2O (9:1, *v*/*v*), 25 °C).

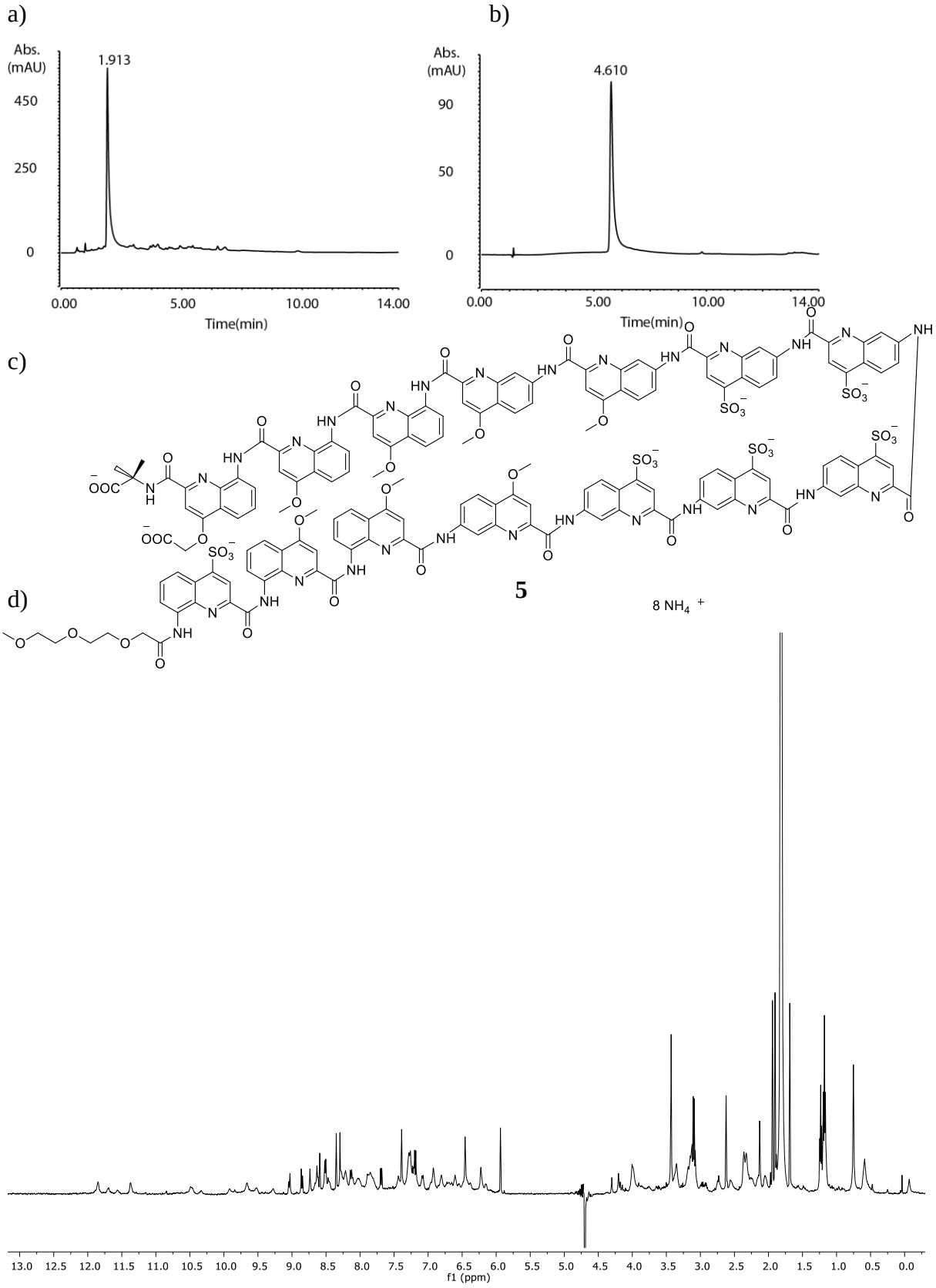
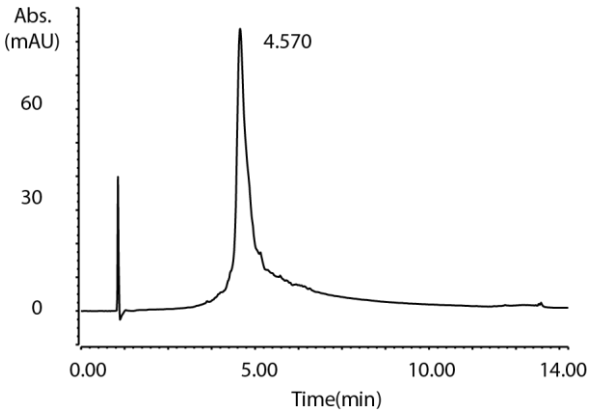
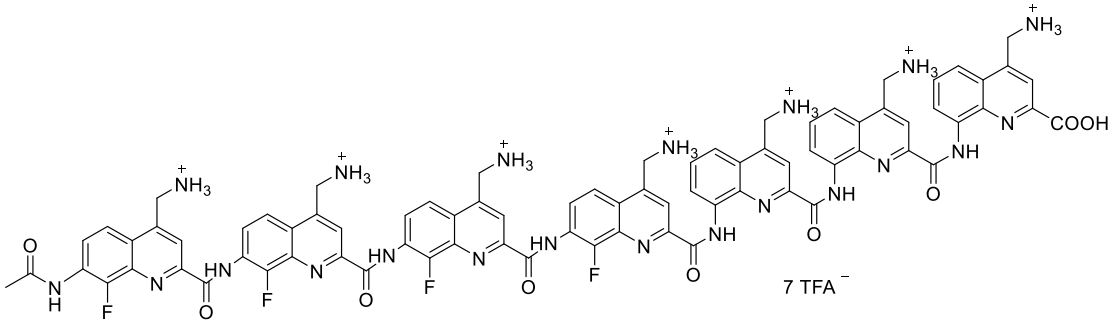


Figure S101. Analytical data of compound **5**. RP-HPLC chromatograms (a) recovered crude from TFA cleavage (C18, 30 to 80 B% over 10 min, λ = 300 nm) and (b) after purification (C18, 10 to 40 B% over 10 min, λ = 300 nm; A: 13 mM ammonium acetate aqueous buffer pH 8.5, B: acetonitrile). (c) Chemical structure of compound **5**. (d) ^{1}H NMR spectrum with water suppression (500 MHz, ammonium acetate aqueous buffer pH 8.5/D_2O (9:1, *v*/*v*), 25 °C).

a)



b)



c)

6

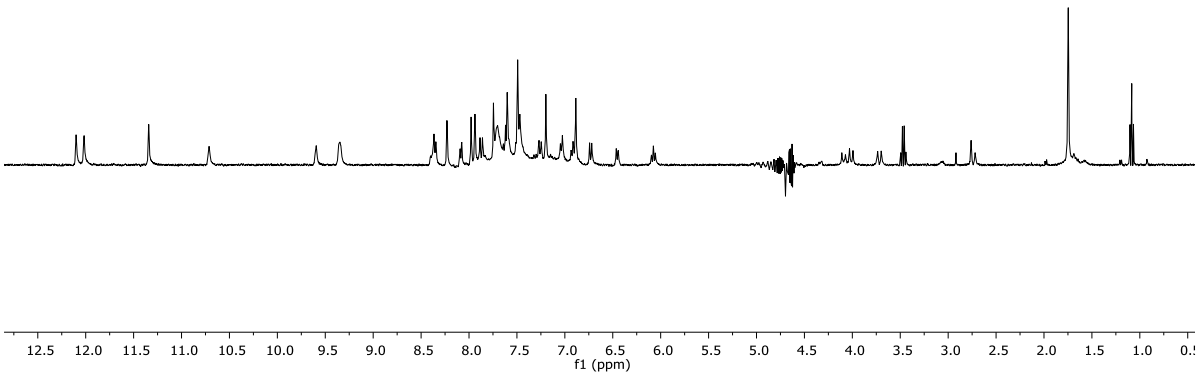
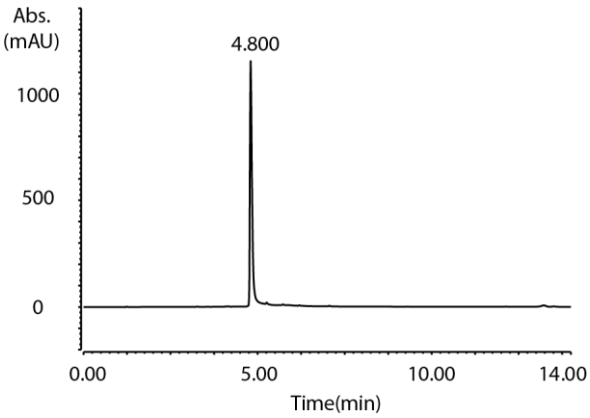
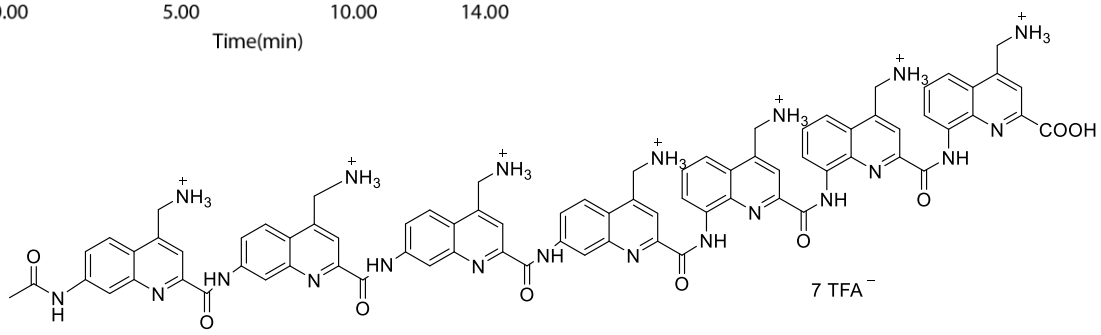


Figure S102. Analytical data of compound **6**. RP-HPLC chromatograms (a) recovered crude from TFA cleavage (C18, 5 to 60 B% over 10 min, λ = 300 nm); A: 0.1% TFA water, B: 0.1% TFA acetonitrile. (b) Chemical structure of compound **6**. (c) ^{1}H NMR spectrum with water suppression (400 MHz, H_2O/D_2O (9:1, *v*/*v*), 25 °C).

a)



b)



7

c)

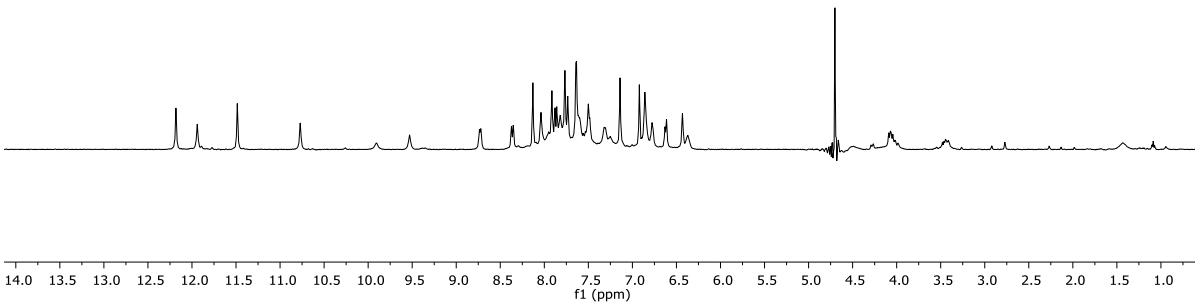
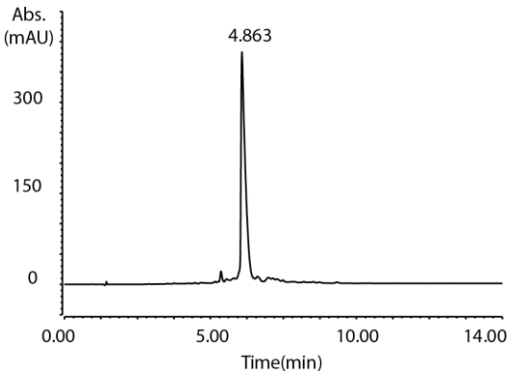
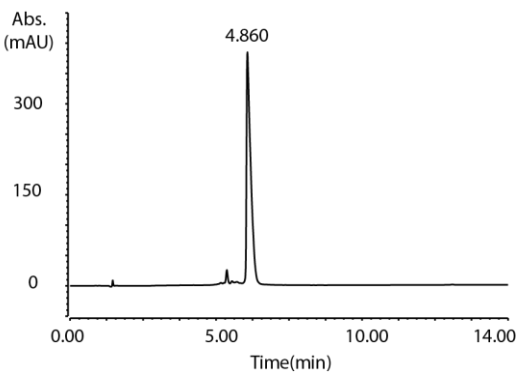


Figure S103. Analytical data of compound **7**. RP-HPLC chromatograms (a) recovered crude from TFA cleavage (C18, 5 to 60 B% over 10 min, λ = 300 nm); A: 0.1% TFA water, B: 0.1% TFA acetonitrile. (b) Chemical structure of compound **7**. (c) ^{1}H NMR spectrum with water suppression (500 MHz, H_2O/D_2O (9:1, *v*/*v*), 25 °C).

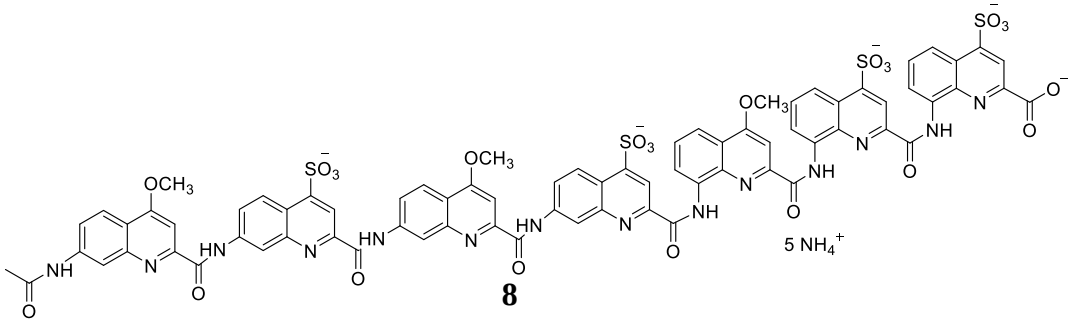
a)



b)



c)



8

d)

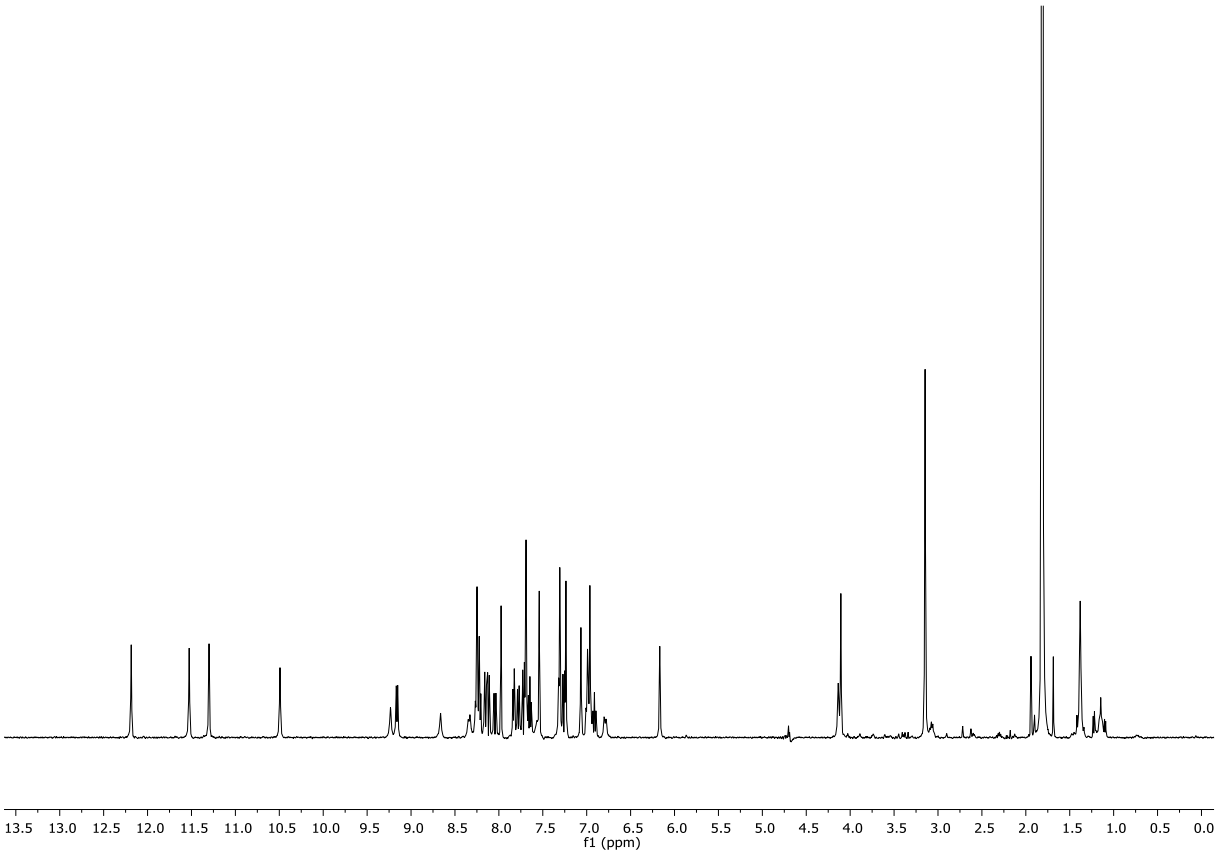


Figure S104. Analytical data of compound **8**. RP-HPLC chromatograms (a) recovered crude from TFA cleavage (C18, 10 to 80 B% over 10 min, λ = 300 nm) and (b) after purification (C18, 10 to 80B% over 10 min, λ = 300 nm; A: 13 mM ammonium acetate aqueous buffer pH 8.5, B: acetonitrile). (c) Chemical structure of compound **8**. (d) ^{1}H NMR spectrum with water suppression (500 MHz, ammonium acetate aqueous buffer pH 8.5/D_2O (9:1, *v*/*v*), 25 °C).

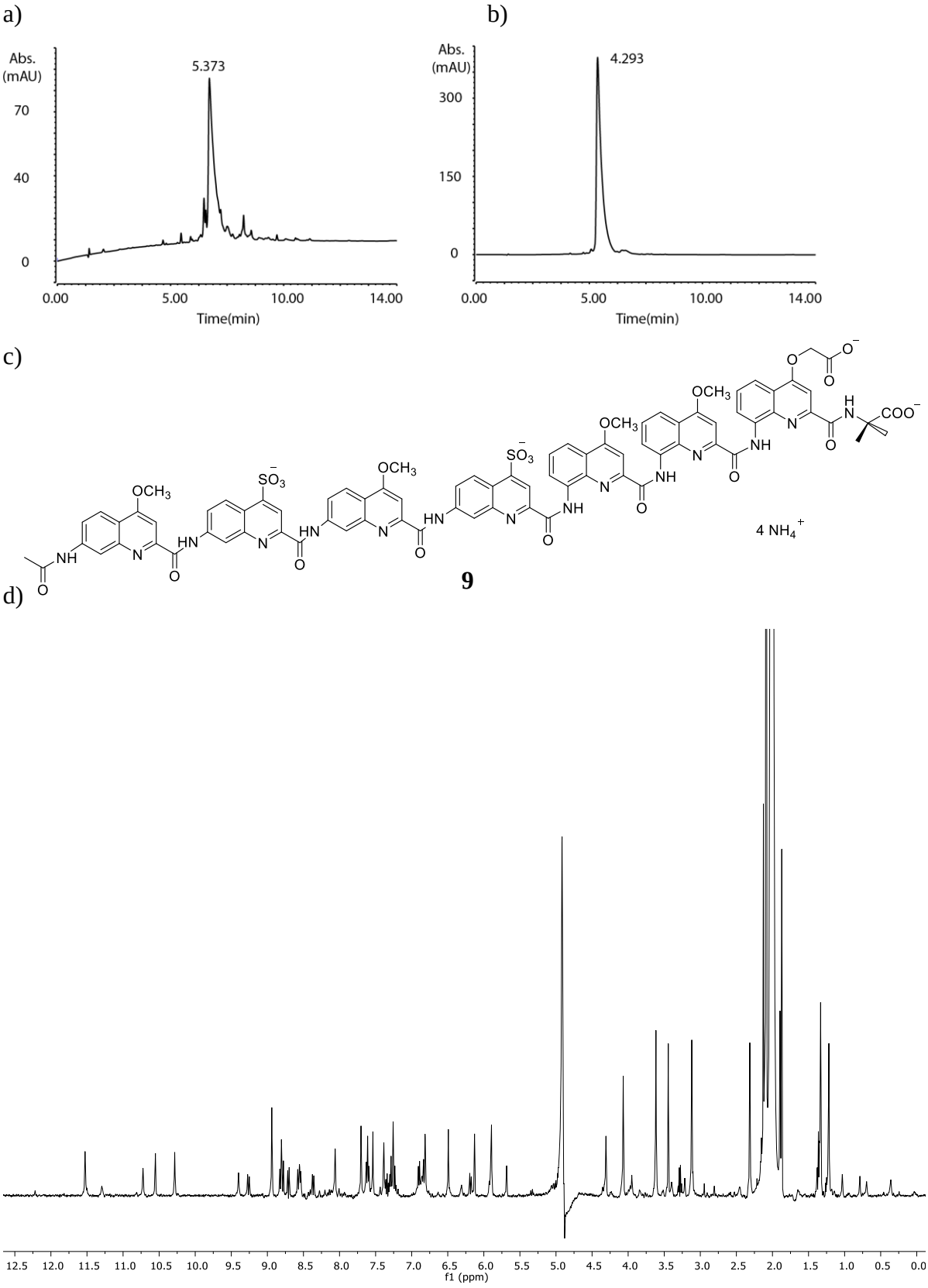


Figure S105. Analytical data of compound **9**. RP-HPLC chromatograms (a) recovered crude from TFA cleavage (C18, 10 to 60 B% over 10 min, λ = 300 nm) and (b) after purification (C18, 10 to 80B% over 10 min, λ = 300 nm; A: 13 mM ammonium acetate aqueous buffer pH 8.5, B: acetonitrile). (c) Chemical structure of compound **9**. (d) ^{1}H NMR spectrum with water suppression (500 MHz, ammonium acetate aqueous buffer pH 8.5/D_2O (9:1, *v*/*v*), 25 °C).

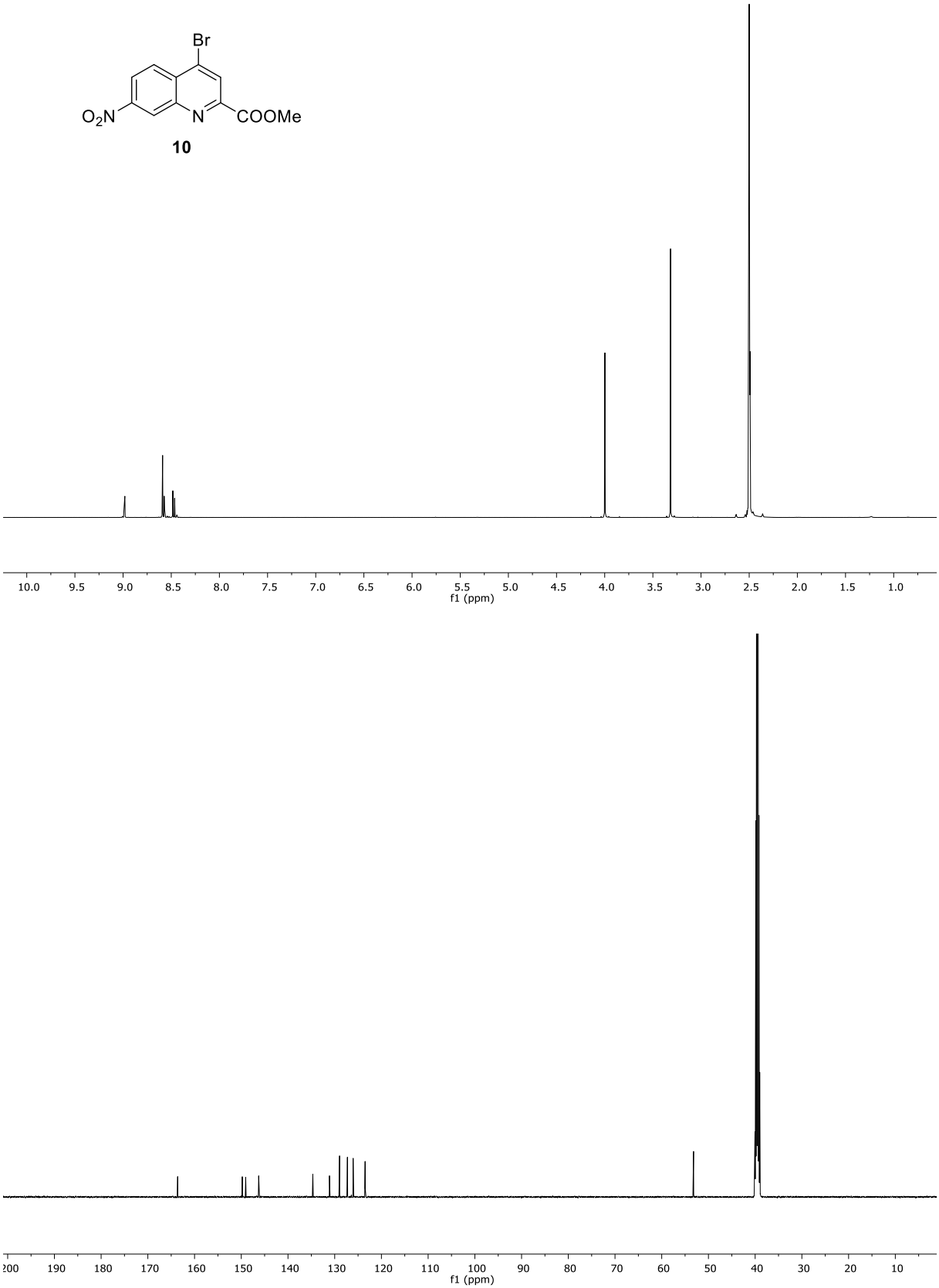


Figure S106. NMR spectra of compound **10** : ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

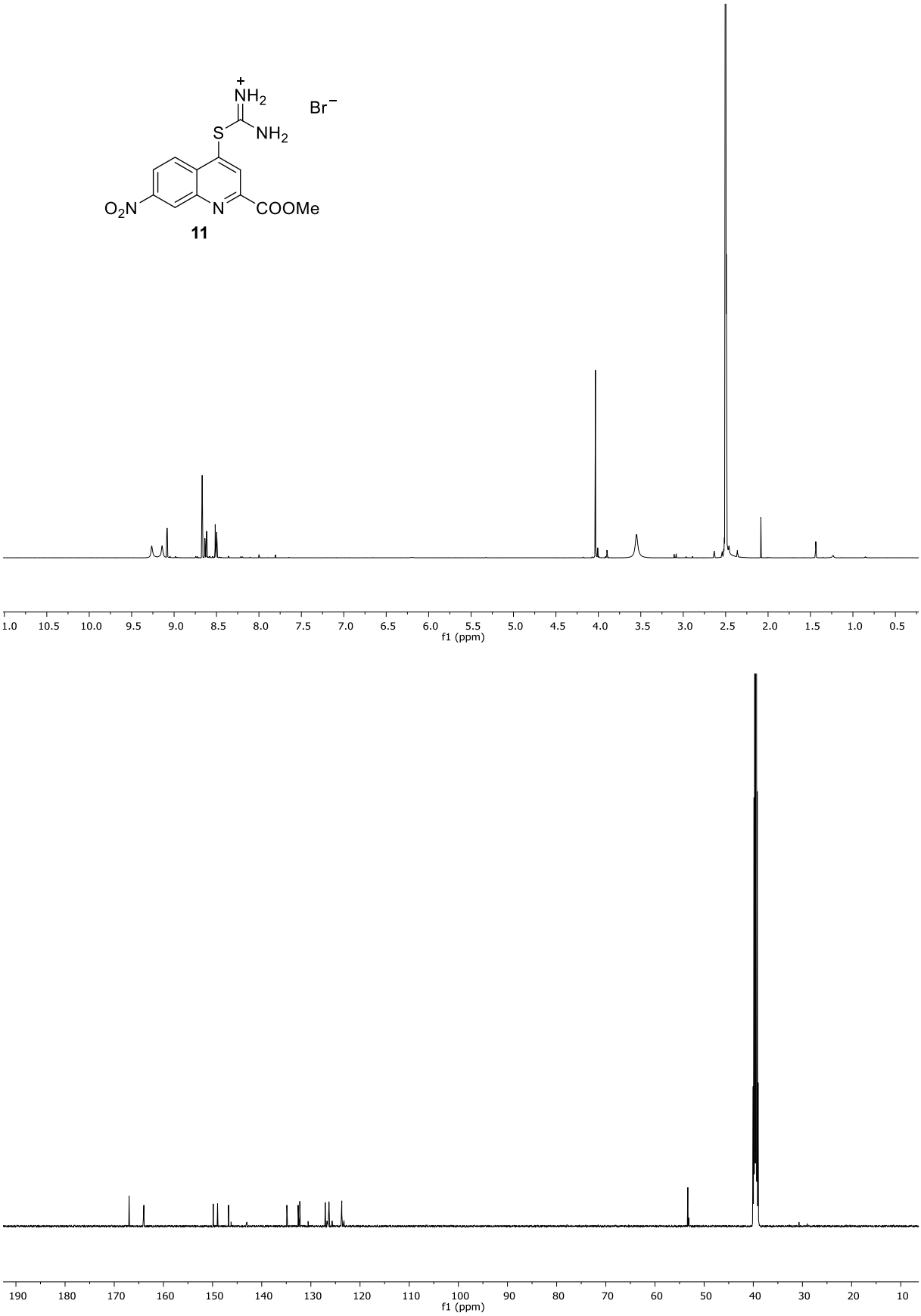


Figure S107. NMR spectra of compound **11** : ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

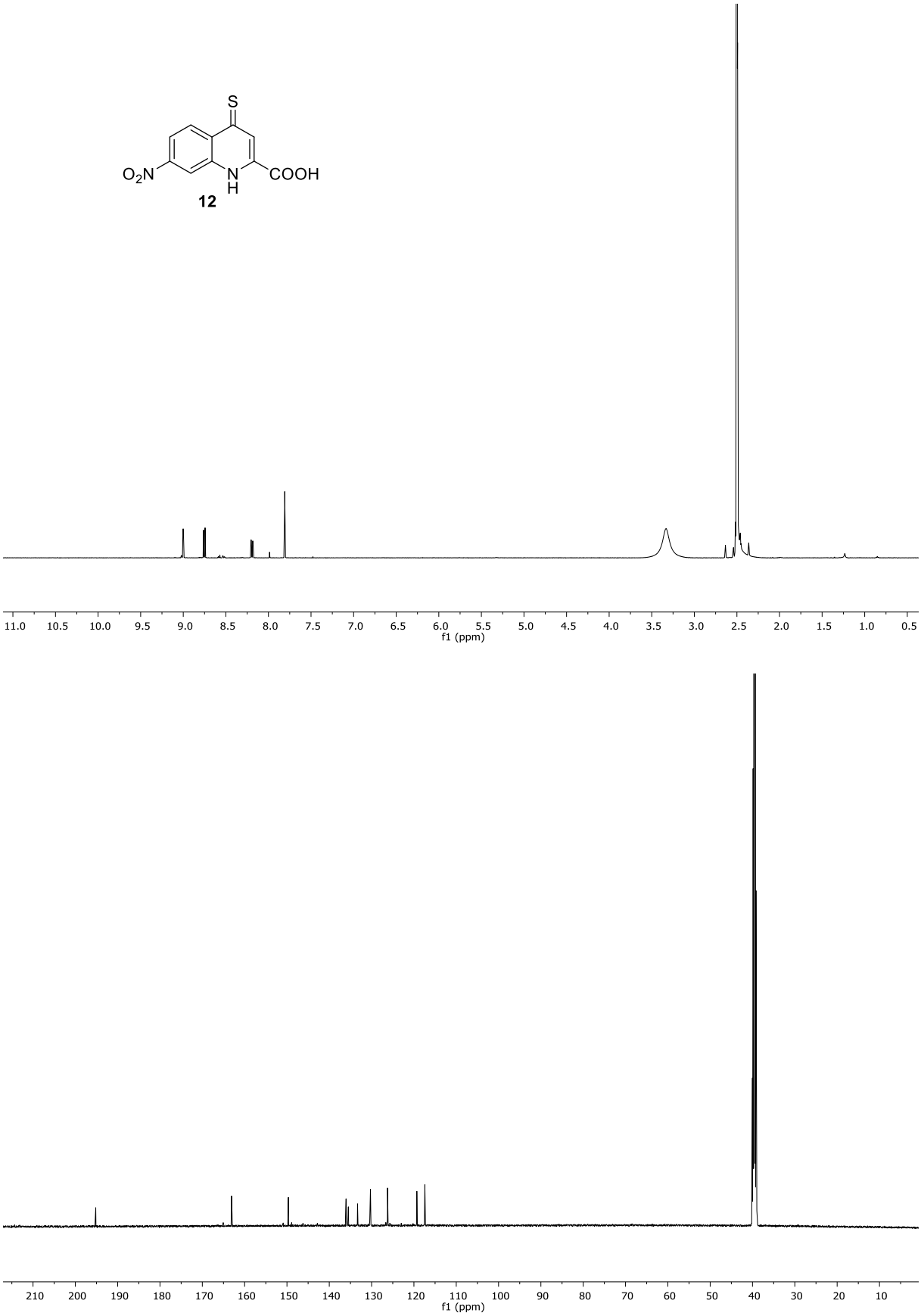


Figure S108. NMR spectra of compound **12** : ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

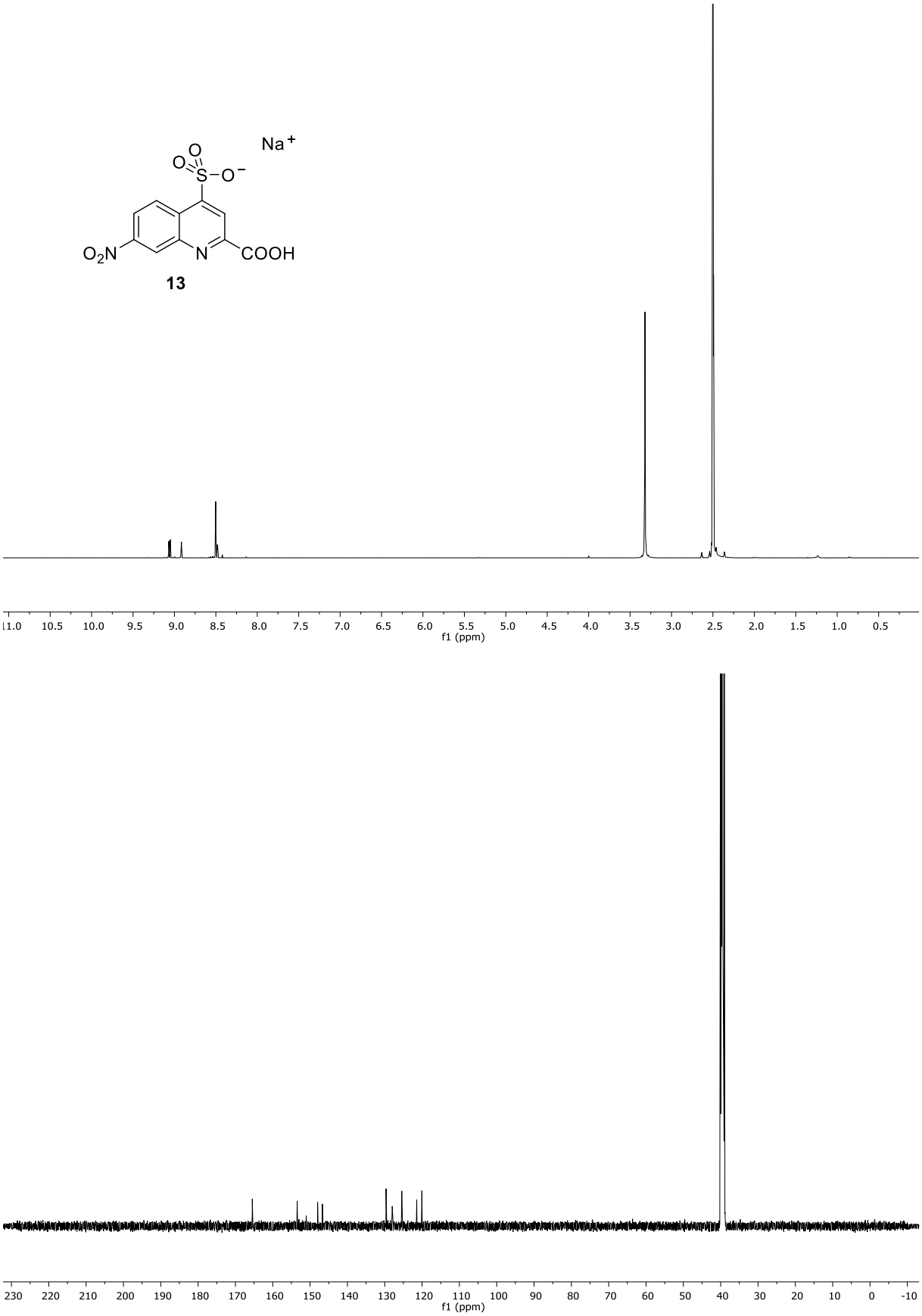


Figure S109. NMR spectra of compound **13** : ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

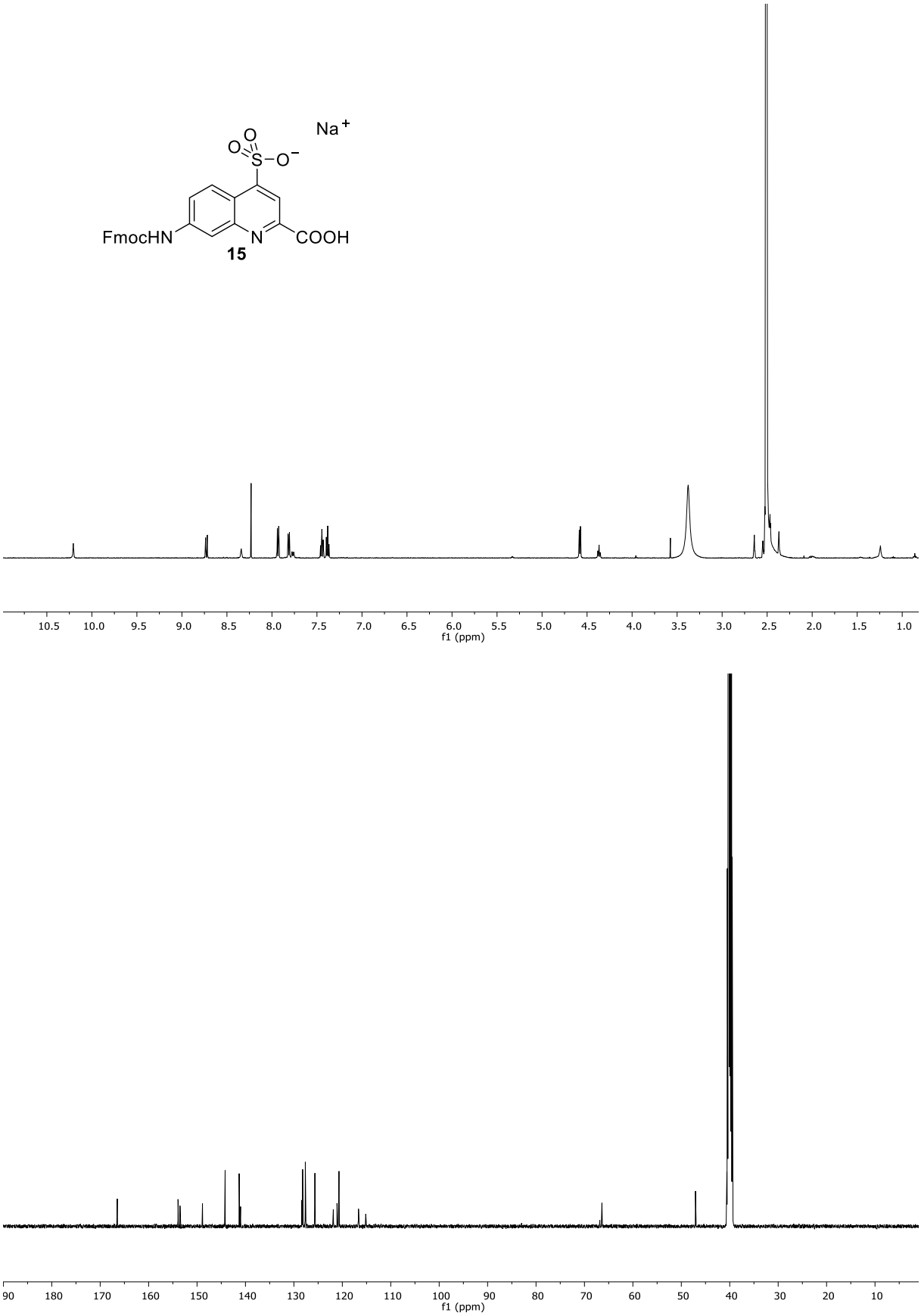


Figure S110. NMR spectra of compound **15** : ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

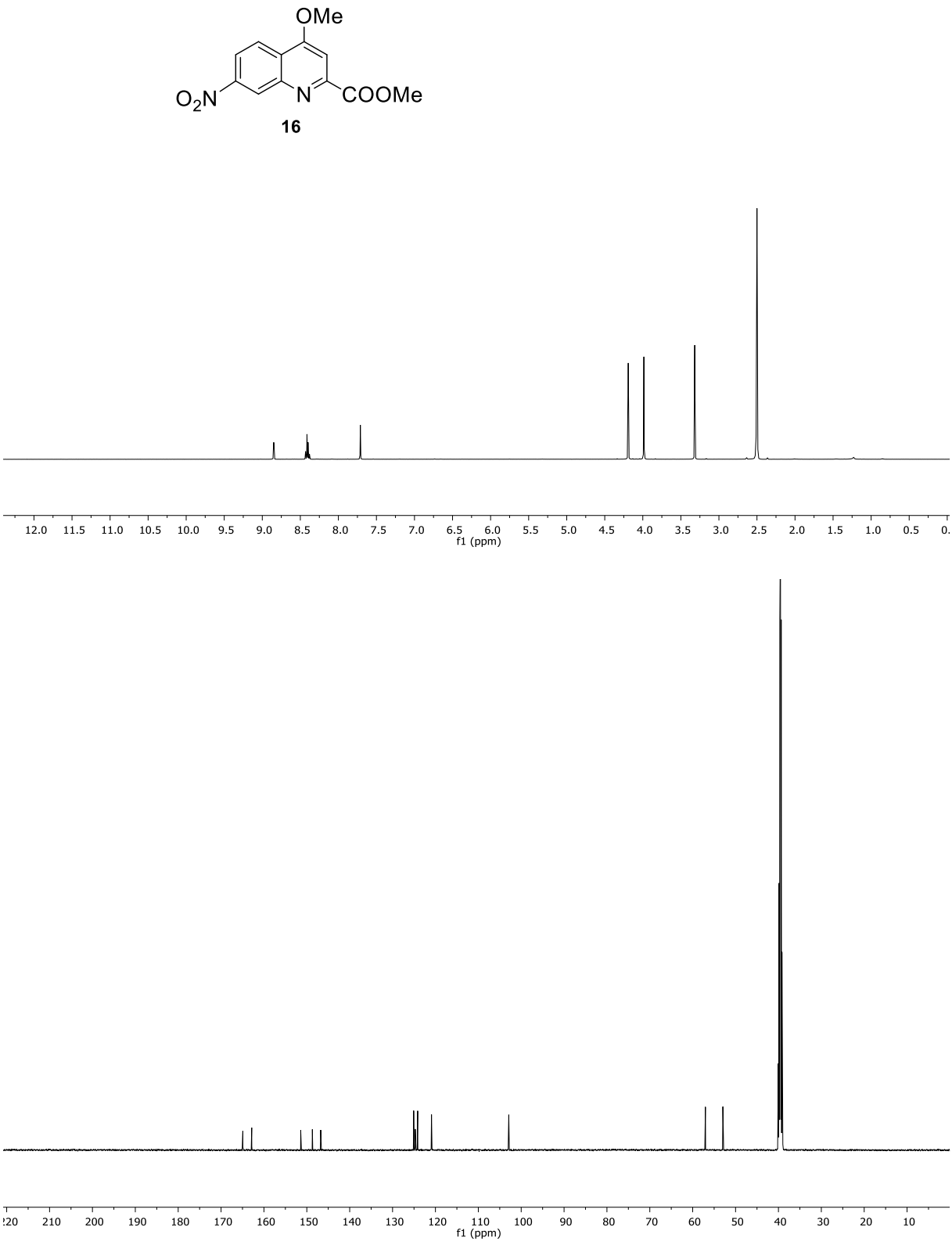


Figure S111. NMR spectra of compound **16** : ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

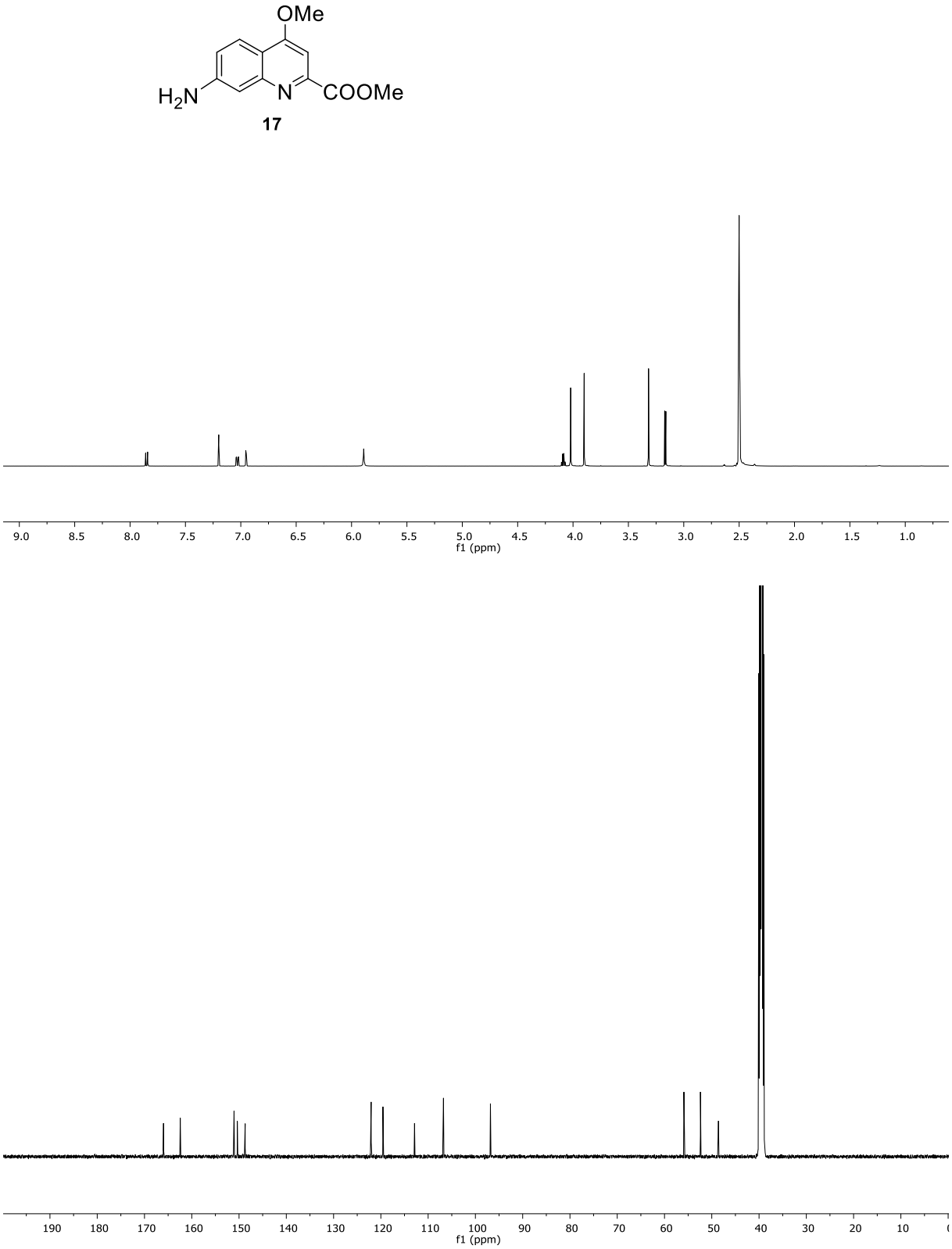


Figure S112. NMR spectra of compound **17** : ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

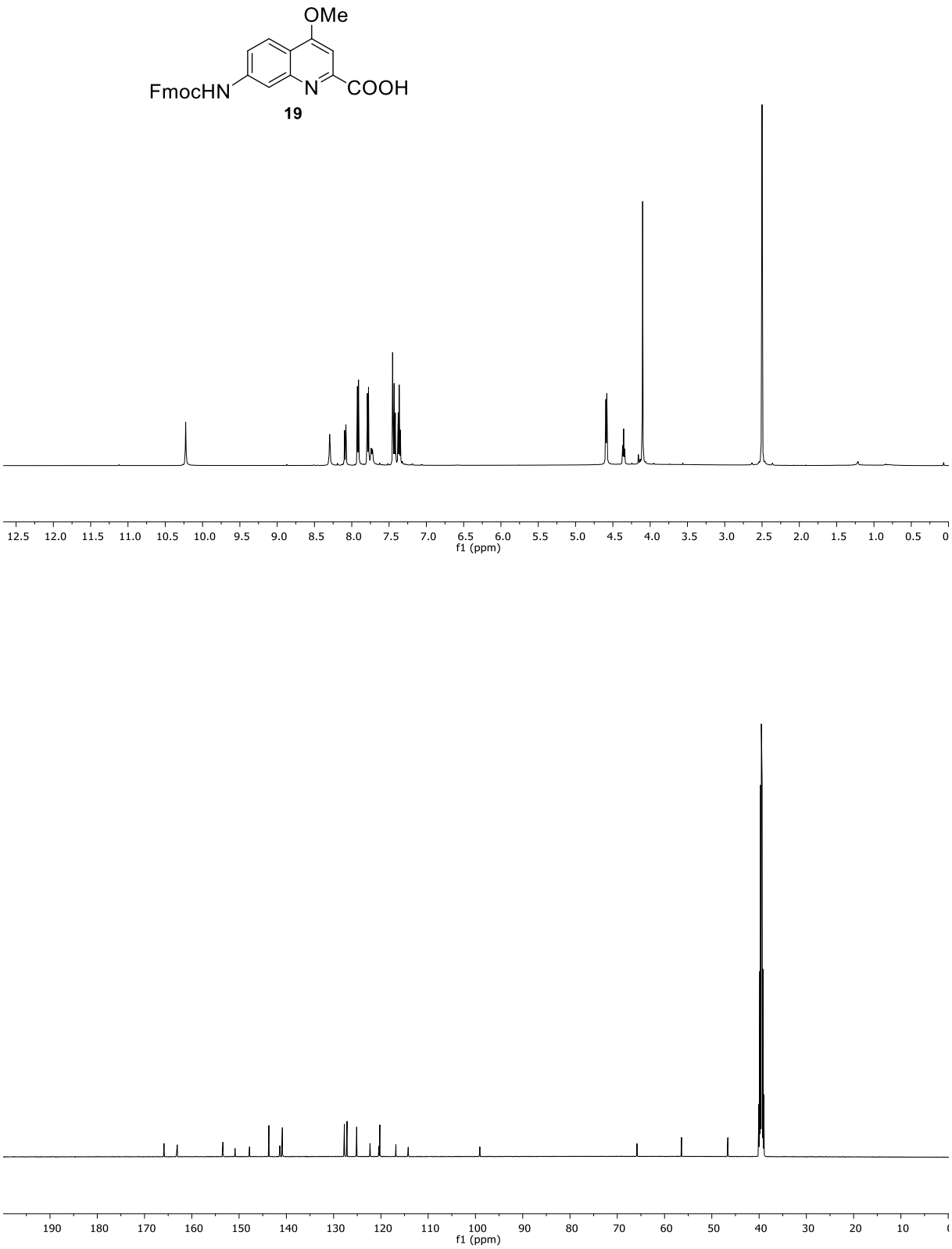


Figure S113. NMR spectra of compound **19** : ^{1}H NMR (500 MHz, DMSO-d_6) and ^{13}C NMR (126 MHz, DMSO-d_6).

5 Reference

1 B. Teng, J. Atcher, L. Allmendinger, C. Douat, Y. Ferrand and I. Huc, *Org. Biomol. Chem.*, 2023, DOI: 10.1039/D3OB00473B.

2 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.

3 X. Hu, S. J. Dawson, P. K. Mandal, X. de Hatten, B. Baptiste and I. Huc, *Chem. Sci.*, 2017, **8**, 3741–3749.

4 B. Baptiste, C. Douat-Casassus, K. Laxmi-Reddy, F. Godde and I. Huc, *J. Org. Chem.*, 2010, **75**, 7175–7185.

5 S. Dengler, P. K. Mandal, L. Allmendinger, C. Douat and I. Huc, *Chem. Sci.*, 2021, **12**, 11004–11012.

6 V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat, and I. Huc, *Chem. Eur. J*, 2023, DOI: 10.1002/chem.202301120.

7 A. A. McCarthy, R. Barrett, A. Beteva, H. Caserotto, F. Dobias, F. Felisaz, T. Giraud, M. Guijarro, R. Janocha, A. Khadrouche, M. Lentini, G. A. Leonard, M. Lopez Marrero, S. Malbet-Monaco, S. McSweeney, D. Nurizzo, G. Papp, C. Rossi, J. Sinoir, C. Sorez, J. Surr, O. Svensson, U. Zander, F. Cipriani, P. Theveneau and C. Mueller-Dieckmann, *J Synchrotron Rad*, 2018, **25**, 1249–1260.

8 W. Kabsch, *Acta Cryst D*, 2010, **66**, 125–132.

9 G. M. Sheldrick, *Acta Cryst A*, 2015, **71**, 3–8.

10 G. M. Sheldrick, *Acta Cryst C*, 2015, **71**, 3–8.

11 O. V. Dolomanov, L. J. Bourhis, R. J. Gildea, J. a. K. Howard and H. Puschmann, *J Appl Cryst*, 2009, **42**, 339–341.

12 P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, *Acta Cryst D*, 2010, **66**, 486–501.

13 A. L. Spek, *Acta Cryst D*, 2009, **65**, 148–155.

7 Conclusion and Perspectives

7.1 Conclusions from published/submitted work

This work resulted in the synthesis of a new 7-amino-2-quinoline carboxylic acid monomer (Q^H) with C and N terminals at a wider angle of 120°, capable of constructing a large diameter cavity. In Q and Q^H synthesis, a new protocol for aminomethyl side chain introduction was implemented to avoid a nitrile reduction that is ineffective in the Q^H synthetic route. The partially conformational bias of Q^H units can nevertheless promote helical folding in water, which is proportional to the length of the helix. Minimum segment length consists of two helical turns, which facilitates π-π stacking between distinct layers. When water was added to foldamer solutions in DMSO, a cooperative conformational transition was observed that has never been reported for foldamers with rigid backbones in organic solvents. The NMR and UV data revealed two well-defined stages relating to the transformation from the stretched to the folded state. Using bidimensional NMR measurements, the helical structure assignment was elucidated; the NOE correlations between distinct aromatic layers and the same layer provide strong evidence for helical order and folded structure. These results demonstrate conclusively the folding potential of non-fully locked building blocks, which inspired the backbone alternatives for future foldamer design.

Utilizing established water-soluble aromatic oligoamide helical capsules, a number of studies on guest binding were conducted in water. As demonstrated by NOESY and titration experiments, a variety of pyridine analogue guests exhibit 1 to 2 host-guest binding models with foldamer capsules in water. Fluorine compounds capable of forming hydrogen bonds with the endo-capsule amide group and highly water-soluble cyclic dipeptides with excellent shape complementarity show affinity for the amphiphilic cavity. In spite of the significant solvation of water-soluble guests, the capsule is still capable of achieving strong binding affinity and selectivity through the cooperation of hydrophobic effects and hydrogen bonds. Due to the release of highly ordered water in the cavity, an uncommon endothermic binding process that was solely driven by entropy was observed. The presence of water molecules in the crystal structure of a capsule with similar properties confirms this hypothesis. In addition, varied helical diameter building blocks were used to execute the rational shape design. In conclusion, our investigations solve the difficult work that binds water-soluble guests in water, which has potential implications in biological systems.

In order to reduce water solubility and facilitate crystal formation from water, the side chains have been iteratively optimized, resulting in a number of different dimerization patterns. After decreasing electron repulsion at the C terminal by substituting the charged aminomethyl side chain with a methoxy side chain, concentration-dependent and DOSY NMR revealed a head-to-head dimer. After preventing the head-to-head dimer with C terminal stereo hindrance, solution and solid-state studies identified a parallel double helix generated by the replacement of two negative side chains with neutral side chains. It was the first time that the crystal structure of a water-growing aromatic oligoamide helical capsule was determined. Furthermore, the backbone flexible Q^H segment performed helical folding in water which constructs a double helix structure as rigid building blocks. The side chain triggered aggregation was also observed in hemi-capsule sequence. Based on the related parallel double helix crystal structure, an energy-minimized model shown that four individual strands stack to produce a unique tetrameric helix. These distinct dimerization patterns illustrate the significant impact of sequence side chains and enable the construction of diverse secondary and tertiary abiotic structures by varying side chains.

7.2 Continuing challenges for foldamer capsules

Common issues in foldamer research include the limited number of functional groups that can be carried out in foldamer synthesis and the lack of commercially available building blocks. For helical capsule foldamers, there are particular requirements that must be met. Although a few monomers with a large helical diameter have been synthesized, the present building blocks all have disadvantages. The Q^F monomer, for instance, will occupy the cavity space with a fluorine atom; the naphthyridine monomer always presents synthetic issues in solid phase synthesis; and the Q^H monomer has a flexible backbone that cannot be folded in organic solvents. In addition, crystal growth for a single helix capsule has not yet been achieved due to the formation of aggregation by side chain alteration, and the capsule's flexible nature may further impede crystal growth or the refinement of X-ray diffusion data. Possible future plans are enhancing structural rigidity. Incorporating side chains into the capsule structure to immobilize the helical structure and prevent the formation of aggregates is a viable method (Figure 16a). Developing a new building block with a complete conformational bias via hydrogen bonding or metal chelation is an alternative strategy that will result in the development of new synthetic routes. In rigid capsule fabrication, additional work must be expended.

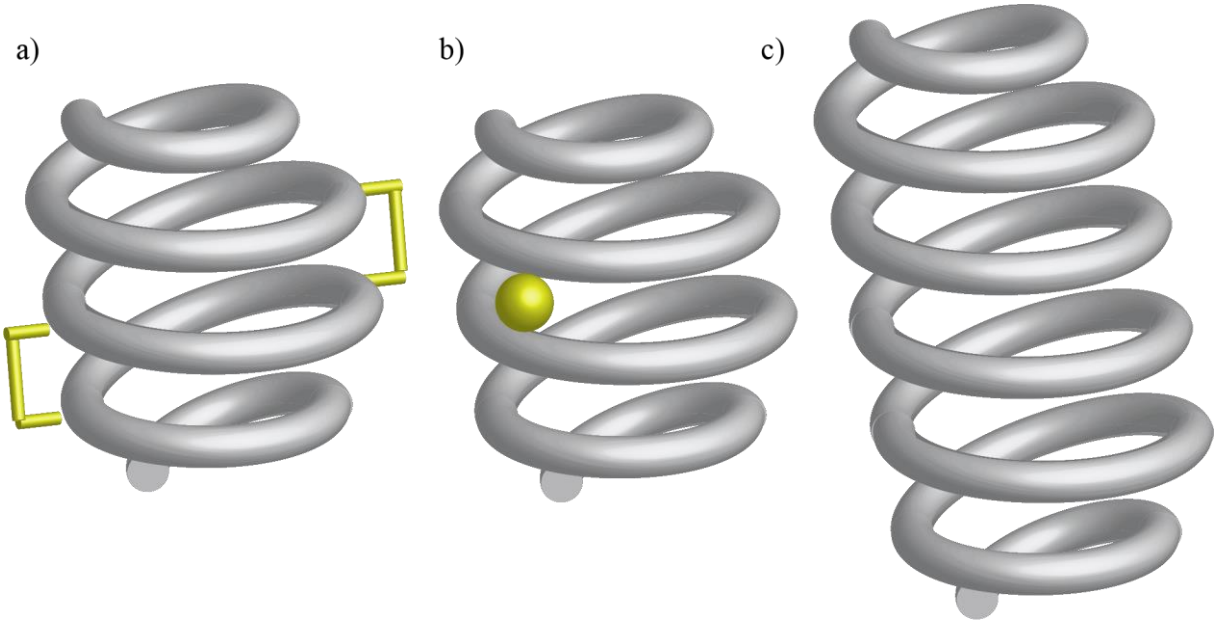


Figure 16. Overview of future helical capsule design. a) Lock capsule structure by rigid side chain to avoid dimerization and immobile structure. b) Introduce the boronic acid group to strongly bind carbohydrates in water. c) Elongate the helix to increase the number of interaction positions and capsulate more water molecules which can be released during the binding process and serve as an entropy driving force.

Carbohydrate binding with water-soluble oligoamide capsules in water is another difficult task that has not been completed. As an essential component of human metabolism, carbohydrate recognition in vivo promotes a variety of biological functions. Despite the fact that natural receptors can bind carbohydrates with high affinity and selectivity, chemical abiotic receptors that can bind carbohydrates specifically in water are still rare due to the strong solvation effect. Utilizing established capsules, NMR titrations were performed to determine binding affinity with carbohydrates in water, but no obvious chemical shift was observed. Extremely high water solubility necessitates enormous energy expenditures when the guest moves inside the cavity, making the operation challenging. As observed in the large cavity capsule, which can bind with the highly water-soluble cyclic dipeptide, the amphiphilic cavity only shows a low affinity. Introducing hydrophilic groups, which can form more hydrogen bonds, is a method for enhancing affinity. Because the cavity side group is tied to the backbone structure, it required multiple new synthetic routes to modify the backbone of the monomer. The introduction of a boronic acid group (Figure 16b), which can develop strong interactions with all types of carbohydrate within the capsule, is an alternative option. It is possible to acquire a high affinity, but the selectivity will be low. As demonstrated by ITC titrations, the binding process could be facilitated by releasing encapsulated water molecules in order to gain entropy. The third

possibility is increasing the length of the Q^H segment (Figure 16c), which can exclude more water molecules while the guest is bonded. In addition, the expanded interface offers more interaction positions, which increases binding affinity. In the future, it may be possible to achieve the extremely selective and strong affinity bonding of carbohydrates with aromatic helical foldamers by combining multiple strategies.

7.3 Potential applications and outlook

Aromatic helical foldamers derived from monomers with conformational bias regularly generate stable, difficult-to-disrupt helices, which is advantageous for rational design and the construction of complex tertiary structures under diverse environments. In contrast to rigid helices, foldamers containing several successive Q^H monomers are capable of entirely reversing their conformational state by modifying the solvent characteristics, which can be utilized for the regulation of guest release. Nitschke's group has described a molecular cargo employing coordination cages[83] dependent on this mechanism; the same application might also be applied

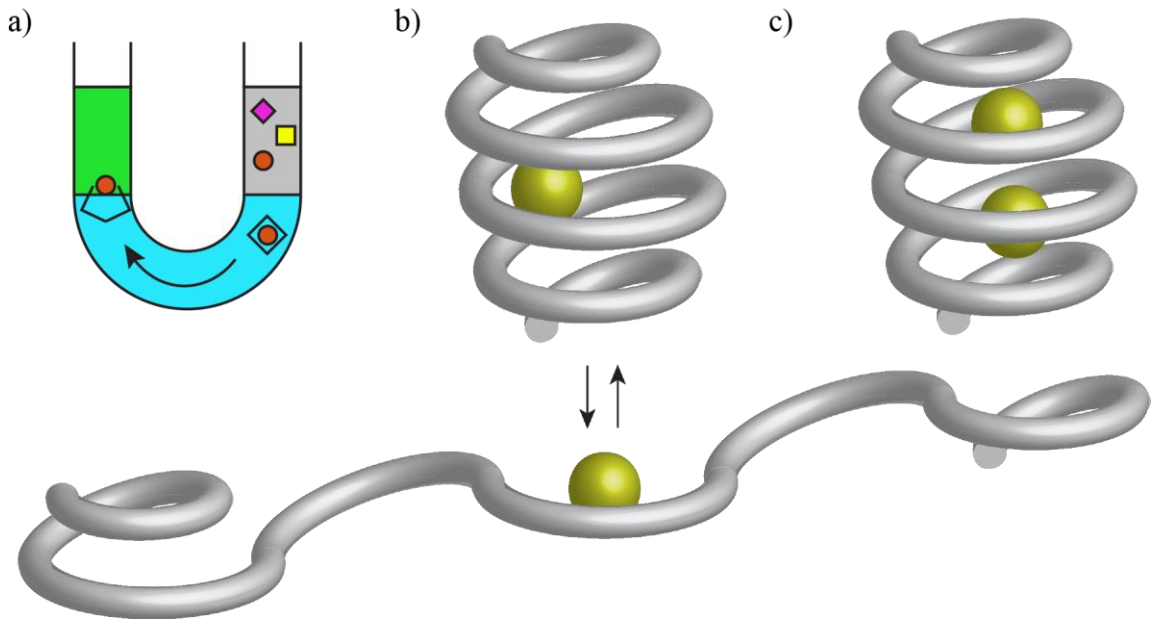


Figure 17. Overview of potential helical capsule applications. a) The triphasic system arranged in a U-shaped tube and the possible mechanism for selective transport via capsules (◇). At the interface of the feedstock and aqueous membrane layers, the target guest is encapsulated. The capsules and their cargos diffuse through the aqueous layer to the phase boundary of the receiving phase. The encapsulated cargoes are subsequently released into the receiving organic phase from the capsule cavities. b) The stretched and folded states control the catalysis components' (represented as a yellow ball) exposure or masking. c) Two guests are encapsulated in the capsule cavity, facilitating intermolecular reactions.

to helical capsules (Figure 17a). In relation to the stretched structure of Q^H capsules, the guest release can be completed in contrast to cage structures, which could result in more efficient transportation. The coordinated conformational transition property can also be exploited to regulate the masking or exposure of catalysis components within the capsule, so constructing a switch that monitors the reaction by varying the solvent (Figure 17b). In addition, the unique 1 to 2 binding model reveals the potential for a reaction container that increases the coupling reaction ratio by pre-locating two molecules together, and a chiral catalysis reaction might be accomplished by the stereo hindrance in the cavity (Figure 17c).

Foldamer is expensive due to its noncommercial building blocks and time-consuming synthesis and purifying operations, which necessitates its recycling. As a recyclable catalyst, affixing foldamers to a solid support is a viable solution (Figure 18a). And if it were utilized as a selective separate column, the foldamer capsule would capture the target compound in an aqueous solution and release the guest when washed with an organic solvent, which efficiently purifies a mixture via the incubation and washing steps. However, the method for installing aromatic helical foldamers to solid support still involves considerable effort. And self-healing materials are a potential future application for foldamers. The two large-diameter helical segments (Figure 18b) that can form a double helix assist polymerization and a rod-like structure can be formed. By combining to form a double helix, the two sticky ends allow for the reconstruction of any portion of the material. Despite the fact that a substantial amount of fundamental mechanism study must be conducted, the possibility exists.

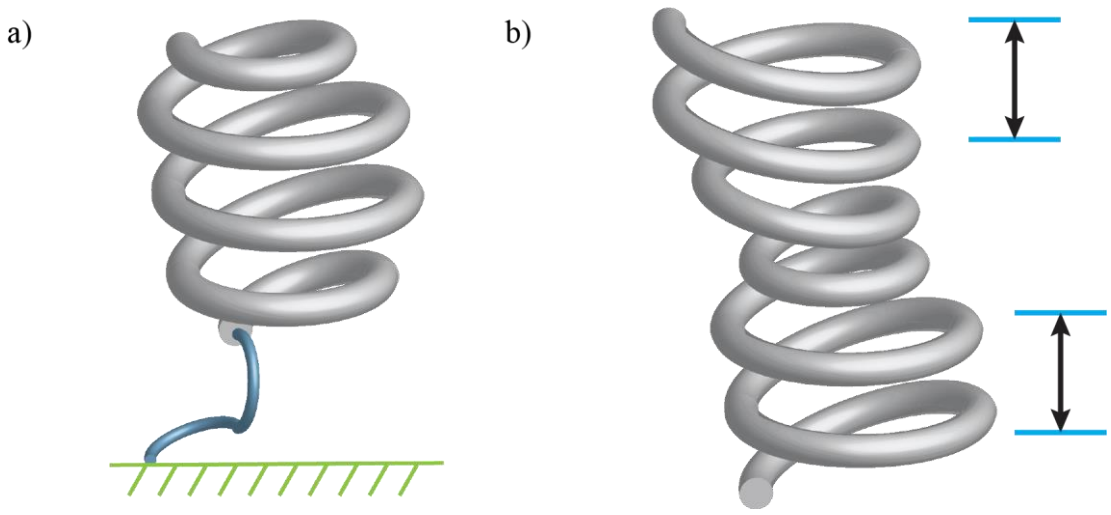


Figure 18. Overview of future foldamer designs. a) Set the capsule on a solid support. b) The building block design for polymer material. The large diameter helical part (pointed out by blue lines) could form two double helix with other units and endlessly elongate the rod-like structure.

8 Reference

1 A. Radzicka and R. Wolfenden, *Science*, 1995, **267**, 90–93.
2 B. P. Callahan and B. G. Miller, *Bioorganic Chemistry*, 2007, **35**, 465–469.
3 A. Stank, D. B. Kokh, J. C. Fuller and R. C. Wade, *Acc. Chem. Res.*, 2016, **49**, 809–815.
4 Q.-H. Guo, L. Zhao and M.-X. Wang, *Angewandte Chemie International Edition*, 2015, **54**, 8386–8389.
5 N. Song, T. Kakuta, T. Yamagishi, Y.-W. Yang and T. Ogoshi, *Chem*, 2018, **4**, 2029–2053.
6 B. W. Purse and J. Rebek, *Proceedings of the National Academy of Sciences*, 2005, **102**, 10777–10782.
7 Q. Wang, Y. Zhong, D. P. Miller, X. Lu, Q. Tang, Z.-L. Lu, E. Zurek, R. Liu and B. Gong, *J. Am. Chem. Soc.*, 2020, **142**, 2915–2924.
8 O. Dumele, N. Trapp and F. Diederich, *Angewandte Chemie International Edition*, 2015, **54**, 12339–12344.
9 J. A. Davies, T. K. Ronson and J. R. Nitschke, *Chem*, 2022, **8**, 1099–1106.
10 N. Chandramouli, Y. Ferrand, B. Kauffmann and I. Huc, *Chem. Commun.*, 2016, **52**, 3939–3942.
11 N. Chandramouli, Y. Ferrand, G. Lautrette, B. Kauffmann, C. D. Mackereth, M. Laguerre, D. Dubreuil and I. Huc, *Nature Chem*, 2015, **7**, 334–341.
12 D.-W. Zhang, X. Zhao and Z.-T. Li, *Acc. Chem. Res.*, 2014, **47**, 1961–1970.
13 S. H. Yoo, B. Li, C. Dolain, M. Pasco and G. Guichard, in *Methods in Enzymology*, ed. E. J. Petersson, Academic Press, 2021, vol. 656, pp. 59–92.
14 M. T. Stone and J. S. Moore, *Org. Lett.*, 2004, **6**, 469–472.
15 D. E. Koshland, *Proceedings of the National Academy of Sciences*, 1958, **44**, 98–104.
16 R. Chowdhury, A. Sau and S. M. Musser, *Nat Cell Biol*, 2022, **24**, 112–122.
17 D. D. Chaplin, *Journal of Allergy and Clinical Immunology*, 2010, **125**, S3–S23.
18 W. Xue, T. K. Ronson, Z. Lu and J. R. Nitschke, *J. Am. Chem. Soc.*, 2022, **144**, 6136–6142.
19 H. Takezawa, Y. Fujii, T. Murase and M. Fujita, *Angewandte Chemie International Edition*, 2022, **61**, e202203970.
20 M.-M. Tang, K. Kanagaraj, J. Rebek Jr. and Y. Yu, *Chemistry – An Asian Journal*, 2022, **17**, e202200466.
21 I. A. Riddell, M. M. J. Smulders, J. K. Clegg and J. R. Nitschke, *Chem. Commun.*, 2010, **47**, 457–459.
22 W. Liu and J. F. Stoddart, *Chem*, 2021, **7**, 919–947.
23 D. Chen, Y. Li, X. Li, X. Hong, X. Fan and T. Savidge, *Chem. Sci.*, 2022, **13**, 8193–8202.
24 H. Takezawa, K. Shitozawa and M. Fujita, *Nat. Chem.*, 2020, **12**, 574–578.
25 V. Martí-Centelles, A. L. Lawrence and P. J. Lusby, *J. Am. Chem. Soc.*, 2018, **140**, 2862–2868.
26 T. Jiao, K. Cai, J. N. Nelson, Y. Jiao, Y. Qiu, G. Wu, J. Zhou, C. Cheng, D. Shen, Y. Feng, Z. Liu, M. R. Wasielewski, J. F. Stoddart and H. Li, *J. Am. Chem. Soc.*, 2019, **141**, 16915–16922.
27 S. Matsuno, M. Yamashina, Y. Sei, M. Akita, A. Kuzume, K. Yamamoto and M. Yoshizawa, *Nat Commun*, 2017, **8**, 749.
28 N. Giri, M. G. Del Pópolo, G. Melaugh, R. L. Greenaway, K. Rätzke, T. Koschine, L. Pison, M. F. C. Gomes, A. I. Cooper and S. L. James, *Nature*, 2015, **527**, 216–220.
29 L. Ma, C. J. E. Haynes, A. B. Grommet, A. Walczak, C. C. Parkins, C. M. Doherty, L. Longley, A. Tron, A. R. Stefankiewicz, T. D. Bennett and J. R. Nitschke, *Nat. Chem.*, 2020, **12**, 270–275.

30 T. A. Goetjen, J. Liu, Y. Wu, J. Sui, X. Zhang, J. T. Hupp and O. K. Farha, *Chem. Commun.*, 2020, **56**, 10409–10418.
31 D. Chandler, *Nature*, 2002, **417**, 491–491.
32 D. Laage and J. T. Hynes, *Science*, 2006, **311**, 832–835.
33 S. Narayan, J. Muldoon, M. G. Finn, V. V. Fokin, H. C. Kolb and K. B. Sharpless, *Angewandte Chemie International Edition*, 2005, **44**, 3275–3279.
34 R. Ludwig, *Angewandte Chemie International Edition*, 2001, **40**, 1808–1827.
35 E. P. Schulz, L. M. Alarcón and G. A. Appignanesi, *Eur. Phys. J. E*, 2011, **34**, 114.
36 J. Trylska, P. Grochowski and J. A. McCammon, *Protein Science*, 2004, **13**, 513–528.
37 B. A. Fields, F. A. Goldbaum, W. Dall'Acqua, E. L. Malchiodi, A. Cauerhff, F. P. Schwarz, X. Ysern, R. J. Poljak and R. A. Mariuzza, *Biochemistry*, 1996, **35**, 15494–15503.
38 E. J. Sundberg, M. Urrutia, B. C. Braden, J. Isern, D. Tsuchiya, B. A. Fields, E. L. Malchiodi, J. Tormo, F. P. Schwarz and R. A. Mariuzza, *Biochemistry*, 2000, **39**, 15375–15387.
39 J. H. Hartley, T. D. James and C. J. Ward, *J. Chem. Soc., Perkin Trans. 1*, 2000, 3155–3184.
40 Y. Ferrand, M. P. Crump and A. P. Davis, *Science*, 2007, **318**, 619–622.
41 E. G. Percástegui, T. K. Ronson and J. R. Nitschke, *Chem. Rev.*, 2020, **120**, 13480–13544.
42 X. Hu, S. J. Dawson, P. K. Mandal, X. de Hatten, B. Baptiste and I. Huc, *Chem. Sci.*, 2017, **8**, 3741–3749.
43 P. Mateus, N. Chandramouli, C. D. Mackereth, B. Kauffmann, Y. Ferrand and I. Huc, *Angewandte Chemie*, 2020, **132**, 5846–5854.
44 M. A. Beatty and F. Hof, *Chem. Soc. Rev.*, 2021, **50**, 4812–4832.
45 F. Biedermann, V. D. Uzunova, O. A. Scherman, W. M. Nau and A. De Simone, *J. Am. Chem. Soc.*, 2012, **134**, 15318–15323.
46 S. Merget, L. Catti, G. Piccini and K. Tiefenbacher, *J. Am. Chem. Soc.*, 2020, **142**, 4400–4410.
47 M. Yoshizawa and M. Fujita, *Pure and Applied Chemistry*, 2005, **77**, 1107–1112.
48 M. M. J. Smulders and J. R. Nitschke, *Chem. Sci.*, 2012, **3**, 785–788.
49 T.-R. Li, F. Huck, G. Piccini and K. Tiefenbacher, *Nat. Chem.*, 2022, **14**, 985–994.
50 X. Yang, A. L. Brown, M. Furukawa, S. Li, W. E. Gardinier, E. J. Bukowski, F. V. Bright, C. Zheng, X. C. Zeng and B. Gong, *Chem. Commun.*, 2003, 56–57.
51 J. Wang, B. Wicher, V. Maurizot and I. Huc, *Chemistry – A European Journal*, 2021, **27**, 1031–1038.
52 M. T. Stone and J. S. Moore, *J. Am. Chem. Soc.*, 2005, **127**, 5928–5935.
53 M. S. Gin, T. Yokozawa, R. B. Prince and J. S. Moore, *J. Am. Chem. Soc.*, 1999, **121**, 2643–2644.
54 S. Saha, B. Kauffmann, Y. Ferrand and I. Huc, *Angewandte Chemie International Edition*, 2018, **57**, 13542–13546.
55 M. Horeau, G. Lautrette, B. Wicher, V. Blot, J. Lebreton, M. Pipelier, D. Dubreuil, Y. Ferrand and I. Huc, *Angewandte Chemie International Edition*, 2017, **56**, 6823–6827.
56 G. Lautrette, B. Kauffmann, Y. Ferrand, C. Aube, N. Chandramouli, D. Dubreuil and I. Huc, *Angewandte Chemie International Edition*, 2013, **52**, 11517–11520.
57 D. Sánchez-García, B. Kauffmann, T. Kawanami, H. Ihara, M. Takafuji, M.-H. Delville and I. Huc, *J. Am. Chem. Soc.*, 2009, **131**, 8642–8648.
58 X. Hu, S. J. Dawson, Y. Nagaoka, A. Tanatani and I. Huc, *J. Org. Chem.*, 2016, **81**, 1137–1150.
59 B. Baptiste, C. Douat-Casassus, K. Laxmi-Reddy, F. Godde and I. Huc, *J. Org. Chem.*, 2010, **75**, 7175–7185.

60 J. Shang, Q. Gan, S. J. Dawson, F. Rosu, H. Jiang, Y. Ferrand and I. Huc, *Org. Lett.*, 2014, **16**, 4992–4995.
61 M. Ueno, A. Murakami, K. Makino and T. Morii, *J. Am. Chem. Soc.*, 1993, **115**, 12575–12576.
62 D. Grueninger, N. Treiber, M. O. P. Ziegler, J. W. A. Koetter, M.-S. Schulze and G. E. Schulz, *Science*, 2008, **319**, 206–209.
63 S. L. Schreiber, *Cell*, 2021, **184**, 3–9.
64 J. C. Sinclair, *Nature Chem*, 2012, **4**, 346–347.
65 D. A. Uhlenheuer, K. Petkau and L. Brunsveld, *Chem. Soc. Rev.*, 2010, **39**, 2817–2826.
66 T. T. T. Truong, C. Cao and D. T. Dang, *RSC Adv.*, 2020, **10**, 29957–29960.
67 D. Bindl, P. K. Mandal, L. Allmendinger and I. Huc, *Angewandte Chemie International Edition*, 2022, **61**, e202116509.
68 S. De, B. Chi, T. Granier, T. Qi, V. Maurizot and I. Huc, *Nature Chem*, 2018, **10**, 51–57.
69 Y. Jiao, H. Mao, Y. Qiu, G. Wu, H. Chen, L. Zhang, H. Han, X. Li, X. Zhao, C. Tang, X.-Y. Chen, Y. Feng, C. L. Stern, M. R. Wasielewski and J. F. Stoddart, *J. Am. Chem. Soc.*, 2022, **144**, 23168–23178.
70 V. Koehler, A. Roy, I. Huc and Y. Ferrand, *Acc. Chem. Res.*, 2022, **55**, 1074–1085.
71 T. Kwon, B. Song, K. W. Nam and J. F. Stoddart, *J. Am. Chem. Soc.*, 2022, **144**, 12595–12601.
72 M. L. Singleton, G. Pirotte, B. Kauffmann, Y. Ferrand and I. Huc, *Angewandte Chemie International Edition*, 2014, **53**, 13140–13144.
73 Y. Suzuki, T. Nakamura, H. Iida, N. Ousaka and E. Yashima, *J. Am. Chem. Soc.*, 2016, **138**, 4852–4859.
74 S. Krykun, M. Dekhtiarenko, D. Canevet, V. Carré, F. Aubriet, E. Levillain, M. Allain, Z. Voitenko, M. Sallé and S. Goeb, *Angewandte Chemie International Edition*, 2020, **59**, 716–720.
75 S. Erbas-Cakmak, S. D. P. Fielden, U. Karaca, D. A. Leigh, C. T. McTernan, D. J. Tetlow and M. R. Wilson, *Science*, 2017, **358**, 340–343.
76 Y. Feng, M. Ovalle, J. S. W. Seale, C. K. Lee, D. J. Kim, R. D. Astumian and J. F. Stoddart, *Journal of the American Chemical Society*, , DOI:10.1021/jacs.0c13388.
77 C. Zeng, C.-Y. Zhang, J.-Y. Zhu and Z.-Y. Dong, *Chin J Polym Sci*, 2018, **36**, 261–265.
78 C. Bao, B. Kauffmann, Q. Gan, K. Srinivas, H. Jiang and I. Huc, *Angewandte Chemie International Edition*, 2008, **47**, 4153–4156.
79 E. Berni, J. Garric, C. Lamit, B. Kauffmann, J.-M. Léger and I. Huc, *Chem. Commun.*, 2008, 1968–1970.
80 Q. Gan, C. Bao, B. Kauffmann, A. Grélard, J. Xiang, S. Liu, I. Huc and H. Jiang, *Angewandte Chemie International Edition*, 2008, **47**, 1715–1718.
81 H. Ito, Y. Furusho, T. Hasegawa and E. Yashima, *J. Am. Chem. Soc.*, 2008, **130**, 14008–14015.
82 Y. Furusho, H. Goto, K. Itomi, H. Katagiri, T. Miyagawa and E. Yashima, *Chem. Commun.*, 2011, **47**, 9795–9797.
83 B.-N. T. Nguyen, J. D. Thoburn, A. B. Grommet, D. J. Howe, T. K. Ronson, H. P. Ryan, J. L. Bolliger and J. R. Nitschke, *J. Am. Chem. Soc.*, 2021, **143**, 12175–12180.

9 Acknowledgements

At this point I would like to thank everyone, who supported me the last four years and make this work possible.

First, I am grateful to Professor Ivan Huc for his guidance and for the chances he has afforded me. He provided me with the necessary guidance and inspiration to flourish in my research activities and realize my career goals.

I would also like to thank Yann Ferrand who provide me a lot of valuable suggestion in guest binding and modeling research.

Thanks to Céline Douat for her enthusiasm and useful advices, especially with HPLC and MS analysis and solid phase synthesis.

I would like to thank to Joan Atcher who supervised me in the beginning of my PhD. The scientific discussion with him really help me and inspire me a lot.

Thanks to Pradeep Mandal for his help in crystallization and X-ray analysis.

I would also like to thank Lars Allmendinger who gave me valuable advices in NMR data analysis and interpretation.

I thank Mathieu Denis for his kind help in scientific discussion.

Thanks to Vasilii Morozov for his guidance of ITC experiments and data analysis.

I would also like to thank Valentina Corvaglia who teach me how to do SPS manually.

Thanks to Susnata Pramanik for his help in NMR titration data analysis.

I would also like to thank Siyuan Wang who introduce the instruments in our lab to me.

Thanks to Daniel Gill and Mariia Palchyk for their precursor synthesis.

I also must thank Niklas Böcher, Lingfei Wang and Agnieszka Breyer for their help in the German affairs treatments.

Thanks to Daniel Bindl, Friedericke Menke, Eva Späh and Jiaojiao Wu for the happy hot pot time.

And I must to thank all the other group members for the collaborative working environment and brainstorm.

9 Acknowledgements

Thanks to my family and friends who encourage me to Germany and support me a lot.

Especially, I must thanks to Luojiao Huang who always motivated me and gave me supports in everywhere. It's her presence that makes achieving my goal possible.