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

Folding properties, handedness control and aggregation behavior of helical aromatic *δ*-amino acid foldamers in water

Daniel Bindl

aus

Burghausen, Deutschland

2022

Erklärung

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

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, den 21.02.2022 _____

Daniel Bindl

Dissertation eingereicht am	22.02.2022
1. Gutachter:	Prof. Dr. Ivan Huc
2. Gutachter:	Prof. Dr. Oliver Trapp
Mündliche Prüfung am	01.04.2022

Contents

List of publications

Accepted:

D. Bindl, E. Heinemann, P. K. Mandal, I. Huc, Quantitative helix handedness bias through a single H *vs.* CH_3 stereochemical differentiation. *Chem. Commun.* **2021**, *57*, 5662–5665. https://doi.org/10.1039/D1CC01452H

(See chapter 5)

D. Bindl, P. K. Mandal, L. Allmendinger, I. Huc, Discrete Stacked Dimers of Aromatic Oligoamide Helices. *Angew. Chem. Int. Ed.* **2022**, e202116509. https://doi.org/10.1002/anie.202116509

(See chapter 6)

Submitted:

D. Bindl, P. K. Mandal, I. Huc, Generalizing the aromatic δ-amino acid foldamer helix. *Chem. Eur. J.* **2022** (submitted).

(See chapter 4)

1 Abstract

Foldamers are an emerging class of molecules inspired by natural biopolymers, that also have the ability to fold into well-defined three-dimensional structures. As shape complementarity dictates many biological processes such as enzyme catalysis, signal transduction, and pathogen recognition, foldamers hold the potential to mimic and even extend the functions observed in proteins and nucleic acids. Over the decades, backbone types have been continually extended from β-peptides, which are still closely related to α-amino acids, to more abiotic oligomers. Aromatic oligoamides and aromatic δ-amino acids, in particular, offer easy synthetic access, amenability to solid phase synthetic methods and high folding propensity leading to helical structures that are very stable in most solvent environments. Thus, side chain positioning and geometry are well predictable, which offers a basis for functional designs. But tertiary and quaternary folds will ultimately be needed to unlock the true scope of these foldamers.

This work describes a new family of aromatic δ-amino acid monomers based on 2-(2-aminophenoxy)acetic acid (B). It was demonstrated that these more flexible units can be combined with the previously known aromatic δ-amino acid building blocks without significantly altering their canonical helical fold. The subtle differences in curvature of these units allow a fine-tuning of side chain positioning and stability of a given oligomer by adjusting its monomer composition and sequence order. Furthermore, a chiral B-unit was developed, which—when incorporated in the middle of aromatic helix sequences—proved to be able to bias handedness to over 99% towards one helicity. The monomer also induced handedness when positioned at the second or penultimate position of a sequence, albeit with weaker bias when close to the N-terminus. Thus, this unit enables designs that rely on both handedness control and free N- and C-termini for binding and/or further functionalization. Finally, the work describes the discovery of a binding interface between C-terminal aromatic helix cross sections leading to homochiral dimers in aqueous solution. Although they are based on aromatic stacking, the dimers are discrete, and their binding strength can be controlled by the nature of the side chains that are positioned close to the interface and the pH environment. By utilizing a primary amide terminus on one binding partner, exclusive heterodimer formation was achieved in the right concentration window. This binding interface can be useful in the future design of larger self-assembled structures.

Conclusively, these findings represent important tools for the development of more sophisticated foldamer designs in aqueous media. Additionally, preliminary results of the formation of a side-to-side helix-aggregate (which has been a guiding goal throughout the research for this thesis) are presented. Strategies for bundle formation that have been utilized and challenges that still remain are discussed and should serve as a starting point for future designs.

2 Introduction

Unlike modern organic synthesis, biochemistry does not rely on harsh reagents or transition metal catalysis—nature operates with shapes. Among many other functions, shape complementarity facilitates enzyme catalysis, signal transduction, and pathogen recognition. For instance, the poisonous effect of *amanita muscaria* also known as fly agaric can be ascribed to several psychoactive small molecules the mushroom produces, muscimol being the most relevant in this context. Muscimol receives its toxicity by mimicking the shape of the natural neurotransmitter GABA, enabling it to bind to the same active site on the $GABA_A$-receptor.[1] To accomplish such structural diversity, nature evolved biopolymers, particularly proteins, with enormous functional complexity, using only a surprisingly small set of building blocks: 20 canonical amino acids. The principle behind creating this diversity is folding, which organizes functional groups in specific spatial orientations. In an endeavor to better understand the underlying principles of folding in natural systems, as well as generating new architectures and functions, synthetic chemists started investigating non-natural (artificial) backbone structures that also had the ability to fold into well-defined structures. A class of molecules that became known as *foldamers*.

Foldamer research is a rapidly expanding field covering wide areas in chemistry, biology, and materials science. In particular, foldamer chemistry and investigation in aqueous solution is of interest due to its relevance for applications in biological systems. To achieve similar complexity compared to biological architectures, more sophisticated, higher-order structures (tertiary and quaternary folding) are needed. This thesis presents steps in this direction while exclusively focusing on aqueous solutions. Folding behavior and aggregation in water is “special” because intra- and intermolecular hydrogen bonding faces strong competition with the solvent and hydrophobic effects are maximized (often leading to unspecific aggregation). This work provides a general introduction into the field of foldamer science, while keeping a focus on aromatic amides as building blocks.

2.1 Introduction to foldamer research

First endeavors to replicate the structures and biological functions of peptides and nucleic acids involved molecules that are still closely related to their natural models.[2] For example, peptoids are intended to mimic peptides, the only difference being that the side chains are located on the amide nitrogen instead of the α-carbon. Although lacking chirality, they bring additional structural complexity through *cis/trans* tertiary amide bond isomerism that is influenced by the nature of the side chains. Similar to α-peptides, peptoids can also form helices, sheets, loops, ribbons, and other structures, while being more resistant so enzymatic degradation and more membrane permeable than α-peptides.[3] In an attempt to replicate DNA on the other hand, chemists substituted the sugar phosphate backbone of nucleic acids with *N*-(2-aminoethyl)glycine units, yielding peptide nucleic acids (PNAs). PNAs are chemically stable and more resistant to enzymatic degradation than their DNA counterpart, and can specifically bind target sequences of DNA or RNA through Watson-Crick base pairing.[4]

β-Peptides were also predicted to adopt well-folded conformations stabilized through hydrogen bonding between backbone amide functions like their α-amino acid counterparts (Fig. 1a). In 1996, Seebach and coworkers reported the synthesis and structural investigation of the β-hexapeptide (β-Val-β-Ala-β-Leu)$_2$, which formed a left-handed helical structure in solution evidenced by NMR and CD spectroscopy.[5] Moreover, it was found to be resistant to Pepsin at pH 2 for at least 60 h at 37 °C, conditions under which the corresponding α-peptide is digested instantaneously. Gellman and coworkers examined the effects of introducing small carbocycles within the backbone of β-peptides using computational models.[6] They identified *trans*-2-aminocyclohexanecarboxylic acid (ACHC) as a rigid building block that should enhance the stability of a 14-helical conformation. Experimental analysis via X-ray crystallography and deuterium exchange in methanol confirmed their assumptions. Amide protons located in the middle of the sequence took more than two days to completely exchange.

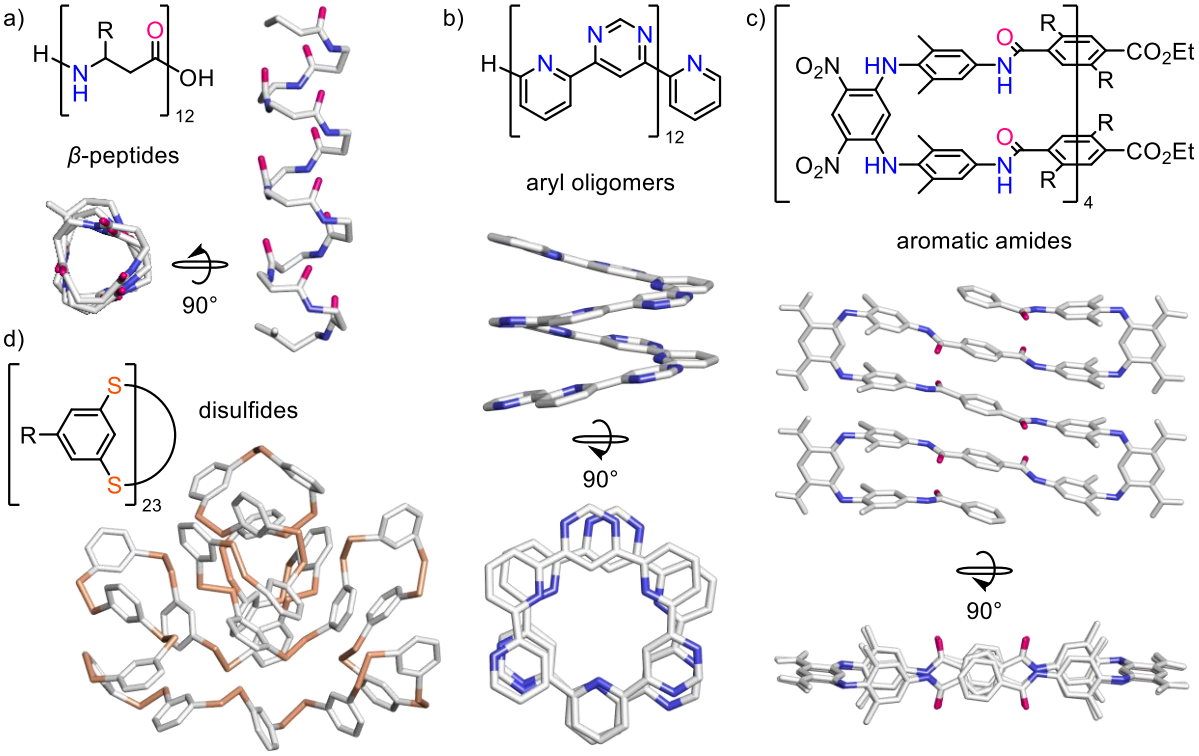


Fig. 1 – Overview of different foldamer families. a) Crystal structure of a β-peptide 12-mer.[7] b) Crystal structure of a helical aryl oligomer.[8] c) Crystal structure of a sheet-forming aromatic oligoamide.[9] d) Crystal structure of an aryl disulfide macrocycle.[10] Disulfide bonds are highlighted in orange. Oxygen and nitrogen atoms are highlighted in purple and blue, respectively. In all structures hydrogen atoms and side chains are omitted for clarity.

These unprecedented results proved that the known biopolymers are not the only oligomers with the ability to adopt well-folded conformations. The findings initiated the field of foldamer research, spurring chemists to look for other backbone types that are more remote from those of biopolymers. Research on the higher homologues (γ- and δ-peptides) and the possibility of exchanging the amide functions by ureas, hydrazides, hydroxyamides, and sulfonamides followed.[11] Because these structures are still very similar to their natural model (α-peptides), they are often considered *biotic* foldamers. To achieve more distinct properties and therefore functions, the field expanded into more diverse, *abiotic* backbone types. In particular, backbones containing aromatic units have received increasing attention since they often bring unique properties and stability to the system, for example, oligo-phenylene-ethynylenes,[12] aromatic electron donor and acceptor systems,[13] aryl oligomers (Fig. 1b),[8] and aromatic oligoamides.[14]

Different criteria may be used to categorize the many foldamer types available today. The type of linkage between units could be one such criteria (e.g., amide, hydrazide, urea, or alkyne groups). Additionally, the nature of the backbone units themselves can be differentiated into aromatic, a mixture of aromatic and aliphatic parts or purely aliphatic. Additionally, foldamers may be a composite of two or more different building blocks, which can be arranged in an alternating[15] or a block-like fashion.[16] On the other extreme, the oligomers may consist of only a single building block. Recent studies by Otto and coworkers showed the selective synthesis of folded oligomers of

specific lengths from dynamic reaction mixtures containing aromatic disulfide monomers. Depending on the peptide side chains that were attached to these building blocks, macrocycles with different chain lengths were preferred, each possessing a unique three-dimensional fold (Fig. 1d).[10] Yet another classification may be the way the monomers are connected: linear, cyclic,[17] branched,[18] or polymeric.[19]

Extensive research has been conducted in the area of aromatic oligoamides, with a multitude of building blocks having been investigated and characterized. Many different heterocyclic as well as multicyclic systems bearing a variety of substituents and side chains have been used. Their structures are strongly influence by the positioning of the amide linkages on the aromatic rings of their monomers. For example, oligomers of *ortho*-substituted amino benzoic acid form a zigzag-shaped linear strand.[20] In contrast, *meta*-substitution tends to yield crescent or helical structures,[21] while *para*-substitution again favors linear strands.[22] However, Huc and coworkers found that oligomers of the quasi *ortho*-substituted 8-amino-2-quinolinecarboxylic acid (Q) fold into stable helical structures with an unusually high curvature for aromatic amide foldamers.[23] Other folds like sheets may be generated through the combination of different building blocks acting as strands and turns (Fig. 1c).[9]

Synthetic approaches to access aromatic oligoamides can be divided into solution and solid phase assisted strategies. As solid phase assisted synthesis has been developed for peptides, reagents and coupling conditions have been optimized for the use of α-amino acids whose amine function is more reactive compared to most aromatic amines. Therefore, a transfer of this chemistry to aromatic amino acids is not trivial. Initially, aromatic oligoamides were exclusively synthesized via solution phase methods. Between the commonly used strategies, acid chloride activation is usually preferred (or necessary) over peptide coupling reagents or reactive ester activation due the low reactivity of aromatic amines. Additionally, helical folding often hampers reactivity drastically once a full helix turn has been synthesized. The synthesis of longer oligomers may then require harsher conditions such as increased temperatures, or the use of polar solvents that disrupt the intramolecular hydrogen bonding. A more sophisticated approach to circumvent this problem is the use of 2,4-dimethoxybenzene (DMB) protecting groups that keep amides in a *cis*-conformation until the end of the synthesis.[24] Once deprotected, the oligomer folds back into its intended (helical) conformation. Employing this strategy, even strained systems not accessible otherwise may be obtained (Fig. 2b).[17]

The synthesis of an oligomer usually consists of multiple deprotection, activation, and coupling steps. Purification might be necessary after each coupling step, severely limiting the scope of solution phase synthesis, which often necessitates the use of convergent schemes or segment doubling condensation strategies. While these approaches may yield long oligomers,[25] the synthesis of sequences bearing a variety of different side chains in arbitrary orders via solution phase chemistry is rarely feasible. Solid phase assisted synthesis, on the other hand, is an ideal tool for these targets and has been a subject of extensive optimization in the Huc group (Fig. 2a).[26] When coupling onto

aromatic amines, Fmoc-protected amino acids are converted to their respective acid chlorides *in situ*, whereas common peptide coupling reagents are used for aliphatic amines. Additionally, folding of the molecule during synthesis becomes an advantage when working on solid support. Aggregation on the resin is a common problem in peptide synthesis, which hampers further couplings. Conversely, the helical shape of quinoline-based oligomers lets them protrude from the solid support eliminating this problem.

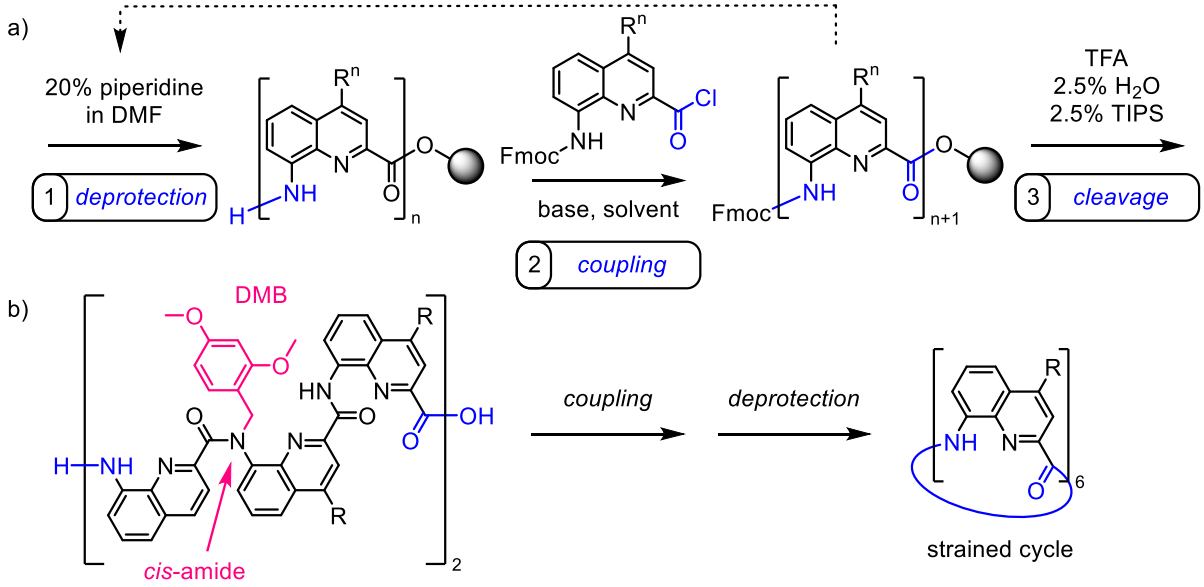


Fig. 2 – Synthetic strategies for the synthesis of aromatic oligoamides. a) Typical scheme for the solid phase synthesis of Q-oligomers consisting of the two repeating steps (1) Fmoc-deprotection and (2) coupling of the next monomer onto the resin-bound amine and (3) a final cleavage from the resin. b) Representative example of a solution phase strategy including the conformational protecting group DMB (highlighted in purple).[17] Moieties taking part in the reactions are highlighted in blue.

2.2 Folding principles and monomer design

The final shape of a molecule is determined by the sum of multiple factors. Folding is driven by internal constraints, the shape and rigidity of the monomers, and the possible attractive and repulsive local and non-local intramolecular interactions between monomers—the latter one being the most difficult to predict. Possible attractive forces include hydrogen bonding, donor–acceptor interactions,[13] and aromatic stacking. Among repulsive forces are steric and electrostatic (dipole–dipole) repulsion.[8] Additionally, external factors like solvent effects (hydrophobic effect), aggregation phenomena, host-guest complexation,[27] and surface interactions are influencing the folding behavior. In flexible backbones, e.g., *β*-peptides, differences in the structure of the monomers can change their conformational preferences. The use of ACHC units favors a 14-helix hydrogen bonding pattern, whereas *trans*-2-aminocyclopentanecarboxylic acid (ACPC) leads to 12-helices.[28] In contrast to aliphatic peptides, hydrogen bonding in aromatic oligoamides typically occurs between neighboring units, leading to higher stability of their secondary structure motives.

One might expect that with the large number of building blocks reported so far, there would be a wide variety of different folding patterns, but this is not the case. The motives common in biopolymers (helices, sheets, linear strands and turns) seem to prevail in non-natural oligomers as well; a universality in folding behavior which has been theoretically predicted for oligomeric molecules.[29] Furthermore, foldamer helices are far more abundant in the literature than sheets. This could stem from the inherent propensity of sheets to aggregate, drastically reducing their solubility and thus hampering their analysis in solution. Apart from this, some motives that are uncommon in biopolymers have been observed in foldamers including cyclic topologies,[30] sheet-like stacks,[13] knots,[31] tail-biters,[32] and non-canonical helices.[33] Structural flexibility between different conformational states is often crucial for biological function. Switching of folded states via external stimuli of foldamers has been achieved through changes in the protonation state,[34] chemical reactions,[35] and binding of small molecules[27] or ions.[36] Conversely, a certain structural stability and predictability is useful in the design of new functional foldamers.

Amides offer simple synthetic accessibility and provide a basis for intramolecular hydrogen bonding: strong and directed interactions that are thus easy to program. In combination with rigid aromatic units, this allows the orientation of monomers in predefined angles. Aromatic amide bonds may form different *cis/trans* and *syn/anti* configurations, which can be influenced by substituents on the aromatic rings and/or the amide nitrogen (Fig. 3).

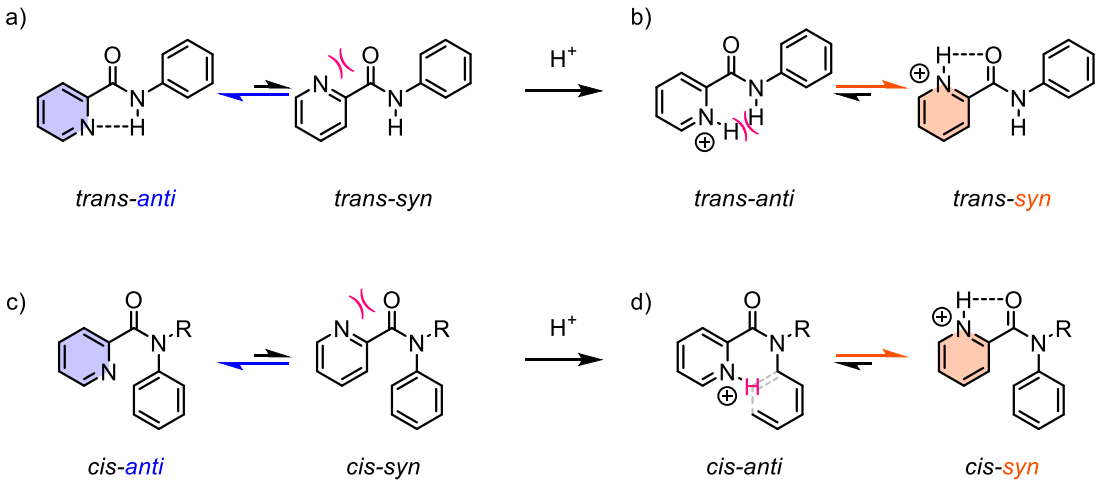


Fig. 3 – Possible conformations for aromatic amide bonds. a) Secondary amide stabilized in *anti*. b) Secondary amide stabilized in *syn*. c) Tertiary amide stabilized in *anti*. d) Tertiary amide stabilized in *syn*. Hydrogen bonding is indicated by dashed lines, steric repulsion is indicated in purple.

Secondary aromatic amides are predominantly *trans*, while tertiary amides prefer a *cis* conformation. Without substituents in *ortho* position, aromatic rings can rotate freely around the Ar–NH(CO)Ar and the Ar–CO(NH)Ar bonds. However, these rotations can be locked by installing hydrogen bond donor or acceptor groups next to the amide. For example, introducing an endocyclic nitrogen *ortho* to the amide-carbonyl will shift the equilibrium towards the *anti*-conformation as it forms a hydrogen bond with the amide-NH (Fig. 3a). Additionally, there would be a dipole–dipole repulsion in *anti*.[37] This preference is inverted via protonation, converting the acceptor into a

donor that will hydrogen bond to the amide carbonyl in the *syn*-conformation (Fig. 3b). The curvature dictated by the amide bond between two aromatic units usually propagates when elongating the sequence leading to either linear, or more or less curved crescent or helical structures. Furthermore, the intramolecular hydrogen bonding bends the molecules leading to higher curvatures at the inner, and lower curvatures when positioned at the outer rim of the strand or helix.

The conformation of Q-monomers is stabilized via two hydrogen bonds between the amide-NH and the neighboring endocyclic quinoline nitrogens (Fig. 4). In Q-oligomers, the theoretical 60°-angle between the units should lead to a helix with 3 units per turn; however, the contraction induced by the hydrogen bonding at the inner rim of the helix results in 2.5 units per turn observed experimentally.[23] Therefore, a quinoline trimer (Q_3) already folds into a helical structure while the dimer is planar. Helices are chiral objects that can exist in a right-handed (*P*) or left-handed (*M*) conformation. In solution, molecules are interconverting between *P* and *M*-helicities via partially unfolded states. The rate of helix handedness inversion of Q-oligomers is influenced by the length of the sequence and the nature of the solvent.[38] An octamer has an inversion half-life of 6 min in chloroform. More polar solvents significantly slow the kinetics resulting in a half-life of 40 min and 900 min for dimethylformamide and methanol, respectively. In water, oligomers consisting of five units or less still invert with a half-life of hours.[39] Longer sequences are kinetically locked and do not interconvert their handedness on a practical timescale.

Q-monomers may be combined with other building blocks like α-amino acids,[26b] or oxazoles.[40] Most importantly, they may be replace by 5-methylaminopyridine-2-carboxylic acid (P) units, which differ from Q only by the removal of one aromatic ring (Fig. 4). Both being aromatic δ-amino acids, P monomers are able to participate in the same hydrogen bonding network and impose the same curvature on the helix, thus not altering the overall structure. The increased flexibility resulting from the aliphatic methylene group and reduced aromatic surface of P leads to a reduction in stability of the helix. Without Q units, P oligomers do not fold into well-defined structures in aqueous media.[26a] However, a couple of P monomers may be used to increase handedness inversion kinetics, which is necessary for example to monitor protein binding through changes in the CD signal of the interacting oligomer.[41] The flexibility may also facilitate distortions of the helix to achieve better binding to target molecules. Furthermore, the removal of one aromatic ring in P changes the surface shape of the helix, which might remove steric hindrance and thus also aid in binding events.[42]

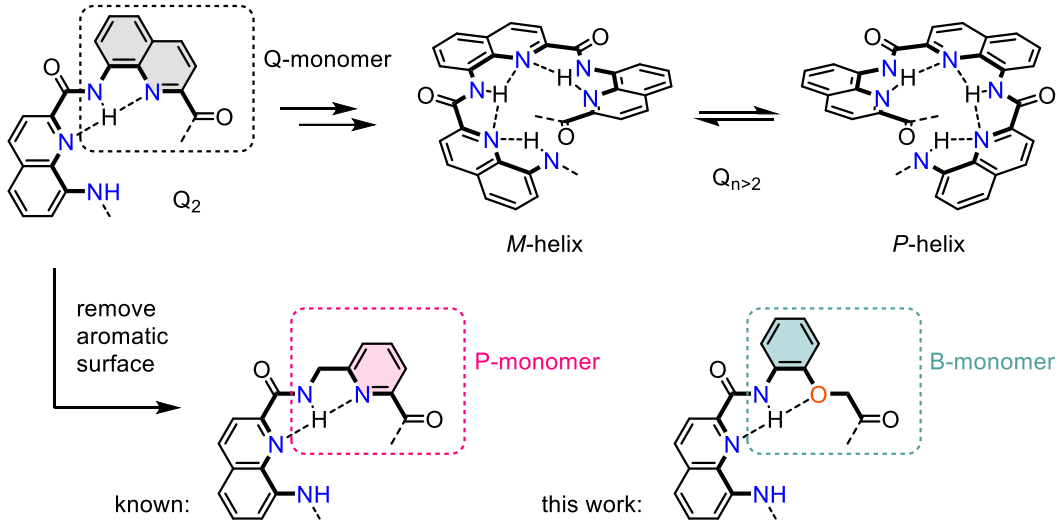


Fig. 4 – Monomer structures of aromatic *δ*-amino acid oligomers. Hydrogen bonds are indicated by dashed lines and atoms involved in hydrogen bonding are highlighted in color (oxygen: orange, nitrogen: blue). Steric repulsion in the trimer leads to helical structures with left-handed (*M*) or right-handed (*P*) helicity. The inner rim of the helix is indicated with bold bonds. Q-monomers may be replaced by P- or B-monomers that are able to form a similar hydrogen bonding pattern.

Another possibility is the removal of the heterocyclic ring of Q leading to 2-(2-aminophenoxy)-acetic acid (B) units. Here the endocyclic nitrogen is replaced by an ether oxygen to preserve the hydrogen bonding pattern. Reports of this monomer are scarce in the literature,[43] and they have never been combined with Q or P units. This work presents progress into this direction, the results are reported in chapter 4.

Helical oligomers consisting of chiral monomers usually have a strong preference for one helix sense. Common helices formed by *α*-peptides (α-helix, 3_{10}-helix, π-helix) are always right-handed. In *β*-peptides, *β*^{3}-substituted units derived from natural L-amino acids were found to prefer a left-handed fold.[5] Since aromatic units are mostly achiral, oligomers thereof are obtained as a racemic mixture of *P*- and *M*-helices. The equilibrium may be biased towards one handedness through external factors such as binding to chiral surfaces[39] or guests.[44] Alternatively, chiral moieties may be covalently attached to side chains, the ends of the helix or within the sequence. Little energy difference between the diastereomers is necessary to affect a small handedness bias, so most chiral functionalizations in proximity to the helix backbone will create a preference.[45] In Q-oligomers, the incorporation of *α*-amino acids into the oligomer leads to a handedness bias, however disturbing the usual curvature.[26b] Groups installed at the terminus do not alter curvature, but they prevent either C- or N-terminus to be functionalized otherwise, or to engage in binding/recognition events. Achieving nearly quantitative bias usually involves bulky groups that themselves engage in hydrogen bonding interactions with the quinoline-backbone.[46] A new way to fully control handedness from within the sequence without strongly altering the canonical curvature of the helix is presented in chapter 5.

2.3 Unique properties and applications

Foldamers are not only intended to mimic the structures of natural systems, but they may also be able to generate new structural features and functions. In contrast to biological constraints like evolutionary pressure, naturally available precursors, and metabolic pathways, the building blocks used in foldamer research are not restricted in terms of chemical diversity. This has led to natural folded oligomers now being a small part of the variety of folded backbones known to chemists. In addition to monomer diversity, more than one kind of building block may be combined in an oligomer. A variety of hybrid sequences has been reported, often expressing unique structural properties: aliphatic hybrids,[47] heterochiral hybrids,[48] aliphatic/aromatic hybrids,[49] or heterogenous aromatic oligoamides, which can form unique structures, for example, capsules (Fig. 5a).[50] Here, monomers coding for different curvatures are combined in such a way that a cavity is generated in the middle of the helix. Both ends of the helix are capped by monomers with a higher curvature. A guest molecule can be bound inside of the capsule, completely isolating it from the solvent. Other general functions of foldamers that have been reported include catalysis,[51] sensors,[52] communication of structural information,[35] self-selection,[53] and molecular machines.[54]

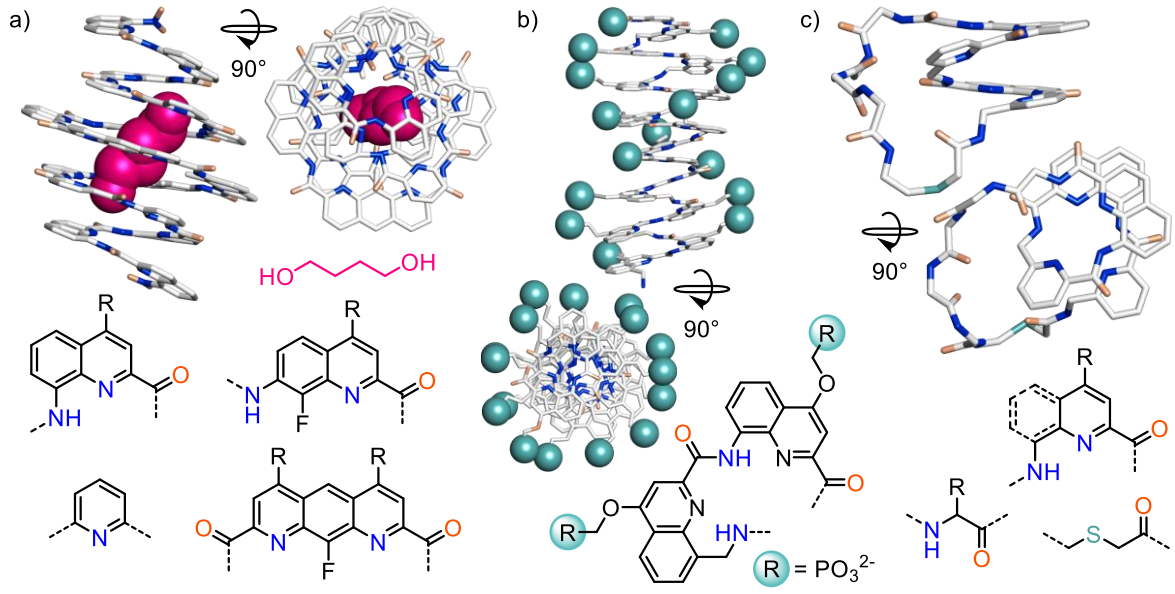


Fig. 5 – Applications of aromatic amide foldamers. a) Crystal structure of a foldamer helix encapsulating 1,4-butanediol.[50] The guest is shown in purple CPK representation. b) DNA mimic foldamer with phosphate groups shown as turquoise balls.[55] c) Crystal structure of an aromatic-aliphatic hybrid foldamer macrocycle with a thioether linkage.[16] The formulas of the utilized building blocks are given next to the respective structures. Nitrogen, amide oxygen, and sulfur atoms are highlighted in blue, orange, and turquoise, respectively.

Foldamers are also promising candidates in the area of chemical biology. While α-peptides and nucleic acids are quickly digested *in vivo*, non-natural oligomers are usually stable. Even β-peptides, despite being structurally very similar to α-peptides, are not degraded.[5] Additionally, some foldamers have been demonstrated to be able to cross cell membranes efficiently, and may even be used as drug carriers.[56] Conversely, foldamers might themselves be pharmacologically active. Their

medium size with large surface areas, structural predictability, and tunability combined with their usually low toxicity and immunogenicity makes them promising candidates for the disruption of protein-protein interactions (PPIs)—a difficult problem to tackle via classical small molecule drugs. Potential strategies for drug discovery include: α-helix stabilization,[57] α-helix mimicry,[58] DNA mimicry (Fig. 5b),[55] and protein surface recognition.[59] Recently, small aromatic amide oligomers could be utilized in ribosomal peptide synthesis amenable for mRNA display screening,[16] leading to hybrid macrocyclic structures (Fig. 5c). In these molecules the aromatic helix forces the α-peptide to adopt a stretched conformation, stabilizing it against proteolytic degradation.[49]

Among aromatic foldamers, Q-oligomers show remarkable properties. Unlike peptides, whose isolated secondary motives are rarely stable, they show unusually high stability in most solvents including water (still folded at 120 °C in DMSO) and thus structural predictability.[23] The high curvature imposed by the ~60° angle between amide linkages and the contraction caused by hydrogen bonding at the inner rim of the helix leads to 2.5 units per turn (or 5 units per 2 turns), resulting in 5 distinct side chain arrays a–e on the side of the helix, in case the typical monosubstituted building blocks are used (Fig. 6c). It also leads to longer helices per unit compared to other aromatic foldamers, which usually have a helix pitch equaling the thickness of an aromatic ring.[20, 60] Another consequence is the very narrow helix cavity, which is occupied by the amido protons blocking solvent molecules from diffusing through the helix. This hydrogen bonding network also provides the molecules with a high mechanical stability.[61] When compared to α-helices (as found in coiled coils), the most striking differences are: hydrogen bonding is occurring only between consecutive units, not neighboring helix turns, Q-oligomers possess a larger diameter, less distinct side chain positions and a smaller helix pitch of 3.4 Å in contrast to 5.4 Å for α-helices (Fig. 6a,b,c).

Furthermore, the synthetic access to quinoline monomers is straightforward (3 steps to methyl 8-nitroquinoline carboxylate with different ether linked side chains in position 4),[23] and the oligomers can be produced via automated solid phase synthesis. The resulting foldamers are typically analyzed in solution by NMR and CD spectroscopy and in the solid state by X-ray crystallography. Because of their compact shape and resulting ability to pack well (Fig. 6e), Q-oligomers tend to be well suited for crystal growth giving insight into their exact structure and side chain positioning. These combined features facilitate the iterative development of more sophisticated designs and functions (for example optimizing protein foldamers interactions). Specialized applications involving quinoline foldamers include circularly polarized luminescence,[62] and charge transport,[63] which could lead to the development of molecular wires.

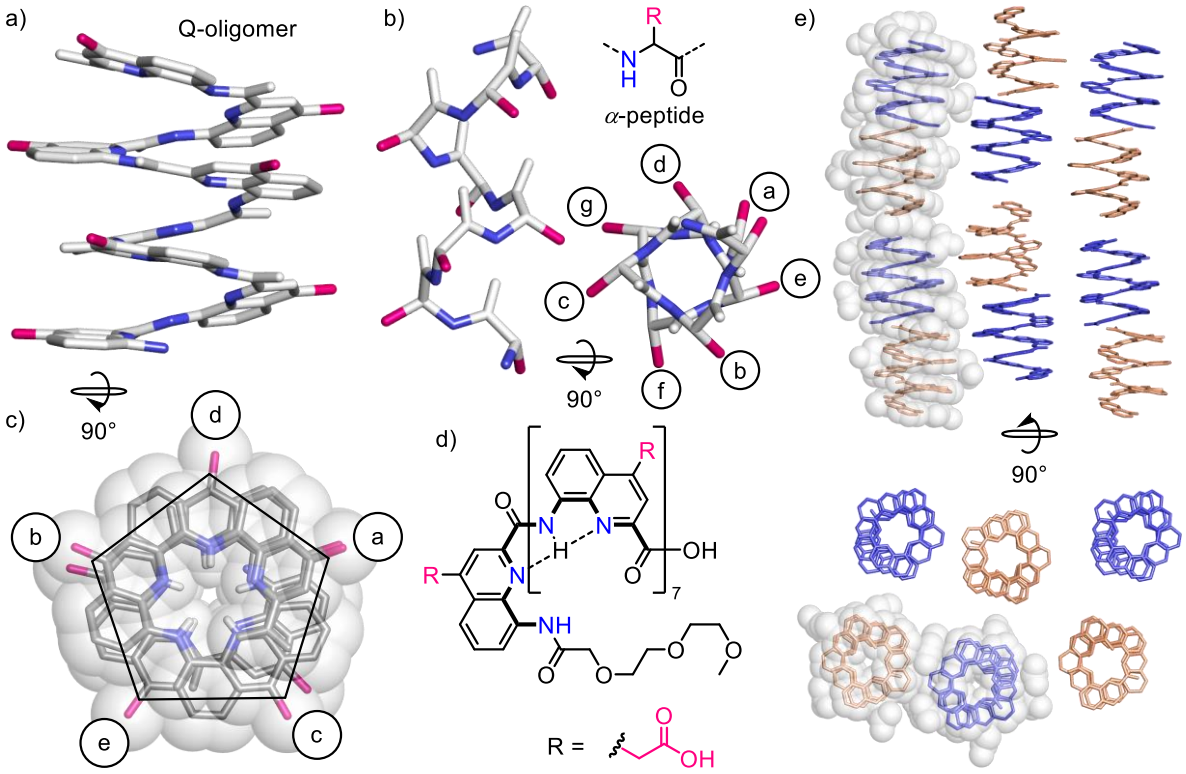


Fig. 6 – Structural properties of Q-oligomers. a) Crystal structure of a Q-octamer in side view.[64] b) Part of the crystal structure of a peptide α-helix (in a coiled coil) in side and top view with the same proportions.[65] Hydrogen atoms are omitted for clarity. c) Q-helix in top view illustrating the five distinct side chain arrays protruding from the side of the helix (numbered according to their positioning in the sequence; overlayed by a transparent CPK representation). Only the first atom of each side chain and amide hydrogen atoms are shown. Side chains and nitrogen atoms are highlighted in purple and blue, respectively. Terminal groups are omitted. d) Chemical formula corresponding to the structures shown in (a), (c) and (e). e) Packing of a Q-helix in the solid state with individual molecules shown in different colors. Part of the structure is overlayed by a transparent CPK representation. Side chains, terminal groups, and hydrogen atoms are omitted for clarity.

2.4 Tertiary folding and self-assembly

The structures of proteins follow a hierarchy from primary sequence information over secondary and tertiary folding to quaternary assemblies. Little can be accomplished with isolated secondary motives: most functions start to arise at the tertiary and quaternary level. A similar trend should be expected for foldamers. Tertiary folding motives may be strongly influenced by external stimuli allowing big structural changes upon binding of a signal molecule for example. Quaternary assembly further adds functionality and also represents a way of drastically increasing the size of the final complex without elongating the primary sequence. For many foldamer types, secondary folding is well understood, but tertiary and quaternary structures are still rare in the literature. Consequently, despite the many applications of foldamers already reported that only rely on secondary motives, there should be many more functions arising at the tertiary and quaternary level. However, the

design of higher order structures is difficult. Attempts to generate tertiary or quaternary structures using foldamers have often been inspired by natural motives, especially α-helix bundles.

Helix bundles of *α*-amino acids, also known as coiled coils, are a binding motive ubiquitous in natural proteins in which two or more α-helices wind around each other, usually forming a left-handed superhelical strand. Although less common, right-handed winding as well as straight orientations are also found. Furthermore, coiled coils do not only form simple strands: many other structures including barrels, sheets, funnels, and spirals have been characterized so far. As already predicted in 1953 independently by Francis Crick[66] and Linus Pauling,[67] coiled coils follow a set of simple and predictable rules, making them one of the most well understood structures in relation to their amino acid sequence—a problem that is still not entirely solved, despite recent advances in machine learning approaches.[68] At the basis of coiled coil structures lies the heptad repeat (Fig. 7a), a pattern that is repeating every seven amino acid units (a–g).

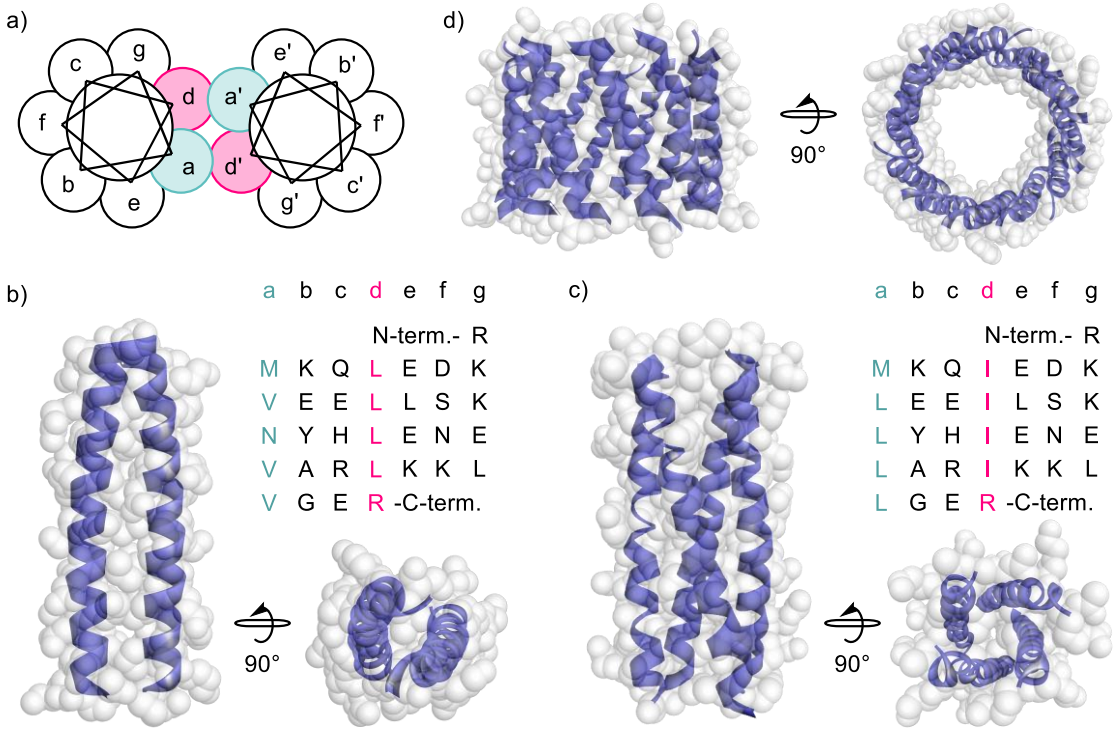


Fig. 7 – Coiled coil structures and design principles. a) Schematic illustration of the heptad repeat pattern. b) Crystal structure of the GCN4 leucine zipper forming a dimeric helix bundle.[65] c) Crystal structure of a mutated GCN4 leucine zipper forming a tetrameric helix bundle. Peptide sequences are shown next to the respective structure. The interface positions a and d are highlighted in turquoise and purple, respectively. d) Part of the crystal structure of the transmembrane protein ToIC.[69] Peptide helices are shown in ribbon representation with a transparent overlay of the CPK structure.

In the heptad repeat, amino acids a and d usually bear apolar side chains, forming a zigzag pattern of hydrophobic groups on one face of the helix. However, a canonical α-helix possesses a periodicity of 3.63 residues per turn on average. To keep the residues parallel, the α-helices must bend to adjust their periodicity to 3.5 (or seven per two turns), leading to a supercoiled structure. Each hydrophobic group protrudes into a vacant spot on the face of the other helix (knob-into-hole packing). The

orientation and multiplicity of the resulting aggregates strongly depends on the nature of the residues in positions a and d as well as the flanking amino acids e and g. For example: β-branched residues (Ile or Val) in position a, and γ-branched residues (Leu) in position d favors dimers (Fig. 7b). Switching this order leads to tetramers, as was illustrated for a mutant of the well-studied leucine zipper domain of the yeast transcriptional factor GCN4 (Fig. 7c).[65] Salt bridges in position e and g contribute to binding selectivity, while hydrophobic residues in these positions can lead to even bigger aggregates. Multiplicities of up to 12-mer helix bundles have been observed, e.g., the bacterial transmembrane protein ToIC, which is responsible for multidrug efflux and protein transport (part of the protein structure: Fig. 7d).[69] Conclusively, the high sequence to structure predictability and functional diversity of coiled coils provides a perfect starting point for protein engineering. A transfer to entirely non-natural oligomers is not trivial though.

The design of tertiary structures using foldamers is difficult, because: long sequences must be synthesized by chemical methods; the backbone must be able to fold into more than one stable secondary motive (e.g., helices and turns); and non-local intramolecular interactions are hard to predict. Foldamers are not limited to the use of only one monomer type in an oligomer. The second problem can thus be solved by the use of heterogenous backbones that contain segments coding for different secondary structures. For example, the combination of *de novo* designed aromatic helical parts that display a hydrogen bonding interface on one side of the helix with a rigid turn-unit led to very stable helix bundles, albeit only in apolar organic solvents.[42, 70] Another possibility is to exchange specific α-amino acid units in an existing tertiary folded protein by non-natural moieties that have a similar structure and should thus retain the overall fold. Using this approach, about 20% of the zinc finger domain SP1-3 was replaced by non-natural units, which could even increase the overall folding stability of the tertiary structure.[71] However, the majority of the molecule still consists of α-amino acids. A fully artificial protein constructed in this fashion has not been reported yet.

A different strategy towards higher-order folding is the design of secondary structure motives capable of self-assembly. Once a tight binding is established, strands may be connected at a later stage to generate true tertiary or quaternary structures. Gellman and coworkers replaced some of the residues in GCN4 (which forms a dimeric parallel coiled coil in its native form) by β-amino acids. The resulting chimeric foldamers were still able to aggregate; however, the changes resulted in trimers and tetramers being formed.[72] At the same time, the Schepartz group reported a *de novo* designed 12-mer entirely consisting of β-amino acids that assembles into stable octameric helix bundles.[7] Similar to coiled coils, aggregation is driven by an array of hydrophobic side chains positioned on one side of the 3_{14}-helix formed by the β-peptide. More recently the concept could be extended to aliphatic oligo-ureas, which usually form helical structures with 2.5 units per turn resulting in five distinct side chain arrays at the side of the helix (note the similarity to oligo-quinolines). The β-urea 11-mer produced in the Guichard laboratory possesses two adjacent arrays of hydrophobic side chains flanked by charges residues promoting the self-assembly into a hexameric helix bundle (Fig. 8a).[73]

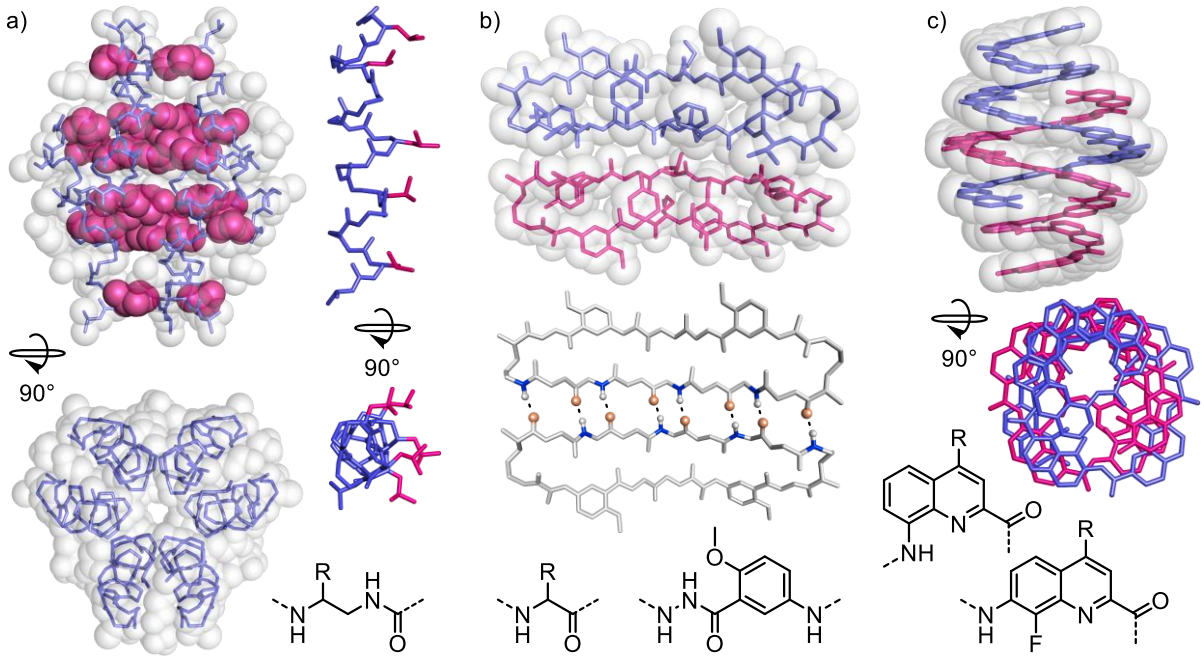


Fig. 8 – Foldamer self-assembly in aqueous media. a) Crystal structure of a hexameric β-urea helix bundle in top and side view with a transparent overlay of the CPK representation and a single urea chain in top and side view.[73] Hydrophobic side chains are highlighted in purple (for the bundle only in the side view in CPK representation). b) Crystal structure of an aromatic foldamer α-peptide hybrid macrocycle dimer.[74] Top: individual molecules are shown in different colors with a transparent overlay of the CPK representation. Bottom: amide hydrogens and oxygens involved in hydrogen bonding are shown as balls (oxygen: orange, nitrogen: blue), and hydrogen bonds are indicated by dashed lines. c) Model of an aromatic oligoamide double helix in top and side view (with a transparent overlay of the CPK structure; software: Maestro, forcefield: OPLS3, method: TNCG, solvent: water).[75] Individual molecules are shown in different colors. In all structures irrelevant side chains and hydrogen atoms are omitted for clarity. Chemical formulas of the building blocks are shown next to the respective structures.

The shape, size, and side chain positioning of aromatic foldamers is more distant from that of α-amino acids, so concepts found in natural peptides are not easily transferred. There are no reports of bundles composed of aromatic helices in aqueous solution so far, only in organic solvent[42] or in the solid state.[58] Other modes of aggregation have been achieved though, such as the dimerization of β-sheet peptides containing rigidifying aromatic segments (Fig. 8b),[74] and the sequence specific association of aromatic strands through complementary hydrogen bonding.[76] However, the most common interaction between aromatic foldamers is stacking, as seen in many crystal structures of these compounds.[64, 77] Different types of aggregation promoted by aromatic stacking have been accomplished, one being multiple helix formation.[75, 78] Fig. 8c shows a hybrid sequence with narrow caps and a wider helical part that is able to intercalate generating double helices. Another type is the end-to-end stacking of helices. Along these lines, Zeng and coworkers used “sticky” end-groups to drive the aggregation of helical pyridine oligoamides. The weak hydrogen bonding interaction between an ester function and aromatic protons was enough to polymerize the helices in aqueous media generating channels capable of transporting water molecules.[79] Ultimately, the strategies to aggregate abiotic foldamer molecules in water are still limited though. This thesis describes a non-polymeric, discrete end-to-end dimerization of oligo-quinoline helices (chapter 6).

3 Guiding objective: design of helix bundles

The development of water soluble helical aromatic foldamers capable of forming strong, discrete side-to-side aggregates served as a guiding objective throughout the research for this thesis. Thus far, helix bundling of foldamer helices in aqueous media has been achieved using β-peptides and β-ureas, but not with aromatic oligomers (see section 2.4). Advances in this direction using quinoline-based foldamers were made by Xiaobo Hu (a Huc group PhD alumnus). In an attempt to replicate coiled coil architectures seen for α-peptide helices, he focused mainly on designs utilizing hydrophobic effects between aliphatic side chains in combination with flanking salt bridges (Fig. 9a). Since there are only five distinct side chain positions in Q-helices when monosubstituted units are used, only a single array of hydrophobic groups on one side of the helix was introduced in most cases. However, no solid evidence for a discrete aggregation in solution was obtained during his investigations.

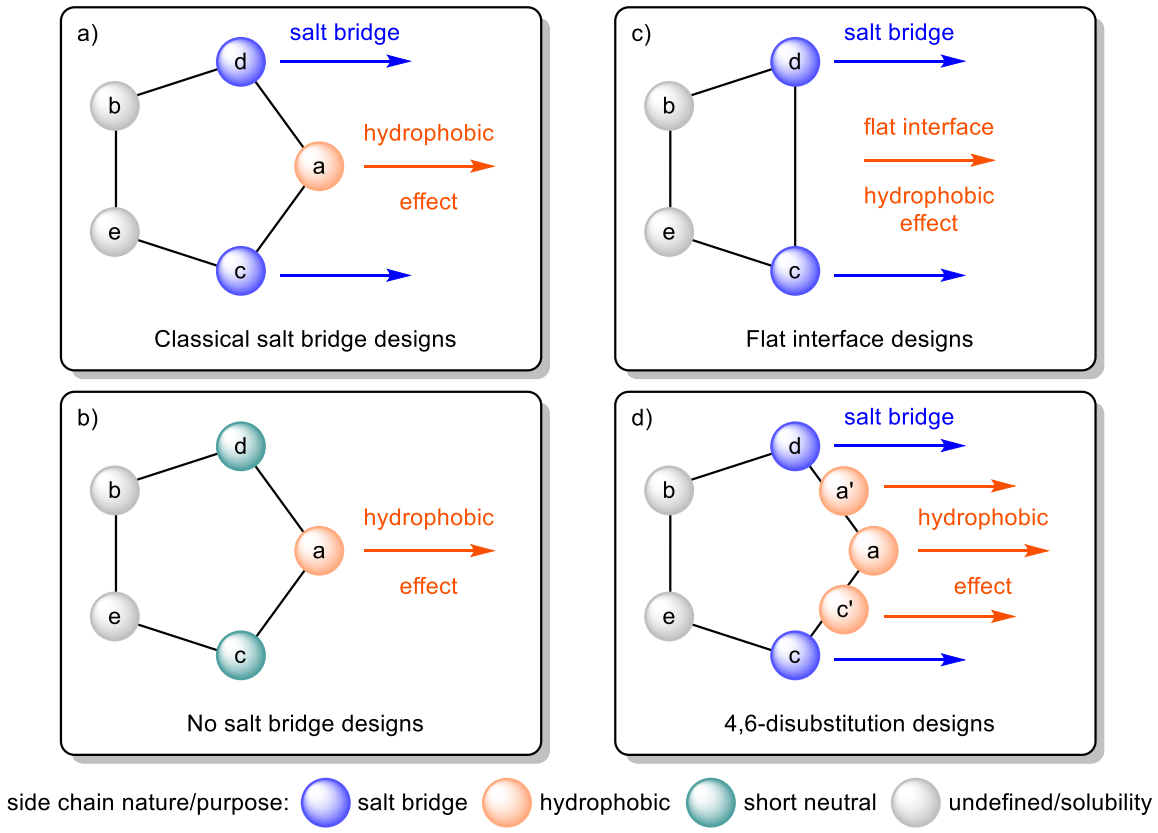


Fig. 9 – Design strategies for side-to-side self-assembly of helical foldamers. a) One array of hydrophobic side chains flanked by salt bridges. b) Aggregation only driven by hydrophobic effects. c) Hydrophobic core is formed by flat surfaces (helix backbone). d) 4,6-Disubstituted monomers are used to increase hydrophobic effects. The nature and purpose of each side chain position is indicated by different colors (color code is given at the bottom).

Considering these previous results, new strategies to obtain side-to-side aggregations were considered (Fig. 9b–d). Since salt bridges appeared to be difficult to design (empiric observations from previously obtained crystal structures), one approach to circumvent this is to only utilize

hydrophobic effects. Another possibility is the introduction of additional hydrophobic side chains at the interface by using disubstituted Q-units. Thereby, hydrophobicity can be increased without having to elongate the helix. A third approach is to create a "flat" surface on one side of the helix, which should act as a binding interface. The advantage in this case is the easier to design shape complementarity as opposed to using hard to predict aliphatic side chains at the interface. One way to create a flat surface is the use of P-monomers at every 5th position of the sequence, which was already utilized in the design of helix bundles in organic solvent.[42] However, we hoped to obtain a more favorable side chain arrangement for salt bridge formation when creating a flat interface with B-monomers (Fig. 10). Here the flanking side chains are closer together, which facilitates the design of dimeric helix bundles.

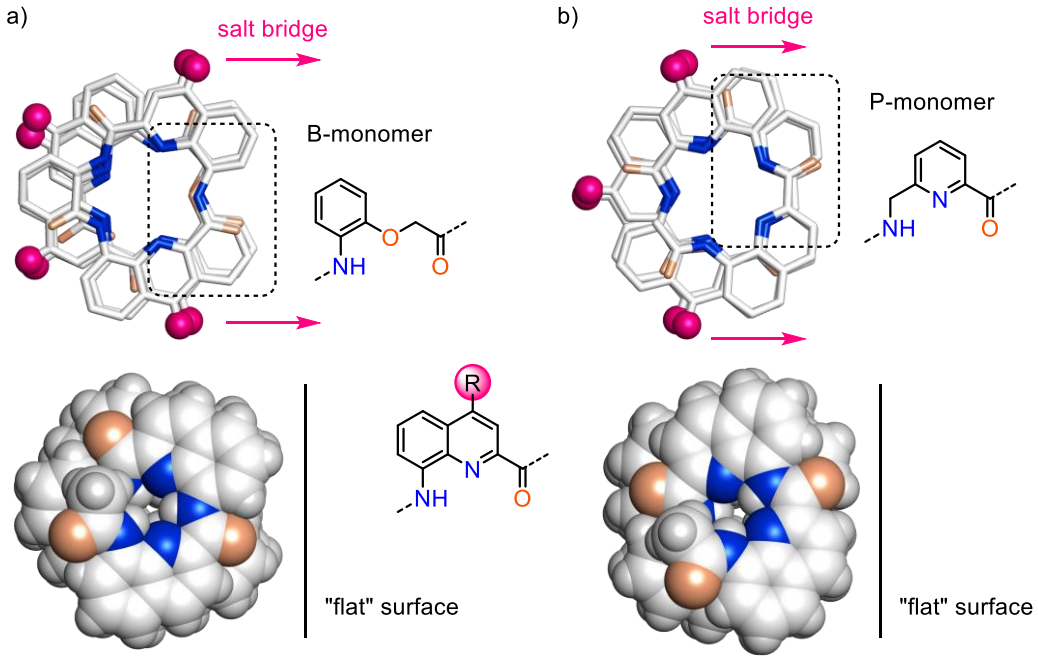


Fig. 10 – Aromatic oligoamide helices with a "flat" surface. At every 5th position of the sequence the Q-units are replaced by either B- (a) or P-monomers (b). Molecular models (Maestro, forcefield: OPLS3, method: TNCG, solvent: water) of the helices are shown in stick (without hydrogen atoms, first atoms of side chains are shown as purple balls) and CPK (without side chains) representation. Oxygen and nitrogen atoms are highlighted in orange and blue, respectively.

The typical workflow of this project was to take a given design approach and plan the exact sequence and thus side chain composition and positioning of a potential oligomer. This process was aided by simple molecular modeling tools (Maestro) that can confirm the feasibility of salt bridges and predict the likelihood of the multiplicity of the aggregates that could be formed (dimers, trimers, tetramers etc.). Promising candidates were synthesized via solid phase synthesis methods and purified via semi-prep. HPLC. After isolation, the products were subjected to mass spectrometry and ^{1}H NMR dilution studies, which give initial hints on the aggregation behavior of the compound. To obtain detailed structural information though, X-ray crystallographic data is necessary. This information is also critical to ultimately plan design improvements and close the iterative cycle of the project.

A multitude of oligomers have been synthesized in the course of this work. Unfortunately, in most cases, NMR dilution experiments revealed mediocre binding affinities and only one set of signals indicating fast exchange of aggregated species on the NMR timescale. This complicates the interpretation of the aggregation phenomena and doesn't provide information on how the binding is mediated without further structural data from X-ray crystallography. Therefore, most of these results are not described in detail in this thesis. However, other interesting findings emerged through these investigations that led to different separate projects: the development of a new monomer family (B-monomers) that was found to be easy to synthesize and compatible with Q-oligomers; the use of a chiral B-unit to control helix handedness of aromatic δ-amino acid foldamers; and the discovery of discrete head-to-head dimers of Q-helices whose binding is mediated by stacking of their C-terminal cross sections. The results from these three projects were published/submitted to peer reviewed journals and are presented in the next chapters. (Note that compound numbering is not continuous and starts again at the beginning of each chapter).

The present challenges and improvement ideas for obtaining strong, discrete side-to-side aggregations are discussed in more detail in chapter 7.

4 A new monomer family for aromatic δ-amino acid foldamers

Single Q-units may be replaced by other monomers (especially P) without disturbing the overall helical fold of Q-oligomers (as already described in the introduction chapter). Based on this knowledge, and with the ultimate goal to obtain side-to-side helix bundling, B-monomers were introduced to create a flat surface on one side of the helix by incorporating B at every 5th position in the sequence. Since little was known about the folding behavior of B-units, except that B_n-sequences fold into zig-zag-helices in organic solvents (Fig. 11a),[43a] the folding shape and stability of oligomers consisting of different combinations of Q-, B- and P-units was investigated first.

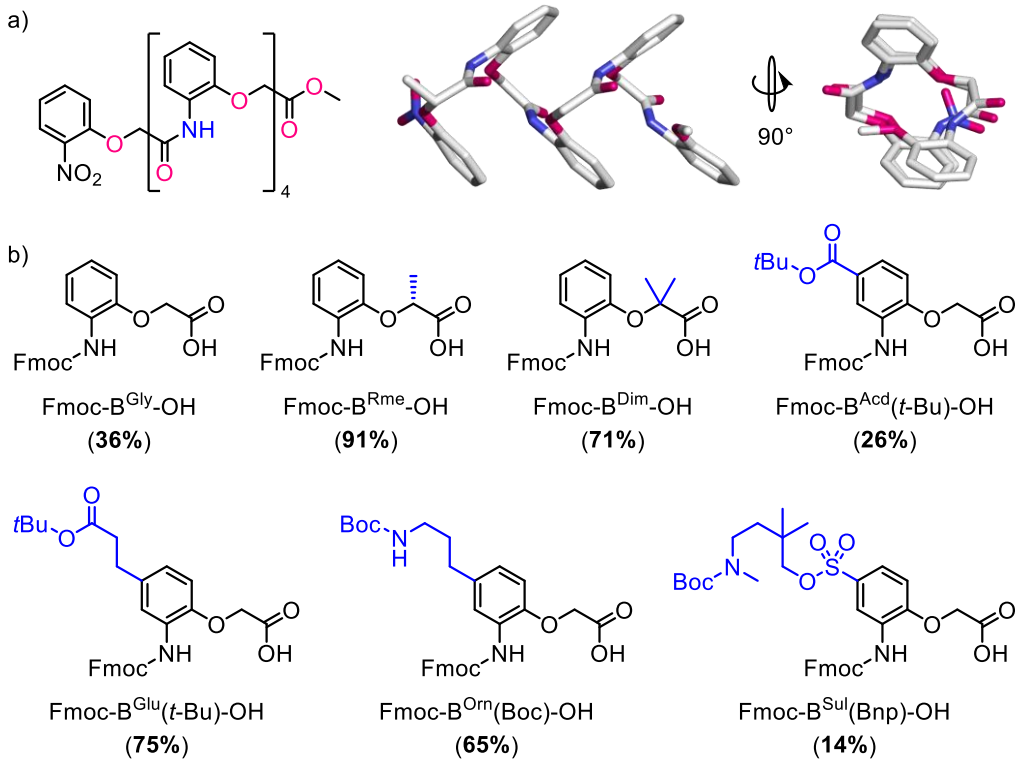


Fig. 11 – Structure and yields of B-monomers. a) Crystal structure of a B-pentamer (crystals grown from organic solvent).[43a] Nitrogen and oxygen atoms are highlighted in blue and purple, respectively. Hydrogen atoms are omitted for clarity. b) Formulas of B-units synthesized in the course of this thesis. Side chains are highlighted in blue, and the overall yields from commercial starting materials are shown.

Fmoc-protected versions of B-units were not described in the literature, thus a new synthetic route had to be developed to be able to utilize them in solid phase synthesis. After initial problems where spontaneous intramolecular cyclization of the amino acid intermediate was observed, a number of B-monomers bearing different substituents either on the 4-position of the aromatic ring or the aliphatic carbon next to the phenyl oxygen could be prepared. When typically starting from cheap commercially available precursors, the synthetic steps were often simple and high yielding, leading

to excellent overall yields for most B-units synthesized in the course of this work (Fig. 11b; for more information on B^{Sul} see section 7.2 and 8.2).

B^{Acd} was preferred for folding investigations in water since its short and charged side chain was expected to provide excellent solubility and crystal growth ability. Our findings, which are summarized in a manuscript that has been submitted to *Chemistry—A European Journal*, confirmed the interchangeability of Q- and B-monomers and demonstrated that two Q-units in a 12-mer are enough to template the canonical helical fold. In contrast, oligomers consisting only of B-monomers do not fold into well-defined structures in aqueous media. Most surprisingly, we also found that $(PB)_n$ does form a canonical helix, despite the lack of Q-units and the long aliphatic part around the B-P-junction. The work expands the basis of backbone-designs for aromatic *δ*-amino acid foldamers allowing a fine-tuning of helix stability and side chain positioning.

Contributions: The project was planned in collaboration with I. Huc. Synthetic routes and conditions for B-monomers were developed by me. B^{Acd} was developed with the help of L. Finger, an internship-student, and was later resynthesized with the help of D. Staudacher, who worked on her Bachelor thesis with me at that time. Oligomer synthesis and analysis was conducted with the help of D. Staudacher. X-ray structure refinement was performed by P. K. Mandal, who also did the X-ray data collection at EMBL Hamburg with the help of Dr. David von Stetten and Dr. Gleb Bourenkov. L. Allmendinger was involved in the design and execution of NMR experiments. The manuscript was written by me in collaboration with I. Huc and P. K. Mandal.

Publication (submitted)

Generalizing the aromatic δ-amino acid foldamer helix

Authors: Daniel Bindl, Pradeep K. Mandal and Ivan Huc

Submitted: *Chem. Eur. J.* **2022**

Abstract: A series of aromatic oligoamide foldamer sequences containing different proportions of three δ-amino acids derived from quinoline, pyridine, and benzene and possessing varying flexibility, *e.g.* due to methylene bridges, were synthesized. Crystallographic structures of two key sequences and ^{1}H NMR data in water concur to show that a canonical aromatic helix fold prevails in almost all cases and that helix stability critically depends on the ratio between rigid and flexible units. Notwithstanding subtle variations of curvature, *i.e.* the numbers of units per turn, the aromatic δ-peptide helix is therefore shown to be general and tolerant of a great number of sp^3 centers. We also demonstrate canonical helical folding upon alternating two monomers that do not promote folding when taken separately: folding occurs with two methylenes between every other unit, not with one methylene between every unit. These findings highlight that a fine-tuning of helix handedness inversion kinetics, curvature, and side chain positioning in aromatic δ-peptidic foldamers can be realized by systematically combining different yet compatible δ-amino acids.

Folding is an essential mechanism used by nature to achieve a wide variety of sophisticated, structurally precise architectures and functions, in particular in proteins.[80] Foldamers are artificial molecules based on non-natural backbones or sequences that also have the ability to fold in a well-defined way, with helical and sheet-like structures being prevalent motifs.[81] Canonical aromatic helices have become a major class of foldamer secondary motifs. They comprise main chain aryl rings that stack upon helix folding, therefore defining a vertical rise per turn, or pitch, equal to the thickness of an aromatic unit (ca. 3.5 Å). A wide range of backbone types adopt aromatic helical conformations, including phenylene-ethynylenes,[12] aryl-aryl linked backbones,[8, 82] aryl hydrazones,[83] aryl ureas,[84] and aryl amides.[20, 60, 85] Depending on the backbone, different forces may contribute to aromatic helix stability, most notably hydrogen bonds and electrostatic repulsions between contiguous units, as well as interactions associated with aromatic stacking such as solvophobic effects. Main chain aromatic units endow aromatic helices with distinct properties and functionalities compared to α-peptidic structures and their homologs,[5-6, 11a-c, 86] *e.g.* structural predictability,[87] generally high conformational stability, photophysical properties,[62, 88] charge transport ability[89] or the possibility to create a cavity within a single helical conformation for guest encapsulation.[90]

Aromatic amide foldamers are of particular interest because of their easy synthetic access and amenability to solid-phase synthesis. Oligoamides of δ-amino acid 8-aminoquinoline-2-carboxylic acid (Q) fold into stable aromatic helices in essentially all solvents (Fig. 12).[91] In this backbone, amide

and aryl groups exclusively consist of sp^2 centers, enabling a high degree of π-stacking in the Q_n helix. Intramolecular hydrogen bonding pinches the main chain leading to a high curvature of five units per two turns. Therefore, when looking down the helix axis, the helix wheel looks like a five-pointed star (Fig. 12c). Positions 4, 5, and 6 of Q may be functionalized by side chains that will diverge from the Q_n helix for the purpose of molecular recognition at the helix surface.[58, 64, 92] The question that we addressed in the present study is the generality of the folding pattern of Q_n oligomers. Other than Q, aromatic δ-amino acids 6-(aminomethyl)pyridine-2-carboxylic acid (P)[26a, 33a] and 2-(2-aminophenoxy)acetic acid (B)[43a, 93] have been introduced (Fig. 13a). P and B correspond to a Q in which the benzene or the pyridine ring has been deleted, respectively, while preserving a hydrogen bond acceptor as a main chain atom, namely the endocyclic nitrogen atom of P and the ether oxygen atom of B. P and B each possess a reduced surface for aromatic stacking, thus reduced hydrophobicity, as well as one sp^3 center and additional rotatable bonds conducive of enhanced flexibility. In the following, we generalize the compatibility of B, P, and Q units to form canonical aromatic helices in water.

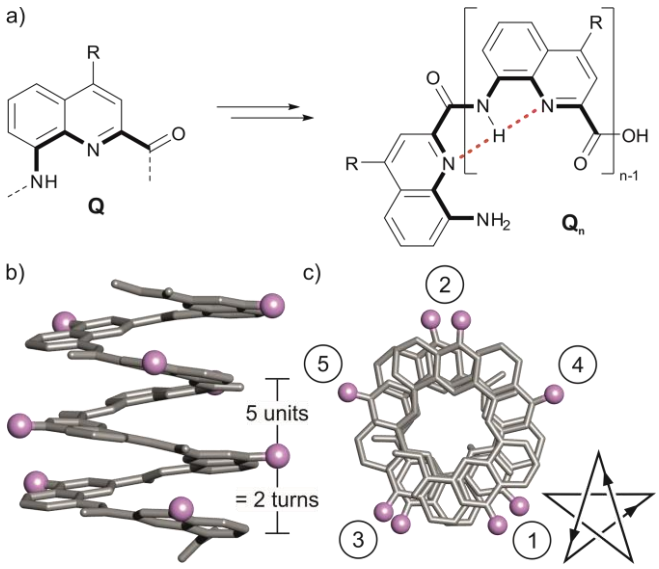


Fig. 12 a) Structure of Q and its oligomer. Hydrogen bonds on the inner rim of the helix are indicated as red dashed lines. Crystal structure of a Q_8 oligomer,[23] representative of the canonical aromatic helix fold, in side (b) and top (c) view. Only the first atom of R side chains is shown as a purple ball to indicate their positions, and hydrogen atoms are omitted for clarity.

We show that a single Q may significantly enhance helix stability. We also report an intriguing sequence dependence, in that a $(PB)_n$ oligomer folds in the absence of any Q unit, whereas P-only or B-only sequences do not.

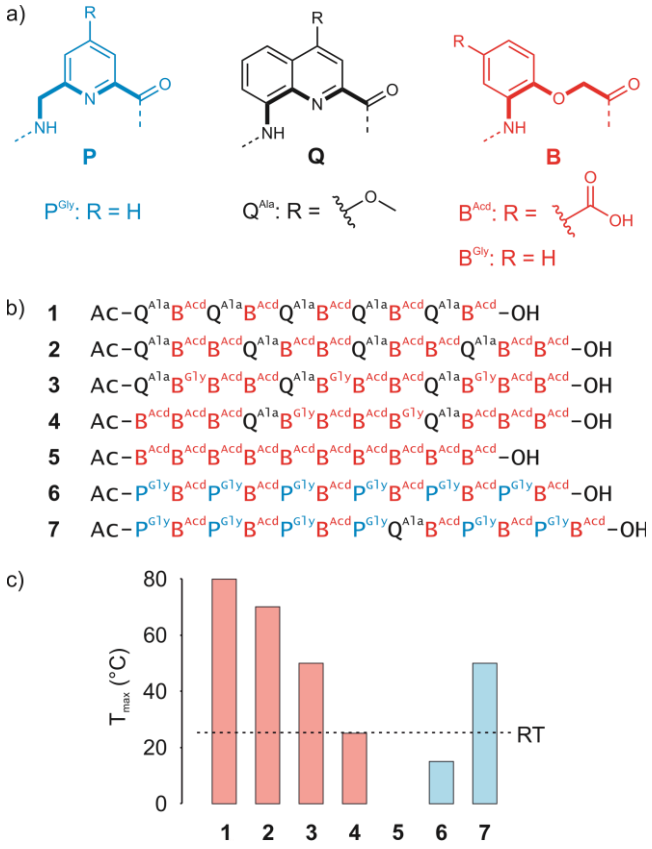


Fig. 13 a) Structures of the building blocks used in this study. The different types of monomer structures and letter codes are indicated by different colors. b) Sequences that were investigated in this study. c) Comparison of T_{max} of diastereotopicity (highest temperature, at which diastereotopic signals were observed for methylene groups by ^{1}H NMR) for **1**–**7**. Sequences without P are represented by red bars, while sequences with P are shown in blue. Room temperature (RT = 25 °C) is indicated with a dashed line.

The additional flexibility of P and B was previously shown to be detrimental to canonical aromatic helix folding. Specifically, the main chain methylene groups were found to promote 90° kinks. These kinks may be disruptive of folding – P_n oligomers fold neither in chloroform[33a] nor in water.[26a] They may also promote different, non-canonical, folding modes – a B_n oligomer formed a so-called herringbone helix in the solid state.[43a] The more rigid Q units have been shown to template the formation of canonical aromatic helices containing otherwise flexible monomers.[25, 94] However, $(PQ)_n$ was also found to fold into a herringbone helix in chloroform,[33a] and to only adopt a canonical aromatic helix conformation in water.[26a] We thus began this investigation with the expectation that a large proportion of Q units may be necessary to stabilize aromatic helices, even in water. The following demonstrates that this is not the case. Instead, we show that B, P, and Q units can be generally mixed without losing the canonical aromatic helix fold in water.

Short side chains were used to preserve the naturally high crystal growth ability of aromatic foldamers.[64] Q^{Ala} was chosen due to its easy synthetic accessibility, however, being non-polar, a solubilizing side chain was needed on the B units. New monomer B^{Acd} was prepared to promote solubility in neutral or basic aqueous medium. Starting from 4-hydroxy-3-nitrobenzaldehyde, a carboxylic acid was obtained through oxidation of the aldehyde and subsequently protected as a *tert*-butyl

ester. The general route to Fmoc-protected B monomers described earlier could then be applied (Fig. 16),[93] and the final monomer was obtained in 26% overall yield over 5 steps. Oligomers **1**–**7** comprised of P^{Gly}, Q^{Ala}, B^{Gly}, and B^{Acd} monomers were then synthesized using low loading Wang or Cl-MPA ProTide™ resins (the latter is advantageous for the synthesis of sequences prone to aggregation) according to solid-phase foldamer synthesis (SPFS) protocols reported previously.[26] Final products were purified by semi-preparative reversed-phase HPLC and were obtained in good overall yields (6.3–31%; section 4.2.2).

Indications of helical folding can easily be obtained by simply measuring ^{1}H NMR spectra and sequences were analyzed in this regard. Typically, ring current effects associated with stacking within canonical aromatic helices cause upfield shifts that greatly vary with the central or peripheral position of the protons considered. It follows that NMR signals of a helically folded oligomer spread over a wide range of chemical shift values even when the sequence has a repetitive nature.[91]

Furthermore, downfield shifted amide signals indicate their involvement in intramolecular hydrogen bonding. Helix stability may also be assessed by ^{1}H NMR through the observation of helix handedness inversion kinetics. Since helices are chiral objects (present as a racemic mixture when the backbone contains no stereogenic center), methylene protons on the molecule are diastereotopic. As the environments above and below a given monomer differ, methylene protons that are close to the helix backbone often become anisochronous, *i.e.* have different ^{1}H NMR chemical values and appear as doublets notwithstanding other couplings. This holds only when *P*- and *M*-helical conformers exchange slowly on the NMR timescale. When the exchange is fast, averaged signals are observed as an indication of faster kinetics and weaker helix stability (yet not as an indication of helix unfolding). In Q_n oligomers, helix handedness kinetics vary with solvent polarity and with the value of n. Exceptionally long half-lives and even kinetic inertness have been reported.[38] Heating enhances handedness inversion dynamics.[91]

With our first design, oligomer **1**, we aimed to test whether the templating effect observed for $(PQ)_n$ sequences in water[26a] is transferable to $(BQ)_n$ sequences, *i.e.* the ability of Q units to impose aromatic helix folding to more flexible B monomers. The ^{1}H NMR spectrum of **1** in aqueous solution has all the characteristics of an aromatic helix structure mentioned above (Fig. 14). Diastereotopic signals (doublets) of the methylene groups of its B monomers in the area of 1.5–4.5 ppm indicate folding into a chiral conformation, of which the enantiomers exchange slowly on the NMR timescale. Upon heating to 80 °C (higher temperatures were not tested), anisochronicity remained indicating that kinetics of handedness inversion remained slow (Fig. 17) and suggesting a very stable structure. An X-ray structure of **1** was obtained and confirmed the canonical aromatic helix fold (Fig. 15a,b). The structure is similar to that of Q_n oligomers except for its curvature. In **1**, 2.33 units are needed to span one helix turn (instead of 2.5 in Q_n), and side chains are ordered in seven distinct arrays when looking down the helix axis instead of five (2.33 units per turn means 7 units for 3 turns). Since the B units contain sp^3 hybridized centers, some bond angles may differ from those of an all sp^2 backbone and thus alter the overall curvature as well as side chain positions. However,

one may object that such an effect was not observed upon mixing Q and P even though the latter also contain methylene groups.[25] Nevertheless, it can be concluded that Q units template the folding of B monomers as they do for P monomers.

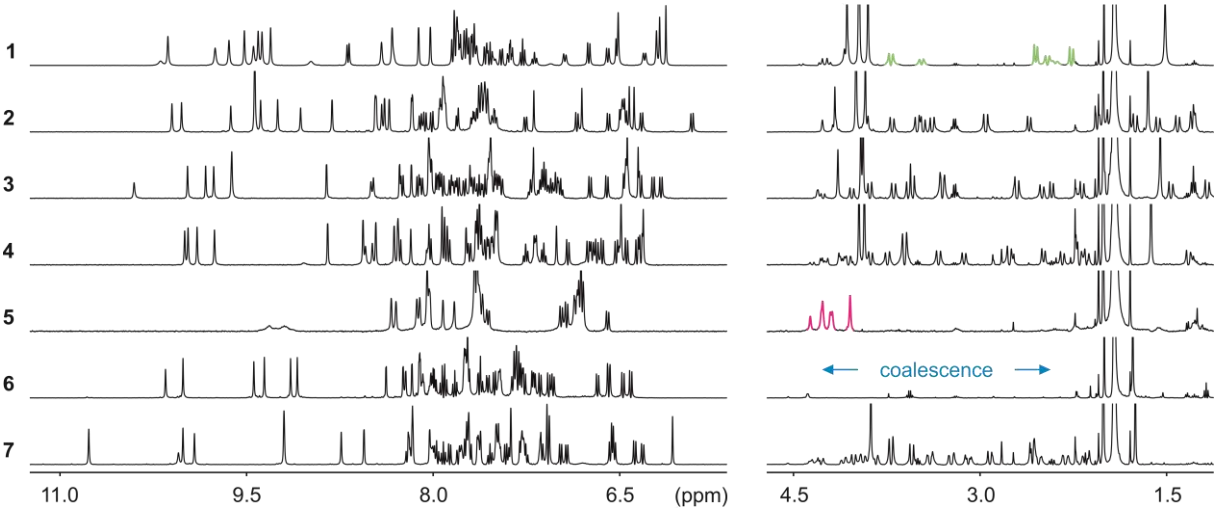


Fig. 14 ^{1}H NMR spectra of compounds **1**–**7** (500 MHz, 12 mM NH_4OAc buffer pH 8.5 with 10% D_2O, water suppression). Anisochronous signals of diastereotopic methylene protons are highlighted in green exemplary in the spectrum of **1**. Singlets of methylene groups are highlighted in purple in the spectrum of **5**.

The high stability of $(BQ)_n$ encouraged the investigation of "B_nQ_m" sequences, that is, oligomers containing n B and m Q units mixed throughout the sequence with a higher B/Q ratio, to see how many Q units are necessary to keep the oligomers well-folded. Sequences **2**–**4** were thus prepared (Fig. 13b). In these sequences, Q monomers were evenly distributed to ensure a homogenous folding behavior. To prevent oligomers from carrying too many negative charges, which might bias the investigation by introducing possibly destabilizing electrostatic repulsions, B^{Acd} was replaced by B^{Gly} in some places. Remarkably, the ^{1}H NMR spectra of all sequences show an extensive chemical shift distribution of signals and slow exchange between *P*- and *M*-conformers (Fig. 14). The similarity of these NMR spectra with that of **1** suggests that all fold into canonical aromatic helices in aqueous solution at room temperature. At least, it can be concluded that the templating effect of Q induces a well-folded structure in "B_nQ_m" oligomers, with an n/m ratio of up to 5/1. To estimate the stability of **2**–**4** in comparison to **1**, variable temperature studies were conducted (Fig. 17–Fig. 20). Oligomers **2**–**4** displayed coalescence of the diastereotopic signals in their ^{1}H NMR spectra at 80, 60, and 40 °C, respectively, which demonstrates a trend of decreasing stability when increasing the B/Q ratio (Fig. 13c). It is a significant observation that only two Q units in a 12-mer are sufficient to generate a stable well-folded structure.

The synthesis of an oligomer containing only one Q and multiple B units was attempted repeatedly. For unclear reasons, these attempts were not successful. One might invoke the conformational behavior of B-rich sequences in organic solvents during SPFS, *i.e.* possible aggregation. Yet, the all-B sequence **5** was successfully prepared. Its ^{1}H NMR spectrum showed clustering of signals in a few groups, implying reduced aromatic stacking (Fig. 14). Some amide NH signals are broad and not

downfield shifted and others have completely disappeared. This disappearance may reflect higher exposure of these protons to the solvent and exchange with water, which would result in some signal suppression. Such a phenomenon would clearly indicate a loss of the aromatic helical structure in which amide protons are normally buried in the helix channel. Furthermore, sequence **5** does not display diastereotopic doublets for its methylene groups (singlets are observed) even at 0 °C (Fig. 13c, Fig. 21), meaning that its conformation is either achiral or that the exchange between different chiral folded states is fast on the NMR time scale. We couldn't obtain crystals of **5** and the exact nature of B_n oligomer conformations in aqueous solution remains unclear. It can only be concluded that at least some Q units are necessary to template B units into the canonical aromatic helix fold.

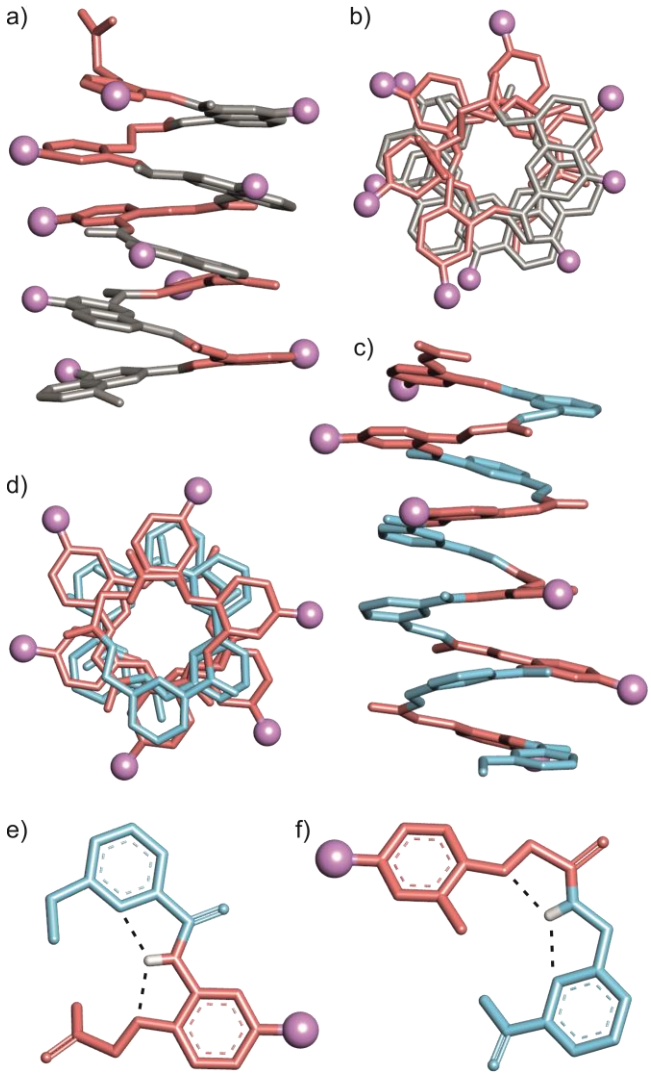


Fig. 15 Crystal structure of **1** in side (a) and top (b) view. Crystal structure of **6** in side (c) and top (d) view. Q, B, and P monomers are shown in gray, red, and blue, respectively. Only the first atom of side chains is shown as a purple ball to indicate their positions, and hydrogen atoms are omitted for clarity. e, f) Parts of the crystal structure of **6** illustrating the rigid part between P and B (e) and the extended flexible part between B and P (f). Amide hydrogen atoms are shown, and hydrogen bonding is indicated with dashed lines.

As stated above, oligomers only comprised of P units also do not fold into well-defined structures in aqueous media.[26a] However, these findings cannot be extrapolated to molecules of the type "B_nP_m", since they contain different flexible and rigid parts than either only B or P containing sequences. "B_nP_m" oligomers comprise an extended flat and rigid aromatic dimer when a B unit follows P and a long flexible -$OCH_2CONHCH_2$- linkage when P follows B (Fig. 15e,f). To investigate the effects of mixing B and P monomers on folding behavior, sequence **6** was designed, which consists of an alternation of B and P units. Like compounds **1**–**4**, oligomer **6** shows again widespread ^{1}H NMR signals and downfield shifted amide signals (Fig. 14). This indicates both more extensive aromatic stacking and stronger intramolecular hydrogen bonding in **6** compared to **5**. In fact, the NMR spectrum of **6** resembles more those of Q-containing sequences **1**–**4** than that of **5**. However, it does not show diastereotopic signals in the aliphatic region. The methylene signals are very broad to invisible, suggesting coalescence is reached at room temperature. Upon cooling to 0 °C the signals split into doublets, demonstrating that well-defined folded states with slow exchange on the NMR timescale can be obtained with a "B_nP_m" sequence at lower temperatures (Fig. 13c, Fig. 22). An X-ray structure of **6** could be obtained, which confirmed a canonical aromatic helix fold, despite the lack of any Q units and the long flexible linkages between B and P monomers (Fig. 15c–f). Again, a curvature change compared to Q_n oligomers can be observed. In the case of **6**, 2.4 units are necessary to propagate the helix by one turn (amounting to 12 units for 5 turns). Side chain positioning is altered as well.

Given that **6** is the least stable folded oligomer investigated in this study (lowest T_{max} of observable anisochronous methylene signals), its canonical fold in the solid state suggests that the aromatic helix also prevails in **2**–**4** for which no crystal structure could be obtained. These results provide evidence that the rigid PB dimers counteract the flexibility of the BP linkage. In other words, the aromatic helix is more tolerant towards larger aromatic units (PB) and more flexible BP linkages than to smaller aromatic units (P or B) separated by less flexible but twice more numerous linkages containing sp^3 centers. The stability of "B_nP_n" oligomers is thus dependent not only on n but also on the sequence. One can speculate that $(PB)_n$ should be the most stable and B_nP_n the least stable since it does not have any PB combination.

Finally, we aimed to test whether the stabilizing effect of Q units seen for "B_nQ_m" can be extended to "B_nP_m" oligomers. For this purpose, sequence **7** with an additional Q monomer compared to **6** was synthesized. The Q unit was placed in such a manner that it enlarges the flat, rigid part between a P and a B. The ^{1}H NMR signals of **7** expectedly indicate canonical folding. Methylene signals appear as doublets at room temperature (Fig. 14) and up to 60 °C (Fig. 13c, Fig. 23). This stability is comparable to that of "B_9Q_3" (*i.e.* **3**), which again illustrates the increased stability of $(PB)_n$ as opposed to B-only sequences: 1 vs. 3 Q units are needed for the same T_{max} of anisochronicity. It can be assumed that introducing additional Q units within "B_nP_m" sequences would further enhance stability as it did in **1**–**4**.

In conclusion, it was shown that B, P, and Q units are essentially interchangeable in helical aromatic δ-amino acid foldamers. "B_nQ_m" oligomers, with an n/m ratio of up to 5/1 still fold into canonical helices. Their stability decreases as n/m increases but appears to remain above that of most aliphatic peptidic helices for which slow exchange of helix handedness inversion on the NMR timescale has rarely been observed.[95] Sequences with alternating B and P units possess an increased folding propensity compared to only B- or P-containing oligomers, and fold into a canonical aromatic helix even without any Q unit within their sequence, albeit cooling was required for $(PB)_6$ to display slow handedness exchange on the NMR timescale. Significant stabilization of $(PB)_6$ was achieved through introducing a single Q monomer within the sequence, leading to an increase of T_{max} of anisochronicity by 35 °C. Furthermore, the introduction of B monomers into Q-oligomers not only influences helix stability, but it also influences curvature and side chain positioning of the resulting foldamers, as shown by two different X-ray structures. Therefore, we demonstrated that adjusting the ratio and/or order of B, P, and Q allows for a fine-tuning of helix stability and side chain positioning without changing the overall fold. This ability will be useful in the context of molecular recognition at the helix surface, *e.g*. for foldamer-protein interactions.

Acknowledgements: This work was supported by the DFG (Excellence Cluster 114, CIPSM). D. Gill is gratefully acknowledged for contributing synthetic precursors, and C. Glas for assistance with NMR measurements. We thank Daniela Staudacher for assistance with synthesis and crystallization experiments. Synchrotron data were collected at beamlines P13 and P14 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany). We thank Dr. David von Stetten and Dr. Gleb Bourenkov for their assistance in using the beamlines.

Conflict of Interest: The authors declare no conflict of interest.

Data Availability Statement: The crystallographic data that support the findings of this study are openly available from the Cambridge Crystallographic Data Centre at https://ccdc.cam.ac.uk, reference numbers 2125508 and 2125515.

Supplementary Information

for:

Generalizing the aromatic δ-amino acid foldamer helix

Daniel Bindl, Pradeep K. Mandal, Lars Allmendinger and Ivan Huc

List of Abbreviations

AcOH	acetic acid
CD	circular dichroism
CyHex	Cyclohexane
DCM	dichloromethane
DIAD	diisopropyl azodicarboxylate
DIPEA	*N*,*N*-diisopropylethylamine
DMAP	4-Dimethylaminopyridine
DMF	*N*,*N*-dimethylformamide
DMSO	dimethyl sulfoxide
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EI	electron ionization
ESI	electrospray ionization
EtOAc	ethyl acetate
Fmoc	fluorenylmethoxycarbonyl
HMBC	heteronuclear multiple bond correlation
HMQC	heteronuclear multiple quantum correlation
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
MeOH	methanol
MW	molecular weight
NMR	nuclear magnetic resonance
RP	reversed phase
RT	room temperature
SPFS	solid phase foldamer synthesis
***t*BuOH**	*tert*-butanol
TEA	triethylamine
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TIPS	triisopropyl silane
TLC	thin layer chromatography
TMSP	3-(trimethylsilyl)propionic-2,2,3,3-d4 acid sodium salt
UV/Vis	ultraviolet-visible

4.1 Supplementary figures

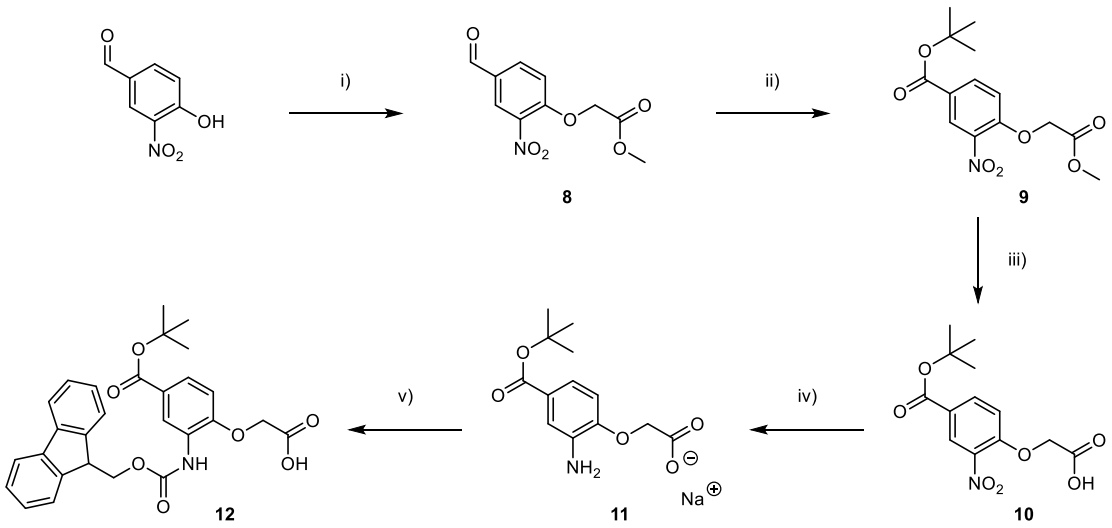


Fig. 16 Synthetic route to the Fmoc-B^{Acd}-OH monomer (**12**): i) methyl bromoacetate, K_2CO_3, acetonitrile, 70 °C (48%); ii) 1) Jones reagent, acetone, 2) *t*BuOH, EDC·HCl, DMAP, DMF (70%); iv) LiOH, H_2O, THF (quant.); v) H_2, Pd/C, Na_2CO_3, MeOH (quant.); vi) Fmoc-Cl, $NaHCO_3$, H_2O, 1,4-dioxane (76%). For detailed synthetic procedures see section 4.2.3.

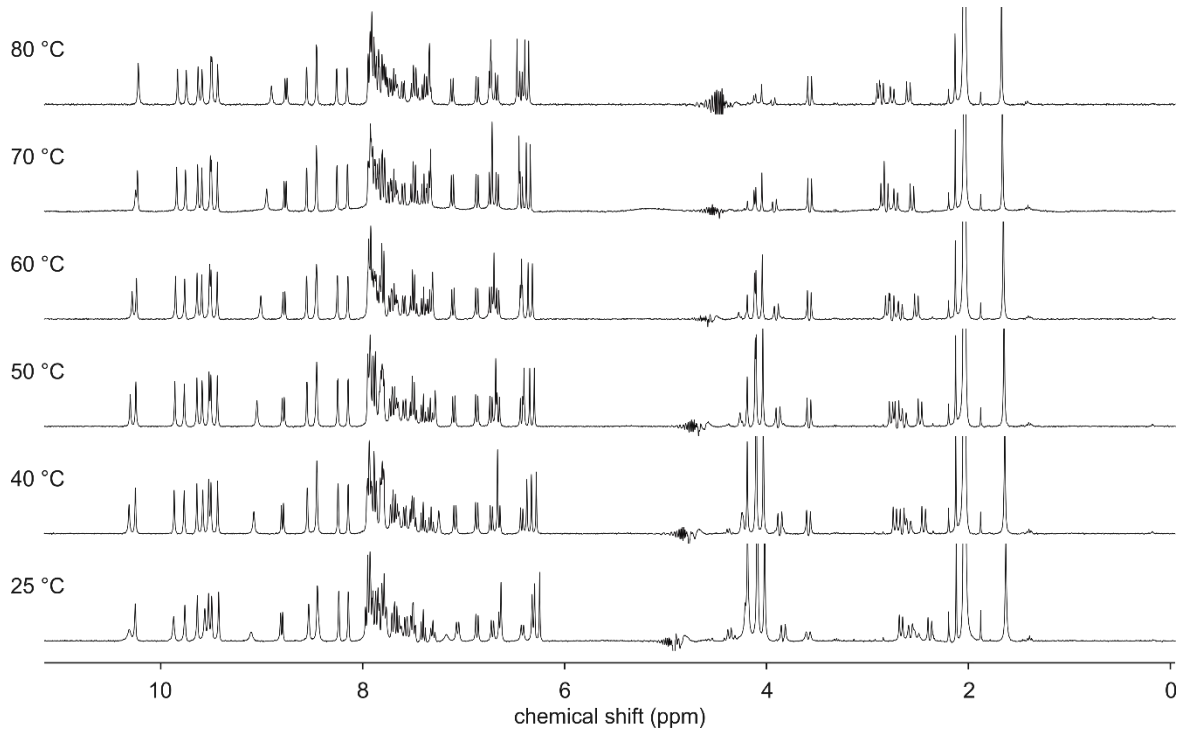


Fig. 17 Variable temperature ^{1}H NMR spectra of oligomer **1** (500 MHz, 0.6 mM in 12 mM NH_4OAc buffer pH 8.5 H_2O/D_2O 9:1, H_2O suppression).

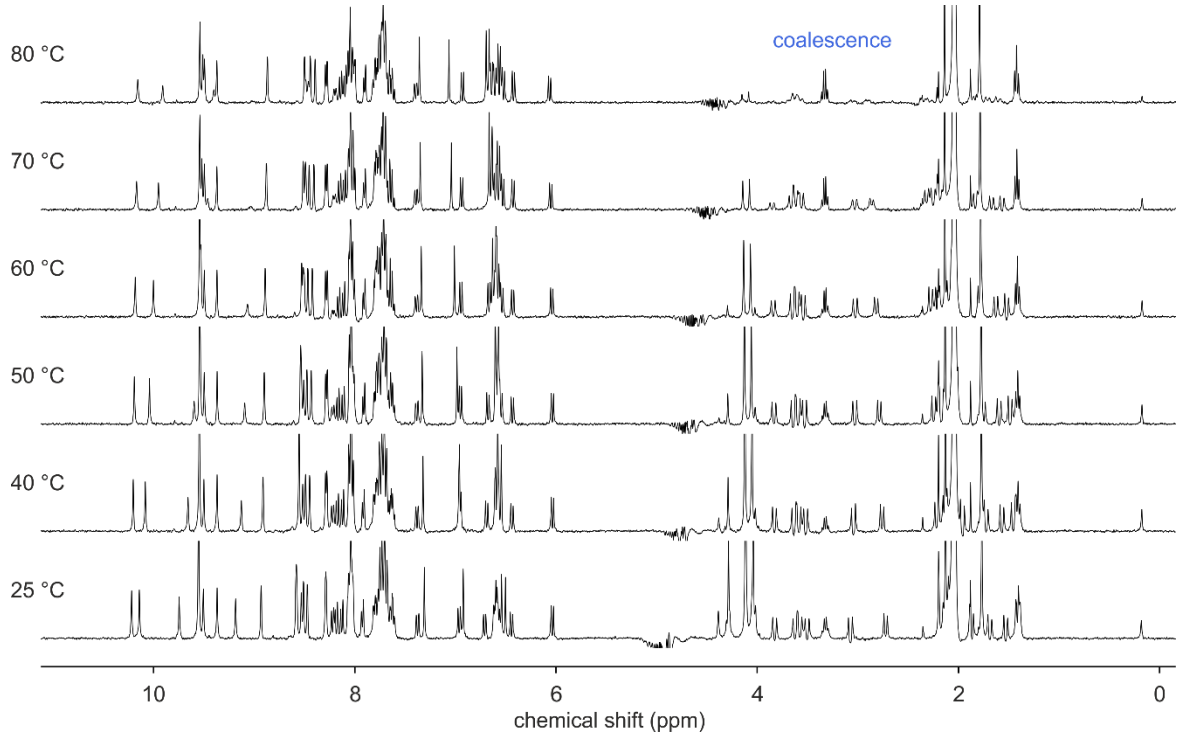


Fig. 18 Variable temperature ^{1}H NMR spectra of oligomer **2** (500 MHz, 0.26 mM in 12 mM NH_4OAc buffer pH 8.5 H_2O/D_2O 9:1, H_2O suppression).

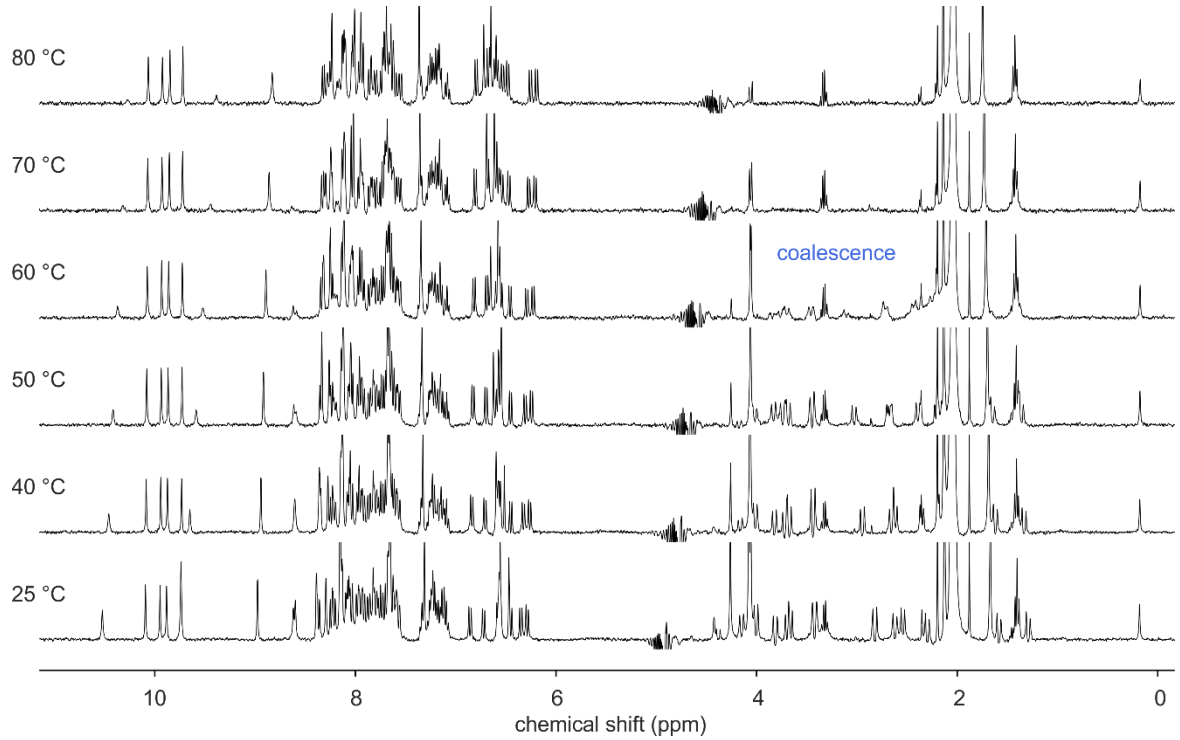


Fig. 19 Variable temperature ^{1}H NMR spectra of oligomer **3** (500 MHz, 0.16 mM in 12 mM NH_4OAc buffer pH 8.5 H_2O/D_2O 9:1, H_2O suppression).

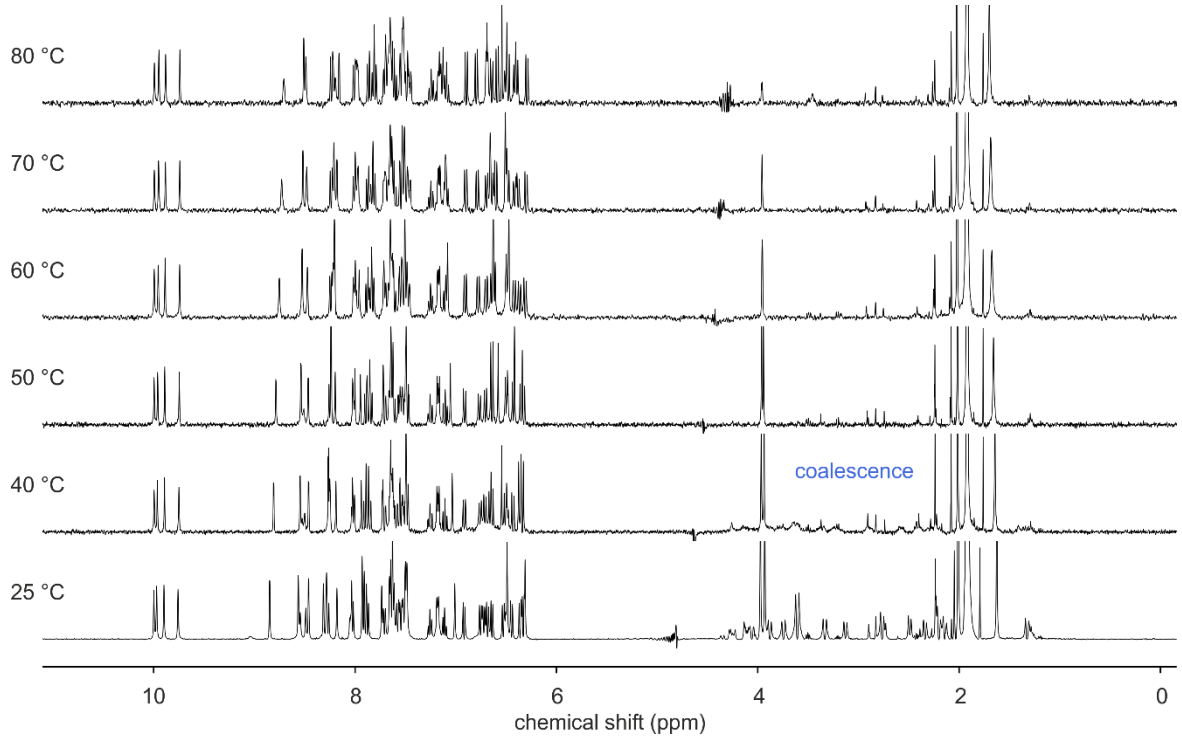


Fig. 20 Variable temperature ^{1}H NMR spectra of oligomer **4** (500 MHz, 0.18 mM in 12 mM NH_4OAc buffer pH 8.5 H_2O/D_2O 9:1, H_2O suppression).

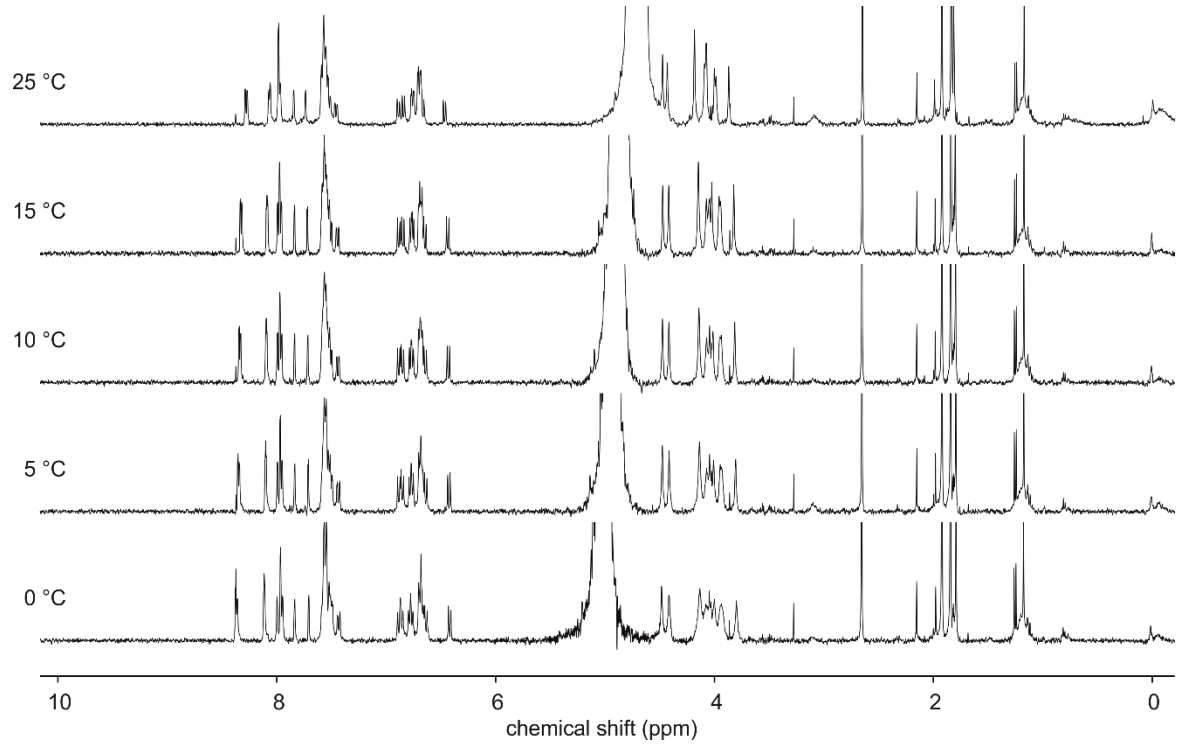


Fig. 21 Variable temperature ^{1}H NMR spectra of oligomer **5** (400 MHz, 0.6 mM in 60 mM ND_4 + 15 mM AcOH-d_4 in D_2O).

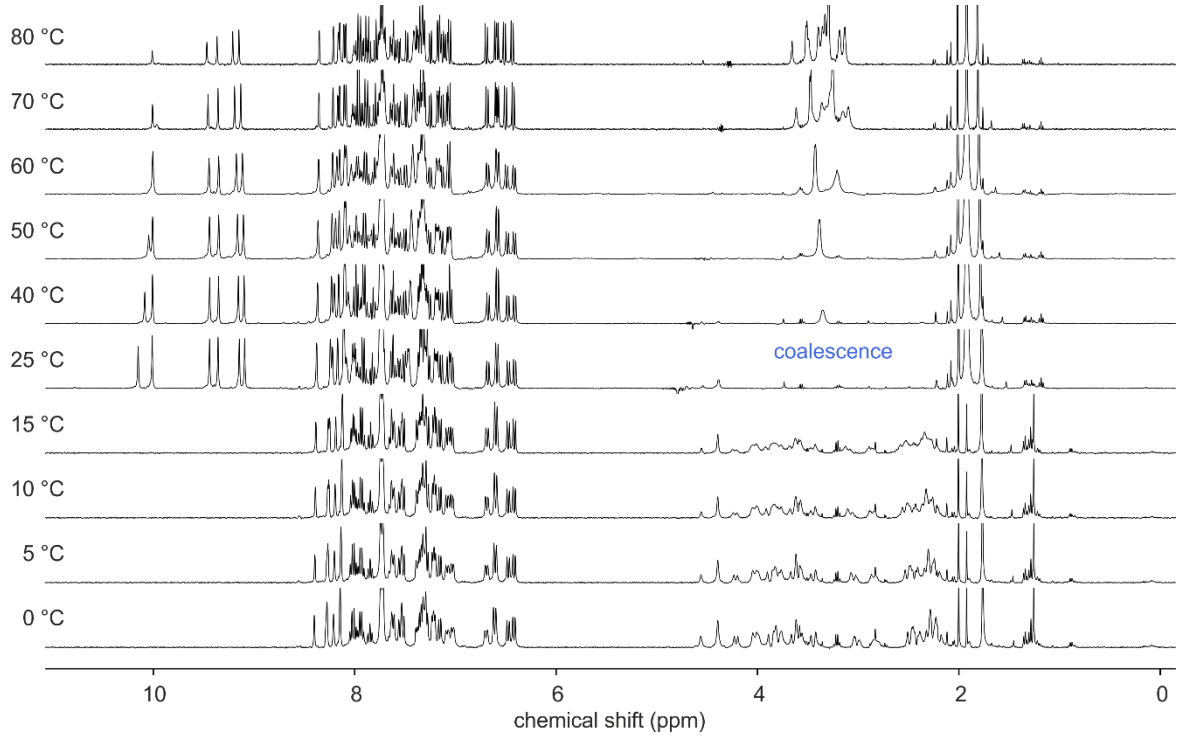


Fig. 22 Variable temperature ^{1}H NMR spectra of oligomer **6** (25–80 °C: 500 MHz, 0.66 mM in 12 mM NH_4OAc buffer pH 8.5 H_2O/D_2O 9:1, H_2O suppression; 0–15 °C: 400 MHz, 0.66 mM in 60 mM ND_4 + 15 mM AcOH-d_4 in D_2O).

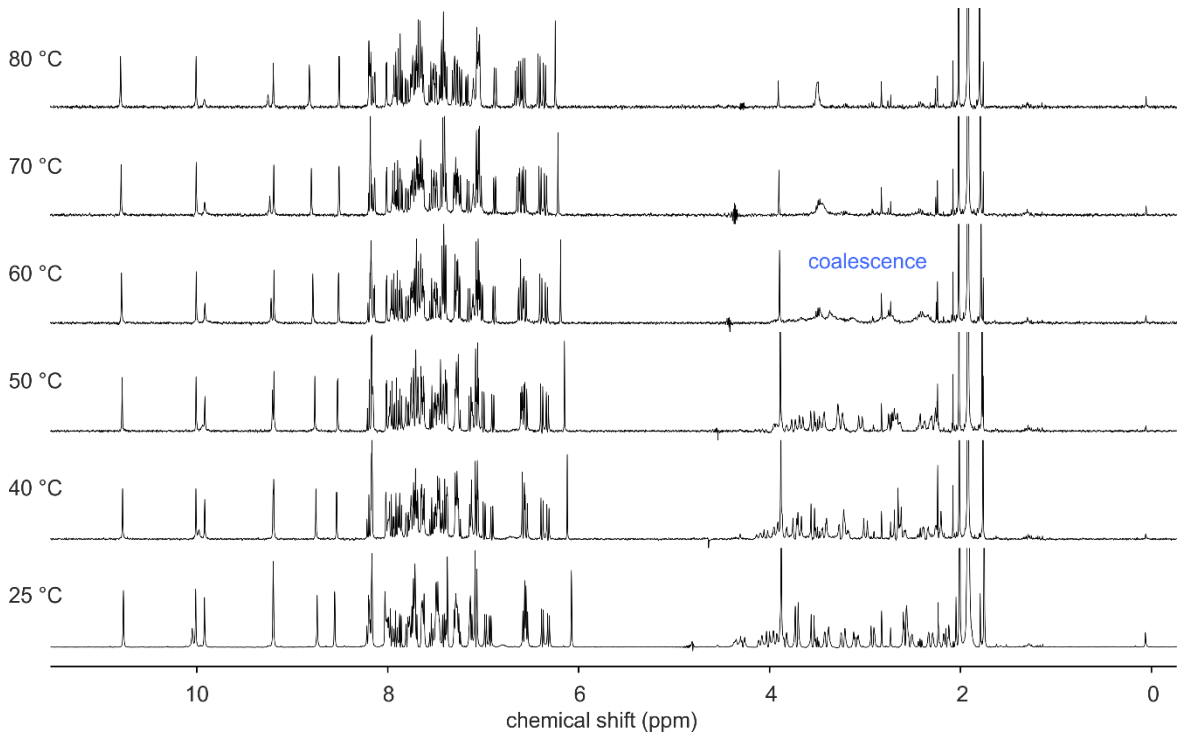


Fig. 23 Variable temperature ^{1}H NMR spectra of oligomer **7** (500 MHz, 0.36 mM in 12 mM NH_4OAc buffer pH 8.5 H_2O/D_2O 9:1, H_2O suppression).

4.2 Materials and Methods

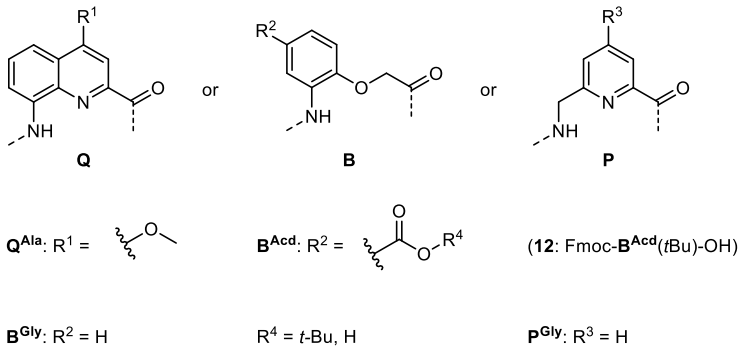


Fig. 24 Fmoc-acid building blocks used in this study. Fmoc-Q^{Ala}-OH,[96] Fmoc-B^{Gly}-OH[93] and Fmoc-P^{Gly}-OH[41] have been described previously. For a detailed procedure to Fmoc-B^{Acd}(*t*Bu)-OH, see section 4.2.3.

4.2.1 General

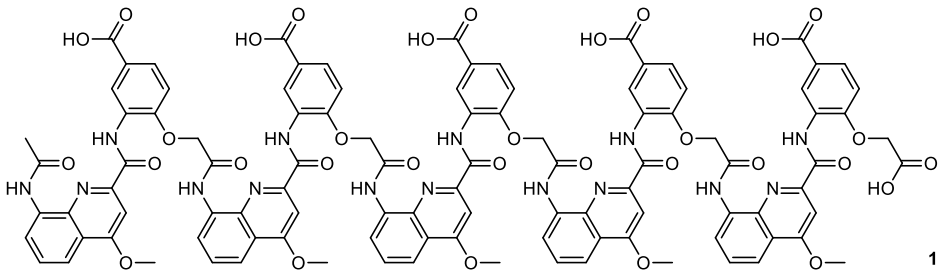
Commercial reagents (Suppliers: Abcr, Fisher Scientific, Merck, Sigma-Aldrich, TCI or VWR) were used without further purification unless otherwise stated. Wang resin LL (100–200 mesh) was purchased from Novabiochem, Cl-MPA ProTide™ resin LL was purchased from CEM. Peptide grade *N*,*N*-dimethylformamide (DMF) was purchased from Carlo Erba. Anhydrous chloroform, triethylamine (TEA) and *N*,*N*-diisopropylethylamine (DIPEA) 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 400MHz Bruker BioSpin spectrometer or an Avance III HD 500MHz Bruker BioSpin spectrometer equipped with a broad band observe 5-mm BB-H&FD CryProbe™ Prodigy. Measurements were performed at 25 °C unless stated otherwise. Water suppression was performed with excitation sculpting. 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.[97] 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). UV absorbance was monitored at 300 nm if not stated otherwise. Simple ultraviolet–visible (UV/Vis) absorbance measurements were done with a Thermo Fisher Scientific Nanodrop One instrument using a 1 cm quartz cuvette. Circular dichroism (CD) spectra were measured on a Jasco J-810 spectrometer. Measurements were performed at 20 °C if not stated otherwise. Manual microwave-assisted solid-phase foldamer synthesis (SPFS) was performed via a CEM® Discover Bio microwave peptide synthesizer. The temperature within the reactor vessel was monitored with an optical fiber probe. Automated SPFS was done via a Gyros Protein Technologies PurePep Chorus synthesizer with induction heating.

4.2.2 Solid phase synthesis procedures

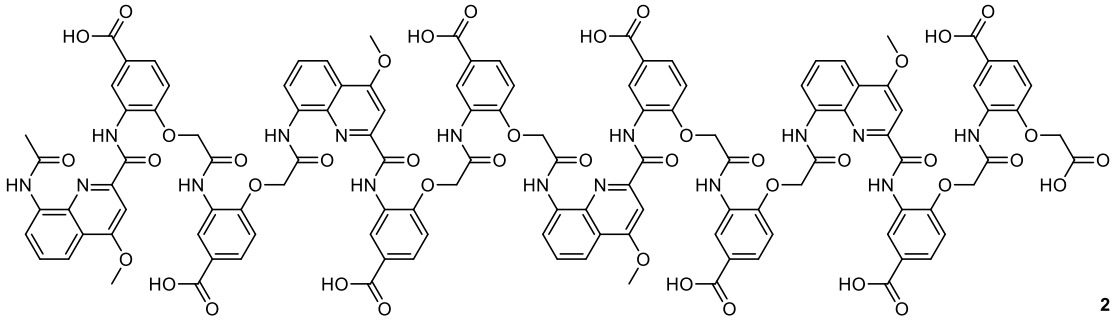
Oligomers were synthesized according to previously reported SPFS protocols,[26] hereafter referred to as standard method. Fmoc acid building blocks were activated *in situ* by generating the respective acid chlorides prior to coupling.

Acetylation: In the microwave vessel: after the resin (1.0 equiv.) was washed with anhydrous THF (4 ×), DIPEA (10.0 equiv.) and acetyl chloride (5.0 equiv.) in anhydrous THF (1 mL per 100 mg resin; not less than 2 mL) were added and the suspension was heated to 50 °C for 15 min (25 W, ramp to 50 °C over 5 min, hold at 50 °C for 15 min). The resin was washed with anhydrous THF (3×) and the coupling step was repeated once. Then, the resin was washed again with anhydrous THF (1×) and DMF (5×), and kept suspended in DMF (if stored longer than 24 h, it was kept at 4 °C).



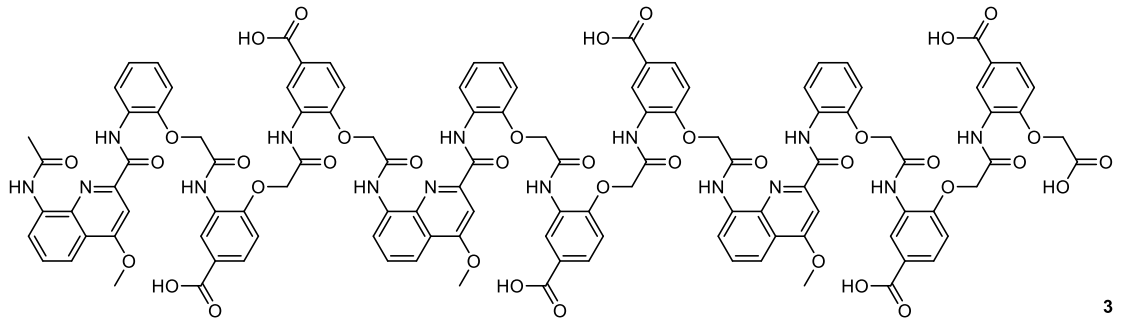
Compound 1: Oligomer **1** was synthesized on Wang resin (0.37 mmol g^{-1}, 27.8 µmol scale) according to the standard method (automated). Loading of the first monomer: 0.25 mmol g^{-1} (68%). The final acetyl group was installed via the general acetylation method. After purification by semi-prep HPLC (C18, 0–30B, 50 °C; A: 13 mmol NH_4OAc buffer pH 8.5, B: acetonitrile), the title compound was obtained as a white solid (13.8 mg, 6.8 µmol, 24%; HPLC-purity: 97.5%). **^{1}H NMR** (500 MHz, 12 mmol NH_4OAc buffer pH 8.5 in H_2O/D_2O 9:1): δ = 10.19 (s, 1H), 10.13 (s, 1H), 9.75

(s, 1H), 9.64 (s, 1H), 9.52 (s, 1H), 9.44 (s, 1H), 9.41 (s, 1H), 9.38 (s, 1H), 9.31 (s, 1H), 8.98 (s, 1H), 8.68 (d, *J* = 8.40 Hz, 1H), 8.42 (s, 1H), 8.33 (s, 1H), 8.12 (s, 1H), 8.02 (d, *J* = 2.01 Hz, 1H), 7.85 (s, 1H), 7.83 (s, 1H), 7.82 – 7.79 (m, 2H), 7.78 (d, *J* = 4.48 Hz, 1H), 7.75 (s, 1H), 7.73 (s, 1H), 7.71 (s, 1H), 7.69 (d, *J* = 4.22 Hz, 1H), 7.67 (s, 1H), 7.65 (s, 1H), 7.59 (s, 1H), 7.58 – 7.55 (m, 1H), 7.54 (s, 1H), 7.53 – 7.48 (m, 1H), 7.48 – 7.42 (m, 1H), 7.38 (td, *J* = 8.61, 3.80 Hz, 1H), 7.28 (t, *J* = 8.64 Hz, 1H), 7.19 (t, *J* = 8.91 Hz, 1H), 7.06 (s, 1H), 6.94 (d, *J* = 10.04 Hz, 1H), 6.75 (d, *J* = 9.33 Hz, 1H), 6.60 (d, *J* = 9.56 Hz, 1H), 6.53 (s, 1H), 6.51 (s, 1H), 6.30 (d, *J* = 9.75 Hz, 1H), 6.20 (s, 1H), 6.18 (s, 1H), 6.13 (s, 1H), 4.28 (d, *J* = 13.23 Hz, 1H), 4.22 (d, *J* = 13.88 Hz, 1H), 4.10 (s, 1H), 4.07 (s, 2H), 3.97 (s, 6H), 3.90 (s, 3H), 3.72 (d, *J* = 16.33 Hz, 1H), 3.47 (d, *J* = 15.24 Hz, 1H), 2.55 (d, *J* = 12.82 Hz, 1H), 2.46 (d, *J* = 16.20 Hz, 1H), 2.40 (d, *J* = 18.53 Hz, 1H), 2.27 (d, *J* = 14.44 Hz, 1H), 1.51 (s, 3H). **HRMS** (ESI⁻) *m/z* calcd. for $C_{102}H_{78}N_{15}O_{32}$: 2024.4943 $(M\text{-}H)^-$; found: 2024.5504.



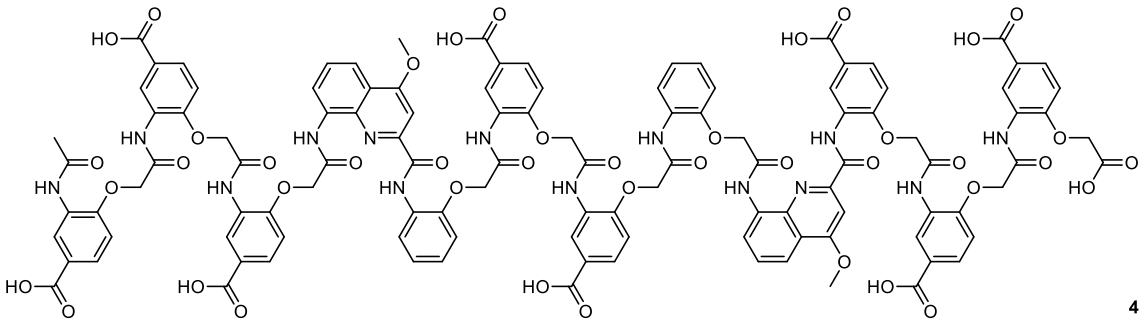
2

Compound 2: Oligomer **2** was synthesized on Wang resin (0.37 mmol g^{-1}, 27.8 µmol scale) according to the standard method (automated). Loading of the first monomer: 0.37 mmol g^{-1} (quant.). The final acetyl group was installed via the general acetylation method. After purification by semi-prep HPLC (C18, 0–15B, 50 °C; A: 13 mmol NH_4OAc buffer pH 8.5, B: acetonitrile), the title compound was obtained as a white solid (14.1 mg, 5.88 µmol, 21%; HPLC-purity: 99.9%). **^{1}H NMR** (500 MHz, 12 mmol NH_4OAc buffer pH 8.5 in H_2O/D_2O 9:1): δ = 10.10 (s, 1H), 10.02 (s, 1H), 9.63 (s, 1H), 9.43 (s, 3H), 9.39 (s, 1H), 9.25 (s, 1H), 9.07 (s, 1H), 8.81 (s, 1H), 8.47 (d, *J* = 2.17 Hz, 1H), 8.46 (d, *J* = 2.10 Hz, 1H), 8.41 (s, 1H), 8.39 (s, 1H), 8.35 (s, 1H), 8.20 – 8.14 (m, 2H), 8.10 (d, *J* = 8.62 Hz, 1H), 8.07 (d, *J* = 9.23 Hz, 1H), 8.01 (d, *J* = 9.28 Hz, 1H), 7.98 – 7.86 (m, 6H), 7.81 (d, *J* = 7.91 Hz, 1H), 7.72 – 7.64 (m, 2H), 7.63 (s, 1H), 7.61 (s, 13H), 7.60 – 7.57 (m, 2H), 7.56 (s, 2H), 7.51 (td, *J* = 8.46, 3.44 Hz, 2H), 7.30 – 7.23 (m, 1H), 7.19 (s, 1H), 6.84 (d, *J* = 9.41 Hz, 1H), 6.80 (s, 1H), 6.59 (d, *J* = 9.39 Hz, 1H), 6.52 – 6.43 (m, 4H), 6.42 (s, 1H), 6.38 (s, 1H), 6.32 (d, *J* = 9.40 Hz, 1H), 5.92 (d, *J* = 9.39 Hz, 1H), 4.27 (s, 1H), 4.17 (s, 2H), 4.00 (s, 5H), 3.92 (s, 3H), 3.90 (s, 1H), 3.71 (d, *J* = 15.54 Hz, 1H), 3.51 (d, *J* = 16.12 Hz, 1H), 3.46 (d, *J* = 15.78 Hz, 1H), 3.39 (d, *J* = 15.74 Hz, 1H), 2.95 (d, *J* = 15.86 Hz, 2H), 2.61 (d, *J* = 13.15 Hz, 1H), 2.06 (d, *J* = 13.51 Hz, 1H), 1.75 (d, *J* = 15.83 Hz, 1H), 1.65 (s, 3H), 1.57 (d, *J* = 15.54 Hz, 1H), 1.41 (d, *J* = 15.05 Hz, 1H), 1.30 (d, *J* = 11.21 Hz, 2H). **HRMS** (ESI⁻) *m/z* calcd. for $C_{118}H_{91}N_{16}O_{42}$: 2403.5482 $(M\text{-}H)^-$; found: 2403.6165.



3

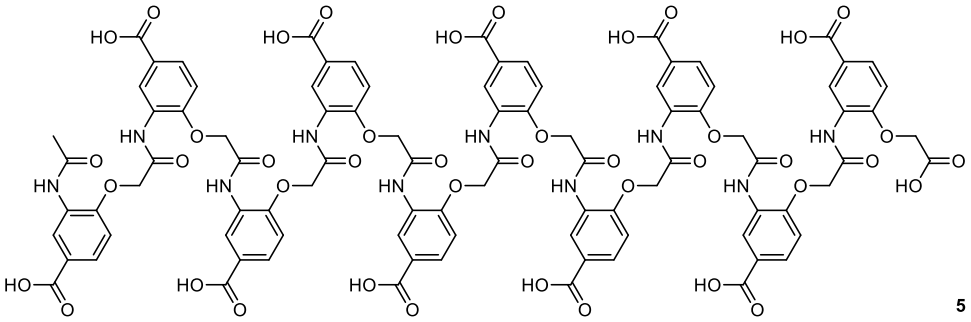
Compound 3: Oligomer **3** was synthesized on Wang resin (0.37 mmol g^{-1}, 27.8 µmol scale) according to the standard method (automated). Loading of the first monomer: 0.37 mmol g^{-1} (quant.). The final acetyl group was installed via the general acetylation method. After purification by semi-prep HPLC (C18, 0–20B, 50 °C; A: 13 mmol NH_4OAc buffer pH 8.5, B: acetonitrile), the title compound was obtained as a white solid (3.99 mg, 1.76 µmol, 6.3%; HPLC-purity: 98.8%). **^{1}H NMR** (500 MHz, 12 mmol NH_4OAc buffer pH 8.5 in H_2O/D_2O 9:1): δ = 10.40 (s, 1H), 9.97 (s, 1H), 9.83 (s, 1H), 9.76 (s, 1H), 9.62 (s, 2H), 8.86 (s, 1H), 8.49 (d, *J* = 10.09 Hz, 1H), 8.27 (d, *J* = 2.16 Hz, 1H), 8.25 (d, *J* = 8.52 Hz, 1H), 8.18 (s, 1H), 8.12 (d, *J* = 8.39 Hz, 1H), 8.09 (d, *J* = 8.62 Hz, 1H), 8.06 – 7.99 (m, 5H), 7.96 (dd, *J* = 9.19, 6.35 Hz, 2H), 7.92 (d, *J* = 9.61 Hz, 1H), 7.86 (d, *J* = 9.93 Hz, 1H), 7.83 (d, *J* = 9.68 Hz, 1H), 7.80 (d, *J* = 9.05 Hz, 1H), 7.78 – 7.73 (m, 1H), 7.72 – 7.68 (m, 2H), 7.68 – 7.64 (m, 1H), 7.63 (s, 1H), 7.61 (d, *J* = 6.30 Hz, 1H), 7.58 (s, 1H), 7.58 – 7.52 (m, 4H), 7.52 – 7.49 (m, 1H), 7.49 – 7.45 (m, 1H), 7.44 (d, *J* = 2.15 Hz, 1H), 7.25 – 7.15 (m, 3H), 7.14 (d, *J* = 1.62 Hz, 1H), 7.13 (s, 1H), 7.11 (s, 1H), 7.09 (s, 1H), 7.07 (s, 1H), 7.05 (d, *J* = 9.56 Hz, 1H), 7.02 (s, 1H), 7.00 (d, *J* = 4.38 Hz, 1H), 6.97 (d, *J* = 8.06 Hz, 1H), 6.74 (d, *J* = 9.37 Hz, 1H), 6.60 (d, *J* = 9.30 Hz, 1H), 6.45 (dt, *J* = 10.81, 5.44 Hz, 4H), 6.38 – 6.31 (m, 3H), 6.23 (d, *J* = 9.52 Hz, 1H), 6.17 (d, *J* = 9.53 Hz, 1H), 4.31 (s, 1H), 4.26 (d, *J* = 14.51 Hz, 1H), 4.14 (s, 2H), 4.03 (d, *J* = 15.33 Hz, 1H), 3.96 (s, 3H), 3.94 (s, 3H), 3.89 (d, *J* = 14.91 Hz, 1H), 3.70 (d, *J* = 15.91 Hz, 1H), 3.56 (t, *J* = 15.60 Hz, 2H), 3.30 (d, *J* = 18.89 Hz, 2H), 2.70 (d, *J* = 15.56 Hz, 1H), 2.50 (d, *J* = 15.66 Hz, 1H), 2.42 (d, *J* = 13.68 Hz, 1H), 2.18 (d, *J* = 14.77 Hz, 1H), 1.55 (s, 3H), 1.47 (d, *J* = 15.82 Hz, 2H), 1.17 (d, *J* = 15.64 Hz, 1H). **HRMS** (ESI$^-$) *m/z* calcd. for $C_{113}H_{90}N_{15}O_{38}$: 2264.5577 $(M-H)^-$; found: 2264.6101.



4

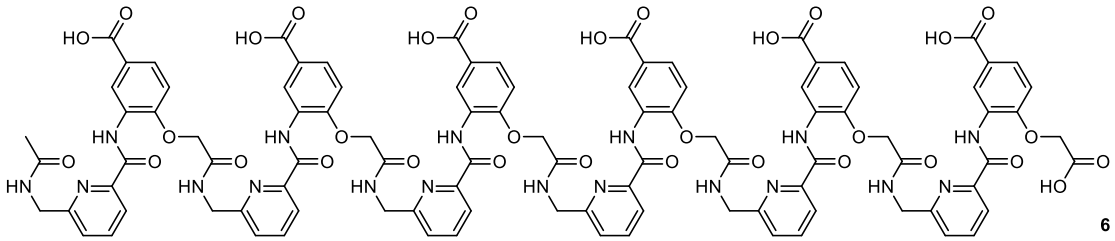
Compound 4: Oligomer **4** was synthesized on ProTide™ resin (0.15 mmol g^{-1}, 25 µmol scale) according to the standard method (automated). Loading of the first monomer: 0.15 mmol g^{-1}

(quant.). The final acetyl group was installed via the general acetylation method. After purification by semi-prep HPLC (C18, 0–5B, 50 °C; A: 13 mmol NH_4OAc buffer pH 8.5, B: acetonitrile), the title compound was obtained as a white solid (4.15 mg, 1.8 µmol, 7.2%; HPLC-purity: 90.31%). **^{1}H NMR** (500 MHz, 12 mmol NH_4OAc buffer pH 8.5 in H_2O/D_2O 9:1): δ = 10.00 (s, 1H), 9.97 (s, 1H), 9.90 (s, 1H), 9.76 (s, 1H), 9.04 (s, 1H), 8.85 (s, 1H), 8.56 (s, 1H), 8.54 (s, 1H), 8.49 (s, 1H), 8.46 (s, 1H), 8.31 (s, 1H), 8.30 – 8.24 (m, 2H), 8.18 (s, 1H), 8.08 – 8.00 (m, 2H), 7.92 (d, *J* = 10.34 Hz, 3H), 7.88 (d, *J* = 10.43 Hz, 1H), 7.74 (s, 1H), 7.71 (d, *J* = 10.15 Hz, 1H), 7.68 – 7.60 (m, 4H), 7.58 (d, *J* = 10.15 Hz, 1H), 7.54 (d, *J* = 10.68 Hz, 1H), 7.53 (s, 1H), 7.52 – 7.46 (m, 4H), 7.26 (t, *J* = 8.48 Hz, 1H), 7.22 – 7.14 (m, 2H), 7.11 (t, *J* = 8.40 Hz, 1H), 7.01 (s, 1H), 6.92 (d, *J* = 9.41 Hz, 1H), 6.75 (d, *J* = 9.51 Hz, 1H), 6.72 (d, *J* = 9.38 Hz, 1H), 6.68 (d, *J* = 9.60 Hz, 1H), 6.64 (d, *J* = 9.32 Hz, 1H), 6.53 (d, *J* = 9.60 Hz, 1H), 6.51 – 6.47 (m, 2H), 6.45 (d, *J* = 9.37 Hz, 1H), 6.36 (d, *J* = 9.42 Hz, 1H), 6.33 (d, *J* = 9.90 Hz, 2H), 6.31 (s, 1H), 4.35 (d, *J* = 15.53 Hz, 1H), 4.30 – 4.20 (m, 1H), 4.16 – 4.06 (m, 1H), 4.03 (d, *J* = 9.62 Hz, 1H), 3.97 (s, 3H), 3.93 (s, 3H), 3.88 (d, *J* = 14.97 Hz, 1H), 3.74 (d, *J* = 16.58 Hz, 1H), 3.61 (d, *J* = 16.36 Hz, 3H), 3.33 (d, *J* = 16.44 Hz, 1H), 3.13 (d, *J* = 15.02 Hz, 1H), 2.76 (d, *J* = 16.02 Hz, 1H), 2.49 (d, *J* = 13.58 Hz, 1H), 2.34 (d, *J* = 15.12 Hz, 1H), 2.25 – 2.22 (m, 1H), 2.20 (d, *J* = 15.54 Hz, 1H), 2.14 (d, *J* = 16.26 Hz, 1H), 1.63 (s, 3H), 1.32 (d, *J* = 16.40 Hz, 1H). **HRMS** (ESI^-) *m/z* calcd. for $C_{112}H_{89}N_{14}O_{42}$: 2301.5264 $(M-H)^-$; found: 2301.3943.



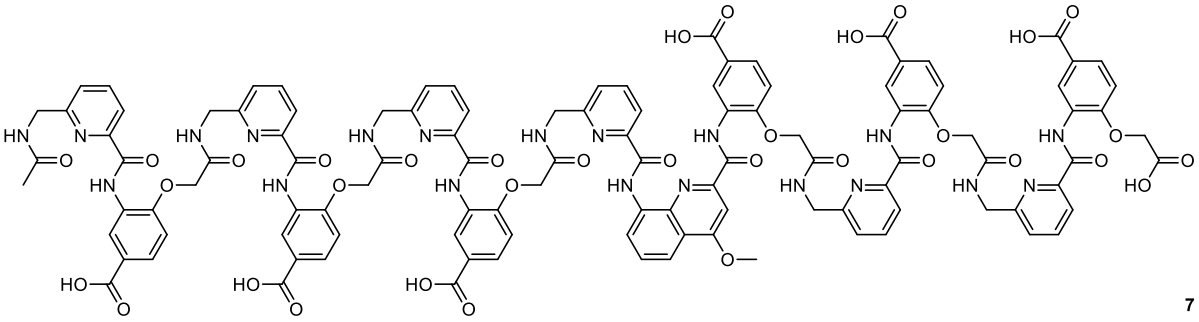
5

Compound 5: Oligomer **5** was synthesized on Wang resin (0.37 mmol g^{-1}, 27.8 µmol scale) according to the standard method (manual). Loading of the first monomer: 0.25 mmol g^{-1} (68%). The final acetyl group was installed via the general acetylation method. After purification by semi-prep HPLC (C18, 0–15B, 50 °C; A: 13 mmol NH_4OAc buffer pH 8.5, B: acetonitrile), the title compound was obtained as a white solid (9.03 mg, 4.46 µmol, 16%; HPLC-purity: 95.4%). **^{1}H NMR** (500 MHz, 12 mmol NH_4OAc buffer pH 8.5 in H_2O/D_2O 9:1): δ = 9.32 (s, 1H), 9.20 (s, 1H), 8.34 (s, 1H), 8.30 (s, 1H), 8.13 (s, 3H), 8.11 (s, 1H), 8.08 – 8.00 (m, 4H), 7.92 (s, 1H), 7.83 (s, 1H), 7.71 – 7.59 (m, 8H), 7.56 (d, *J* = 8.97 Hz, 1H), 6.97 (d, *J* = 9.71 Hz, 1H), 6.93 (d, *J* = 9.53 Hz, 1H), 6.89 – 6.76 (m, 7H), 6.60 (d, *J* = 9.58 Hz, 1H), 4.37 (s, 2H), 4.27 (s, 4H), 4.20 (s, 2H), 4.19 (s, 2H), 4.04 (s, 4H), 2.01 (s, 5H). **HRMS** (ESI^-) *m/z* calcd. for $C_{92}H_{73}N_{10}O_{42}$: 1989.3889 $(M-H)^-$; found: 1989.4469.



6

Compound 6: Oligomer **6** was synthesized on Wang resin (0.37 mmol g^{-1}, 24 µmol scale) according to the standard method (automated). Loading of the first monomer: 0.37 mmol g^{-1} (quant.). The final acetyl group was installed via the general acetylation method. After purification by semi-prep HPLC (C18, 0–10B, 50 °C; A: 13 mmol NH_4OAc buffer pH 8.5, B: acetonitrile), the title compound was obtained as a white solid (14.8 mg, 7.33 µmol, 31%; HPLC-purity: 95.5%). **^{1}H NMR** (500 MHz, 12 mmol NH_4OAc buffer pH 8.5 in H_2O/D_2O 9:1): δ = 10.15 (s, 1H), 10.01 (s, 1H), 9.44 (s, 1H), 9.36 (s, 1H), 9.15 (s, 1H), 9.09 (s, 1H), 8.38 (d, *J* = 2.19 Hz, 1H), 8.23 (dd, *J* = 10.50, 2.12 Hz, 2H), 8.17 (d, *J* = 2.18 Hz, 1H), 8.11 (dd, *J* = 4.03, 2.17 Hz, 2H), 8.08 (t, *J* = 4.25 Hz, 1H), 8.04 – 8.00 (m, 1H), 8.00 (d, *J* = 4.71 Hz, 1H), 7.98 (d, *J* = 4.86 Hz, 0H), 7.94 (d, *J* = 8.14 Hz, 1H), 7.90 (d, *J* = 8.19 Hz, 1H), 7.83 (t, *J* = 8.20 Hz, 1H), 7.77 – 7.70 (m, 6H), 7.66 – 7.59 (m, 2H), 7.59 – 7.54 (m, 1H), 7.51 (d, *J* = 8.40 Hz, 1H), 7.47 (dt, *J* = 7.58, 3.87 Hz, 3H), 7.37 (s, 1H), 7.35 (s, 1H), 7.34 – 7.32 (m, 2H), 7.31 (s, 1H), 7.30 (s, 0H), 7.28 (d, *J* = 3.69 Hz, 1H), 7.26 (s, 1H), 7.20 (d, *J* = 6.03 Hz, 2H), 7.19 (d, *J* = 6.91 Hz, 2H), 7.14 (d, *J* = 8.69 Hz, 1H), 7.07 (d, *J* = 9.18 Hz, 1H), 7.04 (d, *J* = 8.57 Hz, 1H), 6.68 (d, *J* = 9.18 Hz, 1H), 6.59 (d, *J* = 10.98 Hz, 2H), 6.47 (d, *J* = 9.20 Hz, 1H), 6.41 (d, *J* = 9.05 Hz, 1H), 4.54 (s, 0H), 4.39 (s, 0H), 1.77 (s, 4H). **HRMS** (ESI^-) *m/z* calcd. for $C_{98}H_{81}N_{18}O_{32}$: 2021.5270 $(M-H)^-$; found: 2021.6454.

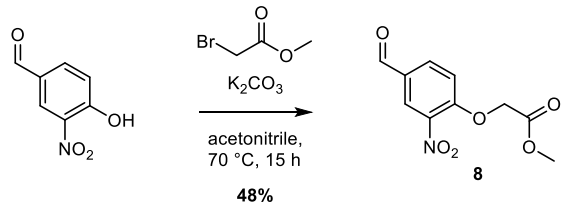


7

Compound 7: Oligomer **7** was synthesized on ProTide™ resin (0.15 mmol g^{-1}, 25 µmol scale) according to the standard method (automated). Loading of the first monomer: 0.15 mmol g^{-1} (quant.). The final acetyl group was installed via the general acetylation method. After purification by semi-prep HPLC (C18, 0–10B, 50 °C; A: 13 mmol NH_4OAc buffer pH 8.5, B: acetonitrile), the title compound was obtained as a white solid (12.0 mg, 5.4 µmol, 22%; HPLC-purity: 95.7%). **^{1}H NMR** (500 MHz, 12 mmol NH_4OAc buffer pH 8.5 in H_2O/D_2O 9:1): δ = 10.77 (s, 1H), 10.05 (s, 1H), 10.01 (s, 1H), 9.92 (s, 1H), 9.20 (s, 2H), 8.74 (s, 1H), 8.56 (s, 1H), 8.23 – 8.15 (m, 4H), 8.05 – 8.01 (m, 1H), 8.00 (d, *J* = 7.09 Hz, 1H), 7.96 (d, *J* = 8.26 Hz, 1H), 7.93 (d, *J* = 8.15 Hz, 1H), 7.90

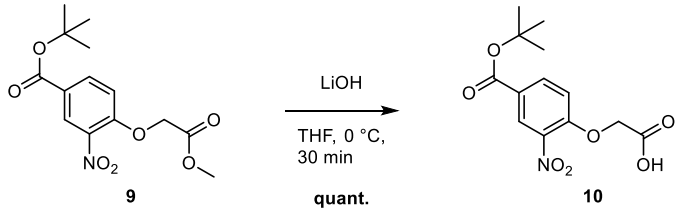
(s, 1H), 7.87 (d, *J* = 8.41 Hz, 1H), 7.81 (d, *J* = 2.03 Hz, 1H), 7.78 (d, *J* = 7.64 Hz, 1H), 7.75 (s, 1H), 7.74 – 7.70 (m, 4H), 7.69 (s, 1H), 7.66 – 7.60 (m, 3H), 7.54 (t, *J* = 8.23 Hz, 1H), 7.51 – 7.45 (m, 3H), 7.42 (d, *J* = 8.50 Hz, 1H), 7.38 (d, *J* = 7.92 Hz, 2H), 7.32 – 7.26 (m, 3H), 7.25 (d, *J* = 8.25 Hz, 1H), 7.17 – 7.11 (m, 2H), 7.08 (d, *J* = 9.50 Hz, 3H), 6.98 (d, *J* = 8.60 Hz, 1H), 6.93 (d, *J* = 8.69 Hz, 1H), 6.57 (d, *J* = 9.05 Hz, 2H), 6.54 (d, *J* = 8.70 Hz, 2H), 6.38 (d, *J* = 9.11 Hz, 1H), 6.32 (d, *J* = 9.38 Hz, 1H), 6.08 (s, 1H), 4.55 (s, 1H), 4.41 – 4.23 (m, 1H), 4.10 (d, *J* = 17.84 Hz, 1H), 4.02 (d, *J* = 15.87 Hz, 1H), 3.94 (d, *J* = 17.40 Hz, 1H), 3.88 (s, 3H), 3.84 (d, *J* = 14.03 Hz, 1H), 3.72 (d, *J* = 15.35 Hz, 2H), 3.55 (d, *J* = 15.01 Hz, 1H), 3.40 (d, *J* = 20.37 Hz, 1H), 3.23 (d, *J* = 20.11 Hz, 1H), 3.17 – 3.04 (m, 1H), 2.92 (d, *J* = 15.61 Hz, 1H), 2.63 – 2.47 (m, 3H), 2.31 (d, *J* = 21.86 Hz, 1H), 2.26 – 2.10 (m, 2H), 1.75 (s, 3H). **HRMS** (ESI⁻) *m/z* calcd. for $C_{109}H_{89}N_{20}O_{34}$: 2221.5856 $(M\text{-}H)^-$; found: 2221.4344.

4.2.3 Monomer synthesis procedures



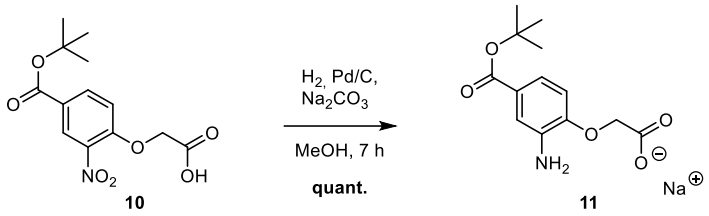
Compound 8: 4-hydroxy-3-nitrobenzaldehyde (8.40 g, 50.3 mmol, 1.0 equiv.) and K_2CO_3 (7.64 g, 55.3 mmol, 1.1 equiv.) was suspended in acetonitrile (300 ml). After the addition of methyl bromoacetate (5.23 ml, 55.3 mmol, 1.1 equiv.), the reaction mixture was stirred at 70 °C for 15 h under N_2 atmosphere. The resulting suspension was filtered, washed with acetonitrile and the filtrate evaporated *in vacuo*. Then, the residue was dissolved in water and $CHCl_3$. After the organic phase was removed, the aqueous phase was extracted with $CHCl_3$ (2×), the combined organic phases were washed with brine, dried over $MgSO_4$, evaporated *in vacuo*, and the residue was dried under high vacuum overnight. Finally, it was redissolved in a minimum amount of DCM. Et_2O was added until precipitation occurred. The solution was kept at 4 °C for 2 h, filtered and washed with cold Et_2O yielding the final compound (5.79 g, 24.2 mmol, 48%) as an off white crystalline solid. ($C_{10}H_9NO_6$; MW = 239.18 g mol^{-1}). $\mathbf{R_f}$ (CyHex/EtOAc 6:4) = 0.34. **^{1}H NMR** (500 MHz, $CDCl_3$): δ = 9.95 (s, 1H, C10-H), 8.39 (d, *J* = 2.05 Hz, 1H, C3-H), 8.06 (dd, *J* = 8.69, 2.10 Hz, 1H, C5-H), 7.09 (d, *J* = 8.68 Hz, 1H, C6-H), 4.90 (s, 2H; C7-H), 3.83 (s, 3H, C9-H). **^{13}C NMR** (126 MHz, $CDCl_3$): δ = 188.7 (C10), 167.5 (C8), 155.4 (C1), 140.4 (C2), 134.5 (C5), 130.2 (C4), 127.8 (C3), 114.9 (C6), 66.2 (C7), 52.9 (C9). **HRMS** (ESI⁻) *m/z* calcd. for $C_{10}H_9NO_6Cl$: 274.0124 $(M+Cl)^-$; Found: 274.0124. (Modified literature procedure[98]; analytical data is in line with the literature).

Compound 9: Compound **8** (19.6 g, 81.9 mmol, 1.0 equiv.) was dissolved in acetone (150 ml). Jones reagent (2 M CrO_3 in 6 M H_2SO_4 $_{(aq.)}$) was added dropwise at 23 °C until the starting material was consumed by TLC analysis (45 ml). Remaining Jones reagent was quenched by adding an excess of *i*PrOH and the suspension was filtered and washed with acetone. The solvent was evaporated under reduced pressure, and the residue was redissolved in DCM (THF may be added to help solubilization). H_2O was added, the organic phase removed, and the aqueous phase extracted with DCM (2×). The combined organic phases were dried over $MgSO_4$ and evaporated *in vacuo* to give the crude carboxylic acid (20.3 g with 20% impurity by NMR). Then, the solid was redissolved in anhydrous DMF (105 ml) under N_2 atmosphere. *t*BuOH (22.4 ml, 239 mmol, 3.0 equiv.), EDC·HCl (22.9 g, 119 mmol, 1.5 equiv.) and DMAP (9.74 g, 79.6 mmol, 1.0 equiv.) were added at 0 °C and the reaction mixture was stirred at 23 °C for 15 h. After evaporating the solvents *in vacuo*, the residue was dissolved in EtOAc and H_2O. The organic phase was removed, and the aqueous phase was extracted with EtOAc (2×). Combined organic phases were washed with sat. NH_4Cl $_{(aq.)}$ and brine, dried over Na_2SO_4 and evaporated *in vacuo*. After purification by column chromatography (SiO_2, CyHex/EtOAc 8:2), the title compound (17.8 g, 57.2 mmol, 70%) was obtained as a white solid ($C_{14}H_{17}NO_7$; MW = 311.29 g mol^{-1}). $\mathbf{R_f}$ (CyHex/EtOAc 8:2) = 0.20. **^{1}H NMR** (500 MHz, $CDCl_3$): δ = 8.45 (d, *J* = 2.15 Hz, C3-H), 8.14 (dd, *J* = 8.75, 2.16 Hz, 1H, C5-H), 6.97 (d, *J* = 8.82 Hz, 1H, C6-H), 4.85 (s, 2H, C7-H), 3.81 (s, 3H, C9-H), 1.59 (s, 9H, C12-H, C13-H, C14-H). **^{13}C NMR** (126 MHz, $CDCl_3$): δ = 167.8 (C8), 163.4 (C10), 154.0 (C1), 139.9 (C2), 135.0 (C5), 127.4 (C3), 126.0 (C4), 114.1 (C6), 82.4 (C11), 66.3 (C7), 52.8 (C9), 28.3 (C12, C13, C14). **HRMS** (ESI$^-$) *m/z* calcd. for $C_{14}H_{17}NO_7Cl$: 346.0699 (M+Cl)$^-$; Found: 346.0702. (Modified literature procedure[98]; analytical data is in line with the literature).

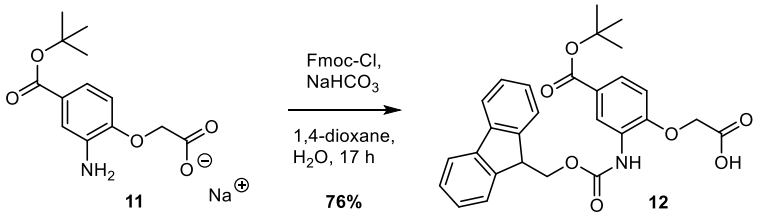


Compound 10: Compound **9** (17.8 g, 57.2 mmol, 1.0 equiv.) was dissolved in THF (800 ml) and cooled to 0 °C. LiOH (1.37 g, 57.2 mmol, 1.0 equiv.) in H_2O (200 ml) was added and the reaction mixture was stirred at 0 °C for 30 min. After the addition of citric acid $_{(aq.)}$ (1 M, 57.2 ml, 1 equiv.) the mixture was extracted with DCM (3×). The combined organic phases were washed with brine, dried over Na_2SO_4 and solvents were removed *in vacuo*. The final product (17.0 g, 57.2 mmol, quant.) was obtained as a white solid. ($C_{13}H_{15}NO_7$; MW = 297.26 g mol^{-1}). $\mathbf{R_f}$ (DCM/MeOH 95:5 + 1% AcOH) = 0.51. **^{1}H NMR** (400 MHz, DMSO-d_6): δ = 13.34 (s, 1H, O9-H), 8.30 (d, *J* = 2.18 Hz,

1H, C3-H), 8.09 (dd, *J* = 8.88, 2.23 Hz, 1H, C5-H), 7.38 (d, *J* = 8.93 Hz, 1H, C6-H), 5.02 (s, 2H, C7-H), 1.55 (s, 9H, C12-H, C13-H, C14-H). **^{13}C NMR** (101 MHz, DMSO-d_6): δ = 168.9 (C8), 163.0 (C10), 153.7 (C1), 139.2 (C2), 134.4 (C5), 125.7 (C3), 123.9 (C4), 115.3 (C6), 81.6 (C11), 65.6 (C7), 27.7 (C12, C13, C14). **HRMS** (ESI$^-$) *m/z* calcd. for $C_{13}H_{14}NO_7$: 296.0776 $(M-H)^-$; Found: 296.0776. (Modified literature procedure[98]; analytical data is in line with the literature).



Compound 11: Compound **10** (17.0 g, 57.2 mmol, 1.0 equiv.) and Na_2CO_3 (6.28 g, 57.2 mmol, 1.0 equiv.) were dissolved in MeOH (750 ml). The solution was quickly degassed by vacuum N_2 cycles (3×), then Pd/C (1.70 g, 10 wt. % loading) was added and the N_2 atmosphere was replaced by H_2. After stirring for 7 h the reaction mixture was filtered over celite©, washed with MeOH and solvents were removed *in vacuo*. The crude product (16.5 g, 57.2 mmol, quant.) was obtained as a slightly brown solid and was used in the next step without further purification. ($C_{13}H_{16}NO_5Na$; MW = 289.26 g mol^{-1}). **R_f** (EtOH) = 0.60. **^{1}H-NMR** (400 MHz, DMSO-d_6): δ = 7.16 (d, *J* = 2.12 Hz, 1H, C3-H), 7.05 (dd, *J* = 8.30, 2.17 Hz, 1H, C5-H), 6.62 (d, *J* = 8.38 Hz, 1H, C6-H), 4.96 (s, 2H, N15-H), 4.11 (s, 2H, C7-H), 1.50 (s, 9H, C12-H, C13-H, C14-H). **^{13}C-NMR** (101 MHz, DMSO-d_6): δ = 170.5 (C8), 165.5 (C10), 150.3 (C1), 138.0 (C2), 123.3 (C4), 118.2 (C5), 113.9 (C3), 111.8 (C6), 79.3 (C11), 69.2 (C7), 27.9 (C12, C13, C14). **HRMS** (ESI$^-$) *m/z* calcd. for $C_{13}H_{16}NO_5$: 266.1034 $(M-H)^-$; Found: 266.1033.



Compound 12: Compound **11** (16.5 g, 57.2 mmol, 1.0 equiv.) was added to $NaHCO_3$ (24.0 g, 286 mmol, 5.0 equiv.) in H_2O (400 ml). After the suspension was cooled to 0 °C, Fmoc-Cl (19.2 g, 74.4 mmol, 1.3 equiv.) in 1,4-dioxane (400 ml) was added dropwise at 0 °C over 1 h. The reaction mixture was stirred at 0 °C for 1 h, then at 23 °C for 15 h. After acidifying to pH 2 using citric acid $_{(aq.)}$ (1 M) the mixture was extracted with DCM (3×). The combined organic phases were washed with brine, dried over $MgSO_4$ and solvents were evaporated *in vacuo*. The crude product was purified by trituration: the residue was dissolved in a minimal amount of THF, a larger amount of Et_2O was added and the suspension was cooled to 4 °C over night to help precipitation. The suspension was filtered and washed with cold Et_2O. This procedure was repeated once yielding the title compound (21.9 g, 44.8 mmol, 76%, HPLC purity: 99%) as a white solid ($C_{28}H_{27}NO_7$; MW = 489.52 g mol^{-1}). **R_f** (EtOAc/MeOH 98:2 + 1% AcOH) = 0.36. **^{1}H NMR** (500 MHz, DMSO-d_6): δ =

13.20 (s, 1H, O9-H), 8.88 (s, 1H, N15-H), 8.28 (s, 1H, C3-H), 7.91 (d, *J* = 7.62 Hz, 2H, C22-H, C23-H), 7.76 (d, *J* = 7.50 Hz, 2H, C19-H, C26-H), 7.62 (dd, *J* = 8.58, 2.16 Hz, 1H, C5-H), 7.43 (t, *J* = 7.50 Hz, 2H, C21-H, C24-H), 7.34 (t, *J* = 7.47 Hz, 2H, C20-H, C25-H), 7.06 (d, *J* = 8.70 Hz, 1H, C6-H), 4.83 (s, 2H, C7-H), 4.41 (d, *J* = 7.26 Hz, 2H, C17-H), 4.31 (t, *J* = 7.18 Hz, 1H, C18-H), 1.52 (s, 9H, C12-H, C13-H, C14-H). **^{13}C NMR** (126 MHz, DMSO-d_6): δ = 169.8 (C8), 164.5 (C10), 153.7 (C16), 152.1 (C1), 143.7 (C18a, C26a), 140.7 (C22a, C22b, 127.7 (C2), 127.4 (C20, C25), 127.1 (C5), 125.7 (C19, C26), 125.4 (C4), 121.8 (C3), 120.2 (C22, C23), 112.5 (C6), 80.3 (C11), 66.3 (C17), 65.7 (C7), 46.5 (C18), 27.8 (C12, C13, C14). **HRMS** (ESI$^+$) *m/z* calcd. for $C_{28}H_{28}NO_7$: 490.1860 $(M+H)^+$; Found: 490.1863.

4.3 X-ray Crystallography

Aqueous solutions of compounds **1** and **6** were prepared from lyophilized samples to final concentrations of 2 mM (in 15 mM NH4OAc buffer pH 8.5) and 8 mM (in H_2O) respectively. Crystallization screening trials were carried out by vapor diffusion method using commercial sparse matrix screens at 293 K.[99] Diffraction-quality crystals of **1** (Fig. 25a) was obtained by sitting drop method by adding 1 µL of **1** and 2 µL of the reservoir solution containing 20% w/v polyethylene glycol 8000, 10 mM TRIS buffer (pH 7.5), and 10 mM calcium chloride. Volume of reservoir solution was 500 µL. Diffraction-quality crystals (Fig. 25b) of **6** was obtained by hanging drop method by adding 1.2 µL of **6** and 1.8 µL of the reservoir solution containing 30% v/v polyethylene glycol 400, 100 mM HEPES buffer (pH 7.5), and 200 mM calcium chloride. 400 µL of the reservoir solution was layered by 100 µL of Silicon oil to slow the rate of vapor diffusion.[100] Single crystals of **1** were fished using MiTeGen microloops, quickly soaked in a cryo-protectant solution of 20% w/v polyethylene glycol 8000 and 40% v/v polyethylene glycol 400 and flash frozen in liquid nitrogen. Single crystals of **6** were fished and directly plunged into liquid nitrogen without cryo-protection.

Synchrotron data for **1** and **6** were collected at P14 and P13[101] beam lines operated by EMBL Hamburg at the Petra III storage ring (DESY, Hamburg, Germany) using EIGER 16M detector. Diffraction data for **1** was processed using *xia2*[102] with *DIALS*[103] for integration and using *Pointless*/*Aimless*[104] for scaling and merging respectively. Diffraction data for **6** was processed using *CrysAlis*^{Pro}.* Both structures were solved using dual space method with the program *ShelxD*[105] and refined by a full-matrix least squares method on F^2 with *ShelxL*-2014[106] within *Olex2* suite.[107] The initial structures of both **1** and **6** revealed most of the main-chain atoms. After each refinement step, visual inspection of the model and the electron-density maps were carried out using *Olex2* and *Coot*.[108] AFIX, DFIX, EADP, SADI and FLAT instructions were used to improve the geometry of molecules and temperature parameters. All non-H atoms were refined with anisotropic displacement parameters. Hydrogen atoms were placed at idealized positions. Restraints on anisotropic displacement parameters were implemented with DELU, SIMU, RIGU and ISOR instructions. In the final stage of refinement *SQUEEZE*[109] procedure from Platon suite was introduced to remove unmodeled electron density.

Statistics of data collection and refinement are described in Table 1. The final cif file was checked using IUCr's checkcif algorithm. Due to large volume fractions of disordered solvent molecules, weak diffraction intensity and poor resolution, a number of A- and B-level remain in the checkcif file. These alerts are inherent to the data and refinement procedures and illustrate the limited practicality of the checkcif tool for medium- size molecule crystallography. They are listed below and have been divided into two groups. The first group illustrates weak quality of the data and refinement statistics if compared to that expected for small molecule structures from highly diffracting crystals. The second group is connected to decisions made during refinement and explained below.

* Agilent, CrysAlisPRO. Agilent Technologies Ltd, Yarnton, Oxfordshire, England (2014).

Atomic coordinates and structure factors for **1** and **6** were deposited in the Cambridge Crystallographic Data Centre (CCDC) with accession codes 2125508 and 2125515 respectively. The data is available free of charge upon request (www.ccdc.cam.ac.uk/).

CheckCIF validation of **1**:

Group 1 alerts:

THETM01_ALERT_3_A	The value of sine(theta_max)/wavelength is less than 0.550
PLAT029_ALERT_3_A _	diffrn_measured_fraction_theta_full value Low. Why?
PLAT084_ALERT_3_A	High wR2 Value (i.e. > 0.25) ..... 0.36 Report
PLAT242_ALERT_2_B	Low 'MainMol' Ueq as Compared to Neighbors
PLAT250_ALERT_2_B	Large U3/U1 Ratio for Average U(I,j) Tensor.. Note
PLAT340_ALERT_3_B	Low Bond Precision on C-C Bonds.. 0.03562 Ang.

Group 2 alerts:

PLAT306_ALERT_2_B Isolated Oxygen Atom Check

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

PLAT412_ALERT_2_A Short Intra XH3 .. XHn Check

These alerts concern H atoms placed geometrically.

PLAT430_ALERT_2_A Short Inter D...A Contact Check

These alerts concern contacts with solvent molecules whose positions were poorly determined.

CheckCIF validation of **6**:

Group 1 alerts:

THETM01_ALERT_3_A	The value of sine(theta_max)/wavelength is less than 0.550
PLAT029_ALERT_3_A _	diffrn_measured_fraction_theta_full value Low. Why?
PLAT082_ALERT_2_B	High R1 Value ..... 0.17 Report
PLAT084_ALERT_3_A	High wR2 Value (i.e. > 0.25) ..... 0.46 Report
PLAT250_ALERT_2_B	Large U3/U1 Ratio for Average U(I,j) Tensor.. Note
PLAT340_ALERT_3_B	Low Bond Precision on C-C Bonds.. 0.01523 Ang.

Group 2 alerts:

PLAT316_ALERT_2_A Too many H on C in C=N Moiety in Main Residue Check

Concerned C in C–N moiety in main residue and H atoms checked.

PLAT306_ALERT_2_B Isolated Oxygen Atom Check

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

PLAT430_ALERT_2_B Short Inter D...A Contact Check

These alerts concern contacts with solvent molecules which positions were poorly determined

Table 1 Crystallographic data and refinement details for **1** and **6**.

Foldamers	**1**	**6**
Chemical formula	$C_{102}H_{71.75}Ca_1N_{15}O_{37.25}$	$C_{98}H_{75}Ca_{2.25}N_{18}O_{49.25}$
Formula weight	2143.57	2382.94
Temperature	100 K	100.15
Wavelength	0.9762 Å	0.8000 Å
Crystal system	Triclinic	Orthorhombic
Space group	P1	Pbcn
Unit cell dimensions	a = 21.833 (4) Å b = 26.901 (4) Å c = 27.093 (3) Å α = 93.876 (1)° β = 108.543 (1)° γ = 111.116 (13)°	a = 48.968 (8) Å b = 43.383 (2) Å c = 27.123 (19) Å α = 90° β = 90° γ = 90°
Volume	13772 (4) Å^3	57620.1 (10) Å^3
Z	4	16
Density (calculated)	1.034 g/cm^3	1.099 g/cm^3
Absorption coefficient	0.215 μ/mm^{-1}	0.230 μ/mm^{-1}
Color and shape	Pale yellow, needles	Colorless, plates
Crystal size	0.15 x 0.02 x 0.02 mm	0.15 x 0.05 x 0.01 mm
Index ranges	-18 ≤ h ≤ 18 -22 ≤ k ≤ 22 -22 ≤ l ≤22	-19 ≤ h ≤ 44 -40 ≤ k ≤ 42 -23 ≤ l ≤26
Reflections collected	57803	96340
R_{int}	0.1337	0.0838
Data/restraints/parameters	29422/675/1376	22635/467/2493
Goodness-of-fit on F^2	1.006	1.786
Final R indexes [$I > 2\sigma(I)$]	R_1 = 0.1210 wR_2 = 0.2838	R_1 = 0.1681 wR_2 = 0.4348
Final R indexes [all data]	R_1 = 0.2104 wR_2 = 0.3613	R_1 = 0.1947 wR_2 = 0.4615
Largest diff. peak and hole	0.69/-0.56 e Å^{-3}	1.07/-0.70 e Å^{-3}
Total potential solvent accessible void volume from SQUEEZE	3592.0 Å^3	16882.8 Å^3
Electron count/cell	1126	4517
CCDC #	2125508	2125515

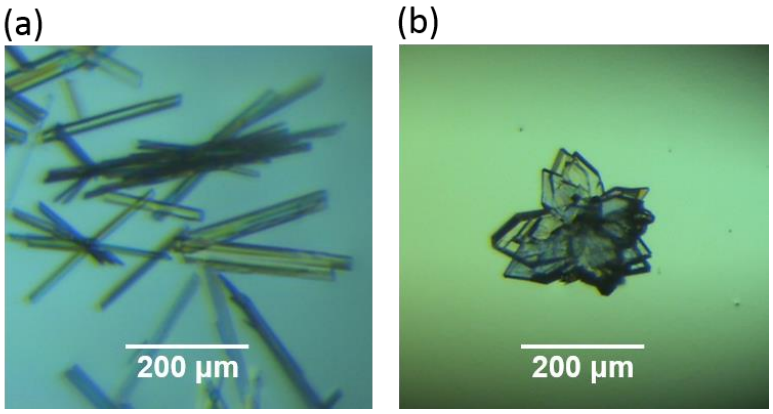


Fig. 25 Crystals of (a) **1** and (b) **6** observed under cross polarizing microscope.

4.4 Spectra and Chromatograms

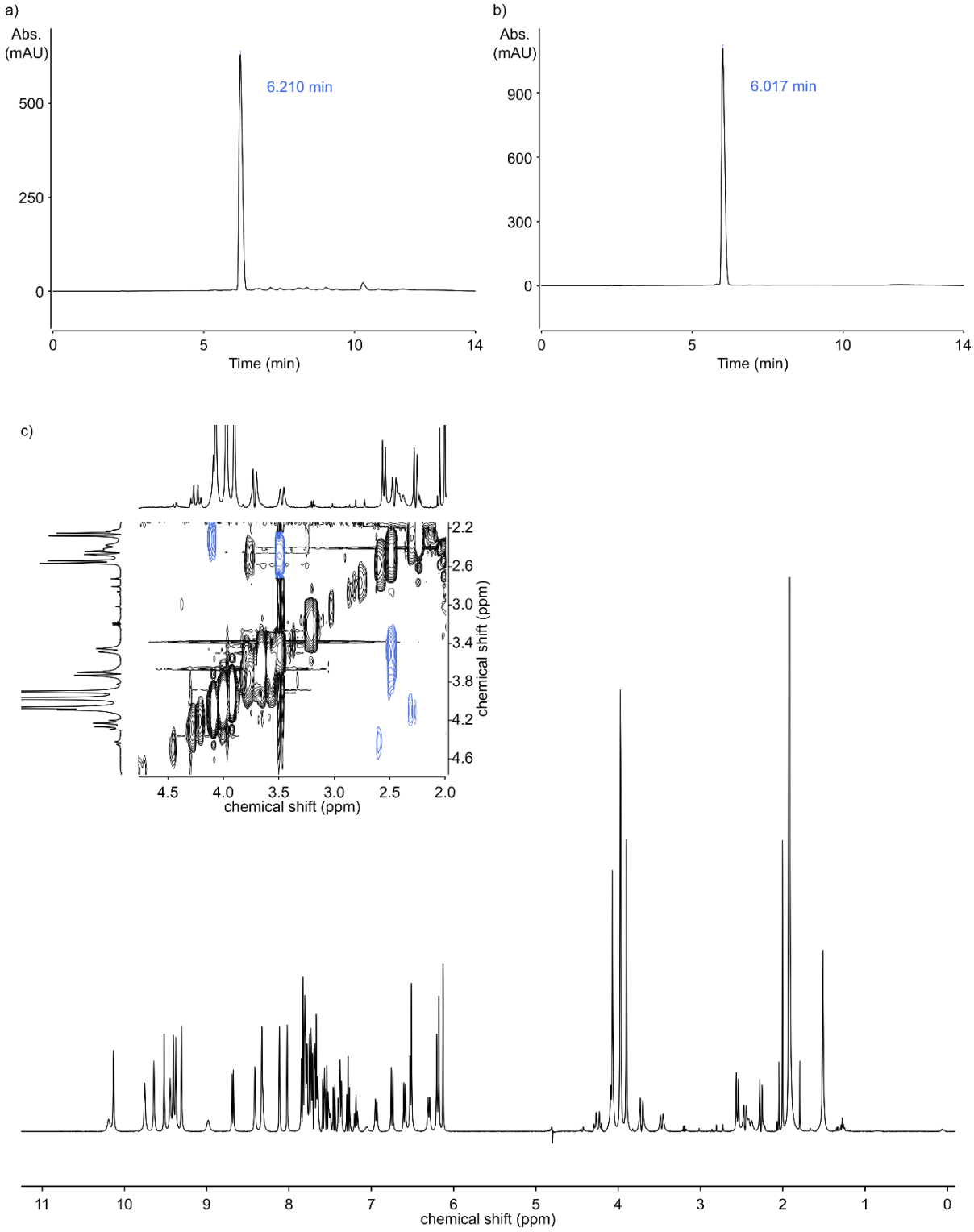


Fig. 26 Analytical data of compound **1**. HPLC chromatograms after cleavage from the resin (a) and after purification (b) (C18, 0–30B, 50 °C; A: 13mM NH_4OAc buffer pH 8.5, B: acetonitrile). c) ^1H NMR and part of the COSY NMR spectrum (500 MHz, 0.6 mM in 12 mM NH_4OAc buffer pH 8.5 H_2O/D_2O 9:1, H_2O suppression), some cross-signals of diastereotopic methylene groups are exemplary highlighted in blue.

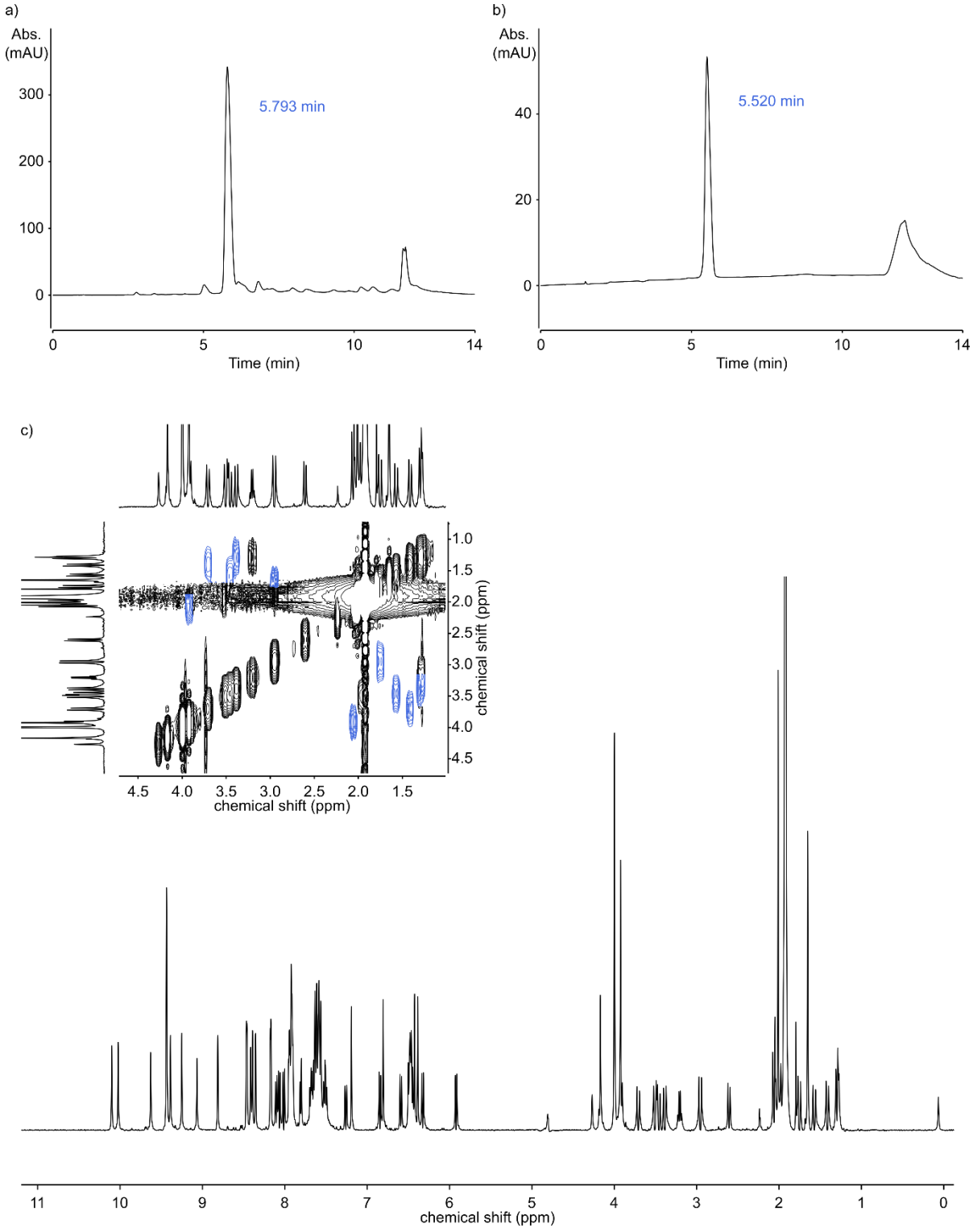


Fig. 27 Analytical data of compound **2**. HPLC chromatograms after cleavage from the resin (a) and after purification (b) (C18, 0–15B, 50 °C; A: 13mM NH_4OAc buffer pH 8.5, B: acetonitrile). c) ^1H NMR and part of the COSY NMR spectrum (500 MHz, 0.26 mM in 12 mM NH_4OAc buffer pH 8.5 H_2O/D_2O 9:1, H_2O suppression), some cross-signals of diastereotopic methylene groups are exemplary highlighted in blue.

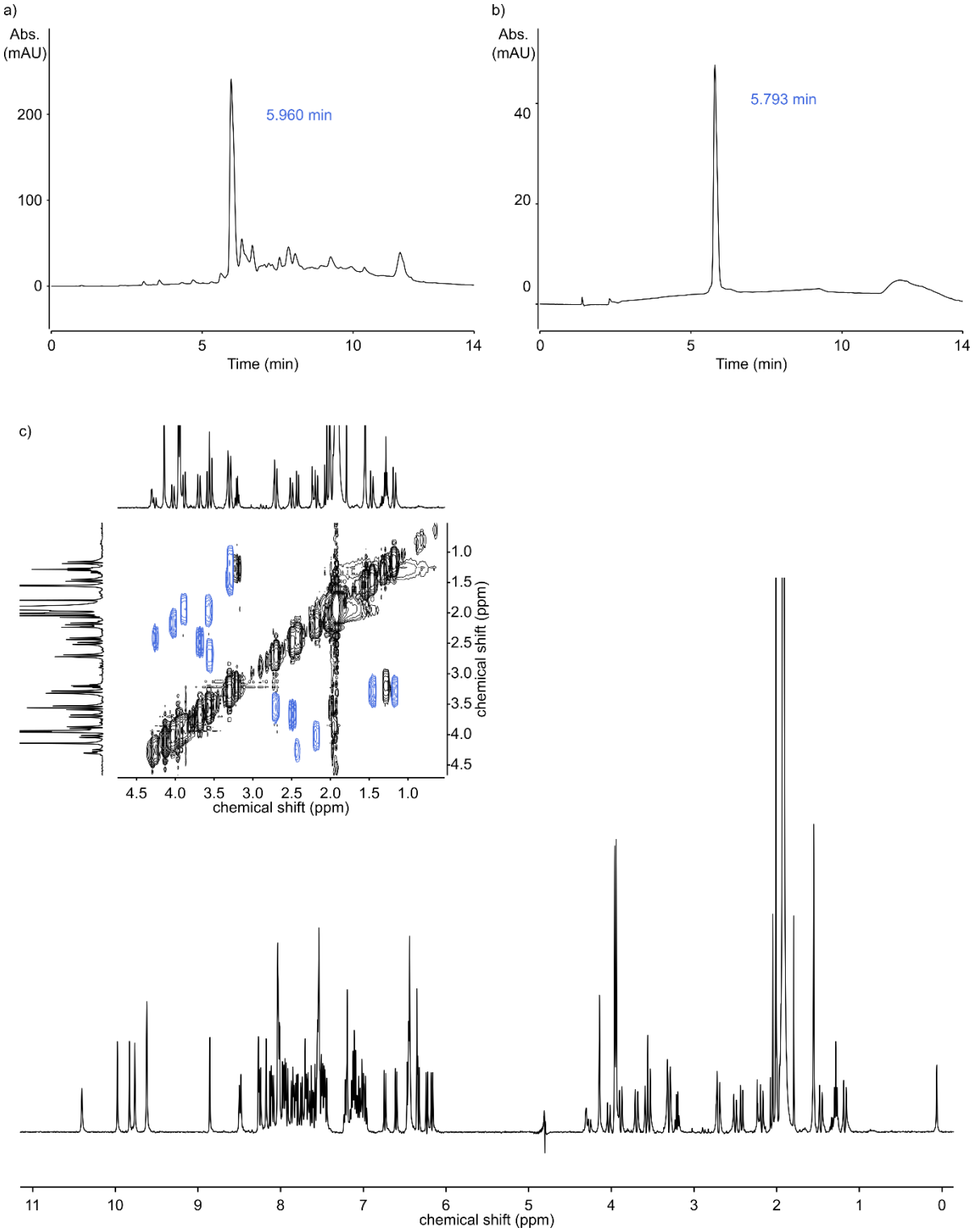


Fig. 28 Analytical data of compound **3**. HPLC chromatograms after cleavage from the resin (a) and after purification (b) (C18, 0–30B, 50 °C; A: 13mM NH_4OAc buffer pH 8.5, B: acetonitrile). c) ^1H NMR and part of the COSY NMR spectrum (500 MHz, 0.16 mM in 12 mM NH_4OAc buffer pH 8.5 H_2O/D_2O 9:1, H_2O suppression), some cross-signals of diastereotopic methylene groups are exemplary highlighted in blue.

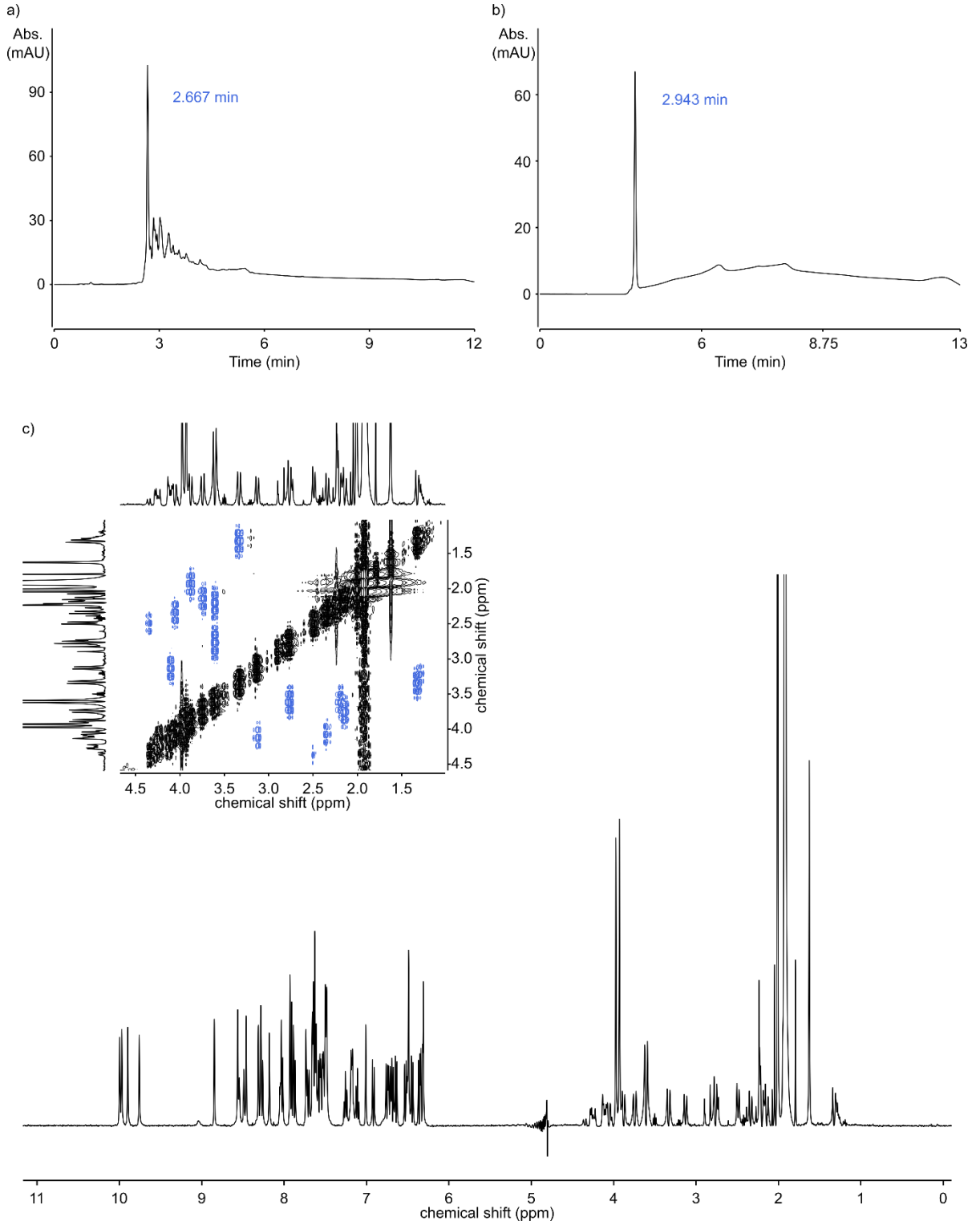


Fig. 29 Analytical data of compound **4**. HPLC chromatograms after cleavage from the resin (a) (C18, 0–100B, 50 °C; A: 13mM NH_4OAc buffer pH 8.5, B: acetonitrile) and after purification (b) (C18, 0–50B, 50 °C; A: 13mM NH_4OAc buffer pH 8.5, B: acetonitrile). c) ^1H NMR and part of the COSY NMR spectrum (500 MHz, 0.18 mM in 12 mM NH_4OAc buffer pH 8.5 H_2O/D_2O 9:1, H_2O suppression), some cross-signals of diastereotopic methylene groups are exemplary highlighted in blue.

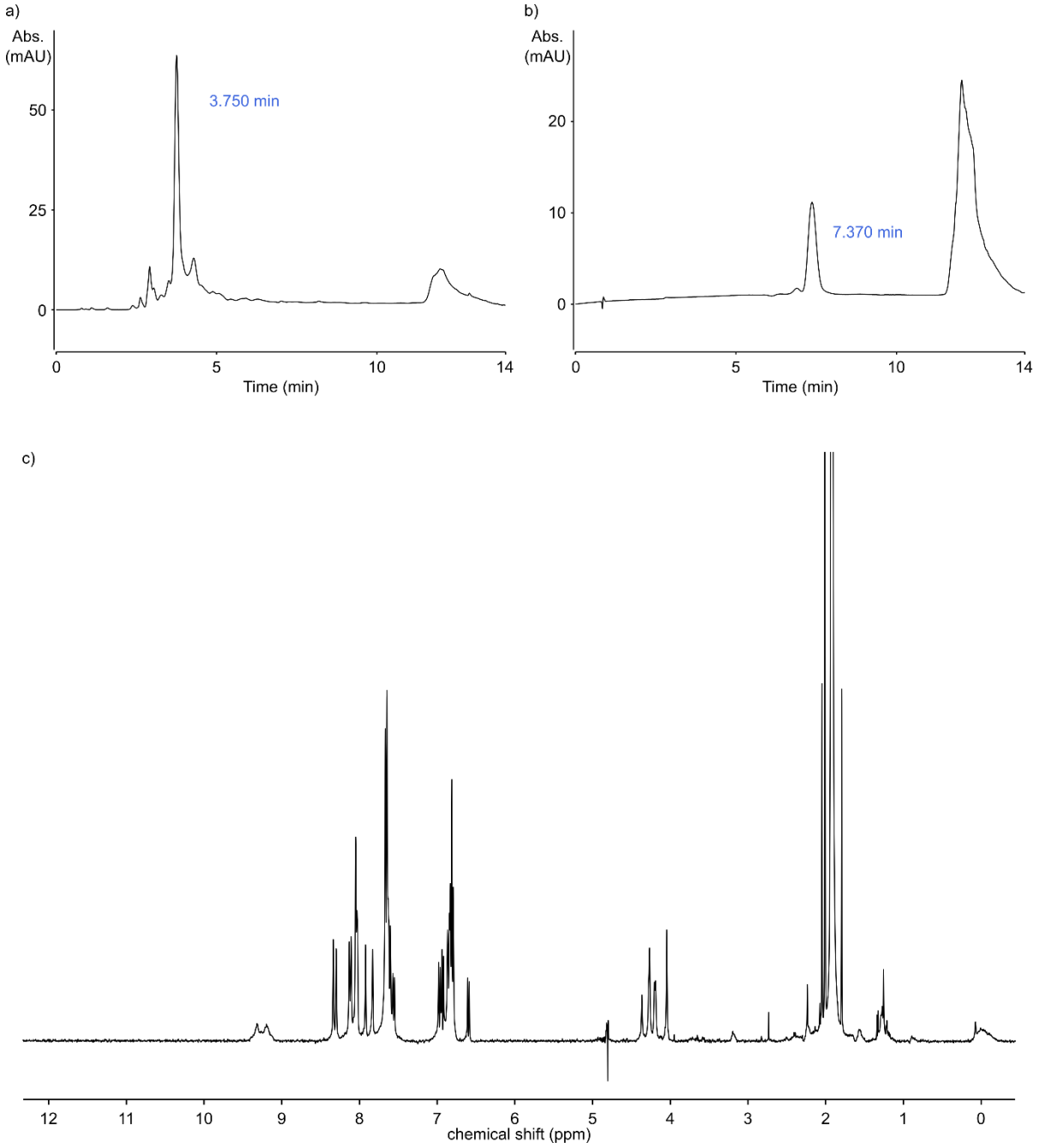


Fig. 30 Analytical data of compound **5**. HPLC chromatograms after cleavage from the resin (a) (C18, 0–15B, 50 °C; A: 13mM NH_4OAc buffer pH 8.5, B: acetonitrile)and after purification (b) (C18, 0–5B, 50 °C; A: 13mM NH_4OAc buffer pH 8.5, B: acetonitrile). c) ^1H NMR spectrum (500 MHz, 0.6 mM in 12 mM NH_4OAc buffer pH 8.5 H_2O/D_2O 9:1, H_2O suppression.

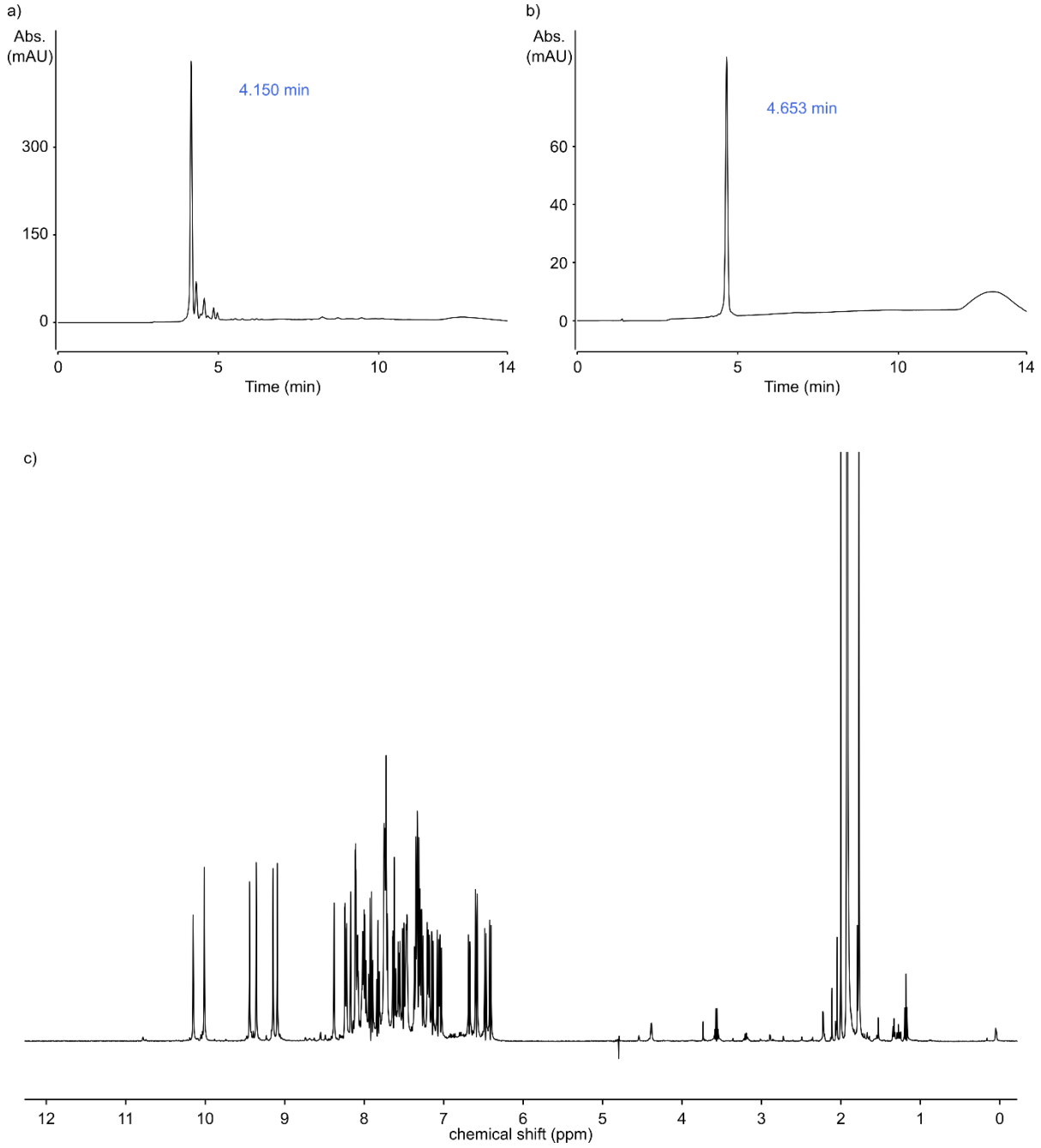


Fig. 31 Analytical data of compound **6**. HPLC chromatograms after cleavage from the resin (a) (C18, 0–30B, 50 °C; A: 13mM NH_4OAc buffer pH 8.5, B: acetonitrile) and after purification (b) (C18, 0–10B, 50 °C; A: 13mM NH_4OAc buffer pH 8.5, B: acetonitrile). c) ^1H NMR spectrum (500 MHz, 0.66 mM in 12 mM NH_4OAc buffer pH 8.5 H_2O/D_2O 9:1, H_2O suppression).

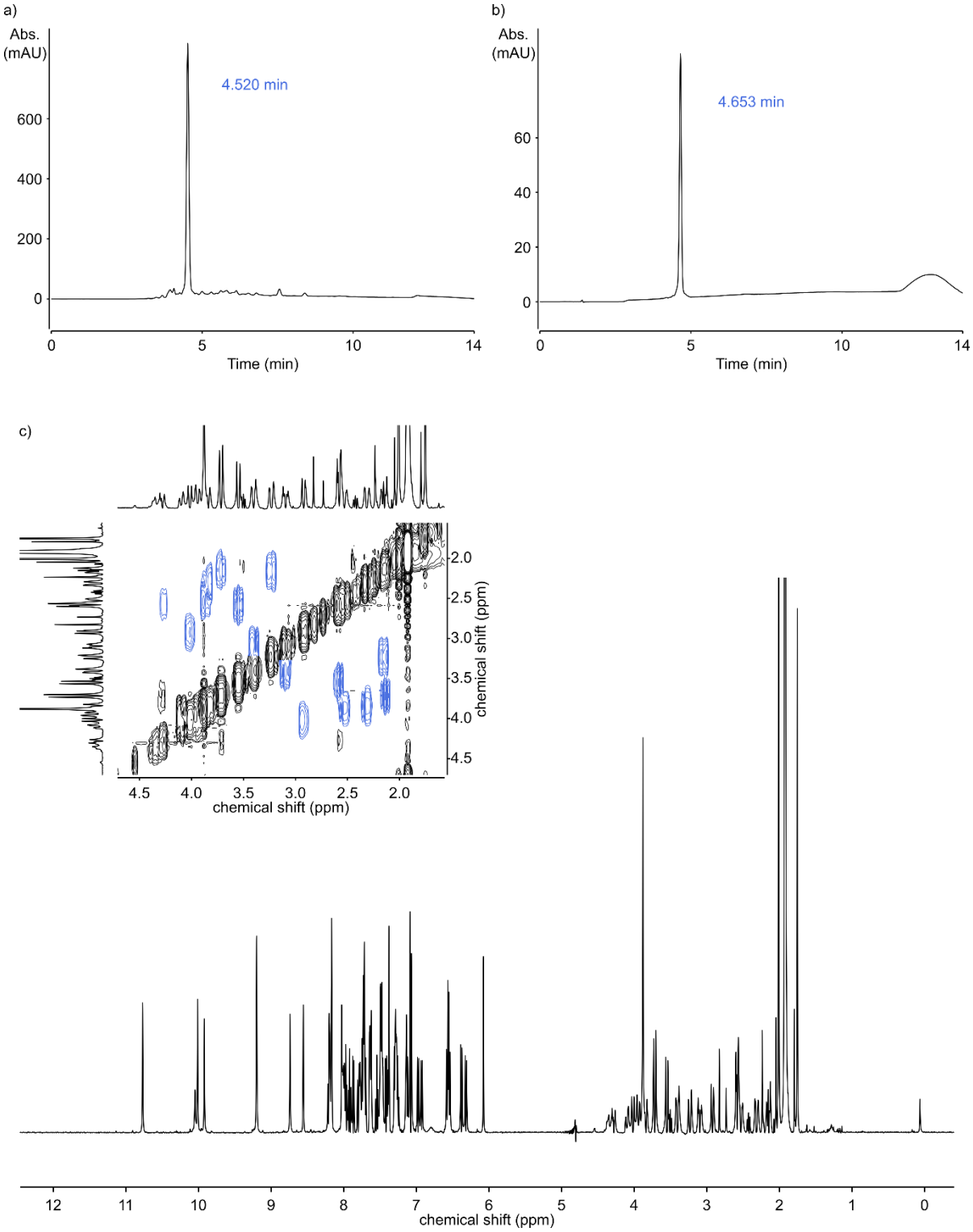


Fig. 32 Analytical data of compound **7**. HPLC chromatograms after cleavage from the resin (a) and after purification (b) (C18, 0–30B, 50 °C; A: 13mM NH_4OAc buffer pH 8.5, B: acetonitrile). c) ^1H NMR and part of the COSY NMR spectrum (500 MHz, 0.36 mM in 12 mM NH_4OAc buffer pH 8.5 H_2O/D_2O 9:1, H_2O suppression), some cross-signals of diastereotopic methylene groups are exemplary highlighted in blue.

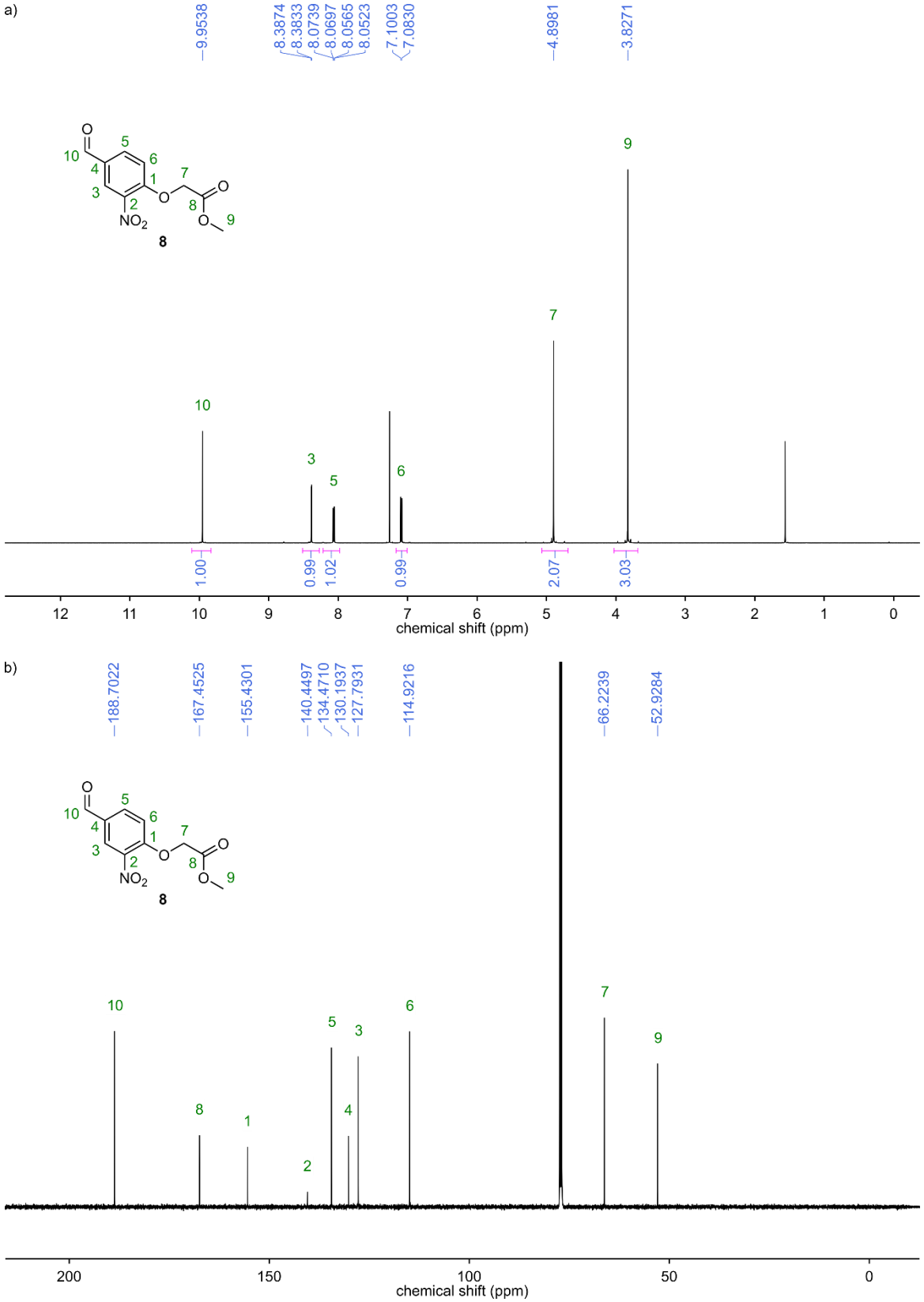


Fig. 33 NMR spectra of compound **8**. a) ^{1}H NMR (500 MHz, $CDCl_3$). b) ^{13}C NHR (126 MHz, $CDCl_3$).

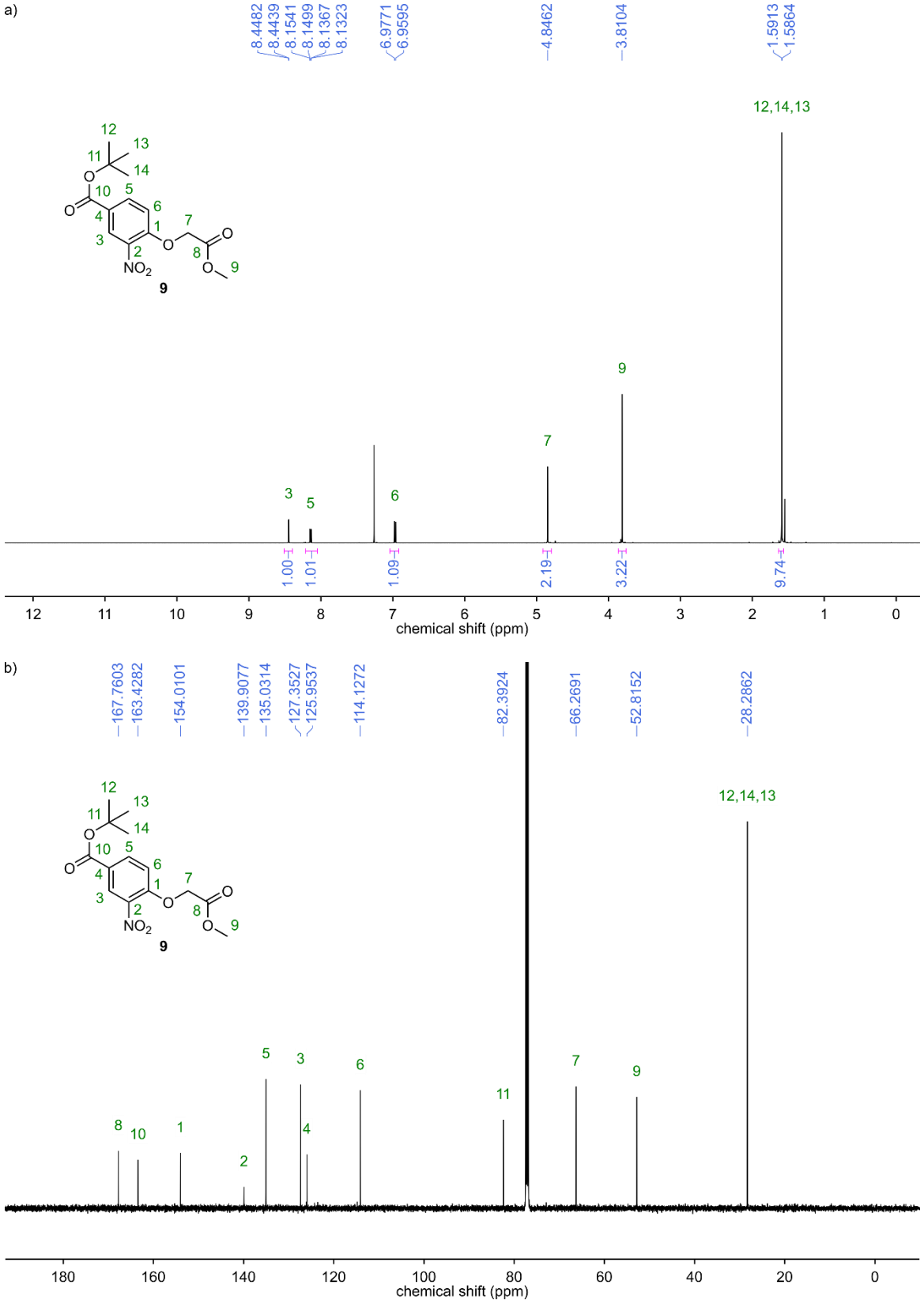


Fig. 34 NMR spectra of compound **9**. a) ^{1}H NMR (500 MHz, $CDCl_3$). b) ^{13}C NHR (126 MHz, $CDCl_3$).

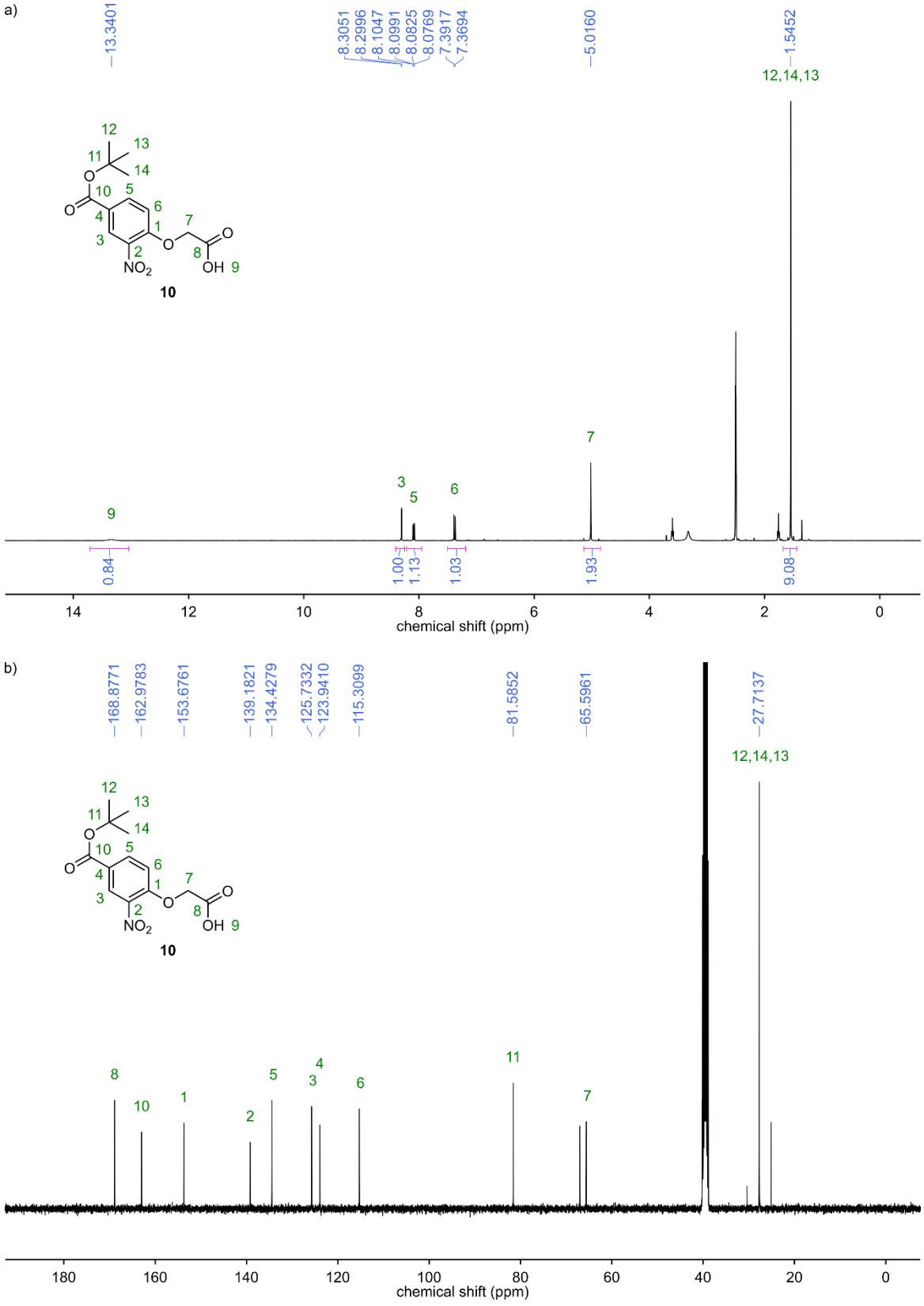


Fig. 35 NMR spectra of compound **10**. a) ^{1}H NMR (400 MHz, DMSO-d$_6$). b) ^{13}C NHR (101 MHz, DMSO-d$_6$).

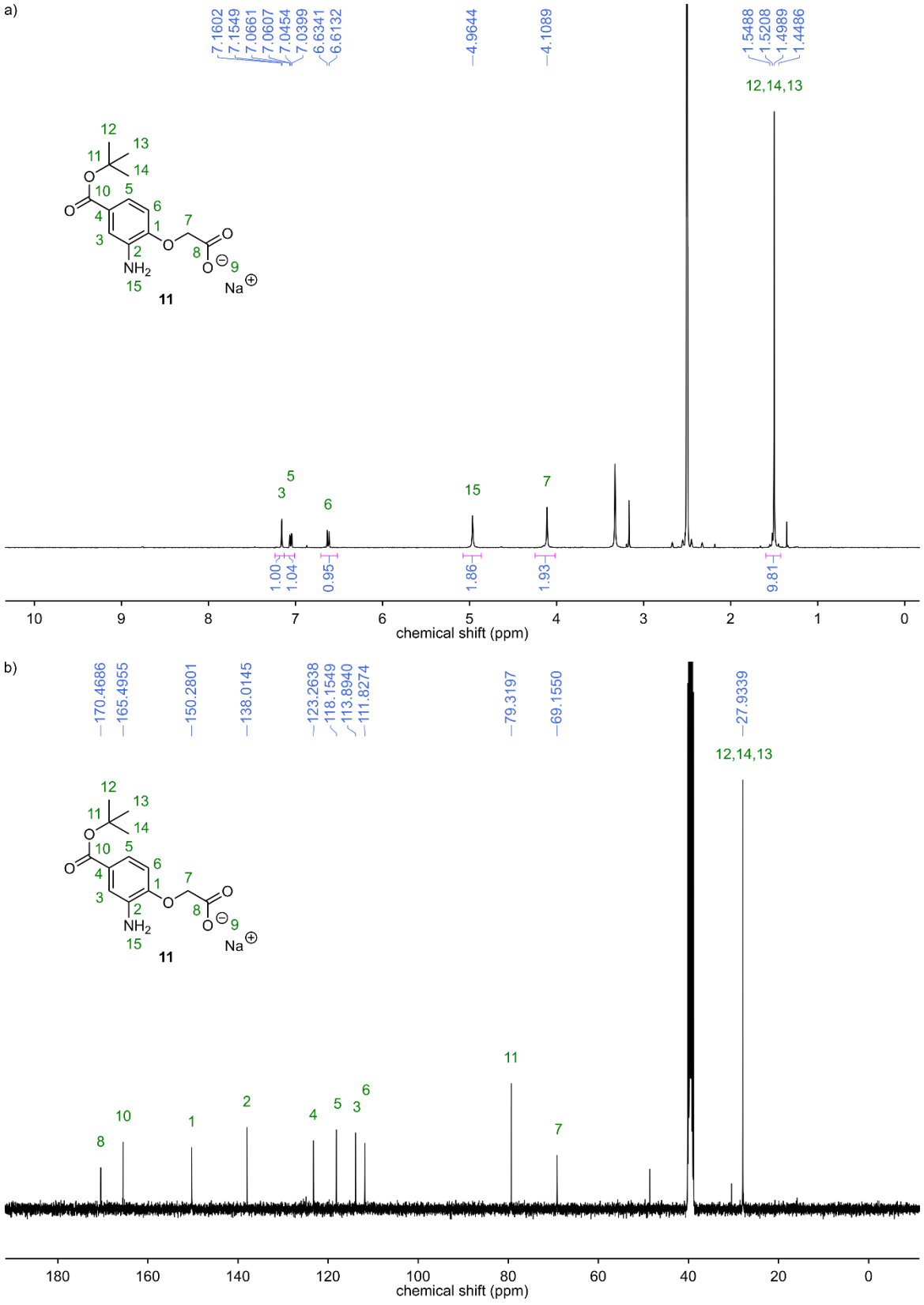


Fig. 36 NMR spectra of compound **11**. a) ^{1}H NMR (400 MHz, DMSO-d$_6$). b) ^{13}C NHR (101 MHz, DMSO-d$_6$).

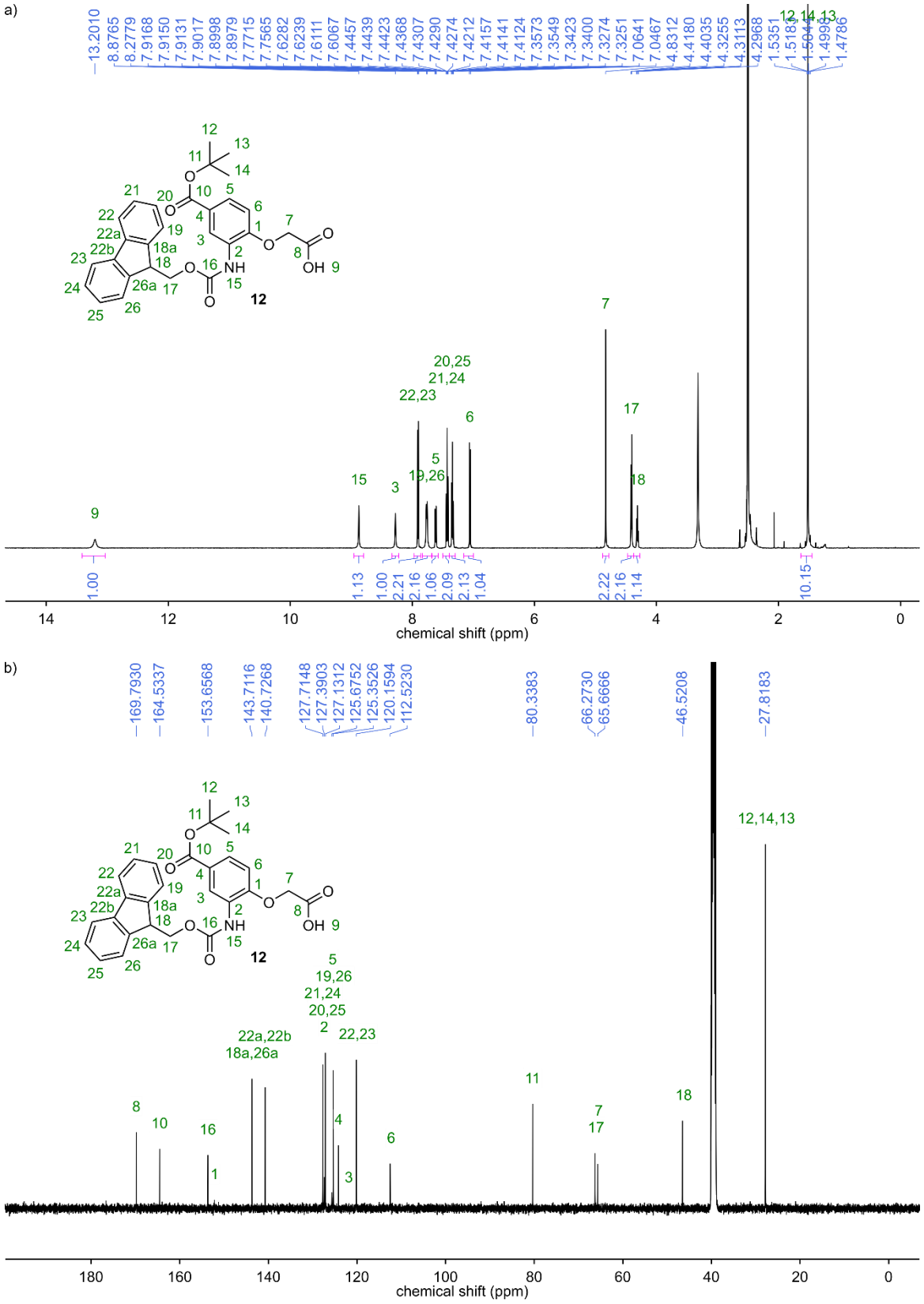


Fig. 37 NMR spectra of compound **12**. a) 1H NMR (500 MHz, DMSO-d_6). b) 13C NHR (126 MHz, DMSO-d_6).

5 Handedness control from within aromatic helices

Several C- and N-terminal moieties that fully bias handedness of Q-based foldamers towards *P*- or *M*-helicity are known. All of these chiral groups have the ability to participate in the quinoline backbone hydrogen bonding network. For example, in the crystal structure of a camphanic acid (Camph) functionalized Q-trimer, the ester oxygen is hydrogen bonded to the N-terminal quinoline amide hydrogen (Fig. 38a).[46a] In the case of the oxazolylaniline unit (Oxaz), the endocyclic nitrogen of the oxazolyl ring is hydrogen bonded to the C-terminal quinoline amide hydrogen (Fig. 38b).[46b] The interactions lock the chiral groups in a specific conformation, providing a large energy gain. With the opposite handedness, these conformations are not possible due to steric hinderance created by the bulky chiral residues, thus preventing the hydrogen bonding. This creates a sufficiently large energy difference between *P*- and *M*-helical diastereomers to shift the equilibrium beyond a 99/1 ratio.

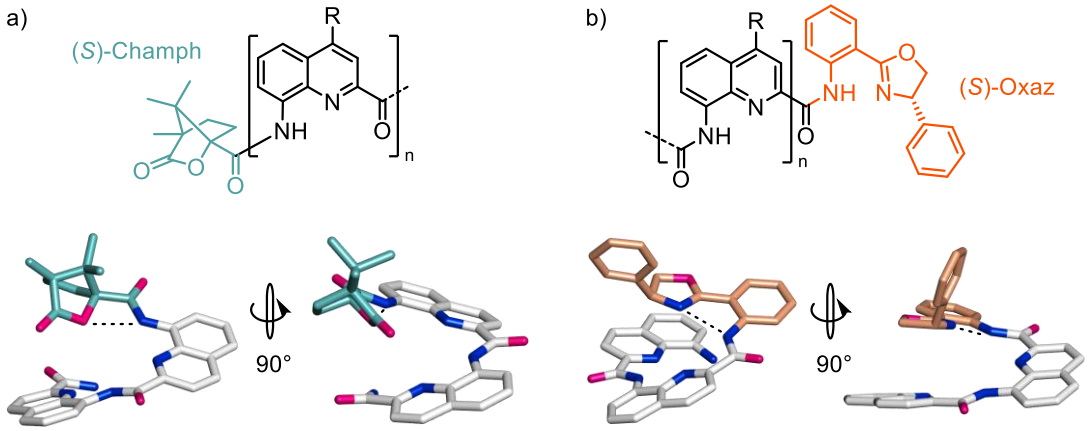


Fig. 38 – Terminal groups for quantitative handedness bias of Q-oligomers. a) Part of the crystal structure of an oligomer functionalized with (*S*)-Camph in front and side view.[46a] b) Part of the crystal structure of an oligomer functionalized with (*S*)-Oxaz in front and side view.[46b] Side chains and hydrogen atoms are omitted for clarity. Oxygen and nitrogen atoms are shown in purple and blue, respectively. The respective formula is given above the structures. Camph and Oxaz moieties are highlighted in turquoise and orange, respectively.

However, these groups also have the disadvantage of blocking either C- or N-terminus for further functionalization. This not only limits design options and influences the properties of the oligomers (such as solubility or binding affinities), but it may also hamper crystal packing, which often involves continuous C-C- and N-N-terminal stacking of the oligomers (see section 2.3, Fig. 6e). Without high resolution structural data from X-ray crystallography, design optimizations are difficult or impossible. The alternative of obtaining structural information via NMR spectroscopy is tedious and not always feasible. New moieties with the ability to fully bias handedness from withing the helix, without significantly altering the fold of the oligomers, are therefore desired. The B-monomer described in chapter 4 provides an ideal starting point: Since it does not drastically alter curvature,

it allows the utilization of different chiral *α*-hydroxy ester precursors (some of them being commercially available such as ethyl (−)-L-lactate) and even a combination with additional side chains on the aromatic ring.

The following results, which were published in *Chemical Communications*, show that a chiral CH_3-substituent on the methylene group of the B-monomer is enough to create a quantitative handedness bias in aqueous media, with (*S*)-chirality favoring *M*-helicity. The monomer not only shows a stronger bias than the previously used Camph-moiety, but it also lacks the disadvantages of blocking one of the termini and significantly altering the shape and solubility of the oligomers. Furthermore, the chiral B-monomer can be functionalized by installing side chains on the aromatic ring, or even by altering the chiral substituent; however, handedness induction would have to be reevaluated in the latter case. It also provides increased handedness inversion kinetics, which would be kinetically locked in oligomers consisting of more than five Q-units. Conclusively, this new method of handedness control is superior compared to previously described groups for foldamers intended to be used in aqueous systems.

Contributions: The ideas for this work were developed by me in collaboration with I. Huc. Experiments were conducted by me with the help of E. Heinemann, who did an internship with me at that time. Crystallization experiments were done in collaboration with P. K. Mandal, who also did the refinement of the obtained structure. X-ray data were collected by P. K. Mandal at EMBL Hamburg with assistance of Dr Saravanan Panneerselvam. The manuscript was written by me in collaboration with I. Huc and P. K. Mandal.

Publication

Quantitative helix handedness bias through a single H *vs.* CH_3 stereochemical differentiation

Authors: Daniel Bindl, Elisabeth Heinemann, Pradeep K. Mandal and Ivan Huc

Published: *Chem. Commun.*, **2021**, *57*, 5662–5665. DOI: 10.1039/d1cc01452h

Abstract: A novel chiral aromatic *δ*-amino acid building block was shown to fully induce handedness in quinoline oligoamide foldamers with the possibility of further increasing the bias by combining multiples of these units in the same sequence. Through its incorporation within the helix, both N- and C-termini are still accessible for further functionalization.

Handedness control in aromatic foldamer helices is of prime importance for their applications in circularly polarized luminescence,[62, 110] enantioselective catalysis,[111] and the diastereoselective recognition of chiral guests either within an internal cavity[112] or at their surfaces as, for example, proteins.[41, 59, 92, 96, 113] Promoting minimal handedness bias is straightforward: most stereogenic centers placed in the vicinity of the helix backbone will create an energy imbalance in favor of one or the other helix sense.[45, 114] Only 1 kJ mol^{-1} is required to elicit a 60/40 ratio. In contrast, quantitative handedness control, i.e. a diastereomeric excess of at least 99%, has rarely been achieved because it requires a large energy difference between *P*- and *M*-helical diastereomers. The few examples reported concern aromatic oligoamide foldamers derived from 8-amino-2-quinolinecarboxylic acid (Q_n) (Fig. 39a) possessing N- or C-terminal chiral moieties such as camphanyl (Camph),[46a] oxazolylaniline (Oxaz)[46b, 111] or β-pinene-derived pyridyl (Pin)[46c] groups (Fig. 39b). All these moieties incorporate stereogenic centers embedded within a cyclic system and form hydrogen bonds with a main chain amide. Terminal functionalization by a chiral moiety is often convenient, but it can hamper other functionalizations, e.g. with a chromophore, a protein ligand, in particular in the context of protein surface recognition.[41, 59, 92, 96] Here we show that 2-(2-aminophenoxy)-propionic acid monomer B^{Rme} (Fig. 40a) promotes quantitative handedness induction in water when embedded within an aromatic oligoamide helix sequence. Handedness bias is thus achieved through the stereochemical differentiation of a single main chain hydrogen atom and a methyl group. This strong effect illustrates that the compactness of aromatic helices allows for stronger stereochemical differentiation than in peptidic helices in which a single alanine residue is unable to achieve complete handedness bias of an otherwise achiral sequence.[115]

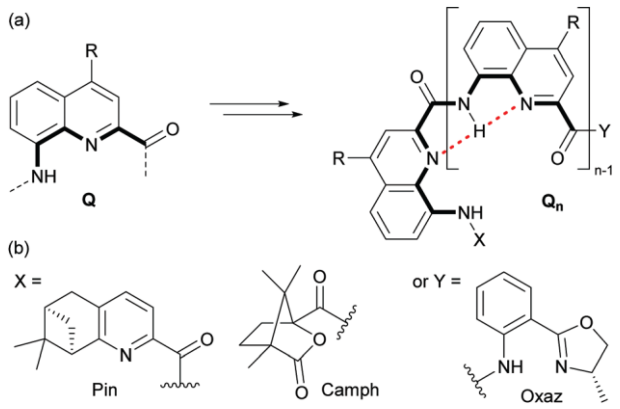


Fig. 39 (a) Chemical structures of Q and its oligomer Qn. Backbone hydrogen bonds are indicated by red dashed lines. (b) N-terminal groups (X) and C-terminal groups (Y), which are known to fully induce one handedness in Q_n oligomers.[46]

We speculated that placing a stereogenic center within an aromatic helix may result in strong helix handedness bias due to stereochemical constraints both above and below the chiral group. Nevertheless, a few earlier attempts failed to reach quantitative handedness control, perhaps due to the fact that the chiral groups were not themselves sufficiently helicogenic; that is, their presence may cause a partial helicity disruption.[116] Chiral monomer B^{Rme} was designed as a δ-peptidic analogue of Q bearing a stereogenic center at a position expected to be at the inner rim of the helix. Its achiral equivalent, 2-(2-aminophenoxy)-acetic acid, has already been described,[43a] and homomeric sequences of this monomer have been shown to fold not into a canonical aromatic helix, but into a herringbone-helical structure.[33a, 43a] However, it is also known that Q_n sequences can template the canonical helical folding of other monomers.[25, 94, 116d] We thus endeavored to prepare an Fmoc protected version of B^{Rme} and to incorporate it into Q_n oligomers to investigate its ability to bias helix handedness.

Two new monomers, Fmoc-B^{Rme}-OH and Fmoc-B^{Gly}-OH, were synthesized from *ortho*-nitrophenol in 91% and 36% overall yield in four steps, respectively (Fig. 42). The stereogenic center of B^{Rme} was installed by condensing ethyl (−)-L-lactate to 2-nitrophenol via a Mitsunobu reaction leading to an inversion of the stereochemistry. A derivatization of the final Fmoc-B^{Rme}-OH with a chiral amine confirmed that the enantiomeric purity of the starting alcohol as given by the supplier (≥97.5%) was preserved throughout the whole synthesis (section 5.2). An unanticipated difficulty had to be overcome during the preparation of Fmoc-B^{Rme}-OH and Fmoc-B^{Gly}-OH. The Mitsunobu product is a nitro ester that must undergo nitro group reduction and ester saponification prior to Fmoc installation. When performing reduction first, we found that the amino-ester readily cyclizes into a six-membered lactam. Even when saponification was carried out first, the amino-acid was quantitatively converted into the same lactam during nitro group hydrogenation (Fig. 42). This unusual reactivity might explain the scarce record of B oligomers in the literature. Two different approaches were applied to circumvent this side reaction. First, for B^{Gly}, the cyclisation was partially prevented by using a bulky *tert*-butyl ester in the reduction step. For B^{Rme}, saponification was carried out first and a base (Na_2CO_3) was added during the hydrogenation to produce the sodium

carboxylate salt. This entirely prevented lactam formation. These synthetic routes should be extendable to a variety of B analogs bearing different chiral functionalities and/or side chains on the aromatic ring.

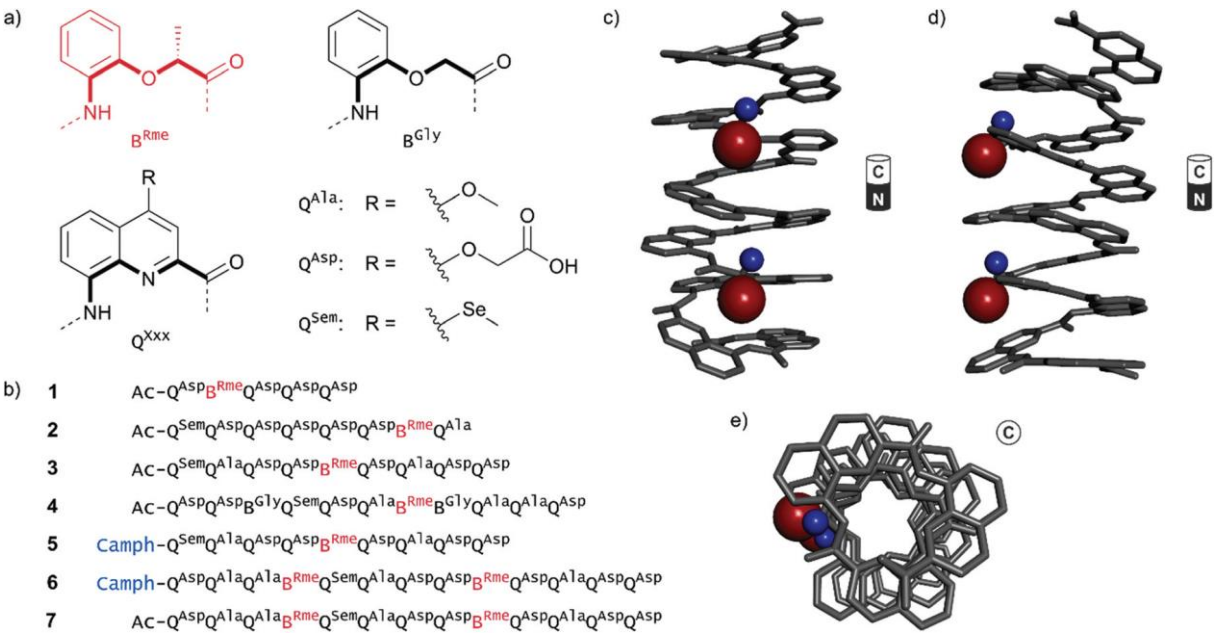


Fig. 40 (a) Foldamer building blocks used in this study. (b) Sequences that were synthesized to investigate the handedness induction properties of B^{Rme}. The chiral units (*S*)-Camph and B^{Rme} are shown in blue and red, respectively. (c–e) Crystal structure of oligomer **7** in side (c and d) and top (e) views showing a canonical aromatic helix structure. Side chains and hydrogen atoms are omitted for clarity. The methyl group (red) and hydrogen atom (blue) on the chiral center of B^{Rme} are represented as balls. The N- and C-terminus orientations are shown next to the respective structure.

Sequences **1**–**7** were designed and synthesized to study the effect of one or two B^{Rme} units placed at various positions in sequences of variable lengths (Fig. 40b). Acidic Q^{Asp} monomers were introduced to provide solubility in aqueous media. To prevent too high solubility that could hamper crystal growth, some Q^{Ala} units were also included in the sequences so as to be positioned on different faces of the helix. In some sequences, one Q^{Sem} moiety containing a selenium atom was included to eventually enable the use of anomalous X-ray scattering, though this proved to be unneeded. Synthesis was performed on low loading Wang resin (100–200 mesh) using previously reported solid phase foldamer synthesis (SPFS) protocols.[26] Fmoc acid building blocks were activated in situ by generating the respective acid chlorides prior to coupling. Both Fmoc-B^{Rme}-OH and Fmoc-B^{Gly}-OH showed excellent coupling reactivity. No noteworthy deletions occurred during SPFS. In a final step, the oligomers were cleaved from the resin and deprotected under acidic conditions. The crude oligomers were purified using semi-preparative reverse phase HPLC with a basic ammonium acetate buffer as the mobile phase. As a result, all compounds were obtained as their respective ammonium salts in good overall purified yields (8.4–51%; section 5.3.2).

As outlined in the introduction, we expected that the strong helicogenic nature of quinoline units would force the B^{Rme} and B^{Gly} monomers into canonical helical folding despite their steric demand

and increased flexibility that results in a herringbone helix for $(B^{Gly})_n$.[43a] NMR and circular dichroism (CD) spectroscopic data suggest that **1**–**7** all adopt a helical fold in water. They show the characteristic pattern of distinct amide and aromatic signals in their ^{1}H NMR spectra and an intense CD band in the 300–450 nm region that is typical for helically folded quinoline oligomers with some handedness bias[46a, 117] (Fig. 41 and Fig. 43). These findings are corroborated by the strongly downfield shifted ^{1}H NMR signals (between –0.18 and –0.33 ppm) of the CH_3 protons of B^{Rme} units in the spectra of **3**–**7** (Fig. 43), as a result of the ring current effects of neighbouring Q units. For comparison, the signal of the same methyl group in Fmoc-B^{Rme}-OH is found at 1.56 ppm. Similar upfield shifts are observed for protons close to aromatic side chains in proteins.[118]

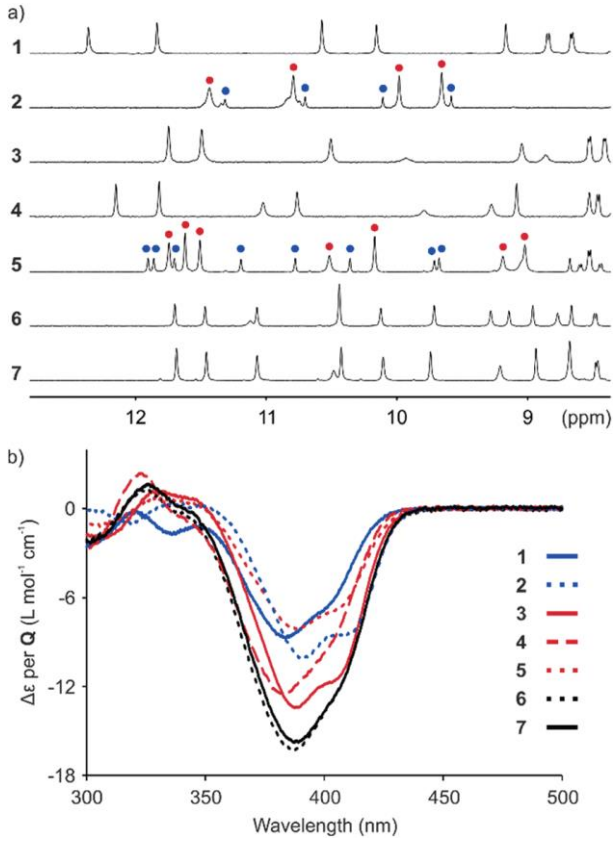


Fig. 41 (a) Amide region of the 1H NMR spectra of sequences **1**–**7** in NH_4OAc buffer pH 8.5. For **2** and **5**, major and minor sets of signals are marked with red and blue circles, respectively. (b) CD spectra of compounds **1**–**7** in NH_4OAc buffer pH 8.5 between 300 and 500 nm. The molar extinction (Δε) is normalized for the number of Q units for better comparability.

The helix conformation was also validated by the solid state structure of **7**, an analogue of **6** bearing an acetyl group instead of (*S*)-Camph (Fig. 40c–e). The structure shows canonical helical folding, with only left-handed helices present in the crystal. This structure was solved despite the low resolution of the data (2.86 Å) by molecular replacement using an energy minimized model (section 5.4). Molecular replacement is a common phasing method in protein crystallography where models are produced from related structures, but it has scarcely been used for smaller molecules.[73] Successful molecular replacement in the case of **7** is a highlight of the effectiveness of molecular

modelling at accurately predicting aromatic foldamer conformations, and represents an important methodological advance.

Helix handedness bias in solution was first assessed when B^{Rme} was placed in the penultimate position to the N- or C-terminus of a sequence, as in **1** and **2**. Since one helix turn contains 2.5 units, the penultimate residue is entirely exposed to the solvent on one of its faces. The ^{1}H NMR spectrum of **1** shows one major set of signals (Fig. 41a), indicating a diastereomeric ratio of (*R*)-*M* to (*R*)-*P* of at least 98/2, assuming that the helix handedness inversion is slow on the NMR timescale, as expected for a pentamer or any longer sequence.[41] The negative band observed by CD in the 300–450 nm region shows that the *M* helicity is dominant.[117] In the *M* helix, the asymmetric methyl group of **1** should 'stick out' of the helix towards the solvent. For **2**, the ^{1}H NMR spectrum shows two sets of signals with a ratio of about 87/13 (Fig. 41a). The CD spectrum indicates that the *M* helix is again dominant (Fig. 41b), which means that the asymmetric methyl this time points towards the helix and not to the solvent, since B^{Rme} is near the C terminus. The handedness preference thus depends on the absolute stereochemistry of B^{Rme} regardless of its position in the sequence, and not on whether the methyl groups point towards the helix or the solvent. We also note that the CD intensity does not correlate well with the diastereomeric excess. The CD band at 380 nm normalized per Q unit is more intense for **2** than for **1** although the handedness bias is stronger for the latter (Fig. 41b). Indeed, the CD intensity also depends on the number of consecutive Q units and on the nature of substituents on each Q unit.

Compounds **3** and **4** contain one B^{Rme} unit flanked with helix segments that span more than one turn. The asymmetric center should thus have close contacts with aryl groups both above and below. Sequence **4** was designed to be more flexible than **3**, due to additional achiral B^{Gly} units, including one adjacent B^{Rme}. The higher flexibility arises from the reduced aromatic stacking surface and additional rotatable bonds in B monomers as compared to Q. The ^{1}H NMR spectra of **3** and **4** both show a single set of sharp signals (Fig. 41a). This indicates quantitative handedness bias as far as NMR can detect. CD shows that **3** and **4** are *M* helical (Fig. 41b), as for **1** and **2**. Quantitative handedness bias achieved by the stereochemical differentiation between a hydrogen atom and a simple methyl group is remarkable and unprecedented. It probably results from the very compact conformation of Q_n helices that create a large energy difference between the diastereomeric conformers. Furthermore, NMR and CD concur to show that the handedness bias for compound **3** is also quantitative in DMSO and MeOH (Fig. 45).

Encouraged by these results, we challenged handedness bias due to B^{Rme} through the introduction of an N-terminal camphanyl group having an (*S*) configuration, that is, a configuration antagonistic to that of B^{Rme}. In the absence of B^{Rme}, the camphanyl group also biases handedness quantitatively and its (*S*) configuration favors *P* helicity.[46a] Sequence **5** is an analog of **3** where the terminal Ac has been replaced with (*S*)-Camph. Its ^{1}H NMR spectrum shows two sets of signals with a ratio of 75/25 corresponding to the presence of two different diastereomeric conformations, (*S*,*R*)-*M* and (*S*,*R*)-*P* (Fig. 41a). The CD spectrum shows a negative band in the range of 300–450 nm; this

indicated that the major conformation has *M* handedness, and thus that B^{Rme} imparts a stronger handedness induction than (*S*)-Camph (Fig. 41b). Because **3** and **5** contain the same sequence of chromophores, their CD spectra should be directly comparable. The relative CD intensities of **5** and **3** at their maxima around 385 nm indicate an 80/20 ratio[†] of *M*- to *P*-diastereomers for **5** (Fig. 44), which matches the ratio observed by ^{1}H NMR. Taking this ratio into consideration, an energy difference between the handedness bias induced by B^{Rme} and Camph of about −3.4 kJ mol^{-1} can be derived.

When two B^{Rme} units cooperate to bias handedness in the same sequence, as in **6**, we find that the effect of a terminal (*S*)-Camph is completely reversed. A main species is observed by ^{1}H NMR, and CD confirms *M* handedness (Fig. 41). The CD intensity of **6** also matches that of **7**, which lacks the camphanyl group. Minor signals in the ^{1}H NMR spectra of **6** and **7** were observed that can be assigned to the incomplete enantiomeric purity of B^{Rme} arising from the enantiomeric purity of the lactate precursor. A small amount of one or the other chiral B unit may have (*S*) stereochemistry opposite to the (*R*) configuration of B^{Rme}. In these cases, the effect of B^{Rme} and that of B^{Sme} would cancel each other and the camphanyl group would favor *P* helicity, leading to small amounts of (*S*,*R*,*S*)-*P* and (*S*,*S*,*R*)-*P* diastereomeric conformers of **6** and (*S*,*R*)-*P*/*M* and (*R*,*S*)-*P*/*M* conformers of **7**, where the bias of Camph is missing.

In conclusion, quantitative handedness bias was achieved in water, methanol, and DMSO by placing the new B^{Rme} monomer within a quinoline helix. The energy difference was greater than that generated by the Camph group. The bias could be further enhanced by incorporating more than one B^{Rme} unit within the same helix. The handedness bias was not complete only when B^{Rme} was placed near the C-terminus. Full handedness control could thus be achieved without any modifications at either the N- or C-terminus, allowing for further functionalization at both ends of the helix. Being able to avoid bulky handedness inducing groups at the N- or C-terminus will also be beneficial to water solubility, as Camph, Pin, and Oxaz are all lipophilic. These combined features will be useful for protein surface recognition using helical foldamers. B units also provide a new means to introduce side chains at the stereogenic center. This prospect and the effect of multiple B units on the helix geometry are being investigated and will be reported in due course.

Acknowledgements: Synchrotron data were collected at beamline P14 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany). We thank Dr Saravanan Panneerselvam for his assistance in using the beamline.

Conflicts of interest: There are no conflicts to declare.

[†] The total CD intensity results from the sum of the CD signals of *P*- and *M*-isomers, which have opposite signs. Therefore, the 40% lower CD intensity observed for **5** in comparison to **3** reflects an 80/20 ratio of *M* to *P*-isomers.

Supplementary Information

for:

Quantitative helix handedness bias through a single H *vs.* CH_3 stereochemical differentiation

Daniel Bindl, Elisabeth Heinemann, Pradeep Kumar Mandal, and Ivan Huc

List of Abbreviations

AcOH	acetic acid
CD	circular dichroism
CyHex	Cyclohexane
DCM	dichloromethane
DIAD	diisopropyl azodicarboxylate
DIPEA	*N*,*N*-diisopropylethylamine
DMF	*N*,*N*-dimethylformamide
DMSO	dimethyl sulfoxide
EI	electron ionization
ESI	electrospray ionization
EtOAc	ethylacetate
Fmoc	fluorenylmethoxycarbonyl
HMBC	heteronuclear multiple bond correlation
HMQC	heteronuclear multiple quantum correlation
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
MeOH	methanol
MW	molecular weight
NMR	nuclear magnetic resonance
RP	reversed phase
SPFS	solid phase foldamer synthesis
TEA	triethylamine
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TIPS	triisopropyl silane
TLC	thin layer chromatography
TMSP	3-(trimethylsilyl)propionic-2,2,3,3-d_4 acid sodium salt
UV/Vis	ultraviolet–visible

5.1 Supplementary Figures

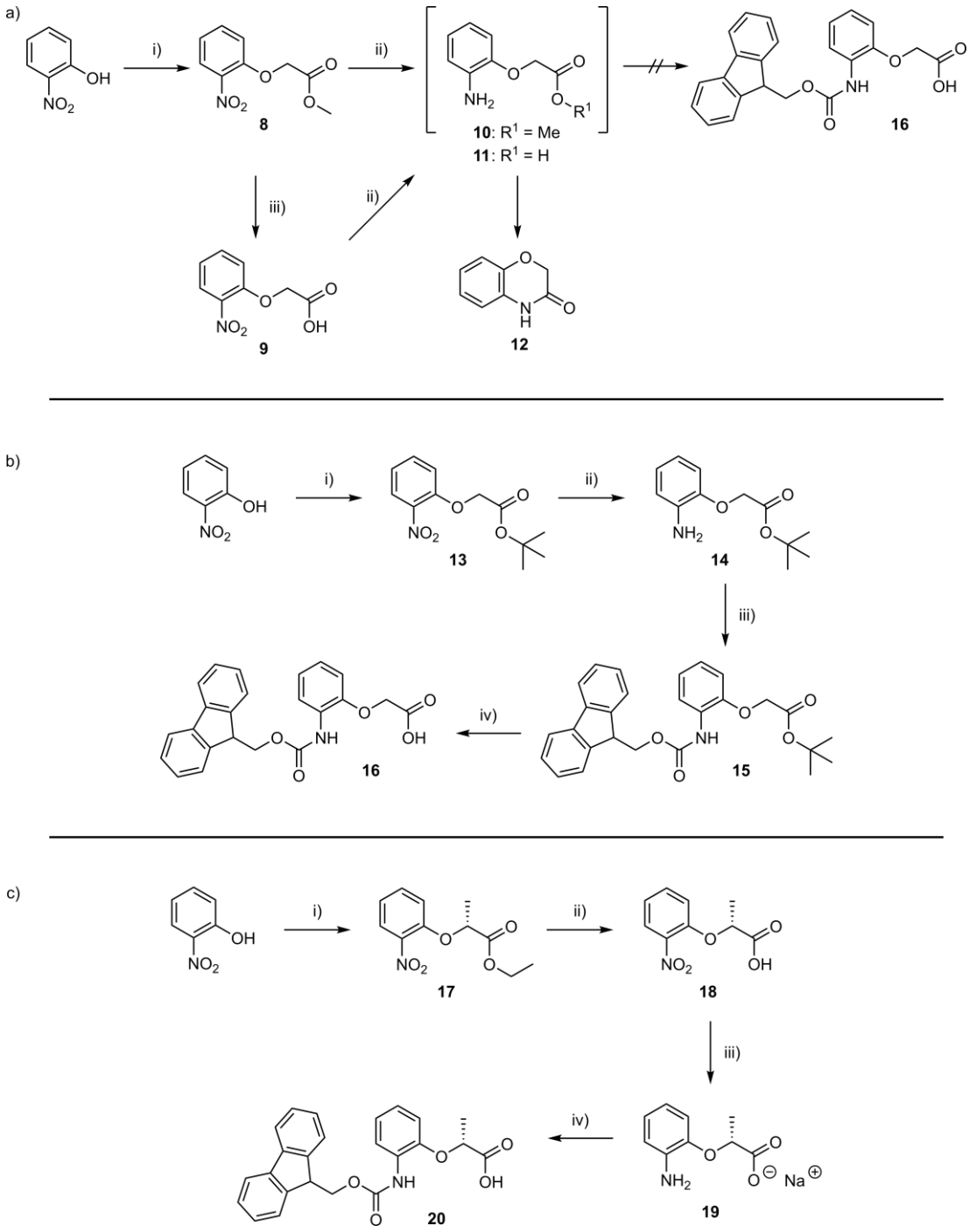


Fig. 42 (a) Failed attempts to synthesize Fmoc-B^{Gly}-OH: i) methyl bromoacetate, K_2CO_3, acetone. ii) H_2, Pd/C, EtOAc or THF iii) LiOH, H_2O, THF. (b) Synthesis route to Fmoc-B^{Gly}-OH: i) tert-butyl bromoacetate, K_2CO_3, acetone. ii) H_2, Pd/C, ethanol. iii) Fmoc-Cl, DIPEA, DCM. iv) TIPS, TFA, DCM. (c) Synthesis route to Fmoc-B^{Rme}-OH: i) (-)-ethyl L-lactate, PPh_3, DIAD, THF. ii) LiOH, H_2O, THF. iii) H_2, Pd/C, Na_2CO_3, MeOH. iv) Fmoc-Cl, $NaHCO_3$, H_2O, 1,4-dioxane. For detailed synthetic procedures see section 5.3.3.

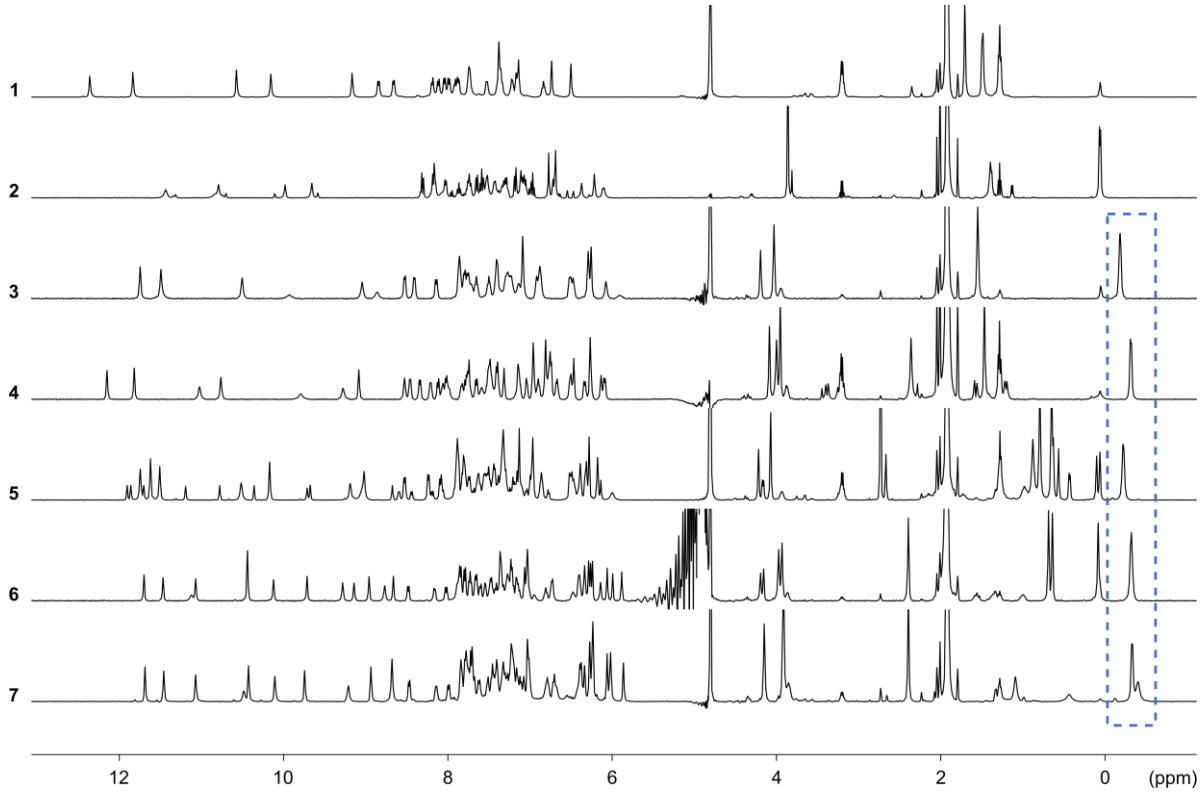


Fig. 43 ^{1}H NMR spectra of oligomers **1**–**7** in 12 mM NH_4OAc buffer with water suppression (**1**: 0.13 mM, **2**: 0.5 mM, **3**: 0.39 mM, **4**: 0.1 mM, **5**: 0.37 mM, **6**: 0.54 mM, **7**: 0.32 mM). The chiral CH_3 groups of B^{Rme} in oligomers **3**–**7** are highlighted with a dashed rectangle.

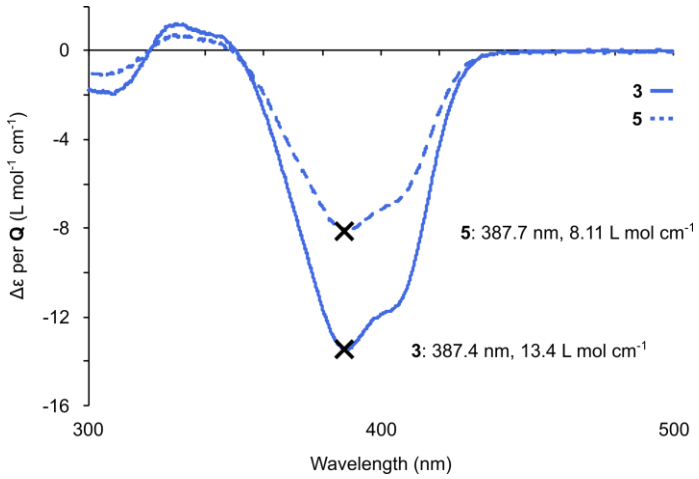


Fig. 44 CD spectra of oligomers **3** and **5**. Maxima are marked with crosses and their respective x and y values are displayed.

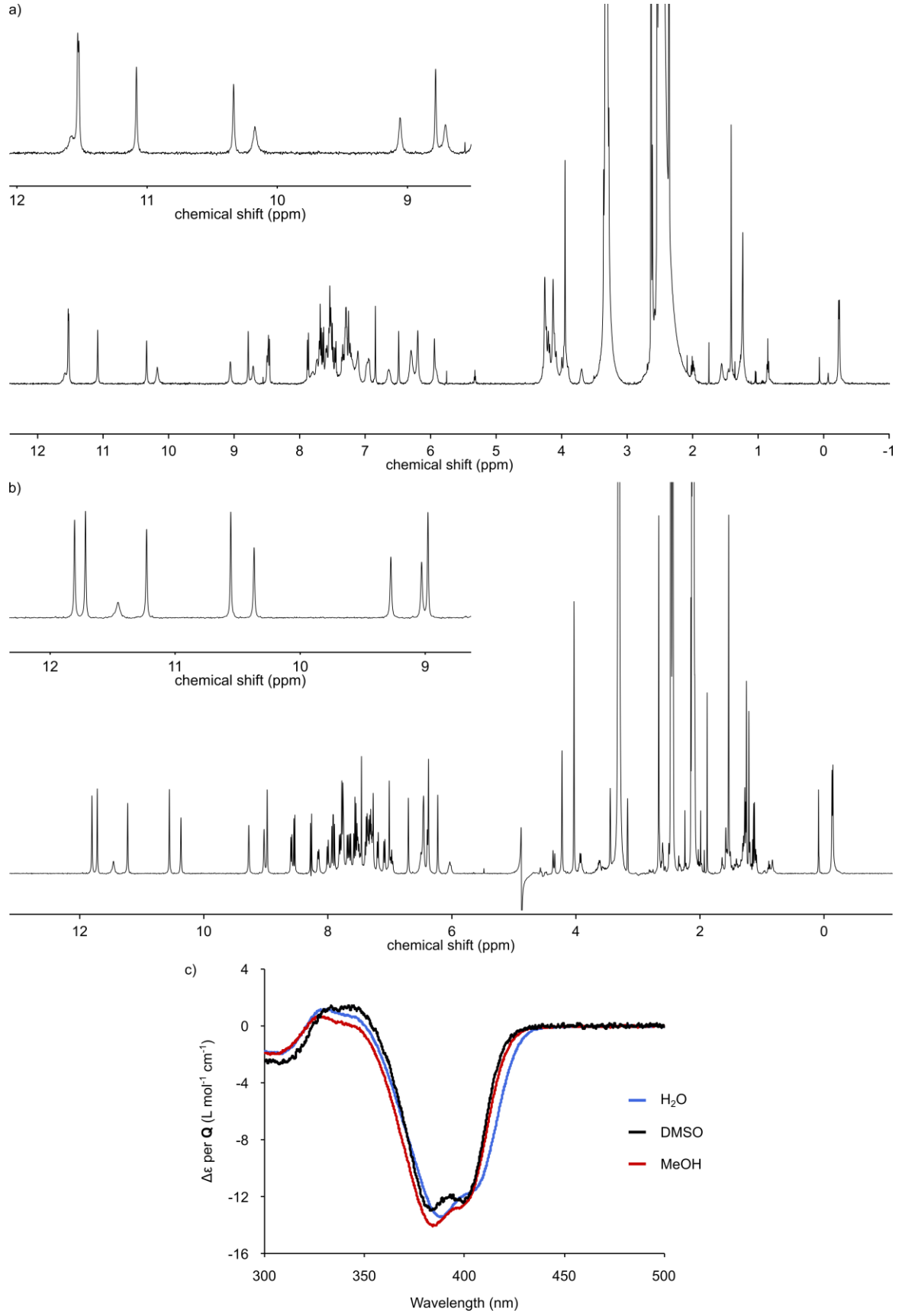


Fig. 45 ^{1}H NMR spectra of oligomer **3**: a) 0.4 mM (500 MHz, DMSO-d_6) and b) 0.4 mM (500 MHz, MeOH-d_3, OH-suppression). A zoom of the amide region is shown above the spectra, respectively. c) CD spectra of oligomer **3**: 22.6 µM in H_2O (blue), 6.85 µM in DMSO (black) and 19.3 µM in MeOH (red).

5.2 Chiral derivatization experiment

The chiral monomer Fmoc-B^{Rme}-OH was coupled to the chiral amine **21** to test if racemization occurred during its synthesis (Fig. 46). Four different products can be expected for this reaction and their ratios are determined by the optical purity of the starting materials. Therefore, knowing the enantiomeric purity of **21** (≥99%), the enantiomeric purity of **20** can be estimated.

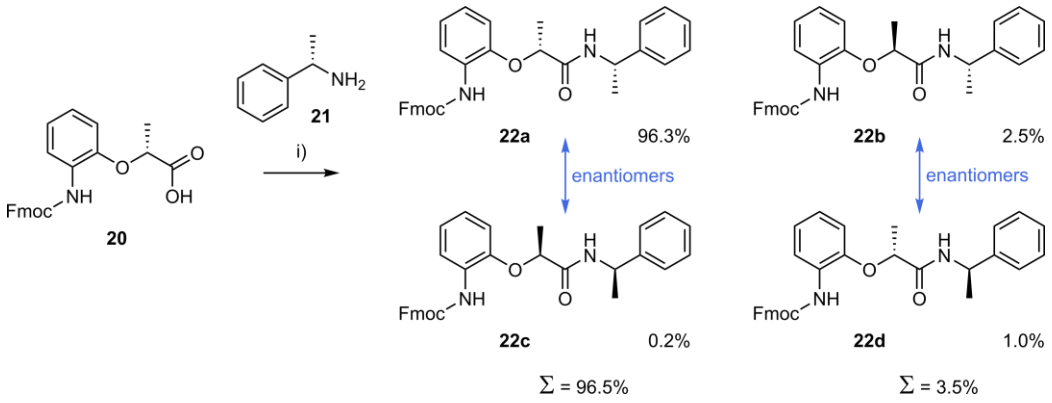


Fig. 46 Reaction scheme for the coupling of monomer **20** to the chiral amine **21** (enantiomeric purity: ≥99%). i) Ghosez's reagent, THF. The expected ratios of products **22a–d**, based on the enantiomeric purities given by the supplier, are shown next to the respective structures.

HPLC analysis of the chiral derivatisation crude shows that two diastereomeric products were formed with a ratio of 98/2, even exceeding the expected ratio of 96.5/3.5 (Fig. 47). This shows that the initial enantiomeric ratio of the chiral precursor of **20** (given by the supplier as ≥97.5/2.5) was preserved throughout its synthesis.

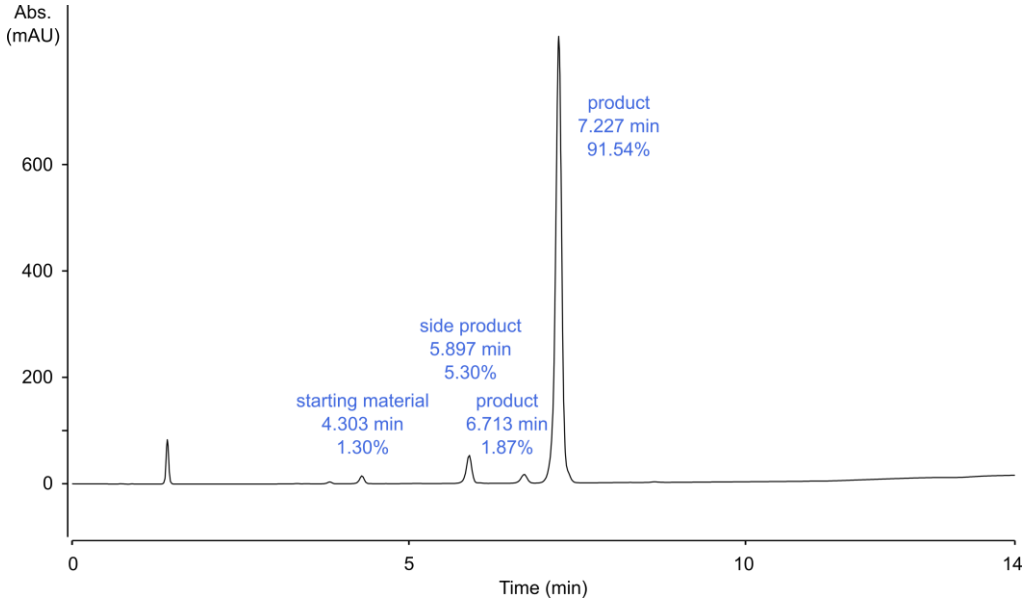


Fig. 47 HPLC analysis (C18, 55–80B, 25 °C; A: H_2O + 0.1% TFA, B: acetonitrile + 0.1% TFA) of the chiral derivatisation of **20**. Peaks were assigned by isolation via semi-prep HPLC (C18, 55–80B, 25 °C; A: H_2O + 0.1% TFA, B: acetonitrile + 0.1% TFA) followed by mass analysis. Assignments, retention times and relative areas are given next to the respective peaks.

5.3 Materials and Methods

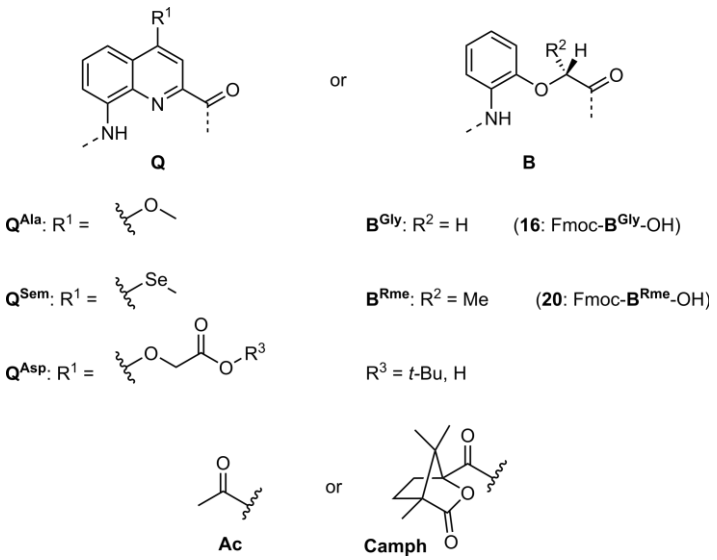


Fig. 48 Summary of monomers and N-terminal groups used in SPFS. Fmoc-Q^{Ala}-OH and Fmoc-Q^{Asp}(*t*-Bu)-OH have been described previously.[26a, 96] The synthesis of Fmoc-Q^{Sem}-OH will be published elsewhere. Fmoc-B^{Gly}-OH (**16**) and Fmoc-B^{Rme}-OH (**20**) are described here (section 5.3.3).

5.3.1 General

Commercial reagents (Suppliers: Abcr, Fisher Scientific, Merck, Sigma-Aldrich, TCI or VWR) were used without further purification unless otherwise stated. Wang resin LL (100–200 mesh) was purchased from Novabiochem. Peptide grade *N*,*N*-dimethylformamide (DMF) was purchased from Carlo Erba. Anhydrous chloroform, triethylamine (TEA) and *N*,*N*-diisopropylethylamine (DIPEA) 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 CryProbeTM Prodigy. Measurements were performed at 25 °C unless stated otherwise. Water suppression was performed with excitation sculpting. 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.[97] 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-NagelNucleodur 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). UV absorbance was monitored at 300 nm if not stated otherwise. Simple ultraviolet–visible (UV/Vis) absorbance measurements were done with a Thermo Fisher Scientific Nanodrop One instrument using a 1 cm quartz cuvette. Circular dichroism (CD) spectra were measured on a Jasco J-810 spectrometer. 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.

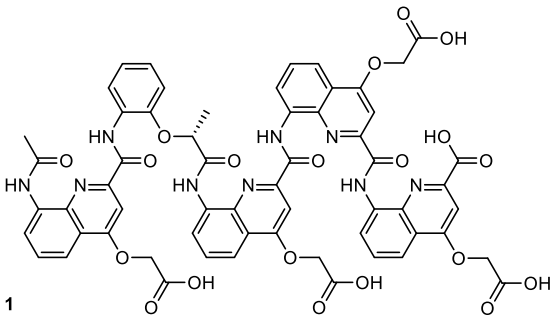
5.3.2 Solid Phase synthesis procedures

Oligomers were synthesized according to previously reported SPFS protocols,[26] hereafter referred to as standard method. Fmoc acid building blocks were activated *in situ* by generating the respective acid chlorides prior to coupling.

Capping: The resin (1.0 equiv.) was washed with DCM (3×) and incubated in acetic anhydride in DCM (50% *v/v*) for 10 min. Then, the resin was washed with DCM (2×), DMF (3×) and kept suspended in DMF (if stored longer than 24 h, it was kept at 4 °C).

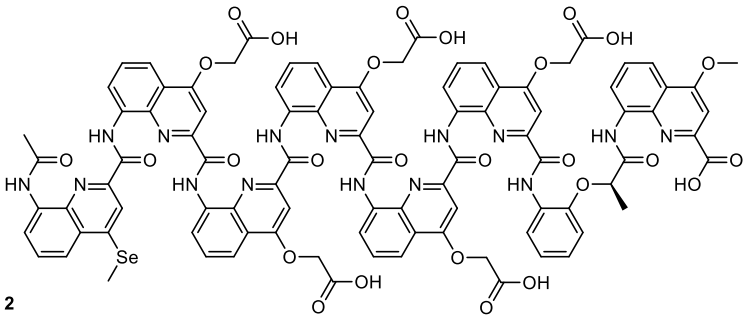
Acetylation: In the microwave vessel: after the resin (1.0 equiv.) was washed with anhydrous THF (4×), DIPEA (10.0 equiv.) and acetyl chloride (5.0 equiv.) in anhydrous THF (1 ml per 100 mg resin; not less than 2 ml) were added and the suspension was heated to 50 °C for 15 min (25 W, ramp to 50 °C over 5 min, hold at 50 °C for 15 min). The resin was washed with anhydrous THF (3×) and the coupling step was repeated once. Then, the resin was washed again with anhydrous THF (1×) and DMF (5×), and kept suspended in DMF (if stored longer than 24 h, it was kept at 4 °C).

Camphanylation: In the microwave vessel: after the resin (1.0 equiv.) was washed with anhydrous THF (4×), DIPEA (10.0 equiv.) and camphanic chloride (5.0 equiv.) in anhydrous THF (1 ml per 100 mg resin; not less than 2 ml) were added and the suspension was heated to 50 °C for 15 min (25 W, ramp to 50 °C over 5 min, hold at 50 °C for 15 min). The resin was washed with anhydrous THF (3×) and the coupling step was repeated once. Then, the resin was washed again with anhydrous THF (1×) and DMF (5×), and kept suspended in DMF (if stored longer than 24 h, it was kept at 4 °C).



1

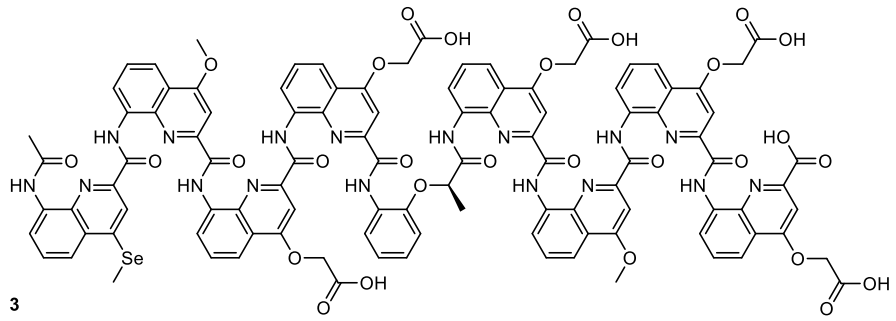
Compound 1: Oligomer **1** was synthesized on Wang resin (0.37 mmol g^{-1}, 9.25 µmol scale) according to the standard method. Loading of the first monomer: 0.21 mmol g^{-1} (56%). Unreacted amines were capped using the general capping method after each coupling step. The final acetyl group was installed via the general acetylation method. After purification by semi-prep HPLC (C8, 0–30B, 25 °C; A: 13 mmol NH_4OAc buffer pH 8.5, B: acetonitrile), the title compound was obtained as a white solid (0.937 mg, 0.781 µmol, 8.4%; HPLC-purity: 98.6%). **^{1}H NMR** (500 MHz, 12 mmol NH_4OAc buffer pH 8.5 in H_2O/D_2O 9:1): δ = 12.35 (s, 1H), 11.82 (s, 1H), 10.56 (s, 1H), 10.14 (s, 1H), 9.15 (s, 1H), 8.83 (d, *J* = 8.79 Hz, 1H), 8.65 (d, *J* = 8.55 Hz, 1H), 8.18 (d, *J* = 9.16 Hz, 1H), 8.10 (d, *J* = 9.47 Hz, 1H), 8.03 (d, *J* = 9.24 Hz, 1H), 7.97 (d, *J* = 9.32 Hz, 1H), 7.94–7.82 (m, 2H), 7.78–7.66 (m, 2H), 7.51 (d, *J* = 8.97 Hz, 1H), 7.42–7.27 (m, 4H), 7.20 (d, *J* = 8.08 Hz, 1H), 7.18–7.07 (m, 2H), 6.82 (t, *J* = 8.64 Hz, 1H), 6.72 (s, 1H), 6.49 (s, 1H), 3.19 (q, *J* = 8.26 Hz, 2H), 1.69 (s, 4H), 1.48 (d, *J* = 6.62 Hz, 4H), 1.27 (t, *J* = 8.06 Hz, 4H). **HRMS** (ESI^-) *m/z* calcd. for $C_{59}H_{44}N_9O_{20}$: 1198.2708 $(M-H)^-$; found: 1198.3269.



2

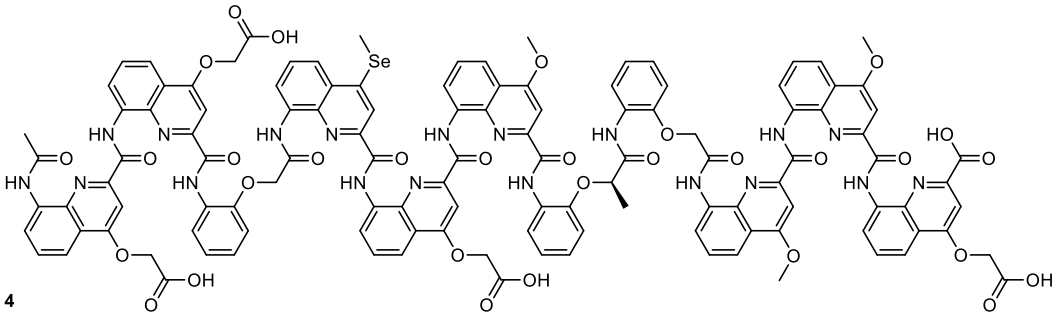
Compound 2: Oligomer **2** was synthesized on Wang resin (0.44 mmol g^{-1}, 17.6 µmol scale) according to the standard method. Loading of the first monomer: 0.44 mmol g^{-1} (100%). The final acetyl group was installed via the general acetylation method. After purification by semi-prep HPLC (C8, 0–20B, 25 °C; A: 13 mmol NH_4OAc buffer pH 8.5, B: acetonitrile), the title compound was obtained as a white solid (17.2 mg, 9.03 µmol, 51%; HPLC-purity: 97.3%). **^{1}H NMR** (500 MHz, 12 mmol NH_4OAc buffer pH 8.5 in H_2O/D_2O 9:1): δ = 11.44 (s, 1H), 11.35 (s, 1H), 11.32 (s, 1H), 10.93–10.73 (m, 1H), 10.70 (s, 1H), 10.11 (s, 1H), 9.98 (s, 1H), 9.66 (s, 1H), 9.58 (s, 1H), 8.37–8.25 (m, 2H), 8.25–8.12 (m, 3H), 8.12–7.98 (m, 2H), 7.95 (t, *J* = 8.56 Hz, 1H), 7.87 (t, *J* = 8.96 Hz, 1H), 7.81 (d, *J* = 9.15 Hz, 1H), 7.79–7.69 (m, 2H), 7.65 (d, *J* = 8.37 Hz, 1H), 7.59 (t, *J* = 9.10 Hz, 1H), 7.53 (d, *J* = 8.32 Hz, 1H), 7.48–7.39 (m, 1H), 7.37 (d, *J* = 8.52 Hz, 1H), 7.36–7.24 (m, 2H), 7.19 (s, 1H), 7.17 (s, 1H), 7.15–7.02 (m, 3H), 7.01 (s, 1H), 6.97 (t, *J* = 8.54 Hz, 1H), 6.77 (s, 2H),

6.72 (s, 1H), 6.69 (s, 2H), 6.64 (d, *J* = 9.48 Hz, 1H), 6.55 (s, 1H), 6.47 (s, 1H), 6.37 (s, 1H), 6.28 (s, 1H), 6.22 (s, 1H), 6.16–6.02 (m, 1H), 3.86 (s, 5H), 3.81 (s, 1H), 1.39 (d, *J* = 9.70 Hz, 5H), 1.13 (d, *J* = 7.20 Hz, 1H), 0.06 (d, *J* = 6.60 Hz, 6H). **HRMS** (ESI⁻) *m/z* calcd. for $C_{93}H_{68}N_{15}O_{27}Se$: 1906.3580 $(M-H)^-$; found: 1906.4064.



3

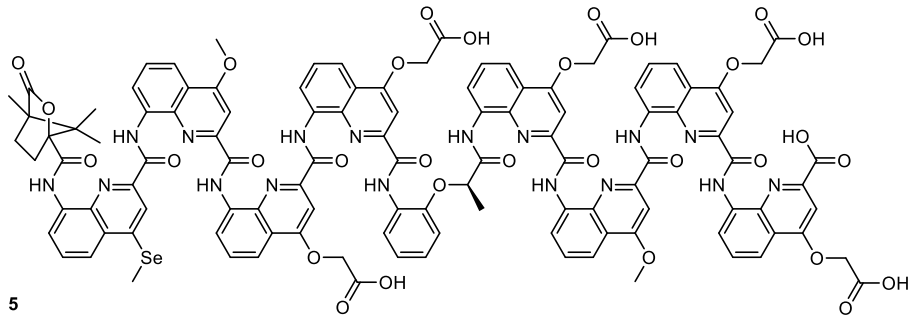
Compound 3: Oligomer **3** was synthesized on Wang resin (0.37 mmol g^{-1}, 22.0 µmol scale) according to the standard method. Loading of the first monomer: 0.34 mmol g^{-1} (92%). Unreacted amines were capped using the general capping method after each coupling step. The final acetyl group was installed via the general acetylation method. After purification by semi-prep HPLC (C18, 0–50B, 25 °C; A: 13 mmol NH_4OAc buffer pH 8.5, B: acetonitrile), the title compound was obtained as a white solid (18.1 mg, 8.58 µmol, 39%; HPLC-purity: 99.2%). **^{1}H NMR** (500 MHz, 12 mmol NH_4OAc buffer pH 8.5 in H_2O/D_2O 9:1): δ = 11.75 (s, 1H), 11.49 (s, 1H), 10.51 (s, 1H), 9.93 (s, 1H), 9.04 (s, 1H), 8.86 (s, 1H), 8.53 (d, *J* = 8.53 Hz, 1H), 8.41 (d, *J* = 8.35 Hz, 1H), 8.14 (d, *J* = 9.79 Hz, 1H), 7.86 (s, 2H), 7.80 (d, *J* = 8.75 Hz, 1H), 7.78–7.68 (m, 1H), 7.66 (d, *J* = 9.21 Hz, 1H), 7.50 (s, 1H), 7.40 (s, 1H), 7.34–7.18 (m, 2H), 7.14 (s, 1H), 7.09 (s, 2H), 6.98–6.79 (m, 3H), 6.58–6.41 (m, 2H), 6.29 (s, 2H), 6.26 (s, 1H), 6.08 (s, 1H), 5.91 (s, 1H), 4.19 (s, 1H), 4.03 (s, 2H), 3.95 (s, 1H), 1.55 (s, 3H), -0.18 (s, 3H). **HRMS** (ESI⁻) *m/z* calcd. for $C_{104}H_{76}N_{17}O_{29}Se$: 2106.4166 $(M-H)^-$; found: 2106.4162.



4

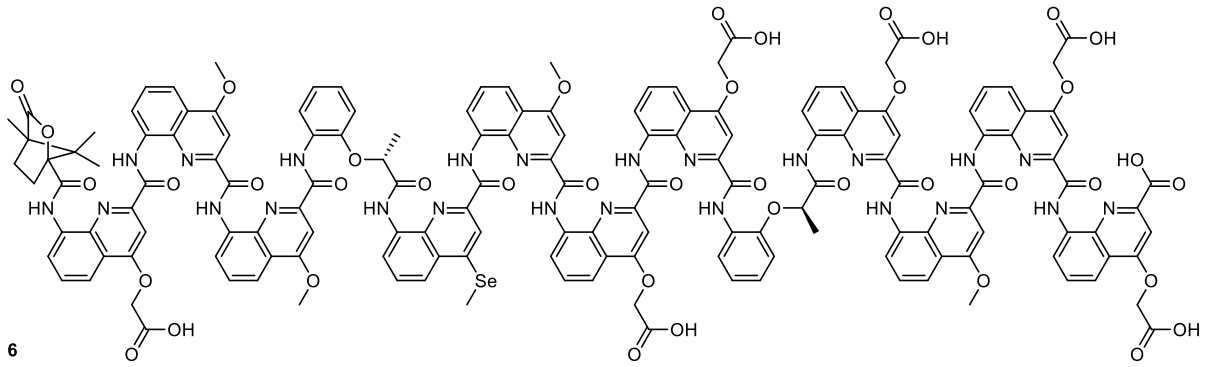
Compound 4: Oligomer **4** was synthesized on Wang resin (0.37 mmol g^{-1}, 18.5 µmol scale) according to the standard method. Loading of the first monomer: 0.33 mmol g^{-1} (89%). Unreacted amines were capped using the general capping method after each coupling step. The final acetyl group was installed via the general acetylation method. After purification by semi-prep HPLC (C18, 10–20B, 25 °C; A: 13 mmol NH_4OAc buffer pH 8.5, B: acetonitrile), the title compound was obtained as a white solid (9.75 mg, 4.13 µmol, 22%; HPLC-purity: 99.9%). **^{1}H NMR** (500 MHz, 12 mmol NH_4OAc buffer pH 8.5 in H_2O/D_2O 9:1): δ = 12.15 (s, 1H), 11.82 (s, 1H), 11.02 (s, 1H),

10.76 (s, 1H), 9.79 (s, 1H), 9.28 (s, 1H), 9.08 (s, 1H), 8.53 (s, 1H), 8.46 (d, *J* = 8.29 Hz, 1H), 8.34 (d, *J* = 8.25 Hz, 1H), 8.21 (d, *J* = 8.27 Hz, 1H), 8.12 (d, *J* = 9.13 Hz, 1H), 8.06 (s, 1H), 8.05–7.93 (m, 1H), 7.82 (s, 1H), 7.81–7.69 (m, 2H), 7.65 (d, *J* = 9.10 Hz, 1H), 7.59 (s, 1H), 7.56–7.44 (m, 2H), 7.40 (d, *J* = 8.47 Hz, 1H), 7.32 (s, 1H), 7.20–7.09 (m, 2H), 7.05 (s, 1H), 6.96 (s, 2H), 6.89 (t, *J* = 7.97 Hz, 1H), 6.81 (s, 1H), 6.79–6.72 (m, 3H), 6.68 (d, *J* = 7.64 Hz, 1H), 6.50 (s, 1H), 6.47 (s, 1H), 6.33 (d, *J* = 9.08 Hz, 1H), 6.27 (s, 2H), 6.13 (s, 1H), 6.09 (d, *J* = 9.00 Hz, 1H), 4.09 (s, 2H), 4.00 (s, 2H), 3.95 (s, 2H), 3.87 (d, *J* = 7.59 Hz, 1H), 3.38 (d, *J* = 15.15 Hz, 1H), 3.31–3.11 (m, 3H), 2.36 (s, 3H), 1.57 (d, *J* = 15.40 Hz, 1H), 1.47 (s, 3H), 1.28 (t, *J* = 7.92 Hz, 3H), 1.21 (d, *J* = 13.39 Hz, 1H), −0.32 (d, *J* = 6.97 Hz, 3H). **HRMS** (ESI⁻) *m/z* calcd. for $C_{119}H_{90}N_{19}O_{31}Se$: 2360.5221 (M-H)⁻; found: 2360.5251.



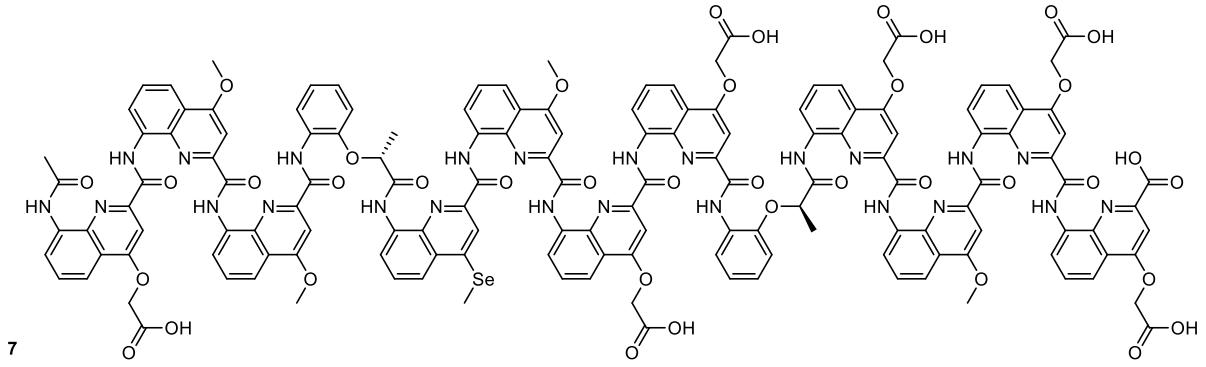
5

Compound 5: Oligomer **5** was synthesized on Wang resin (0.37 mmol g^{-1}, 22.0 µmol scale) according to the standard method. Loading of the first monomer: 0.34 mmol g^{-1} (92%). Unreacted amines were capped using the general capping method after each coupling step. The final camphanyl group was installed via the general camphanylation method. After purification by semi-prep HPLC (C18, 0–50B, 25 °C; A: 13 mmol NH_4OAc buffer pH 8.5, B: acetonitrile), the title compound was obtained as a white solid (12.7 mg, 5.64 µmol, 26%; HPLC-purity: 98.0%). **^{1}H NMR** (500 MHz, 12 mmol NH_4OAc buffer pH 8.5 in H_2O/D_2O 9:1): δ = 11.90 (s, 1H), 11.86 (s, 1H), 11.74 (s, 4H), 11.70 (s, 1H), 11.62 (s, 3H), 11.51 (s, 3H), 11.19 (s, 1H), 10.78 (s, 1H), 10.52 (s, 3H), 10.36 (s, 1H), 10.17 (s, 3H), 9.71 (s, 1H), 9.68 (s, 1H), 9.19 (s, 3H), 9.02 (s, 4H), 8.68 (s, 1H), 8.60 (d, *J* = 8.31 Hz, 1H), 8.53 (d, *J* = 8.36 Hz, 3H), 8.44 (d, *J* = 8.41 Hz, 1H), 8.24 (d, *J* = 8.22 Hz, 4H), 8.19 (d, *J* = 9.22 Hz, 1H), 8.14–8.00 (m, 4H), 7.97–7.85 (m, 11H), 7.85–7.77 (m, 7H), 7.74 (d, *J* = 8.99 Hz, 2H), 7.69 (d, *J* = 8.76 Hz, 1H), 7.67–7.59 (m, 3H), 7.59–7.48 (m, 3H), 7.43 (d, *J* = 8.01 Hz, 3H), 7.41–7.26 (m, 12H), 7.24 (d, *J* = 8.56 Hz, 1H), 7.22–7.08 (m, 7H), 7.05 (d, *J* = 8.11 Hz, 1H), 7.02–6.91 (m, 8H), 6.86 (t, *J* = 8.11 Hz, 4H), 6.77 (d, *J* = 8.26 Hz, 1H), 6.58–6.42 (m, 7H), 6.39 (s, 3H), 6.32 (d, *J* = 8.23 Hz, 4H), 6.28 (s, 4H), 6.18 (s, 4H), 6.14 (s, 1H), 6.00 (s, 2H), 4.22 (s, 4H), 4.16 (d, *J* = 8.38 Hz, 2H), 4.07 (s, 6H), 3.93 (s, 2H), 2.73 (s, 13H), 2.67 (s, 3H), 0.98 (s, 1H), 0.88 (s, 10H), 0.79 (s, 12H), 0.64 (d, *J* = 10.55 Hz, 16H), 0.57 (s, 3H), 0.43 (d, *J* = 7.18 Hz, 3H), 0.10 (s, 4H), 0.06 (s, 4H), −0.22 (d, *J* = 7.00 Hz, 9H). **HRMS** (ESI⁻) *m/z* calcd. for $C_{112}H_{86}N_{17}O_{31}Se$: 2244.4846 (M-H)⁻; found: 2244.5082.



6

Compound 6: Oligomer **6** was synthesized on Wang resin (0.37 mmol g^{-1}, 11.0 µmol scale) according to the standard method. Loading of the first monomer: 0.34 mmol g^{-1} (92%). Unreacted amines were capped using the general capping method after each coupling step. The final camphanyl group was installed via the general camphanylation method. After purification by semi-prep HPLC (C18, 10–35B, 25 °C; A: 13 mmol NH_4OAc buffer pH 8.5, B: acetonitrile), the title compound was obtained as a white solid (4.65 mg, 1.52 µmol, 14%; HPLC-purity: 96.6%). **^{1}H NMR** (500 MHz, 12 mmol NH_4OAc buffer pH 8.5 in H_2O/D_2O 9:1): δ = 11.70 (s, 1H), 11.47 (s, 1H), 11.12 (s, 1H), 11.07 (s, 1H), 10.44 (s, 2H), 10.12 (s, 1H), 9.71 (s, 1H), 9.28 (s, 1H), 9.14 (s, 1H), 8.96 (s, 1H), 8.77 (s, 1H), 8.66 (s, 1H), 8.48 (d, *J* = 8.48 Hz, 1H), 8.16 (d, *J* = 8.20 Hz, 1H), 8.02 (d, *J* = 9.25 Hz, 1H), 7.93–7.82 (m, 2H), 7.79 (d, *J* = 8.93 Hz, 1H), 7.73 (t, *J* = 7.95 Hz, 1H), 7.66 (d, *J* = 9.17 Hz, 1H), 7.60 (d, *J* = 8.11 Hz, 1H), 7.55 (s, 1H), 7.48 (d, *J* = 7.48 Hz, 1H), 7.42 (d, *J* = 8.32 Hz, 1H), 7.40–7.31 (m, 2H), 7.27 (s, 1H), 7.25–7.19 (m, 1H), 7.19–7.09 (m, 1H), 7.07 (s, 1H), 7.06–6.98 (m, 3H), 6.95 (s, 1H), 6.80 (t, *J* = 8.58 Hz, 1H), 6.73 (s, 1H), 6.48 (s, 1H), 6.40 (d, *J* = 8.41 Hz, 2H), 6.34 (s, 1H), 6.29 (s, 1H), 6.26 (s, 1H), 6.24 (s, 1H), 6.14 (s, 1H), 6.06 (s, 1H), 5.99 (s, 1H), 5.88 (s, 1H), 4.17 (d, *J* = 18.58 Hz, 3H), 3.95 (d, *J* = 20.88 Hz, 5H), 2.39 (s, 3H), 0.69 (s, 4H), 0.64 (s, 3H), 0.07 (s, 4H), -0.32 (d, *J* = 7.74 Hz, 5H). **HRMS** (ESI$^-$) *m/z* calcd. for $C_{155}H_{118}N_{24}O_{41}Se$: 1525.3531 $(M\text{-}2H)^{2-}$; found: 1525.3591.

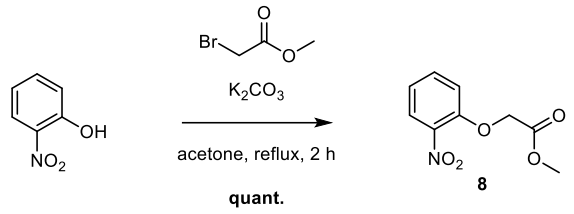


7

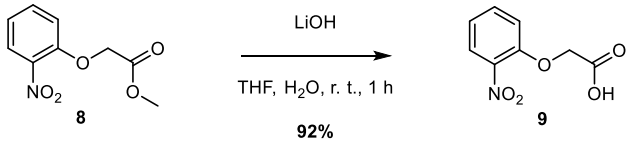
Compound 7: Oligomer **7** was synthesized on Wang resin (0.37 mmol g^{-1}, 11.0 µmol scale) according to the standard method. Loading of the first monomer: 0.34 mmol g^{-1} (92%). Unreacted amines were capped using the general capping method after each coupling step. The final acetyl group was installed via the general acetylation method. After purification by semi-prep HPLC (C18, 10–35B, 25 °C; A: 13 mmol NH_4OAc buffer pH 8.5, B: acetonitrile), the title compound was

obtained as a white solid (8.33 mg, 2.86 µmol, 26%; HPLC-purity: 97.6%). **^{1}H NMR** (500 MHz, 12 mmol NH_4OAc buffer pH 8.5 in H_2O/D_2O 9:1): δ = 11.69 (s, 1H), 11.46 (s, 1H), 11.07 (s, 1H), 10.48 (s, 1H), 10.43 (s, 1H), 10.11 (s, 1H), 9.74 (s, 1H), 9.21 (s, 1H), 8.94 (s, 1H), 8.68 (s, 2H), 8.47 (d, *J* = 8.38 Hz, 1H), 8.14 (d, *J* = 8.11 Hz, 1H), 7.99 (d, *J* = 9.29 Hz, 1H), 7.84 (t, *J* = 6.87 Hz, 1H), 7.81–7.74 (m, 1H), 7.71 (d, *J* = 9.28 Hz, 1H), 7.68 (s, 1H), 7.62 (d, *J* = 9.15 Hz, 1H), 7.51 (s, 1H), 7.47 (t, *J* = 8.76 Hz, 1H), 7.40 (d, *J* = 6.59 Hz, 1H), 7.37–7.30 (m, 1H), 7.28 (d, *J* = 7.44 Hz, 1H), 7.26–7.18 (m, 3H), 7.16 (d, *J* = 7.14 Hz, 1H), 7.12 (d, *J* = 8.01 Hz, 1H), 7.08 (s, 1H), 7.05–6.95 (m, 3H), 6.78 (t, *J* = 8.29 Hz, 1H), 6.75–6.65 (m, 1H), 6.46–6.36 (m, 2H), 6.34 (s, 1H), 6.27 (s, 2H), 6.24 (s, 2H), 6.06 (s, 1H), 6.02 (s, 1H), 5.86 (s, 1H), 4.15 (s, 3H), 3.92 (s, 4H), 3.85 (s, 1H), 2.39 (s, 3H), 1.09 (s, 2H), -0.33 (d, *J* = 7.07 Hz, 3H), −0.40 (s, 2H). **HRMS** (ESI^-) *m/z* calcd. for $C_{147}H_{109}N_{24}O_{39}Se$: 2913.6454 $(M-H)^-$; found: 2913.6924.

5.3.3 Monomer synthesis procedures

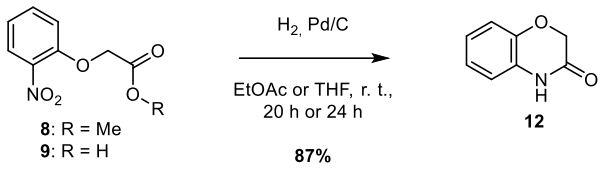


Compound 8: Compound **8** was prepared according to previously described methods.[43a] Analytical data is in line with the literature. ($C_9H_9NO_5$, MW = 211.17 g mol^{-1}). $\mathbf{R}_f$ (CyHex/EtOAc 9:1) = 0.10. **^{1}H NMR** (500 MHz, $CDCl_3$): δ = 7.88 (dd, *J* = 8.11, 1.71 Hz, 1H, C3-H), 7.52 (ddd, *J* = 8.29, 7.44, 1.69 Hz, 1H, C5-H), 7.11 (ddd, *J* = 8.33, 7.49, 1.16 Hz, 1H, C4-H), 6.99 (dd, *J* = 8.43, 1.11 Hz, 1H, C6-H), 4.79 (s, 2H, C7-H), 3.81 (s, 3H, C9-H). **^{13}C NMR** (126 MHz, $CDCl_3$): δ = 168.4 (C8), 151.4 (C1), 140.6 (C2), 134.2 (C5), 126.1 (C3), 122.0 (C4), 115.3 (C6), 66.7 (C7), 52.7 (C9). **HRMS** (ESI^+) *m/z* calcd. for $C_9H_{10}NO_5$: 212.0553 $(M+H)^+$; found: 212.0556.



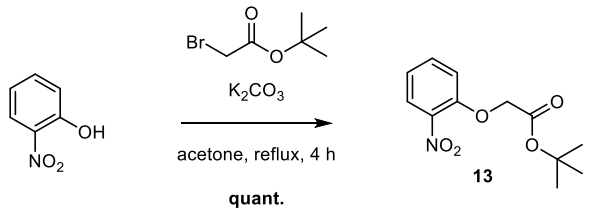
Compound 9: Compound **8** (8.21 g, 38.9 mmol, 1.0 equiv.) was dissolved in THF (120 ml). Then, LiOH (1.86 g, 77.8 mmol, 2.0 equiv.) in H_2O (30 ml) was added and the reaction mixture was stirred at room temperature for 1 h. After acidifying the solution to pH 1 using 1 M $HCl_{(aq.)}$, it was extracted with DCM (3×), dried over $MgSO_4$, and solvents were removed *in vacuo* (50 °C). The title compound was obtained as an off white solid (7.02 g, 35.6 mmol, 92%). ($C_8H_7NO_5$,MW = 197.15 g mol^{-1}). **^{1}H NMR** (500 MHz, DMSO-d_6): δ = 13.21 (s, 1H, O9-H), 7.87 (dd, *J* = 8.04, 1.71 Hz, 1H, C3-H), 7.62 (ddd, *J* = 8.87, 7.38, 1.71 Hz, 1H, C5-H), 7.26 (dd, *J* = 8.55, 1.07 Hz, 1H, C6-H), 7.13 (ddd, *J* = 8.27, 7.43, 1.10 Hz, 1H, C4-H), 4.91 (s, 2H, C7-H). **^{13}C NMR** (101 MHz, DMSO-d_6): δ = 169.3 (C8), 150.4 (C1), 139.7 (C2), 134.1 (C5), 124.9 (C3), 121.1 (C4), 115.0 (C6),

65.3 (C7). **HRMS** (ESI⁻) *m/z* calcd. for $C_8H_6NO_5$: 196.0251 (M-H)⁻; found: 196.0250. (Modified literature procedure;[43a] analytical data is in line with the literature).



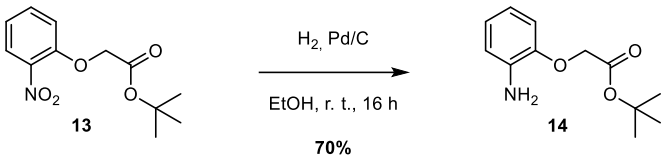
Compound 12: From the ester: Compound **8** (6.85 g, 32.4 mmol, 1.0 equiv.) was dissolved in EtOAc (350 ml). The solution was quickly degassed by vacuum N_2 cycles (3×), then Pd/C (685 mg, 10% *m/m* to **9a**) was added and the N_2 atmosphere was replaced by H_2. After stirring for 20 h the reaction mixture was filtered over celite©, washed with EtOAc and solvents were removed *in vacuo* (50 °C). The title compound was obtained as a light brown solid (4.22 g, 28.3 mmol, 87%).

From the acid: Compound **9** (7.02 g, 35.6 mmol, 1.0 equiv.) was dissolved in THF (600 ml). The solution was quickly degassed by vacuum N_2 cycles (3×), then Pd/C (702 mg, 10% *m/m* to **9e**) were added and the N_2 atmosphere was replaced by H_2. After stirring for 24 h the reaction mixture was filtered over celite©, washed with THF and solvents were removed *in vacuo* (50 °C). The crude product was not further purified. ($C_8H_7NO_2$, MW = 149.15 g mol^{-1}). $\mathbf{R_f}$ (CyHex/EtOAc 7:3) = 0.30. **^{1}H NMR** (500 MHz, $CDCl_3$): δ = 8.46 (s, 1H, C9-H), 7.06–6.89 (m, 3H, C3-H, C4-H, C6-H), 6.87–6.78 (m, 1H, C5-H), 4.63 (s, 2H, C7-H). **^{13}C NMR** (126 MHz, $CDCl_3$): δ = 166.0 (C8), 143.8 (C1), 126.2 (C2), 124.4 (C3), 122.9 (C6), 117.0 (C4), 116.1 (C5), 67.4 (C7). **HRMS** (ESI⁻) *m/z* calcd. for $C_8H_6NO_2$: 148.0404 (M-H)⁻; found: 148.0403. (Known compound; analytical data is in line with the literature).[119]

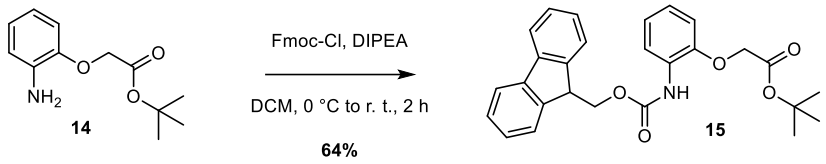


Compound 13: 2-nitrophenol (3.0 g, 21.6 mmol, 1.0 equiv.), *tert*-butyl bromoacetate (3.50 ml, 23.7 mmol, 1.1 equiv.) and K_2CO_3 (3.28 g, 23.7 mmol, 1.1 equiv.) were suspended in anhydrous acetone (60 ml) under an N_2 atmosphere. Then, the reaction mixture was refluxed for 4 h (bath temperature 75 °C). After cooling to room temperature the mixture was filtered, washed with acetone and solvents were removed *in vacuo* (50 °C). Without further purification, the title compound was obtained as a brown solid (5.46 g, 21.6 mmol, quant.). ($C_{12}H_{15}NO_5$, MW = 253.25 g mol^{-1}). $\mathbf{R_f}$ (CyHex/EtOAc 9:1) = 0.21. **^{1}H NMR**(500 MHz, $CDCl_3$): δ = 7.86 (dd, *J* = 8.09 Hz, 1.72 Hz, 1H, C3-H), 7.51 (ddd, *J* = 8.43, 7.45, 1.77 Hz, 1H, C5-H), 7.08 (ddd, *J* = 8.36, 7.44, 1.16 Hz, 1H, C4-H), 6.96 (dd, *J* = 8.47, 1.14 Hz, 1H, C6-H), 4.67 (s, 2H, C9-H), 1.46 (s, 9H, C10-H, C11-H, C12-H). **^{13}C NMR** (126 MHz, $CDCl_3$): δ = 166.9 (C8), 151.5 (C1), 140.4 (C2), 134.0 (C5), 126.0 (C3), 121.5 (C4), 114.9 (C6), 83.2 (C9), 66.8 (C7), 28.1 (C10, C11, C12). **HRMS** (EI⁺) *m/z* calcd. for

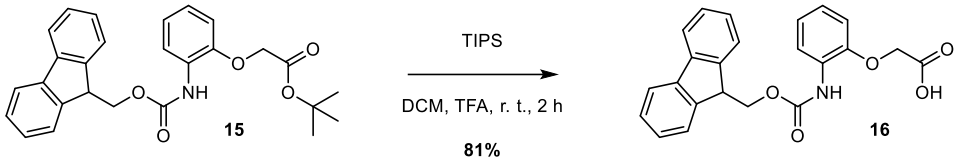
$C_{12}H_{15}NO_5$: 253.0945 $(M)^+$; found: 253.0955. (Known compound; analytical data is in line with the literature).[120]



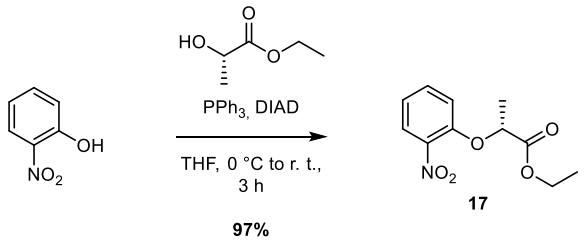
Compound 14: Compound **13** (5.46 g, 21.6 mmol, 1.0 equiv.) was dissolved in ethanol (350 ml). The solution was quickly degassed by vacuum N_2 cycles (3×), then Pd/C (546 mg, 10% *m/m* to **9b**) were added and the N_2 atmosphere was replaced by H_2. After stirring for 16 h the reaction mixture was filtered over celite©, washed with ethanol and solvents were removed *in vacuo* (50 °C) (product was found to be unstable, especially at elevated temperatures). The title compound was obtained as a brown solid (4.60 g, contaminated with 30% cyclic side product). The crude product was directly used in the next step without further purification. ($C_{12}H_{17}NO_3$, MW = 223.27 g mol^{-1}). $\mathbf{R_f}$ (CyHex/EtOAc 7:3) = 0.59. **^{1}H NMR** (500 MHz, $CDCl_3$): δ = 8.24 (s, 1H), 7.03–6.91 (m, 1H), 6.87–6.77 (m, 1H), 6.77–6.63 (m, 3H), 4.62 (s, 1H), 4.53 (s, 2H), 1.49 (s, 9H).



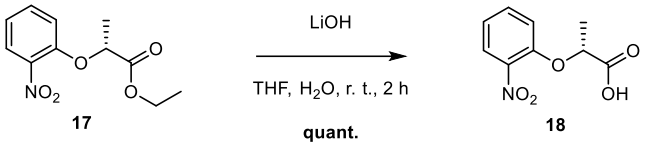
Compound 15: Compound **14** (4.60 g, contaminated with 30% lactam, 15.1 mmol, 1.0 equiv.) was dissolved in anhydrous DCM (100 ml) under an N_2 atmosphere. Then, DIPEA (3.94 ml, 22.6 mmol, 1.5 equiv.) was added and the solution was cooled to 0 °C. After the addition of Fmoc-Cl (5.08 g, 19.6 mmol, 1.3 equiv.) at 0 °C the reaction mixture was stirred at room temperature for 2 h. The mixture was acidified by adding citric acid $_{(aq.)}$ (1 M, 15.1 ml, 15.1 mmol, 1.0 equiv.), extracted with DCM (2×), the organic phases were dried over $MgSO_4$ and solvents were removed *in vacuo* (50 °C). After purification by column chromatography (SiO_2, CyHex/EtOAc 95:5) the title compound was obtained as a colorless oil (4.32 g, 9.70 mmol, 64%). ($C_{27}H_{27}NO_5$, MW = 445.52 g mol^{-1}). $\mathbf{R_f}$ (CyHex/EtOAc 9:1) = 0.31. **^{1}H NMR** (500 MHz, $CDCl_3$): δ = 8.13 (s, 1H, C3-H), 7.97 (s, 1H, N13-H), 7.78 (d, *J* = 7.70 Hz, 2H, C20-H, C21-H), 7.70 (d, *J* = 7.46 Hz, 2H, C-17-H, C24-H), 7.41 (t, *J* = 7.54 Hz, 2H, C19-H, C22-H) 7.33 (t, *J* = 7.41 Hz, 2H, C18-H, C23-H), 7.03 (t, *J* = 8.32 Hz, 1H, C4-H), 6.98 (td, *J* = 7.72, 7.53, 1.76 Hz, 1H, C5-H), 6.83 (d, *J* = 8.06 Hz, 1H, C6-H), 4.60 (s, 2H, C7-H), 4.49 (d, *J* = 7.20 Hz, 2H, C15-H), 4.32 (t, *J* = 7.24 Hz, 1H, C16-H), 1.51 (s, 9H, C10-H, C11-H, C12-H). **^{13}C NMR** (126 MHz, $CDCl_3$): δ = 168.5 (C8), 153.6 (C14), 147.1 (C2), 144.5 (C1), 144.1 (C16a, C24a), 141.5 (C20a, C20b), 129.2 (C4), 127.9 (C19, C22), 127.3 (C18, C23), 125.4 (C17, C24), 123.0 (C5), 120.1 (C20, C21), 119.2 (C3), 113.7 (C6), 83.0 (C9), 68.0 (C7), 67.2 (C15), 47.3 (C16), 28.2 (C10, C11, C12). **HRMS** (ESI^+) *m/z* calcd. for $C_{27}H_{27}NO_5Na$: 468.1781 $(M+Na)^+$; found: 468.1790.



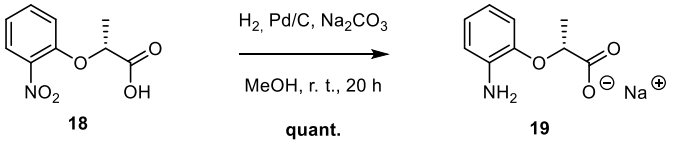
Compound 16: Compound **15** (4.32 g, 9.70 mmol, 1.0 equiv.) and triisopropyl silane (500 µl, 5% *v/v*) were dissolved in DCM (10 ml). After the addition of TFA (10 ml) the reaction mixture was vigorously stirred under an N_2 atmosphere for 2 h. The solution was extracted with DCM (3×) (some precipitate appeared, so the organic phases could not be dried) and solvents were removed *in vacuo* (50 °C). The crude product was triturated with Et_2O, filtered and washed with cold Et_2O to yield the title compound as a white solid (3.06 g, 7.87 mmol, 81%). ($C_{23}H_{19}NO_5$, MW = 389.41 g mol^{-1}). $\mathbf{R_f}$ (CyHex/EtOAc 8:2 + 1% AcOH) = 0.20. **^{1}H NMR** (500 MHz, DMSO-d_6): δ = 13.10 (s, 1H, O9-H), 8.67 (s, 1H, N10-H), 7.91 (d, *J* = 7.63 Hz, 2H, C17-H, C18-H), 7.75 (d, *J* = 7.51 Hz, 2H, C14-H, C21-H), 7.64 (s, 1H, C3-H), 7.43 (t, *J* = 7.49 Hz, 2H, C16-H, C19-H), 7.34 (td, *J* = 7.44, 1.18 Hz, 2H, C15-H, C20-H) 7.03 (ddd, *J* = 8.73, 7.20, 1:65 Hz, 1H, C5-H), 6.98 (dd, *J* = 8.26, 1.51 Hz, 1H, C6-H), 6.93 (t, *J* = 7.70 Hz, 1H, C4-H), 4.73 (s, 2H, C7-H), 4.42 (d, *J* = 7.10 Hz, 2H, C12-H), 4.31 (t, *J* = 7.01 Hz, 1H, C13-H). **^{13}C NMR** (126 MHz, DMSO-d_6): δ = 170.5 (C8), 153.6 (C11), 148.59 (C2), 148.58 (C1), 143.8 (C13a, C21a), 140.7 (C17a, C17b), 127.7 (C16, C19), 127.1 (C15, C20), 125.3 (C14, C21), 124.1 (C5), 121.4 (C4), 120.2 (C17, C18), 113.5 (C6), 66.11 (C12), 66.06 (C7), 46.6 (C13). **HRMS** (ESI$^+$) *m/z* calcd. for $C_{23}H_{20}NO_5$: 390.1336 (M+H)$^+$; found: 390.1330.



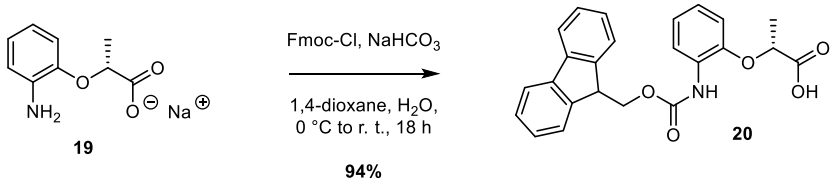
Compound 17: 2-nitrophenol (6.00 g, 43.1 mmol, 1.0 equiv.), (-)-ethyl L-lactate (4.93 ml, 43.1 mmol, 1.0 equiv.) and PPh_3 (13.6 g, 51.8 mmol, 1.2 equiv.) were dissolved in anhydrous THF (80 ml) under an N_2 atmosphere. Then, DIAD (10.2 ml, 51.8 mmol, 1.2 equiv.) was slowly added at 0 °C and the resulting mixture was stirred at room temperature for 3 h. Solvents were evaporated *in vacuo* (50 °C), and the residue was purified by column chromatography (SiO_2, CyHex/EtOAc 9:1) yielding the title compound as a white solid (10.0 g, 41.8 mmol, 97%). ($C_{11}H_{13}NO_5$, MW = 239.23 g mol^{-1}). $\mathbf{R_f}$ (CyHex/EtOAc 8:2) = 0.32. **^{1}H NMR** (500 MHz, $CDCl_3$): δ = 7.83 (dd, *J* = 8.09, 1.71 Hz, 1H, C3-H), 7.48 (ddd, *J* = 8.48, 7.48, 1.75 Hz, 1H, C5-H), 7.07 (ddd, *J* = 8.18, 7.45, 1.12 Hz, 1H, C4-H), 6.96 (dd, *J* = 8.40, 1.12 Hz, 1H, C6-H), 4.84 (q, *J* = 6.84 Hz, 1H, C7-H), 4.21 (qd, *J* = 7.16, 3.98 Hz, 2H, C10-H), 1.69 (d, *J* = 6.80 Hz, 3H, C8-H), 1.24 (t, *J* = 7.31 Hz, 3H, C11-H). **^{13}C NMR** (101 MHz, $CDCl_3$): δ = 171.1 (C9), 151.2 (C1), 140.9 (C2), 133.9 (C5), 125.8 (C3), 121.6 (C4), 116.0 (C6), 74.8 (C7), 61.8 (C10), 18.5 (C8), 14.2 (C11). **HRMS** (ESI$^+$) *m/z* calcd. for $C_{11}H_{14}NO_5$: 240.0866 (M+H)$^+$; found: 240.0866.



Compound 18: Compound **17** (6.00 g, 25.1 mmol, 1.0 equiv.) was dissolved in THF (100 ml). After the addition of LiOH (1.20 g, 50.2 mmol, 2.0 equiv.) in H_2O (20 ml), the mixture was stirred at room temperature for 2 h. Then, the solution was acidified by adding HCl $_{(aq.)}$ (1 M, 62.7 ml, 62.7 mmol, 2.5 equiv.) and extracted with DCM (3×). Drying over $MgSO_4$ and removing the solvents *in vacuo* (50 °C) yielded the title compound as a white solid (5.35 g, 25.1 mmol, quant.). ($C_9H_9NO_5$, MW = 211.17 g mol^{-1}). **^{1}H NMR** (500 MHz, DMSO-d_6): δ = 13.24 (s, 1H, O10-H), 7.85 (dd, *J* = 8.06, 1.68 Hz, 1H, C3-H), 7.61 (ddd, *J* = 8.61, 7.38, 1.71 Hz, 1H, C5-H), 7.19 (dd, *J* = 8.62, 1.09 Hz, 1H, C6-H), 7.12 (ddd, *J* = 8.36, 7.40, 1.09 Hz, 1H, C4-H), 5.11 (q, *J* = 6.83 Hz, 1H, C7-H), 1.52 (d, *J* = 6.80 Hz, 3H, C8-H). **^{13}C NMR** (126 MHz, DMSO-d_6): δ = 172.0 (C9), 150.1 (C1), 140.0 (C2), 134.1 (C5), 124.9 (C3), 121.0 (C4), 115.5 (C6), 72.8 (C7), 18.0 (C8). **HRMS** (ESI$^-$) *m/z* calcd. for $C_9H_8NO_5$: 210.0408 (M-H)$^-$; found: 210.0408. (Modified literature procedure;[43a] analytical data is in line with the literature).



Compound 19: Compound **18** (4.85 g, 23.0 mmol, 1.0 equiv.) and Na_2CO_3 (2.43 g, 23.0 mmol, 1.0 equiv.) were suspended in methanol (400 ml). The solution was quickly degassed by vacuum N_2 cycles (3×), then Pd/C (485 mg, 10% *m/m* to **6c**) were added and the N_2 atmosphere was replaced by H_2. After stirring for 20 h the reaction mixture was filtered over celite©, washed with methanol and solvents were removed *in vacuo* (50 °C). The crude product was used in the next step without further purification. ($C_9H_{10}NO_3Na$, MW = 203.17 g mol^{-1}). **^{1}H NMR** (500 MHz, DMSO-d_6): δ = 6.65 (dd, *J* = 7.94, 1.30 Hz, 1H, C6-H), 6.62–6.57 (m, 1H, C4-H), 6.56 (dd, *J* = 7.73, 2.01 Hz, 1H, C3-H), 6.39 (ddd, *J* = 7.84, 6.94, 2.02 Hz, 1H, C5-H), 4.91 (s, 2H, N11-H), 4.05 (q, *J* = 6.72 Hz, 1H, C7-H), 1.36 (d, *J* = 6.71 Hz, 3H, C8-H). **^{13}C NMR** (126 MHz, DMSO-d_6): δ = 175.2 (C9), 147.2 (C1), 139.9 (C2), 121.3 (C4), 116.3 (C5), 115.7 (C6), 114.4 (C3), 78.4 (C7), 20.2 (C8). **HRMS** (ESI$^+$) *m/z* calcd. for $C_9H_{12}NO_3$: 182.0812 (M+H)$^+$; found: 182.0809.



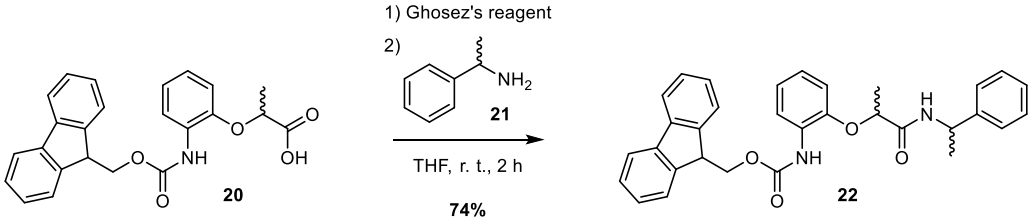
Compound 20: Compound **19** (4.16 g, 23.0 mmol, 1.0 equiv.) and $NaHCO_3$ (9.64 g, 115 mmol, 5.0 equiv.) were dissolved in H_2O (300 ml). Then, Fmoc-Cl (7.72 g, 29.8 mmol, 1.3 equiv.) in 1,4-dioxane (300 ml) was added at 0 °C over 1 h. After stirring at 0 °C for 1 h, the reaction mixture was

stirred at room temperature for an additional 16 h. The mixture was acidified below pH 4 using HCl $_{(aq.)}$ (1 M), extracted with DCM (3×), the organic phases were dried over $MgSO_4$ and solvents were removed *in vacuo* (50 °C). The residue was purified by column chromatography (SiO_2, CyHex/EtOAc 7:3 + 1% DIPEA → CyHex/EtOAc 7:3 → CyHex/EtOAc 7:3 + 1% AcOH → EtOAc + 1% AcOH) to yield the title compound as a white solid (8.73 g, 21.6 mmol, 94%). ($C_{24}H_{21}NO_5$, MW = 403.43 g mol^{-1}). $\mathbf{R_f}$ (CyHex/EtOAc 8:2 + 1% AcOH) = 0.10. **^{1}H NMR** (500 MHz, DMSO-d_6): δ = 13.16 (s, 1H, O10-H), 8.72 (s, 1H, N11-H), 7.91 (d, *J* = 7.53 Hz, 2H, C18-H, C19-H), 7.74 (dd, *J* = 7.69, 3.35 Hz, 2H, C15-H, C22-H), 7.61 (s, 1H, C3-H), 7.43 (t, *J* = 7.44 Hz, 2H, C17-H, C20-H), 7.34 (t, *J* = 7.41 Hz, 2H, C16-H, C21-H), 7.07–6.98 (m, 1H, C5-H), 6.99–6.89 (m, 2H, C4-H, C6-H), 4.81 (q, *J* = 6.77 Hz, 1H, C7-H), 4.49–4.35 (m, 2H, C13-H), 4.32 (t, *J* = 7.07 Hz, 1H, C14-H), 1.56 (d, *J* = 6.75 Hz, 3H, C8-H). **^{13}C NMR** (126 MHz, DMSO-d_6): δ = 173.5 (C9), 153.5 (C12), 148.4 (C1), 143.8 (C14a, C22a), 140.7 (C18a, C18b), 128.1 (C2), 127.7 (C17, C20), 127.1 (C16, C21), 125.29 (C15, C22), 125.27 (C3), 124.1 (C5), 121.6 (C4), 120.2 (C18, C19), 114.6 (C6), 74.0 (C7), 66.2 (C13), 46.5 (C14), 18.5 (C8). **HRMS**(ESI$^-$) *m/z* calcd. for $C_{24}H_{22}NO_5$: 404.1492 (M-H)$^-$; found: 404.1491.

The synthesis of B^{Rme} (**20**) and its (*S*)-enantiomer B^{Sme} was repeated with the protocol reported above. This time precursors with an enantiomeric purity given by the supplier as ≥99% were used. The enantiomeric ratio is retained during synthesis as shown by chiral derivatisation (section 5.2). Therefore, an enantiomeric purity of ≥99% is expected for the final Fmoc-monomers as well. Their specific rotation was measured and is shown in Table 2.

Table 2 Specific rotation values of chiral B monomers.

B^{Rme}	$[\alpha]_D^{20}$ = −26.9 (*c* 2.24, MeOH)
B^{Sme}	$[\alpha]_D^{20}$ = +26.8 (*c* 0.56, MeOH)



Compound 22a: Compound **8** (100 mg, 0.248 mmol, 1.0 equiv.) was dissolved in anhydrous THF (2 ml). Then, 1-Chlor-*N*,*N*,2-trimethyl-1-propenylamin (Ghosez's reagent) (66.0 µl, 0.496 mmol, 2.0 equiv.) was added, and the reaction mixture was stirred at room temperature for 1.5 h. Solvents were evaporated at room temperature under high vacuum for 4 h, and the residue was redissolved in anhydrous THF (2 ml). After the addition of (*S*)-1-phenyl-ethylamin (96.0 µl, 7.44 µmol, 3.0 equiv.), the mixture was stirred at room temperature for another 30 min. The mixture was diluted with HCl $_{(aq.)}$ (1 M), extracted with DCM (2×), and the organic phases were dried

over $MgSO_4$. Evaporating the solvents *in vacuo* (50 °C) yielded the crude product as a white solid (93.0 mg, 184 µmol, 74%). ($C_{32}H_{30}N_2O_4$, MW = 506.60 g mol^{-1}). $\mathbf{R_f}$ (CyHex/EtOAc 8:2) = 0.20. **^{1}H NMR** (500 MHz, DMSO-d_6): δ = 8.98 (s, 1H, C19-H), 8.51 (d, *J* = 8.26 Hz, 1H, C10-H), 7.91 (d, *J* = 7.57 Hz, 2H, C26-H, C27-H), 7.72 (dd, *J* = 7.71, 3.70 Hz, 2H, C23-H, C30-H), 7.47 (s, 1H, C3-H), 7.43 (t, *J* = 7.49 Hz, 2H, C25-H, C28-H), 7.32 (t, *J* = 8.75 Hz, 1H, C24-H, C29-H), 7.25 (d, *J* = 4.30 Hz, 3H, C14-H, C15-H, C17-H, C18-H), 7.18 (h, *J* = 4.41, 4.34 Hz, 1H, C16-H), 7.06 (t, *J* = 7.76 Hz, 1H, C5-H), 6.95 (d, *J* = 8.53 Hz, 1H, C6-H), 6.93–6.87 (m, 1H, C4-H), 4.92 (p, *J* = 7.23 Hz, 1H, C11-H), 4.80 (q, *J* = 6.60 Hz, 1H, C7-H), 4.44 (d, *J* = 7.63 Hz, 2H, C21-H), 4.32 (t, *J* = 6.91 Hz, 1H, C22-H), 1.44 (d, *J* = 6.59 Hz, 3H, C8-H), 1.24 (d, *J* = 7.00 Hz, 3H, C12-H). **^{13}C NMR** (101 MHz, DMSO-d_6): δ = 170.4 (C9), 154.0 (C20), 148.5 (C1), 144.1 (C13), 143.7 (C22a, C30a), 140.8 (C26a, C26b), 128.2 (C15, C17), 127.7 (C25, C28), 127.4 (C2), 127.1 (C24, C29), 126.6 (C16), 125.7 (C14, C18), 125.1 (C23, C30), 124.8 (C5), 123.2 (C3), 121.3 (C4), 120.2 (C26, C27), 113.7 (C6), 74.5 (C7), 66.0 (C21), 47.4 (C11), 46.6 (C22), 22.0 (C12), 18.7 (C8).

5.4 X-ray Crystallography

After preparative HPLC, the ammonium salt of compound **7** was dissolved in pure water to final concentration of 3.5 mM. Crystallization trials were performed at 20 °C in standard aqueous sitting drop vapour diffusion method and subsequently optimized by hanging drop vapour diffusion method at 20 °C. Hexagonal crystals (Fig. 49) were obtained by mixing 1 µl of **7** and 2 µl of crystallization reagent 5% *v/v* 2-propanol, 50 mM TRIS buffer, pH 7.5, 10 mM Magnesium chloride; equilibrated against 500 µl of crystallization reagent in the reservoir. For data collection, a single crystal was fished using a MiTeGen microloop, cryo-protected in a solution composed of 33% *v/v* Glycerol, 3.3% 2-propanol, 33% TRIS buffer, pH 7.5, 6.6 mM Magnesium chloride and flash frozen in liquid nitrogen.

Synchrotron data was collected on beam line P14 operated by EMBL Hamburg at the Petra III storage ring (DESY, Hamburg, Germany) using 0.9808 Å wavelength. During data collection, the crystal was cooled to 100 K. The crystal was exposed for 0.008 s and 0.1° oscillation per frame and a rotation pass of 360° was measured using an EIGER 16M detector. Data were processed using xia2[102] with DIALS[103] for integration and using Pointless/Aimless[104] for scaling and merging respectively. The crystal belonged to the space group $P3_121$ (or $P3_221$) with unit cell parameters a = b = 79.06 Å and c = 39.86 Å, V = 212837 Å^3 and 4 molecules in the asymmetric unit. The structure was solved by Molecular Replacement (MR) method using *PHASER*[121] from the *CCP4* suite.[122] The model was built in *Maestro* (Version 11.5.011) in the following way: First, a Q helix was built, and was minimized using the parameters shown in Table 3. Then, respective Q units were converted to B units and side chains were added. Finally, the structure was minimized again using the same parameters. The side chain atoms were omitted from the model to remove flexibility and to allow minimum number of clashes.[123] The best MR solution (Fig. 50) had log likelihood gain (LLG) of 284 and translation function Z-score (TFZ) of 7.4 in $P3_121$ space group.

Geometric restraints for maximum likelihood restrained refinement was generated using *PRODRG*.[124] Model building and restrained refinement were performed in *Coot*[108] and *Refmac5*[125] respectively. Thermal displacement parameters were refined isotropically. Fig. 51 shows the sigma weighted 2*Fσ-Fc* electron density map superimposed on one of the four helices. The final model was refined to a resolution of 2.86 Å, with R_{work} and R_{free} factors of 24.72% and 30.37% respectively. Statistics of data collection and refinement can be found in Table 4, with statistics for the highest-resolution shell shown in parentheses.

The final cif file was checked using IUCr's checkcif algorithm. Due to large volume fractions of disordered solvent molecules, weak diffraction intensity and poor resolution, a number of A- and B-level remain in the checkcif file. These alerts are inherent to the data and refinement procedures and illustrate the limited practicality of the checkcif tool for medium-size molecule crystallography. Atomic coordinates and structure factors for **7** was deposited in the Cambridge Crystallographic Data Centre (CCDC) with accession code 2070816. The data is available free of charge upon request (www.ccdc.cam.ac.uk/).

Table 3 Parameters used to build the model of **7**.

Forcefield	OPLS3
Solvent	Water
Charges from	Force Field
Cutoff	Extended
Constraints	0
Method	TNCG
Max. Iterations	2500
Converge on	Gradient
Converge Threshold	0.05
Minimization Mode	Minimization of non-conformers

Table 4 Crystallographic data collection and refinement statistics for **7**.

Data collection	
Space group	$P3_121$
Unit cell	
a, *b*, *c* (Å)	79.06, 79.06, 39.86
α, *β*, *γ* (°)	90.0, 90.0, 120.0
Resolution (Å)	39.53 – 2.86 (2.93 – 2.86)
R_{meas}	0.260 (1.944)
CC half	0.992 (0.921)
I / σ	4.8 (1.1)
Completeness (%)	100 (100)
Reflections (total)	68739
Reflections (unique)	3481
Redundancy	19.7 (20.8)
Refinement	
Resolution (Å)	39.53 – 2.86
R_{work} / R_{free} (%)	24.72 / 30.37
No. of non-H atoms	852
Overall B-factor (Å^3)	52.30
R.m.s. deviations	
Bond lengths (Å)	0.019
Bond angles (°)	3.139
CCDC entry	2070816*

*As key information were lost during the .pdb file to .cif file conversion process, the entire .pdb file was inserted into the 'comments' section of .cif file available from the Cambridge Crystallographic Data Centre via https://www.ccdc.cam.ac.uk/getstructures/.

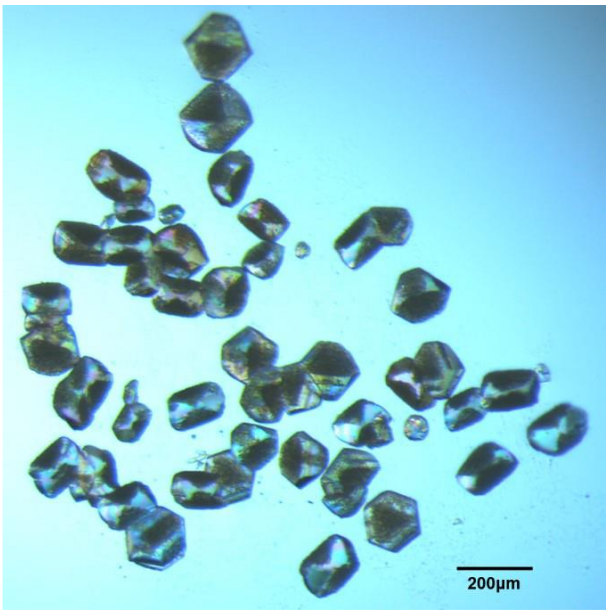


Fig. 49 Crystals of **7** observed under crossed polarizing microscope.

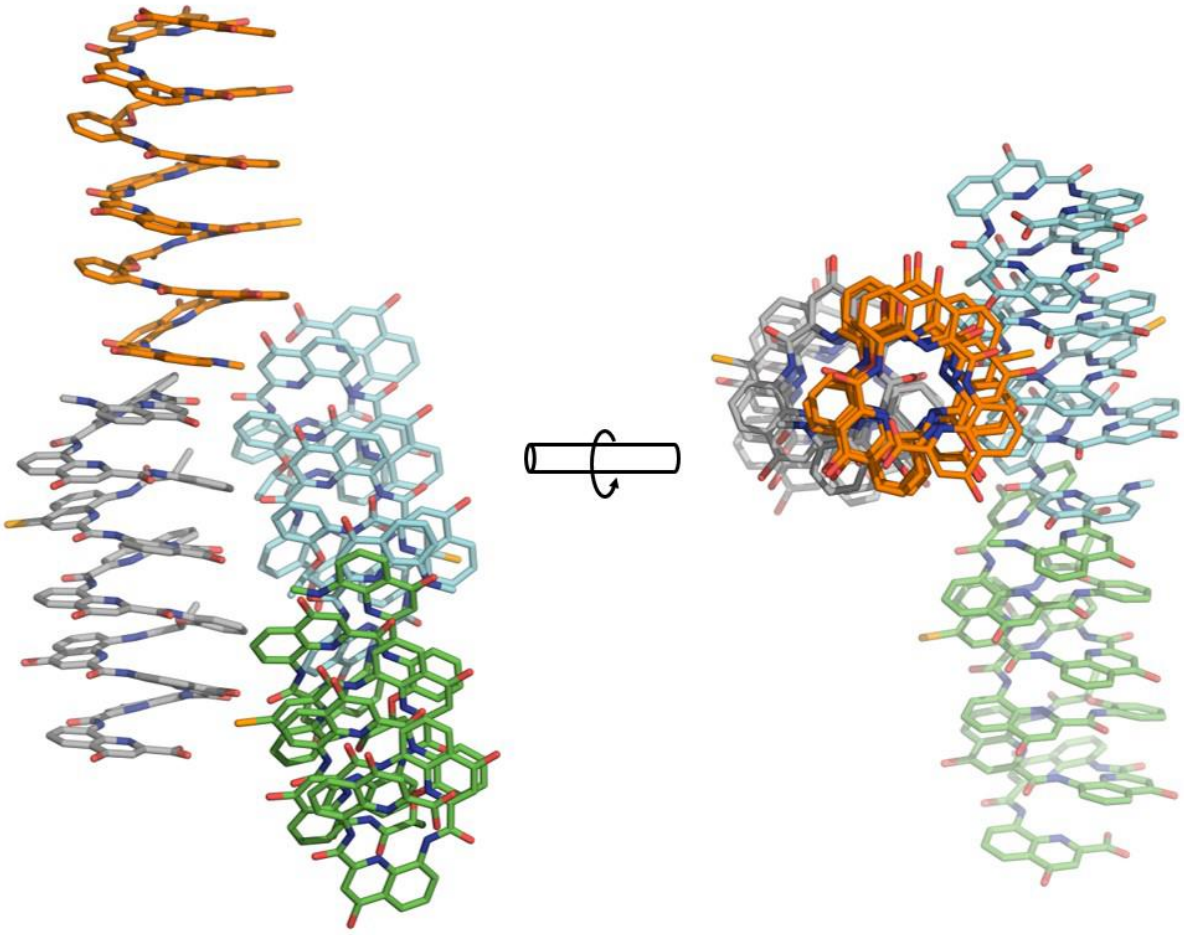


Fig. 50 The crystallographic asymmetric unit of **7** consisted of four independent molecules.

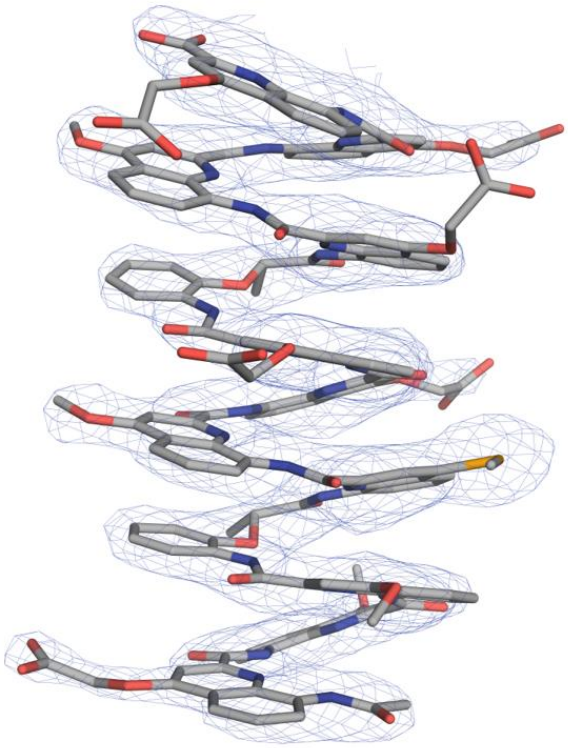


Fig. 51 Sigma weighted $2F_o$-F_c electron density map (contoured at 1.6 σ, carved within 2.0 Å) superimposed on one **7** helix.

5.5 Spectra and Chromatograms

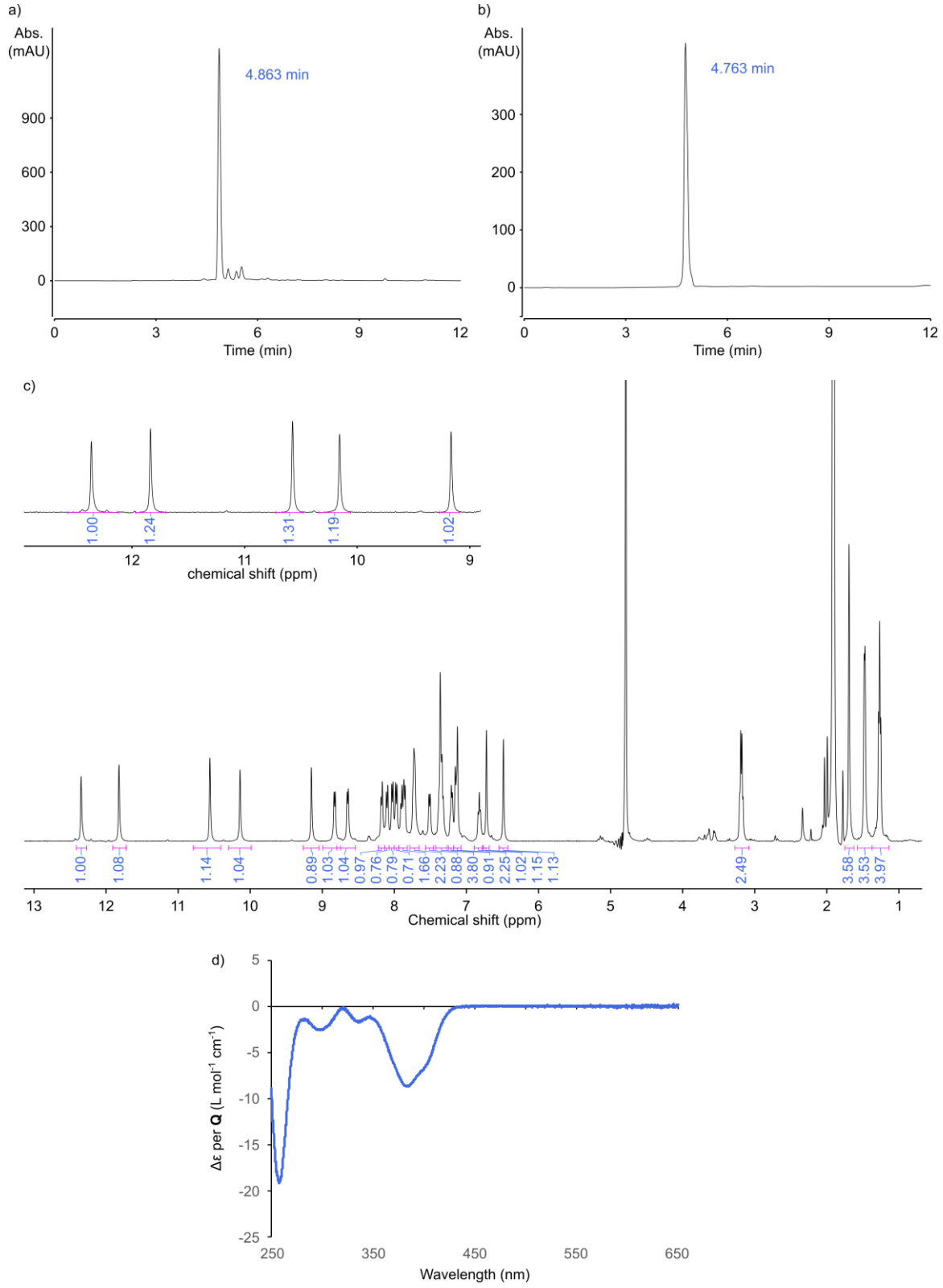


Fig. 52 Analytical data of compound **1**. HPLC chromatograms after cleavage from the resin (a) and after purification (b) (C18, 0–60B, 25 °C; A: 13 mM NH_4OAc buffer pH 8.5, B: acetonitrile). c) ^1H NMR (500 MHz, 0.13 mM in 12 mM NH_4OAc buffer pH 8.5 H_2O/D_2O 9:1, H_2O suppression). d) CD spectrum (41.2 µM in 13 mM NH_4OAc buffer pH 8.5).

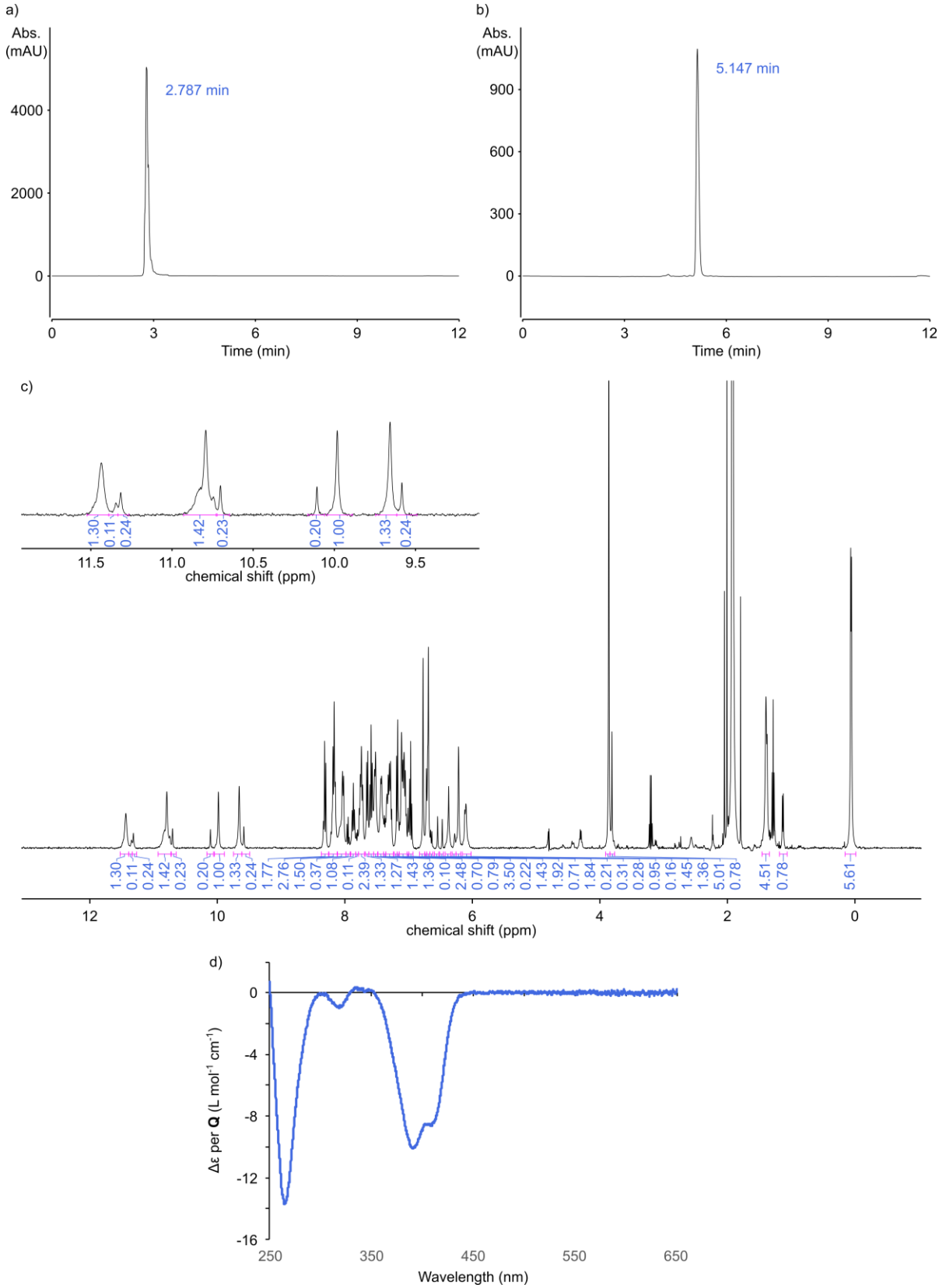


Fig. 53 Analytical data of compound **2**. HPLC chromatograms after cleavage from the resin (a) (C8, 0–60B, 50 °C; A: 13 mM NH_4OAc buffer pH 8.5, B: acetonitrile) and after purification (b) (C8, 0–40B, 50 °C; A: 13 mM NH_4OAc buffer pH 8.5, B: acetonitrile). c) ^1H NMR (500 MHz, 0.13 mM in 12 mM NH_4OAc buffer pH 8.5 H_2O/D_2O 9:1, H_2O suppression). d) CD spectrum (15.4 µM in 13 mM NH_4OAc buffer pH 8.5).

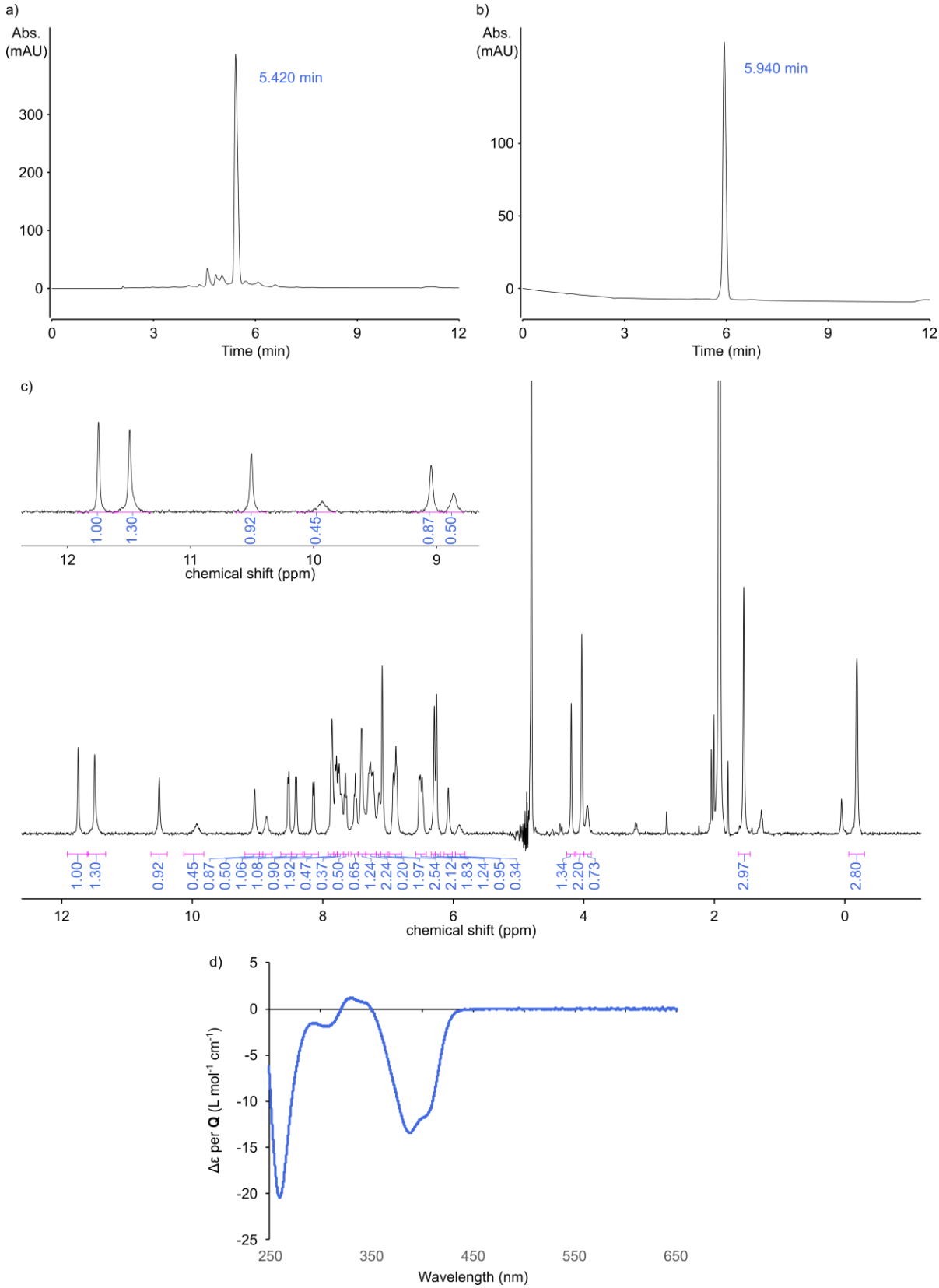


Fig. 54 Analytical data of compound **3**. HPLC chromatograms after cleavage from the resin (a) and after purification (b) (C18, 0–60B, 25 °C; A: 13 mM NH_4OAc buffer pH 8.5, B: acetonitrile). c) ^1H NMR (500 MHz, 0.13 mM in 12 mM NH_4OAc buffer pH 8.5 H_2O/D_2O 9:1, H_2O suppression). d) CD spectrum (22.6 µM in 13 mM NH_4OAc buffer pH 8.5).

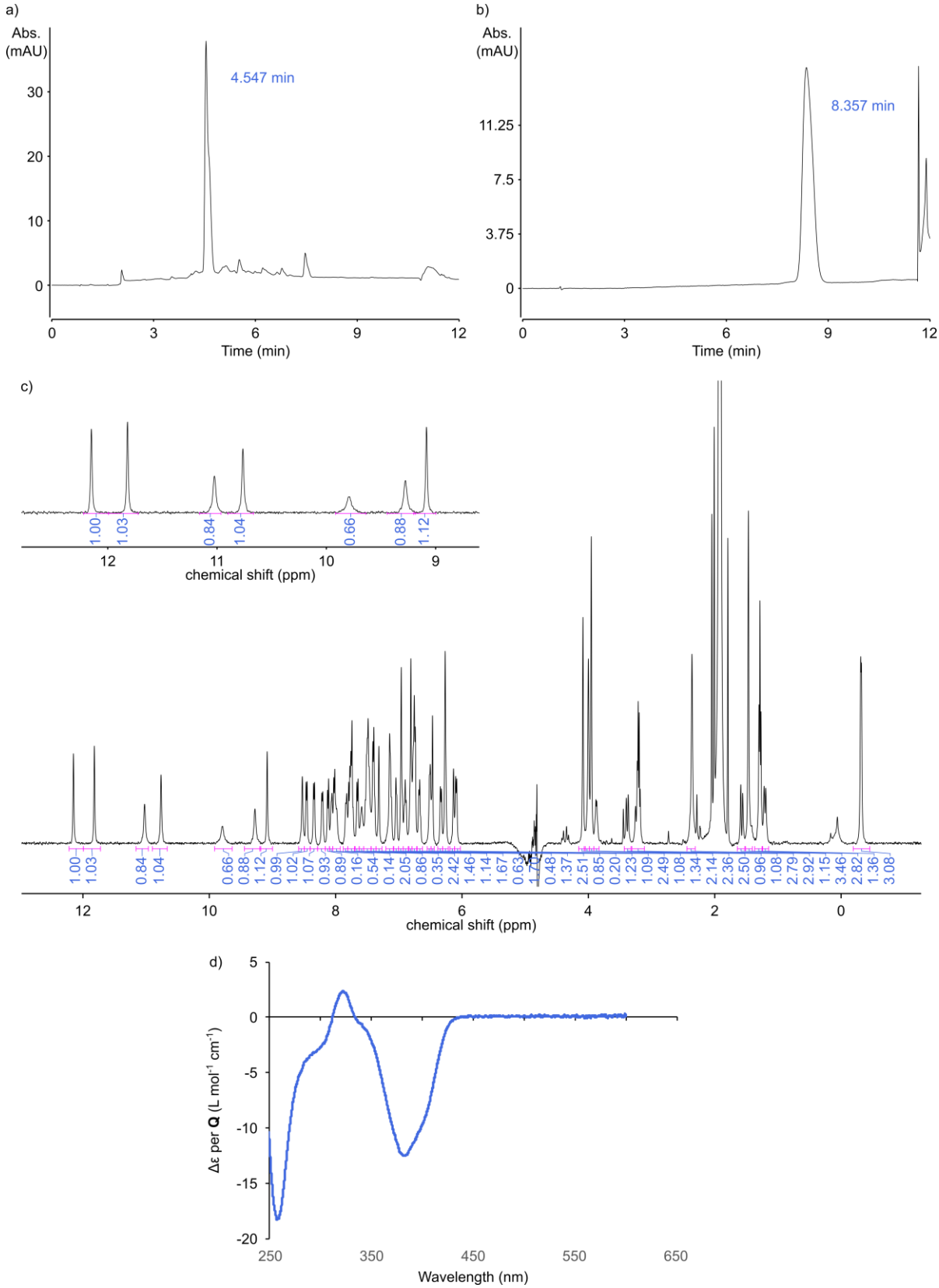


Fig. 55 Analytical data of compound **4**. HPLC chromatograms after cleavage from the resin (a) (C18, 0–60B, 25 °C; A: 13mM NH_4OAc buffer pH 8.5, B: acetonitrile) and after purification (b) (C18, 10–20B, 25 °C; A: 13 mM NH_4OAc buffer pH 8.5, B: acetonitrile). c) ^1H NMR (500 MHz, 0.13 mM in 12 mM NH_4OAc buffer pH 8.5 H_2O/D_2O 9:1, H_2O suppression). d) CD spectrum (10.2 µM in 13 mM NH_4OAc buffer pH 8.5).

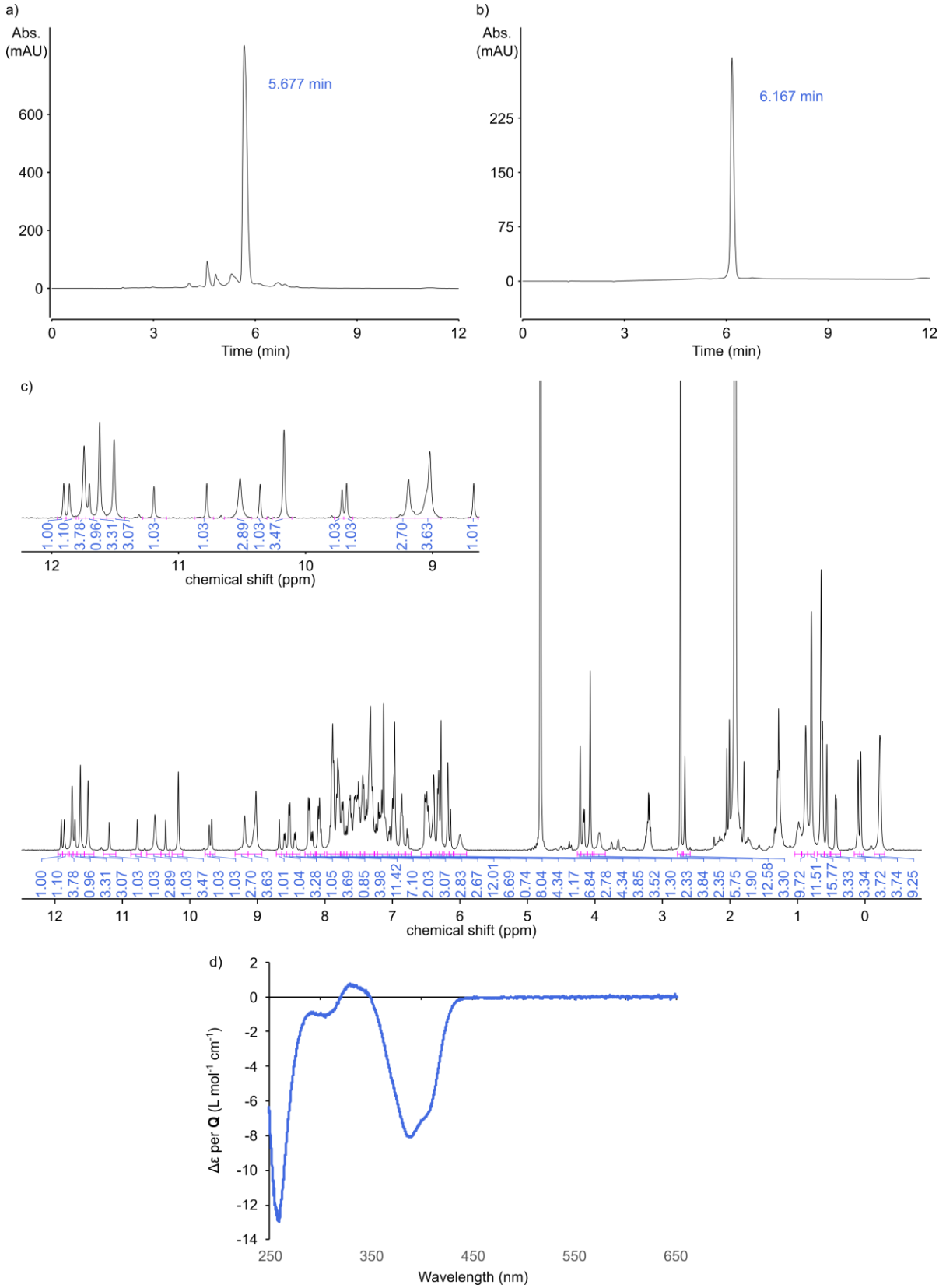


Fig. 56 Analytical data of compound **5**. HPLC chromatograms after cleavage from the resin (a) and after purification (b) (C18, 0–60B, 25 °C; A: 13 mM NH_4OAc buffer pH 8.5, B: acetonitrile). c) ^1H NMR (500 MHz, 0.13 mM in 12 mM NH_4OAc buffer pH 8.5 H_2O/D_2O 9:1, H_2O suppression). d) CD spectrum (31.1 µM in 13 mM NH_4OAc buffer pH 8.5).

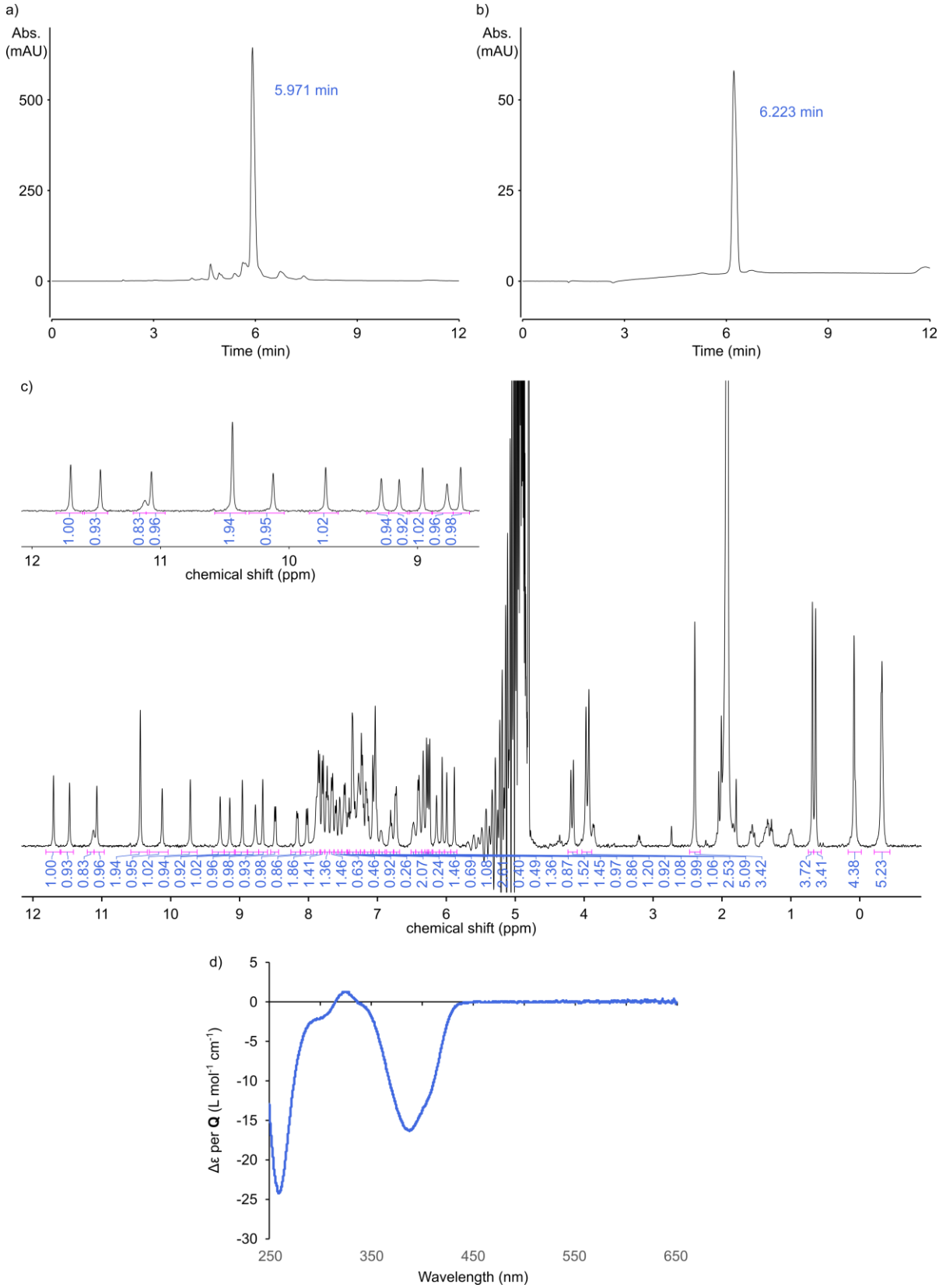


Fig. 57 Analytical data of compound **6**. HPLC chromatograms after cleavage from the resin (a) and after purification (b) (C18, 0–60B, 25 °C; A: 13 mM NH_4OAc buffer pH 8.5, B: acetonitrile). c) ^1H NMR (500 MHz, 0.13 mM in 12 mM NH_4OAc buffer pH 8.5 H_2O/D_2O 9:1, H_2O suppression). d) CD spectrum (8.73 µM in 13 mM NH_4OAc buffer pH 8.5).

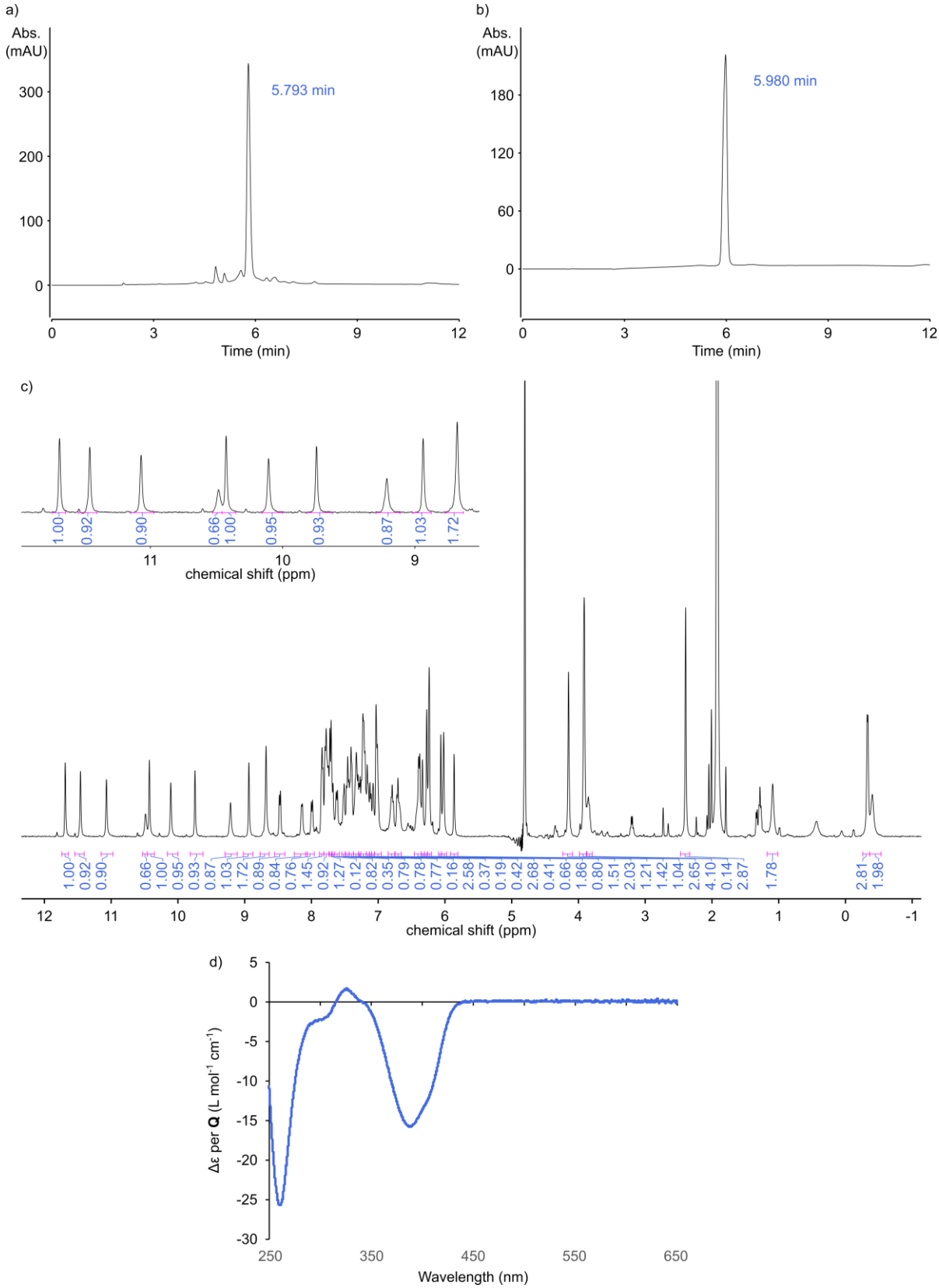


Fig. 58 Analytical data of compound **7**. HPLC chromatograms after cleavage from the resin (a) and after purification (b) (C18, 0–60B, 25 °C; A: 13 mM NH_4OAc buffer pH 8.5, B: acetonitrile). c) ^1H NMR (500 MHz, 0.13 mM in 12 mM NH_4OAc buffer pH 8.5 H_2O/D_2O 9:1, H2O suppression). d) CD spectrum (9.04 µM in 12 mM NH_4OAc buffer pH 8.5).

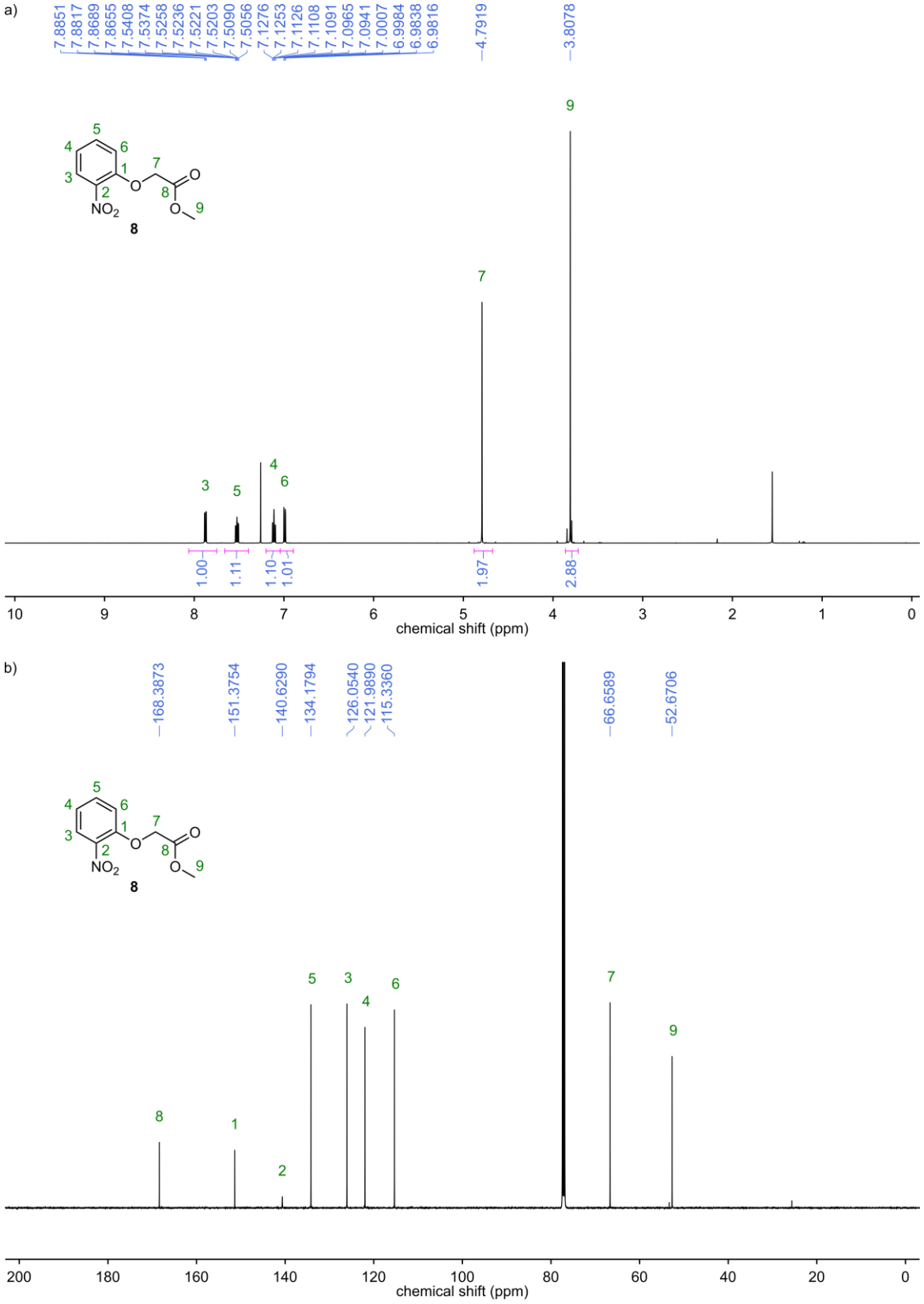


Fig. 59 NMR spectra of compound **8**. a) ^{1}H NMR (500 MHz, $CDCl_3$). b) ^{13}C NMR (126 MHz, $CDCl_3$).

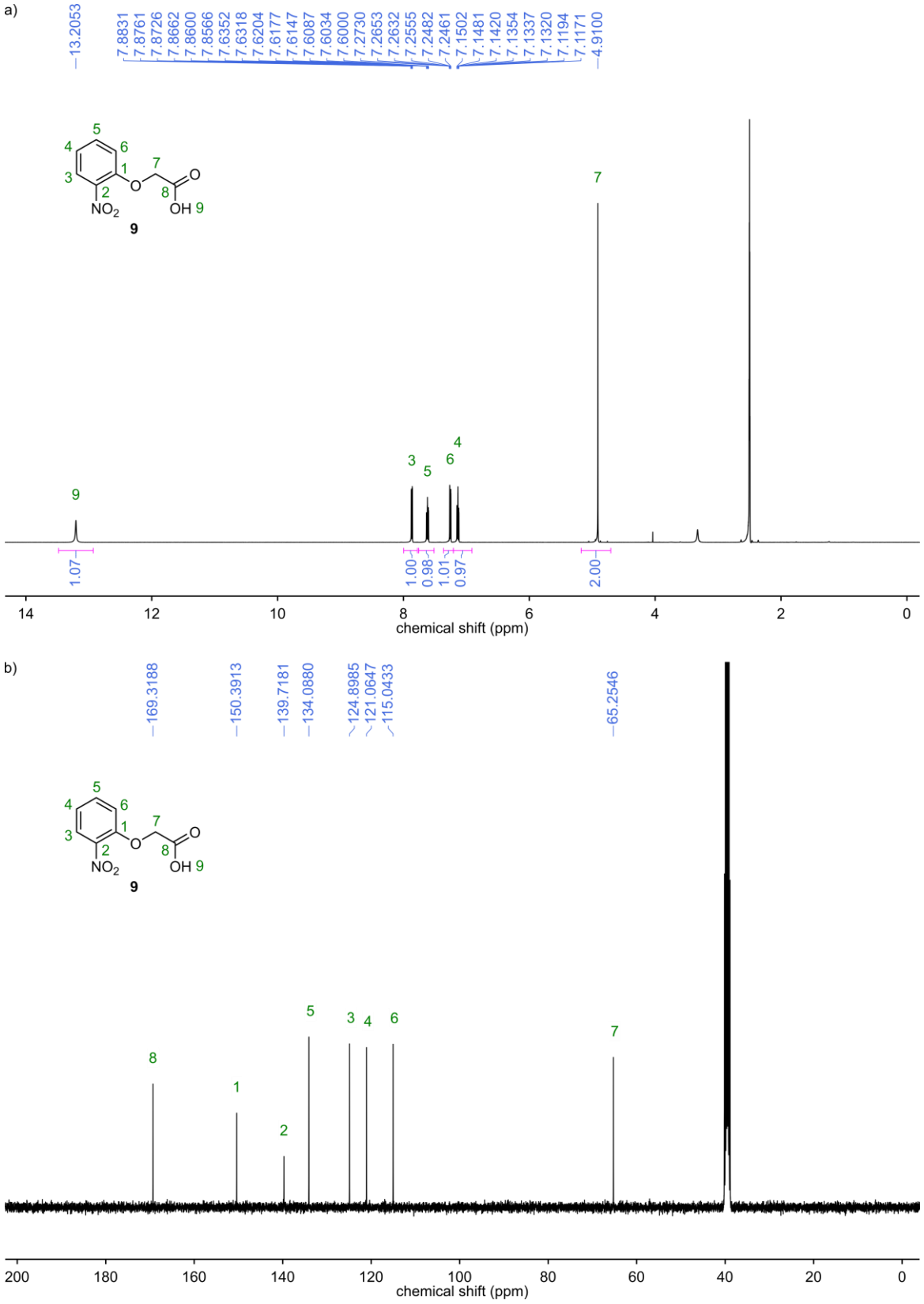


Fig. 60 NMR spectra of compound **9**. a) ^{1}H NMR (500 MHz, DMSO-d$_6$). b) ^{13}C NMR (101 MHz, DMSO-d$_6$).

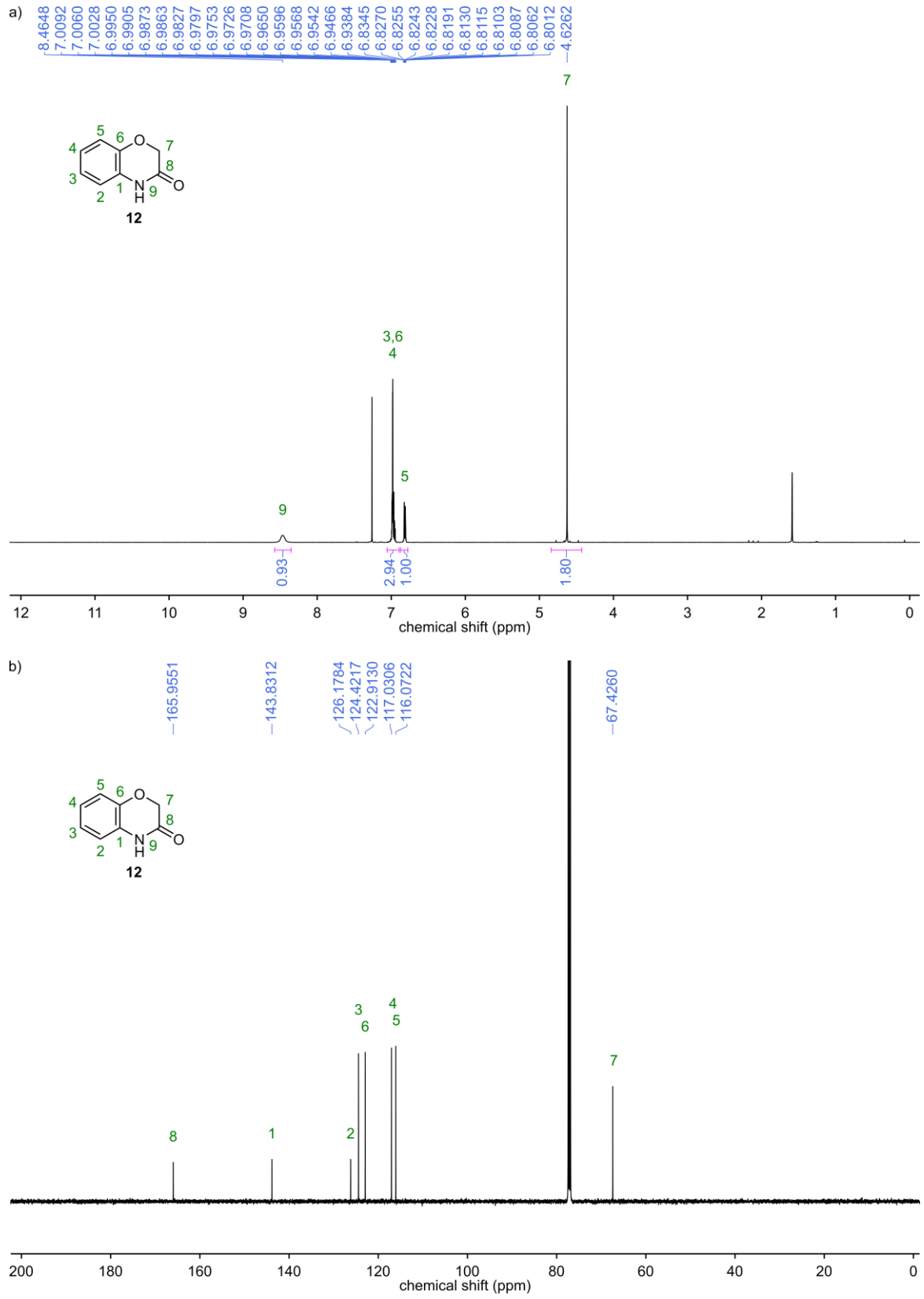


Fig. 61 NMR spectra of compound **12**. a) ^{1}H NMR (500 MHz, $CDCl_3$). b) ^{13}C NMR (126 MHz, $CDCl_3$).

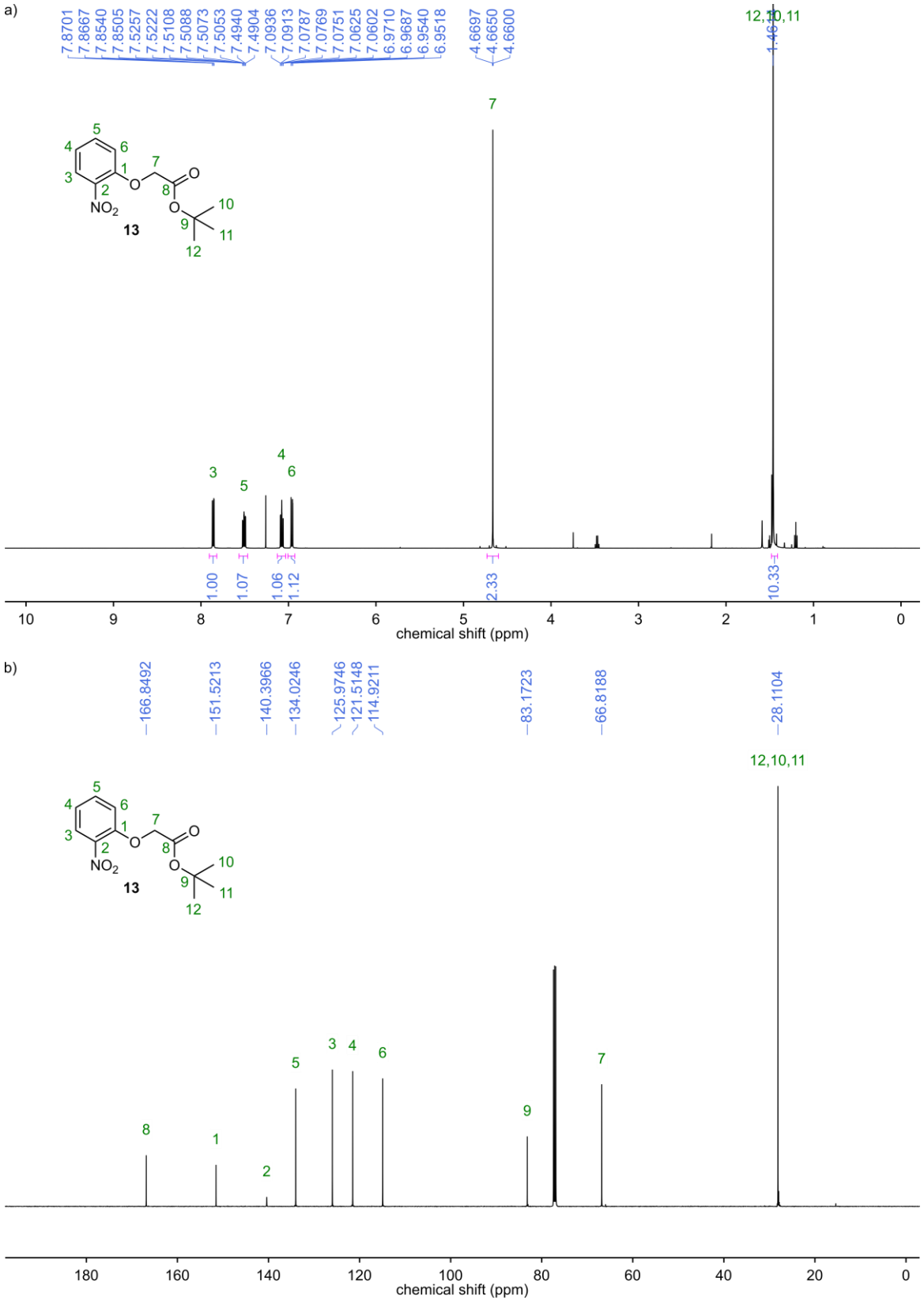


Fig. 62 NMR spectra of compound **13**. a) ^{1}H NMR (500 MHz, $CDCl_3$). b) ^{13}C NMR (126 MHz, $CDCl_3$).

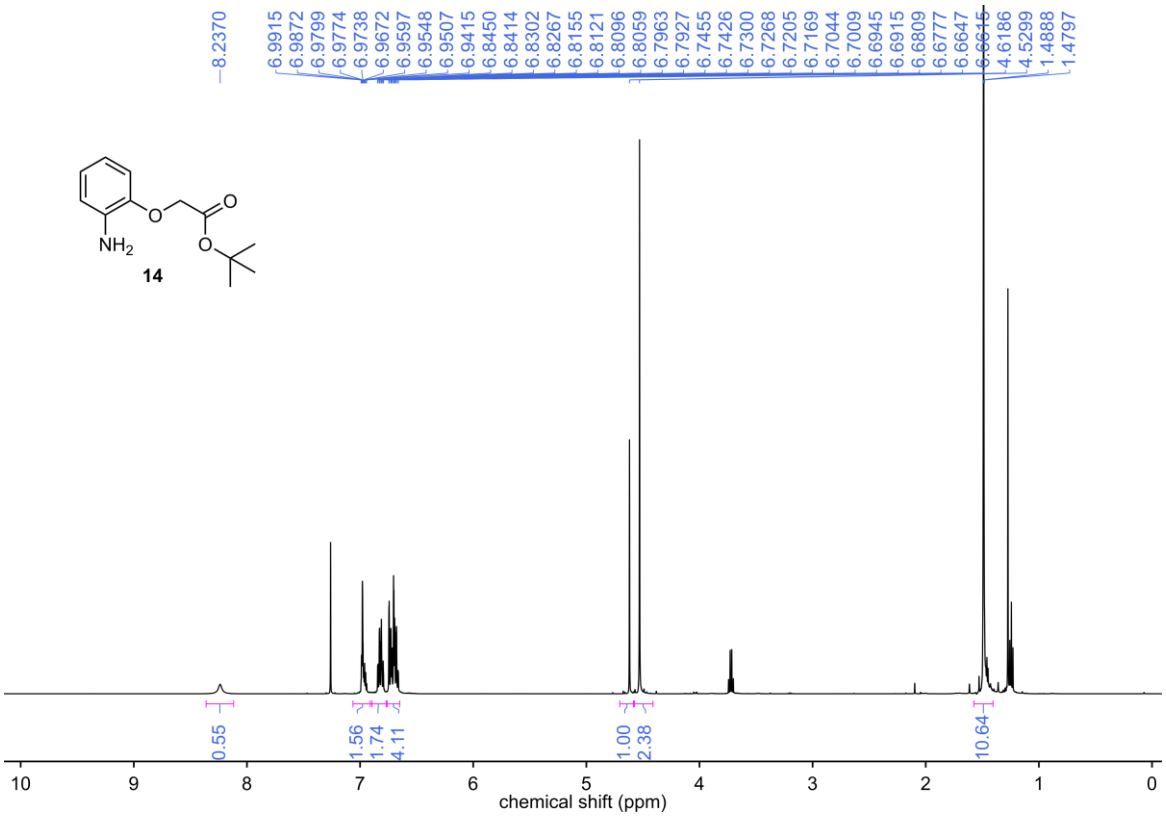


Fig. 63 ^{1}H NMR spectrum of compound **14** (500 MHz, $CDCl_3$).

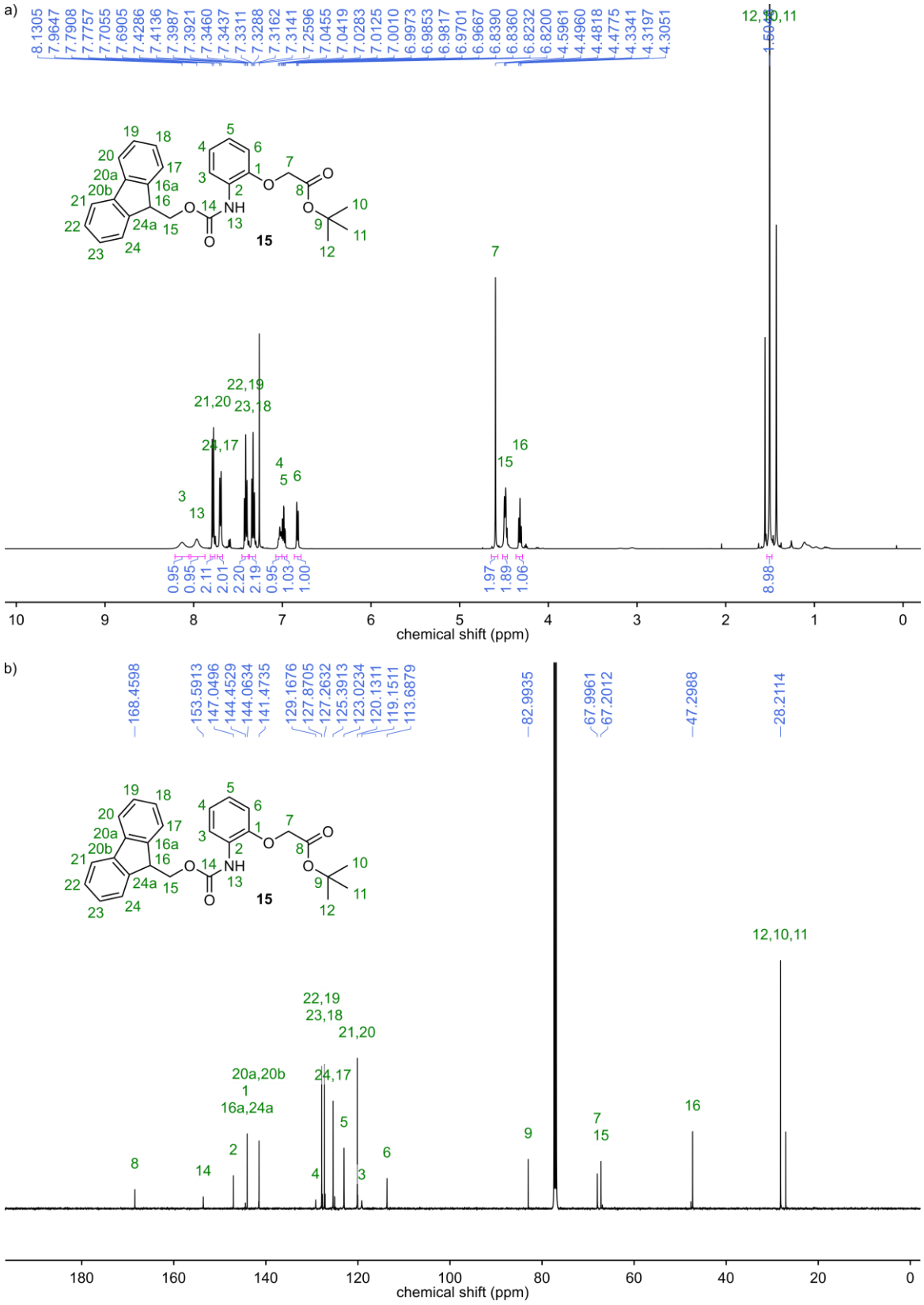


Fig. 64 NMR spectra of compound **15**. a) ^{1}H NMR (500 MHz, $CDCl_3$). b) ^{13}C NMR (126 MHz, $CDCl_3$).

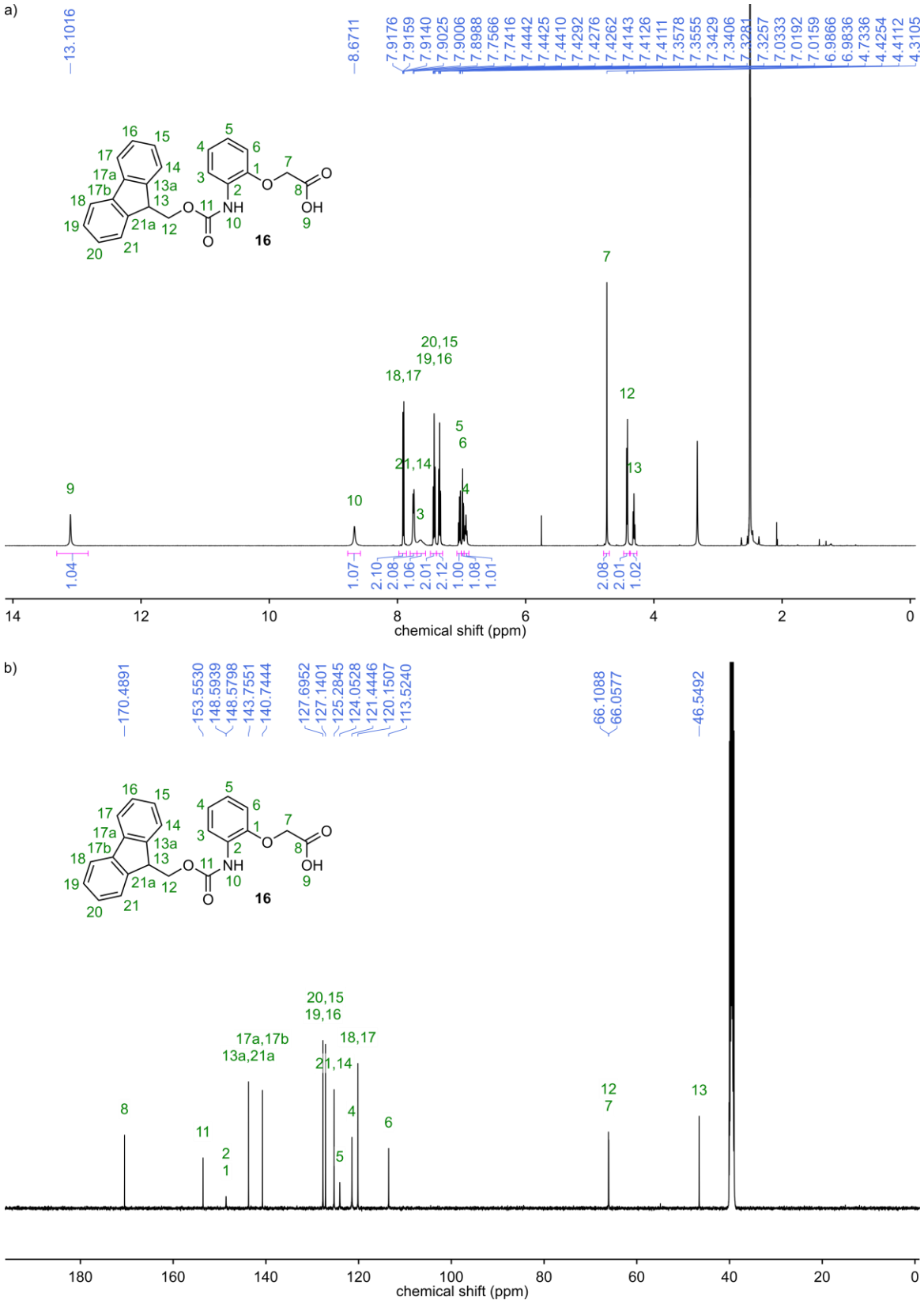


Fig. 65 NMR spectra of compound **16**. a) ^{1}H NMR (500 MHz, DMSO-d$_6$). b) ^{13}C NMR (126 MHz, DMSO-d$_6$).

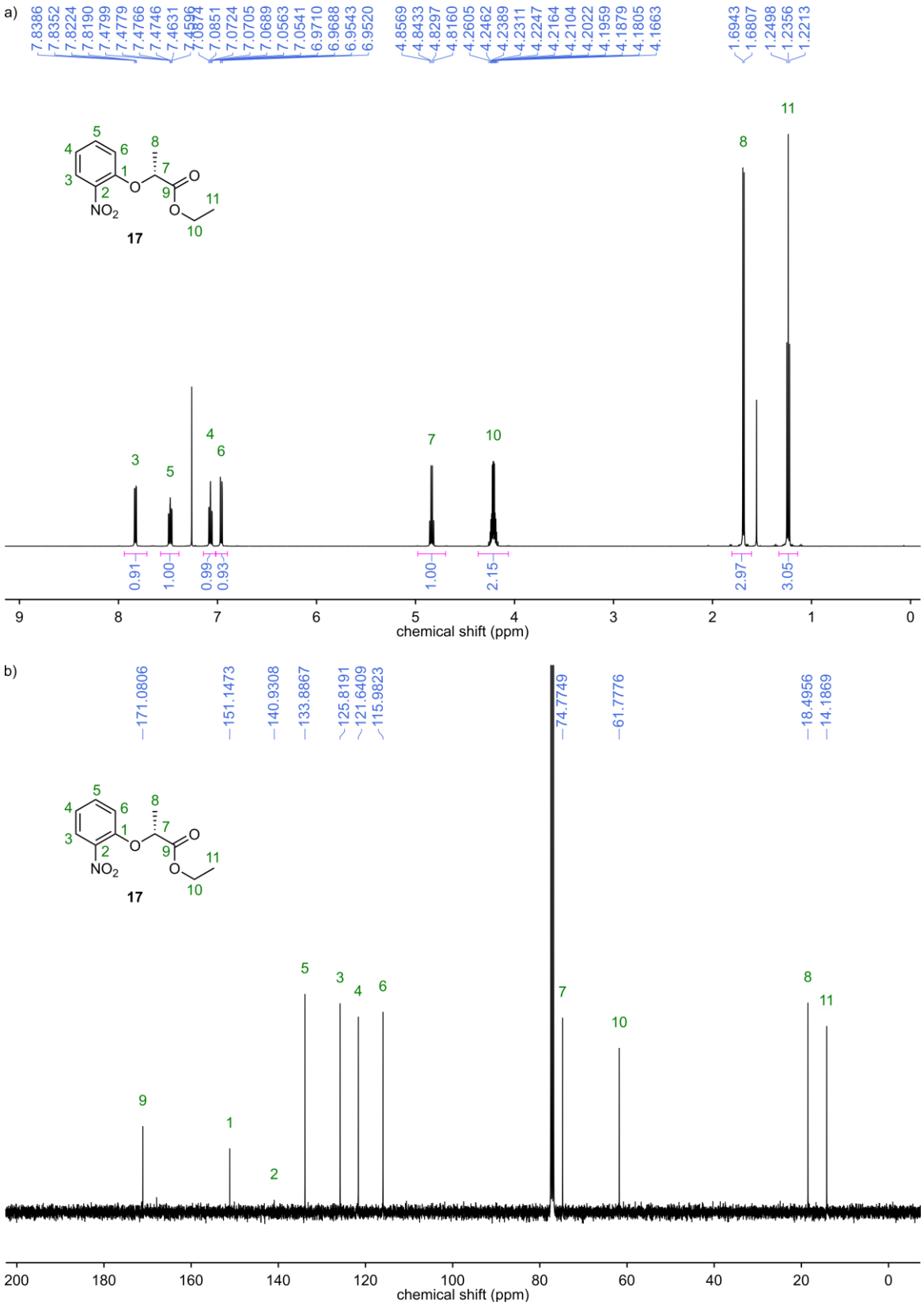


Fig. 66 NMR spectra of compound **17**. a) ^{1}H NMR (500 MHz, $CDCl_3$). b) ^{13}C NMR (101 MHz, $CDCl_3$).

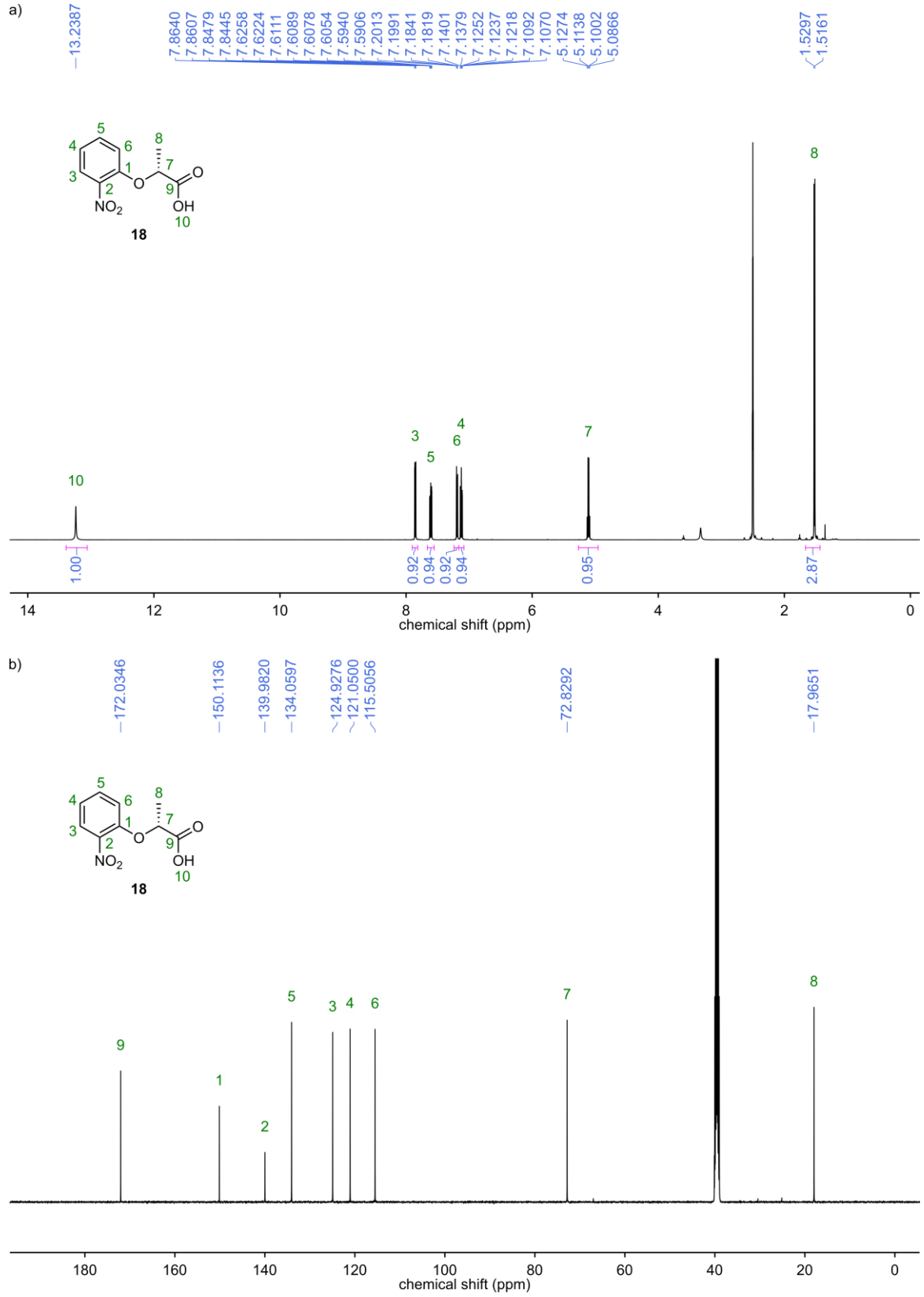


Fig. 67 NMR spectra of compound **18**. a) ^{1}H NMR (500 MHz, DMSO-d$_6$). b) ^{13}C NMR (126 MHz, DMSO-d$_6$).

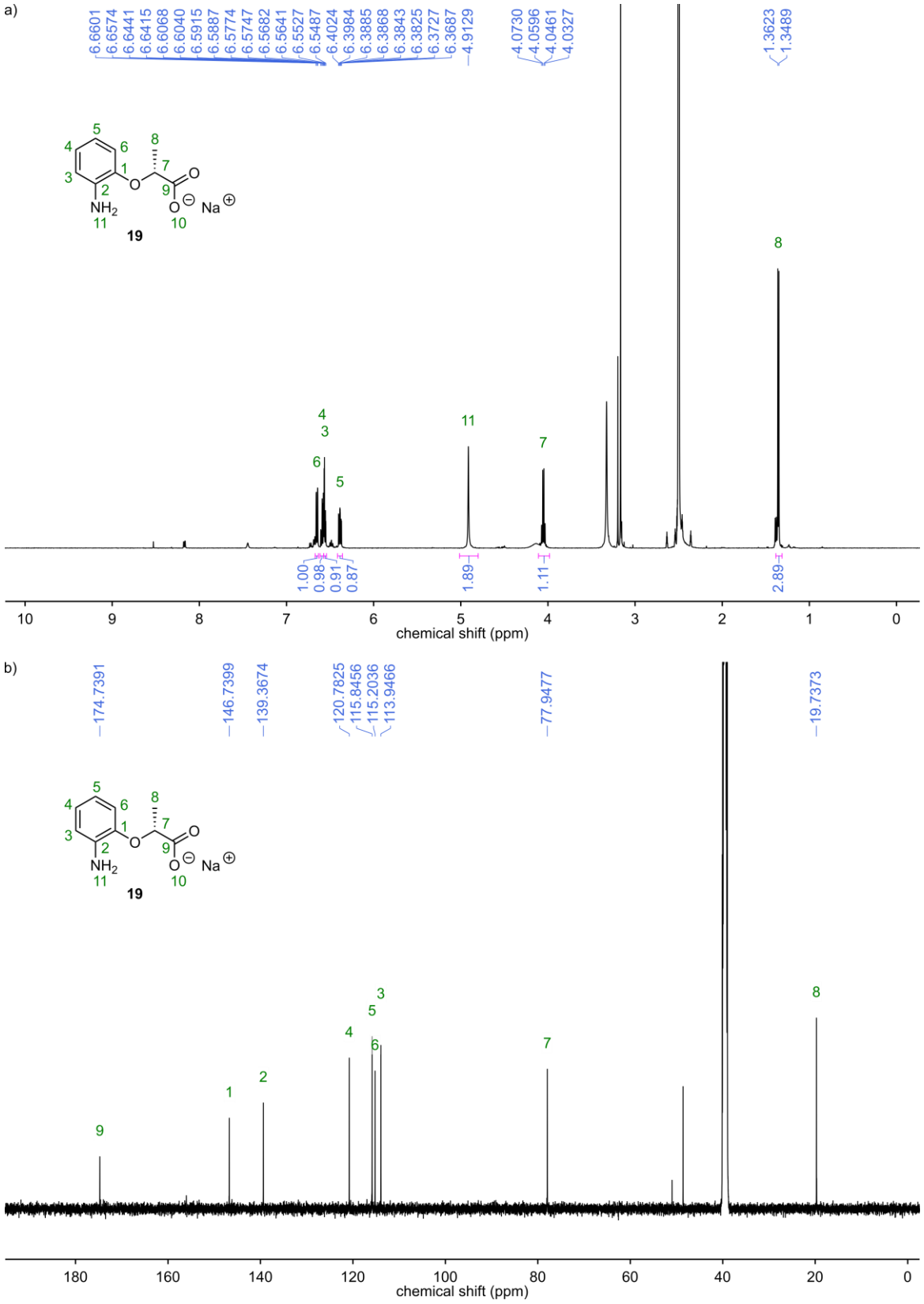


Fig. 68 NMR spectra of compound **19**. a) ^{1}H NMR (500 MHz, DMSO-d$_6$). b) ^{13}C NMR (126 MHz, DMSO-d$_6$).

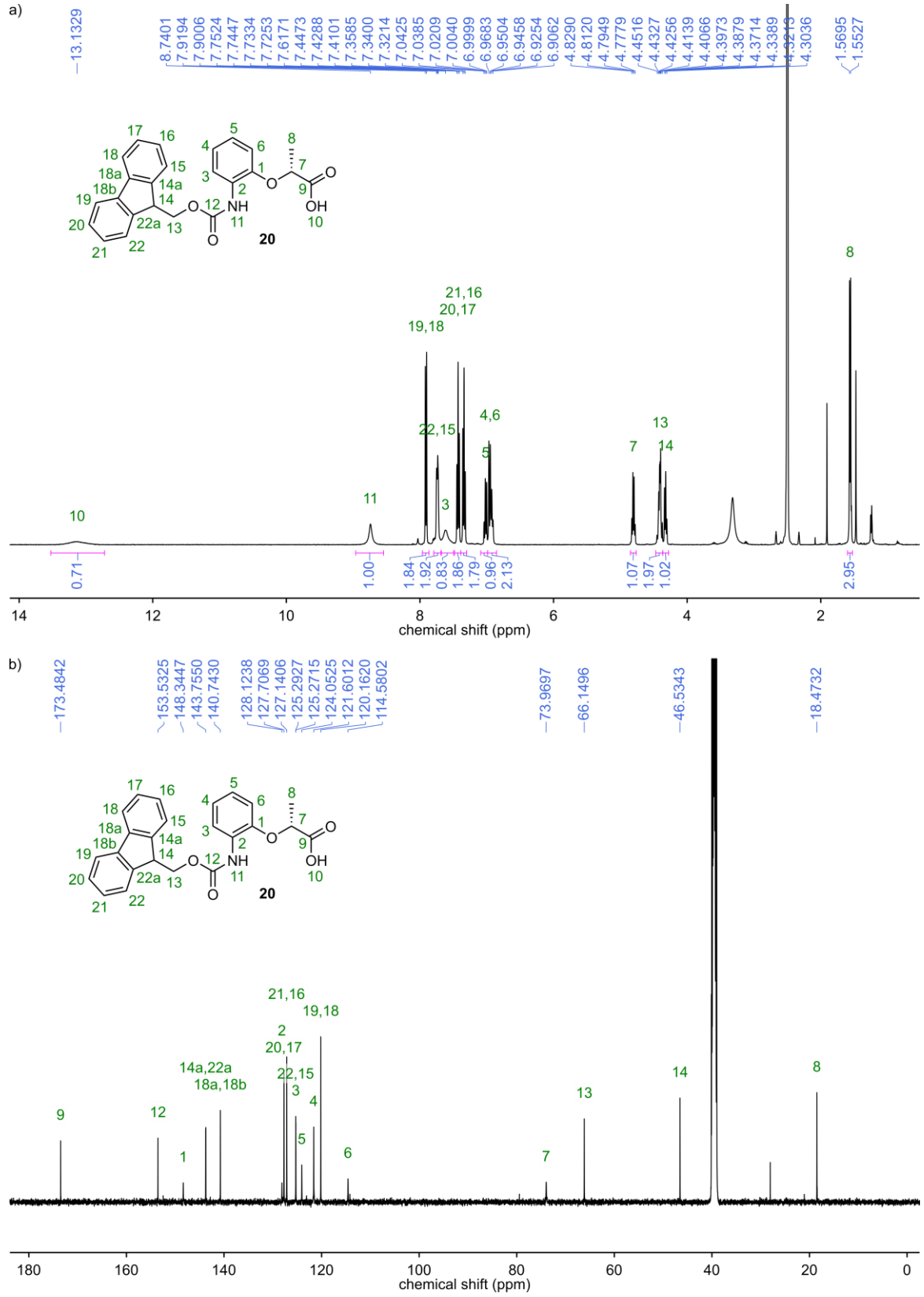


Fig. 69 NMR spectra of compound **20**. a) ^{1}H NMR (500 MHz, DMSO-d$_6$). b) ^{13}C NMR (126 MHz, DMSO-d$_6$).

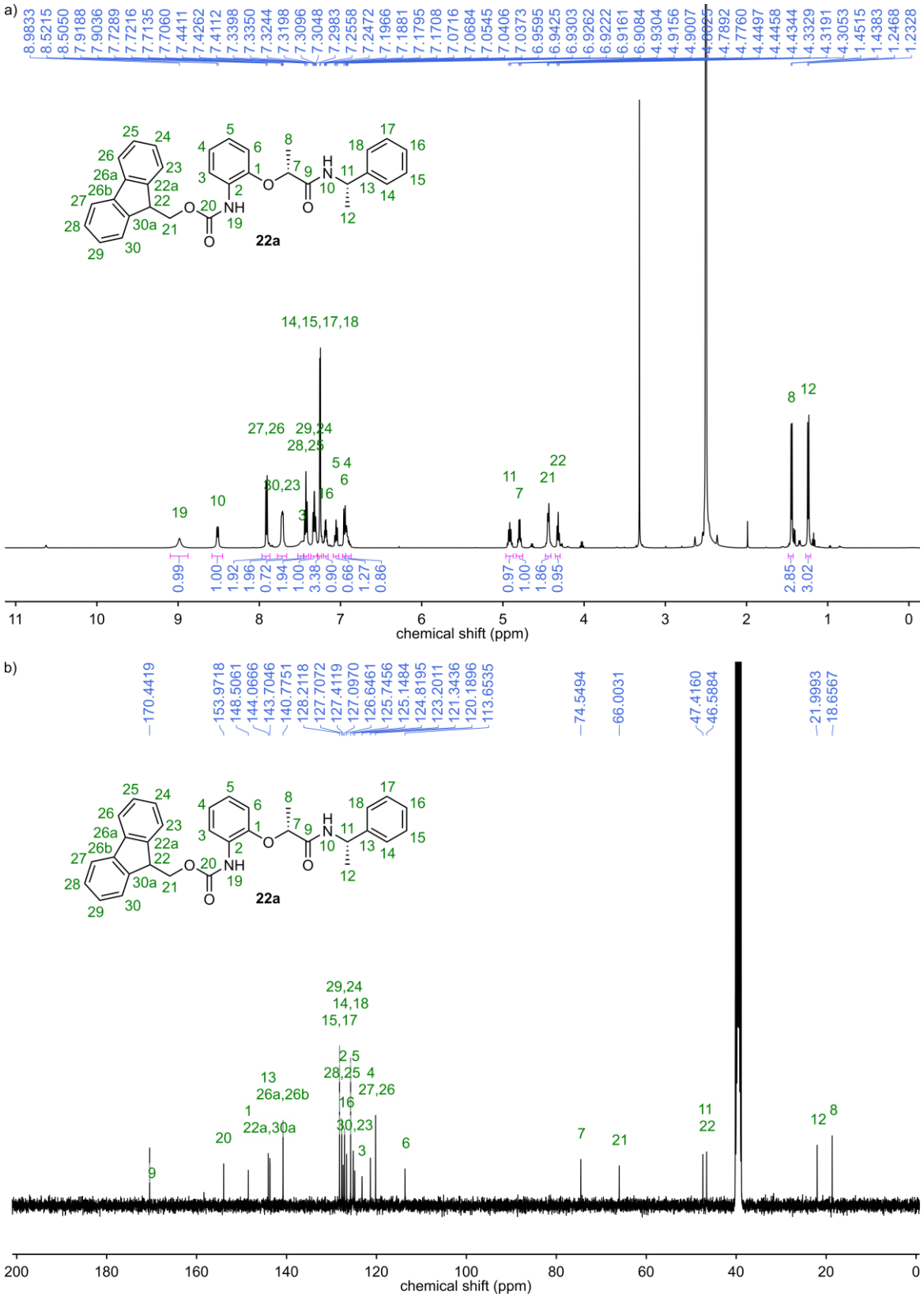


Fig. 70 NMR spectra of compound **22**. a) ^{1}H NMR (500 MHz, DMSO-d$_6$). b) ^{13}C NMR (101 MHz, DMSO-d$_6$).

6 Discrete dimerization of aromatic helices in water

Controlled self-assembly of foldamers in aqueous media through discrete binding interfaces is rare in the literature. Most examples involve biotic foldamers such as β-peptides[7, 72] and ureas[73] and are inspired by coiled coils. We adopted a similar approach but instead using aromatic quinoline-based foldamers. Fig. 71 shows the model of oligomer **1**, which was designed to dimerize via salt bridge interactions based on the "flat" interface approach. In the model, these interactions take place between positively charged ammonium and negatively charges sulfonate side chains as well as backbone carbonyl oxygens. Additionally, the "flat" surface generated by B-monomers should support this aggregation through hydrophobic effects. Because of the convincing shape complementarity of the oligomer, we proceed with the synthesis and analysis of the compound in aqueous media. Indeed, a dimerization phenomenon could be detected in dilution studies monitored by NMR spectroscopy. However, as described in the following publication in *Angewandte Chemie*, the aggregation was facilitated by stacking of the C-terminal cross sections of the oligomers, not by side-to-side salt bridges.

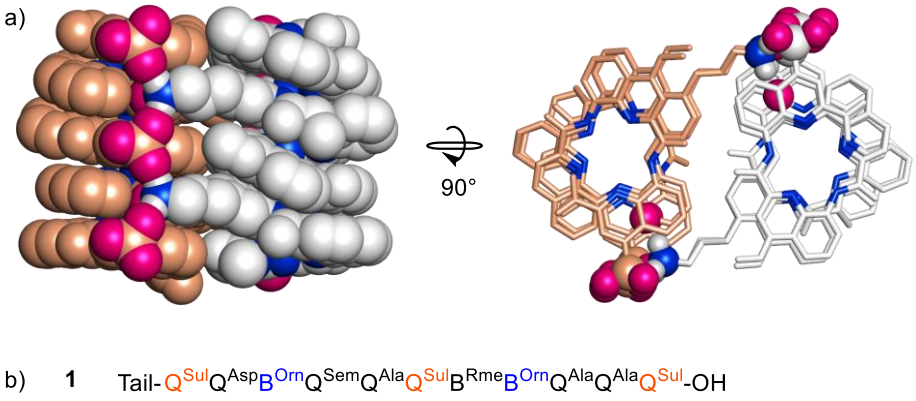


Fig. 71 – Model of a "flat" interface bundle supported by salt bridges. a) Side view shown in CPK representation. In the top view, groups that are involved in hydrogen bonding are shown as balls. Nitrogen atoms are shown in blue, and oxygen atoms involved in hydrogen bonding are highlighted in purple. Irrelevant side chains and hydrogen atoms are omitted for clarity. b) Sequence **1** corresponding to the model. Units involved in salt bridges are shown in color.

A similar type of aggregation was previously investigated by Zeng and coworkers using pyridine oligoamides.[79] In 2012, they reported a pentamer functionalized with "sticky" end groups at its termini: an ester and a Cbz group that can interact via weak hydrogen bonds (Fig. 72a, b).[79a] The foldamer showed oligomerization into stacked columns in the solid state leading to chiral crystallization. Furthermore, the oligomers are able to encapsulate guests like methanol or dichloromethane in their helix cavity. By replacing Cbz with a benzoyl moiety, the water affinity of the pore formed by the helical stack could be fine-tuned to allow a proton gradient induced water transport when incubated with large unilamellar vesicles (Fig. 72c).[79b] Later, it was shown that the Cbz

functionalized pentamer—albeit having a lower binding affinity for water molecules—has an even higher water transport capability of 3×10^9 H_2O s^{-1} channel^{-1} when incorporated into membranes.[126] This self-assembly is interesting for the use in nanotechnology; however, it is not discrete and possesses only a weak binding strength in solution.

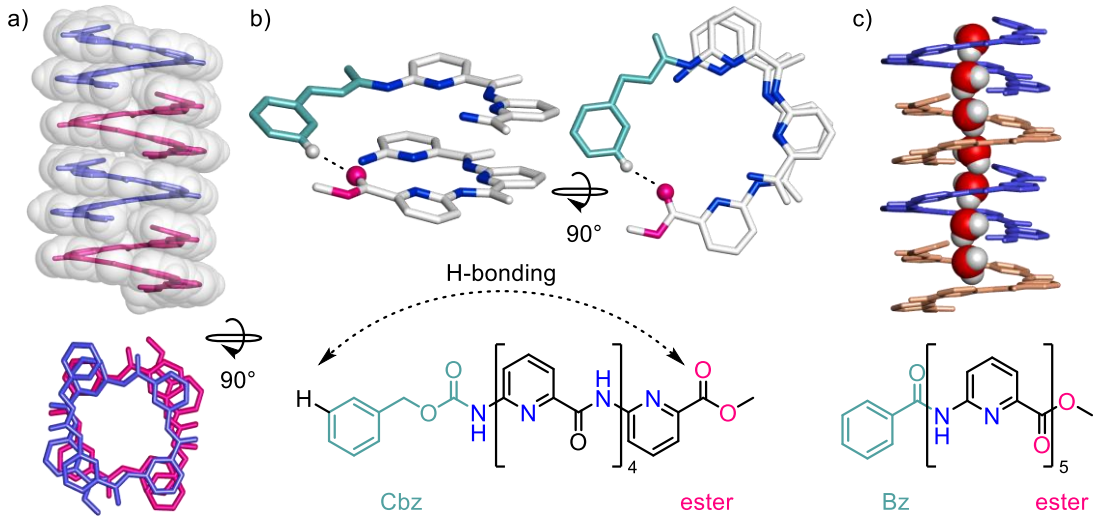


Fig. 72 – Self-assembly through stacking of pyridine oligoamides. a) Crystal packing of a pyridine oligoamide pentamer with an N-terminal Cbz and a C-terminal ester function in side (overlayed with a transparent CPK representation) and top view.[79a] Individual oligomers are shown in different colors. b) Part of the crystal structure shown in (a). Hydrogen bonds are indicated by dashed lines and atoms involved in hydrogen bonding are shown as balls. N- and C-terminal groups are highlighted in turquoise and purple, respectively. Nitrogen atoms are shown in blue. c) Crystal packing of a pyridine oligoamide pentamer with an N-terminal Bz and a C-terminal ester function in side view.[79b] Individual oligomers are shown in different colors. Water molecules are shown in CPK representation (oxygen: red, hydrogen: gray). The chemical formula is given next to the respective structures. Irrelevant hydrogen atoms are omitted for clarity.

The stacking interaction described here is special since it is both strong and discrete, as opposed to many other stacking interfaces reported so far, that lead to polymeric aggregation.[79a-c, 127] This results from the C_2-symmetry of the interaction leading to homodimer formation (either *P-P* or *M-M* chirality). Additionally, binding strength can be influenced by the nature of the side chains close to the interface and pH variation. Most importantly, it can be programmed to exclusively form heteromeric aggregates through the use of an amide terminus on one binding partner in the right concentration and pH range. These features make the interaction an excellent starting point for the design of bigger self-assembled structures in combination with other binding motives.

Contributions: The initial model was designed by me. Subsequent, design plans were developed in collaboration with I. Huc. Chemical syntheses and most analyses were conducted by me. L. Allmendinger helped with the design and interpretation of NMR experiments. Crystallization experiments were conducted in collaboration with P. K. Mandal. X-ray measurements were done at EMBL Hamburg with assistance of Dr Saravanan Panneerselvam, and the structure was refined by P. K. Mandal. The manuscript was written by I. Huc and me in collaboration with L. Allmendinger and P. K. Mandal.

Publication

Discrete Stacked Dimers of Aromatic Oligoamide Helices

Authors: Daniel Bindl, Pradeep K. Mandal, Lars Allmendinger and Ivan Huc

Published: *Angew. Chem. Int. Ed.*, **2022**, e202116509, DOI: 10.1002/anie.202116509

Abstract: Tight binding was observed between the C-terminal cross section of aromatic oligoamide helices in aqueous solution, leading to the formation of discrete head-to-head dimers in slow exchange on the NMR timescale with the corresponding monomers. The nature and structure of the dimers was evidenced by 2D NOESY and DOSY spectroscopy, mass spectrometry and X-ray crystallography. The binding interface involves a large hydrophobic aromatic surface and hydrogen bonding. Dimerization requires that helices have the same handedness and the presence of a C-terminal carboxy function. The protonation state of the carboxy group plays a crucial role, resulting in pH dependence of the association. Dimerization is also influenced by neighboring side chains and can be programmed to selectively produce heteromeric aggregates.

Structurally precise and designable interaction interfaces are a key component of large self-assembled architectures. In this respect, folded molecules constitute building blocks of unmatched sophistication. The assemblies formed by proteins and nucleic acids in living systems, e.g. virus capsids or ribosomes, provide spectacular illustrations of the structures and functions enabled by assembling folded building blocks. There has thus been strong interest for programming interaction interfaces in artificial folded molecules. Great advances have been made using non-natural DNA[128] and protein[129] sequences. Using smaller molecules, the bundling of α-helical peptides is now so well understood that it permits reliable programming and function.[130] By extension, helix bundles have been produced from β-peptides[7, 131] and urea-based γ-peptide isosteres.[73, 132] Assemblies made of completely abiotic folded building blocks could bring advantages of their own including biochemical and thermal resistance as well as unrestricted functionalization. However, this line of research is less advanced because well-defined interaction interfaces have been lacking. Examples include the bundling of aromatic helices[42, 70] and the formation of multistranded helices[45a, 75, 78, 133] and sheets[134] primarily in organic solvents. Here we report the serendipitous discovery of stable discrete dimers of aromatic oligoamide helices in water. We find that aromatic stacking and hydrogen bonding mediate the dimerization of the helix C-terminal cross-section in a pH and side-chain-dependent manner. Aggregation can also be made heteromeric but, in all cases, it remains discrete. This contrasts with the stacking of many other aromatic objects,[127] including helices,[79a-c, 126] that tend to form polymeric assemblies.

Sequence **1** (Fig. 73) comprises Q and B aromatic monomers that code for the formation of stable helices according to well-established design principles.[87, 135] The chiral B^{Rme} unit was introduced to quantitatively bias handedness towards the *M* (left-handed) helix.[93] The various positively and

negatively charged side chains provide solubility in water and were originally designed to promote helix bundling via side-chain mediated salt bridges. As shown in the following, helices do aggregate but not in the way initially intended.

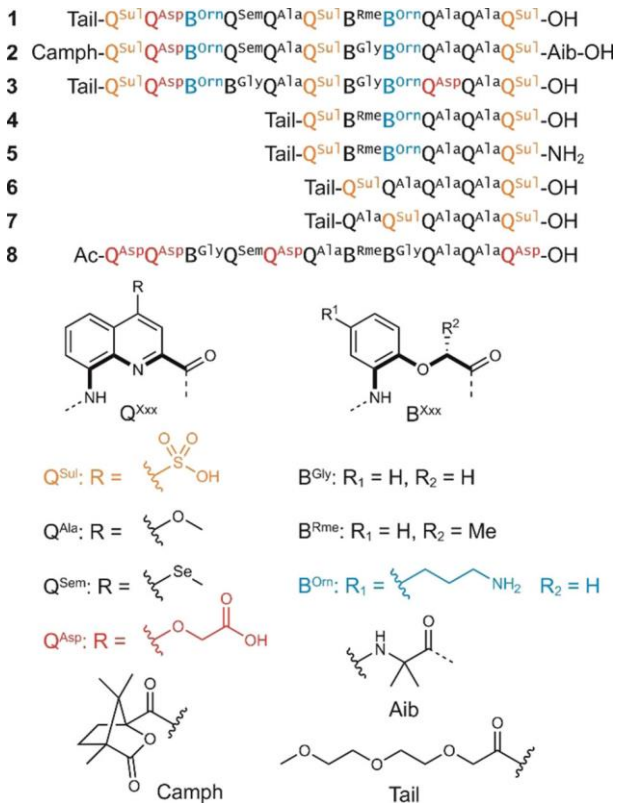


Fig. 73 Sequences and building blocks that were investigated in this study. Monomers possessing charged side chains are highlighted with color (sulfonate: orange, carboxylate: red, ammonium: blue).

The ^{1}H NMR spectrum of **1** in water shows two sets of sharp signals the proportion of which change with concentration (Fig. 74a), indicating a reversible aggregation phenomenon in slow exchange on the NMR timescale. The signals coalesce upon heating to 50 °C (Fig. 77). The number of signals indicate that the aggregate is on average symmetrical, i.e. its helical subcomponents are in the same environment. The signals of the aggregated species, including that of the CH_3 group of the B^{Rme} monomer below 0 ppm, are upfield-shifted. Upfield shifts associated with ring current effects are typically observed upon elongating helical sequences,[136] and suggest that helices of **1** may stack via their aromatic cross-section. The ESI-MS shows a large population of $[2M-2H]^{2-}$ and $[2M-3H]^{3-}$ dimeric species in the gas phase (Fig. 78).

A crystal structure of **1** confirmed the expected helical structure (Fig. 74b).[‡] Packing in the lattice shows extended head-to-head stacks of helices and revealed several pairwise helix-helix contacts (Fig. 99–Fig. 104). Rare intermolecular salt-bridges were dismissed as the possible driving force of

[‡] Deposition Numbers 2122518 contain the supplementary crystallographic data for this paper. These data are provided free of charge by the joint Cambridge Crystallographic Data Centre and Fachinformationszentrum Karlsruhe Access Structures service.

aggregation in water. Stacking of the helix N-terminal cross sections involved a reduced aromatic surface and was also dismissed. In contrast aromatic contacts between the C-terminal cross sections were extensive (total buried surface of 431 Å^2)[§] and accompanied by a close proximity of the terminal carboxy groups that can be attractive only if one of the two is in its protonated form (Fig. 74d). In addition, tight side-by-side pair-wise contacts were observed that involve multiple hydrophobic aromatic and aliphatic CH groups that could not be dismissed without further experiment as the possible reason for aggregation in solution.

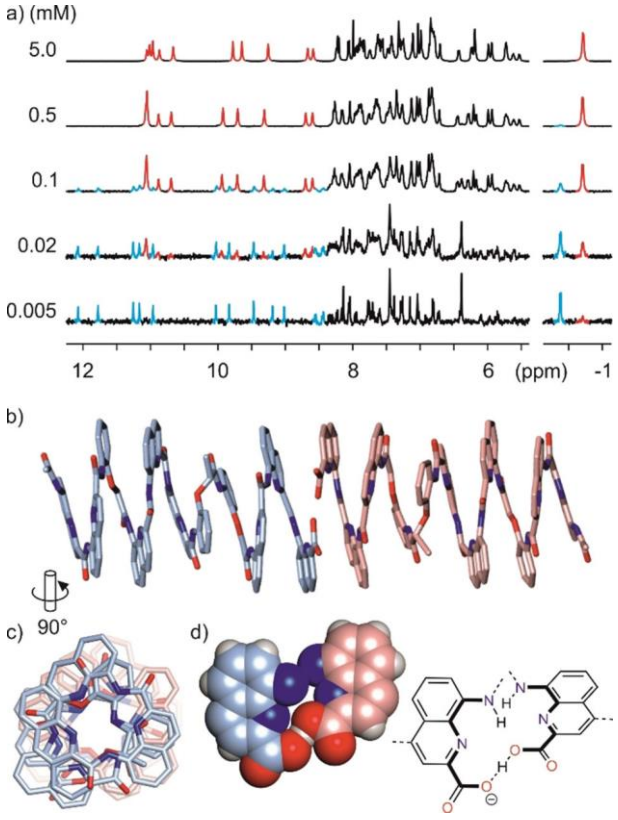


Fig. 74 a) ^{1}H NMR spectra of compound **1** at different concentrations in 27 mM sodium phosphate aqueous buffer pH 7.0. Selected distinct signals belonging to either the monomeric or dimeric form are highlighted in cyan and red, respectively. X-ray structure of **1** showing the stacking of two helices at their C-terminus in side view (b) and top view (c). The two crystallographically distinct molecules engaged in binding are colored in cyan and pink, respectively. Side chains, solvent molecules and hydrogen atoms are omitted for clarity. d) Fragment of the helix–helix binding interface showing the C-terminal monomers in space-filling representation. Side chains and solvent molecules are omitted. The acidic proton is not visible in the electron density map. The assignment of the carboxylic acid and carboxylate function is tentative, based on orientation.

The ^{1}H NMR spectrum of the aggregate of **1** was assigned using bidimensional NMR experiments (Supporting Information section 6.3.1). NOE correlations were for most compatible with intramolecular contacts but at least two correlations could be explained only when the C-terminal cross-section of two helices are stacked (Fig. 75a). In addition, analogous sequence **2**, which has the same

[§] Solvent accessible surface of two monomers minus the solvent accessible surface of one dimer, as calculated in PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.

side chains as **1** and a C-terminal Aib extension, does not aggregate and its signals are not upfield-shifted, i.e. they appear in the same range as those of the monomer of **1**. When mixed with **1**, compound **2** does not interfere with the aggregation of **1** and a DOSY spectrum shows that **2** is a smaller species despite having the additional Aib (Fig. 75b), thus hinting at a monomeric state. In contrast shorter sequence **4** does aggregate in a similar manner as **1** (Fig. 84).

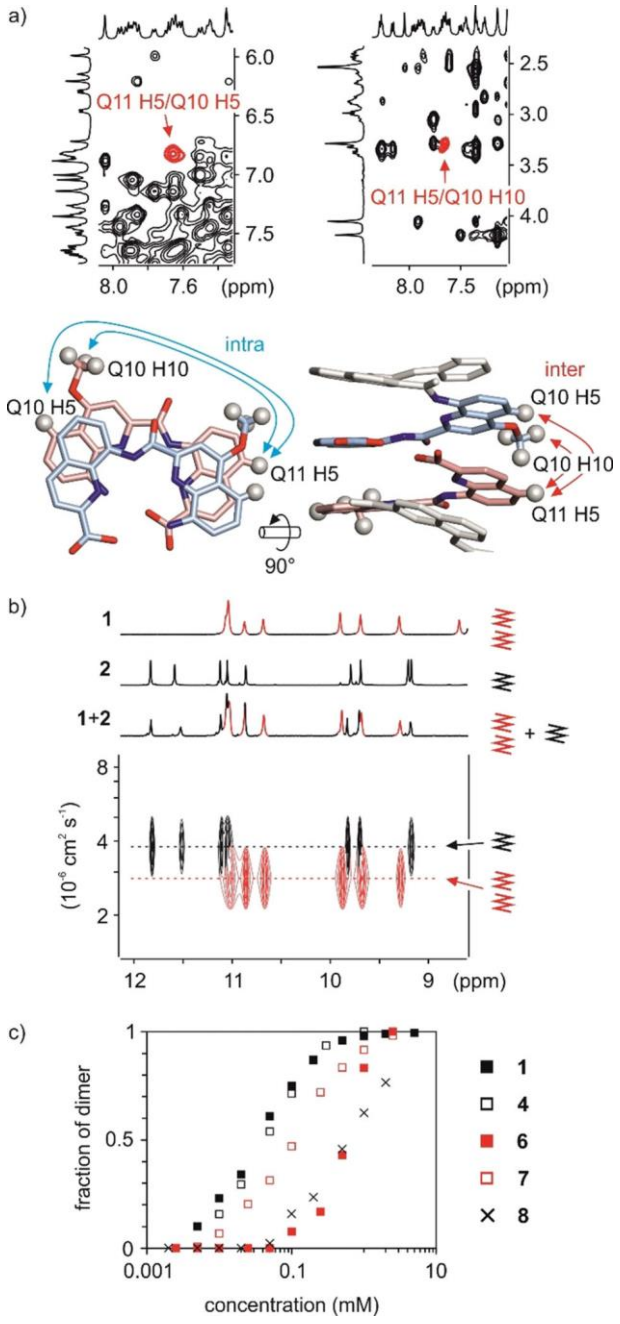


Fig. 75 a) Selected parts of the NOESY spectrum of **1**. The intermolecular NOE correlations between Q11 H5/Q10 H5 and Q11 H5/Q10 H10 are highlighted in red. Structural models indicate the location of the protons involved in these correlations showing why these are likely to be intermolecular correlations and not intramolecular contacts. b) Amide-region of the ^{1}H NMR spectra of **1** (1 mM) in 27 mM sodium phosphate aqueous buffer pH 7.0, **2** (1 mM) and part of the DOSY spectrum of a mixture of **1** (1 mM) and **2** (1 mM) in 27 mM sodium phosphate aqueous buffer pH 7.0. Signals corresponding to compound **1** are highlighted in red. The two different levels of signals in the DOSY spectrum are indicated by dashed lines. c) ^{1}H NMR titration data of compounds **1**, **4**, **6**–**8**

in 27 mM sodium phosphate aqueous buffer pH 7.0. Relative integrals of selected signals are plotted against sample concentration.

Altogether, these data clearly support the formation of discrete dimers of **1** in aqueous solution via head-to-head stacking of the C-terminal cross-section. That simple stacking and one carboxyl-carboxylate hydrogen bond give rise to slow exchange on the NMR timescale is quite remarkable. Discrete aggregation mediated by aromatic stacking has been reported for some macrocycles[137] but it remains rare in aromatic systems which more frequently form polymeric aggregates. This discovery made us realize that several water-soluble aromatic helices that we have reported in the past presumably dimerize in the same way as they show the exact same stacking motif of the C-termini in the solid state.[64] This had however been overlooked.

The relative position of the carboxylate and carboxylic acid in the solid-state results in dissymmetry within the dimer (Fig. 74d). In solution, NMR signals presumably reflect fast exchange between two degenerate dissymmetrical dimers upon proton exchange between the carboxylic acid and the carboxylate. The involvement of a carboxylic acid and a carboxylate was supported by different observations. Unlike sequence **4**, amide terminated analogous sequence **5** does not aggregate. In addition, the dimerization of **1** is hampered at higher pH (Fig. 81). Dimerization at pH 7 in fact indicates a significantly increased apparent p*K*a within the dimer. In the absence of external factors, the acid form is not expected at pH 7. These different effects explain why slightly different dissociation constants are calculated from the proportions of **1** and $(\mathbf{1})_2$ measured at different concentrations (Fig. 74a): the extent of protonation also slightly depends on concentration and this impacts dimerization. Nevertheless, an average K_d value of 30 ± 8 µmol can be extracted at neutral pH (27 mM phosphate buffer).

Dimerization was also investigated in sequences **6**–**8** (Fig. 75c, Fig. 85–Fig. 87) and was found to vary by up to 20-fold from compound to compound. These results hint at possible intermolecular charge repulsions between side chain within the dimers and at possible effects of the electron richness of the quinoline rings at the helix-helix interface. For example, the difference between **6** and **7** is a simple change of position of a sulfonate. Based on the crystal structure of $(\mathbf{1})_2$, one can speculate about shorter intermolecular distances between anions in $(\mathbf{6})_2$ than in $(\mathbf{7})_2$ (Fig. 88). From these results, one can envisage to tailor attractive interactions as well, e.g. intermolecular salt bridges between side chains.

In additional experiments, we explored the possibility to form heterodimers. We shall point again that each “homodimer” in fact consists of an acid and a carboxylate undergoing proton exchange. The “homodimer” is thus an average. Upon mixing $(\mathbf{1})_2$ and $(\mathbf{6})_2$ whose NMR signals differ due to the different lengths of **1** and **6**, a new species formed whose NMR chemical shift values suggest an intermediate length (Fig. 89). This species could thus be reasonably assigned to heterodimer **1**.**6**. Depending on the sequences involved, the proportions between homo and heterodimers were found to vary as a reflection of their respective stability. Quantitative heterodimerization was achieved by mixing **1** and **5**. Sequence **5** cannot dimerize, but its primary amide can act as a

hydrogen bond donor. Conversely, at higher pH, **1** is exclusively monomeric because its C-terminus is entirely deprotonated. Upon mixing the two, a single new species formed that was assigned to heterodimer **1.5** (Fig. 76a).

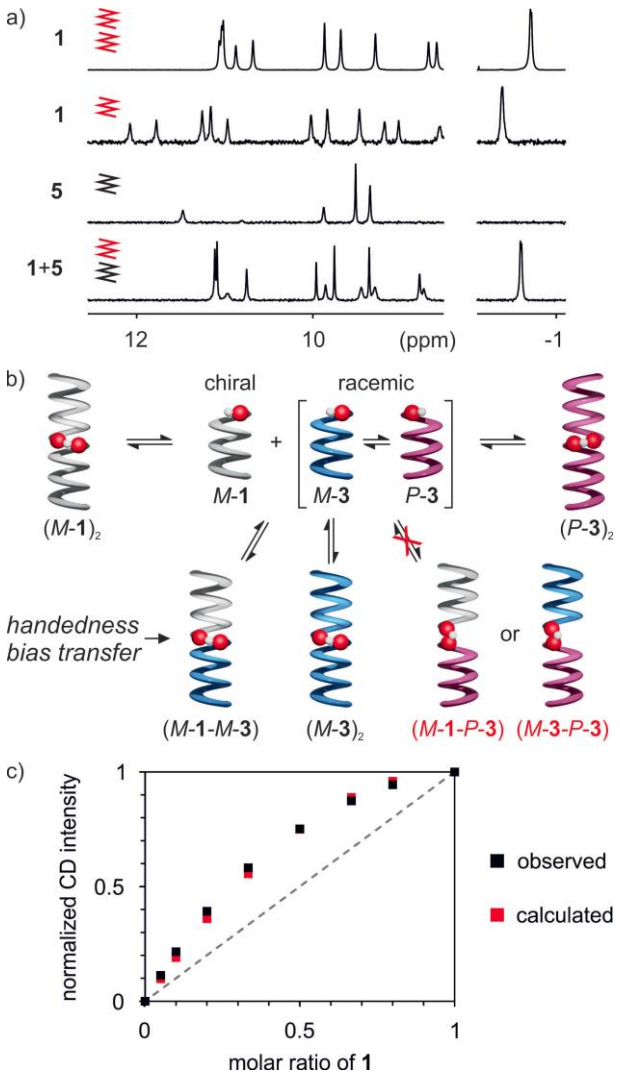


Fig. 76 a) Parts of the ^{1}H NMR spectra of **1** (dimeric: 2 mM in 27 mM sodium phosphate aqueous buffer pH 7.0; monomeric: 0.1 mM in 12 mM ammonium acetate aqueous buffer pH 8.5), **5** (0.1 mM in 12 mM ammonium acetate aqueous buffer pH 8.5) and a mixture of **1** and **5** (0.1 mM and 0.1 mM in 12 mM ammonium acetate aqueous buffer pH 8.5). b) Scheme of equilibria present in a mixture of compounds **1** and **3**. C-termini are indicated by red and white spheres. Impossible aggregates are highlighted in red. c) Relative CD intensities of **1** mixed with **3** at different ratios (total c = 0.5 mM in water).

Finally, we investigated stereochemical aspects of the dimerization. The crystal structure of $(\mathbf{1})_2$ (Fig. 74b) and the presence of a single dimer in the NMR spectra of achiral sequences **3**, **6**, and **7** show that dimerization occurs between helices that have the same handedness and not between a *P* and an *M* helix. We reasoned that upon mixing an exclusively *M* dimer such as $(\mathbf{1})_2$ from chiral sequence **1** with a racemic *P*/*M* mixture of dimers such as $(\mathbf{3})_2$ from achiral sequence **3**, heterodimers **1.3** would form only with the *M*-helix of **3**. Heterodimerization would thus bias the handedness of **3** in favor of the *M* helix. Sequence **3** was therefore designed with additional flexibility (two

pairs of contiguous B units) to allow for its helix handedness reversal to take place.[41], ** Circular dichroism spectra of mixtures of **1** and **3** in different proportions were recorded and demonstrated a deviation from linearity that perfectly matches with the predicted contribution of heterodimer **1.3**, assuming all three dimers have the same stability (Fig. 76b, c, Fig. 90).

In summary we have characterized a binding interface between the C-terminal cross section of aromatic helices in water that is thermodynamically stable and undergoing slow exchange on the NMR timescale. Association is strictly dependent on the presence of a C-terminal main chain acid function and of helix handedness. In contrast, helix length likely has little influence. Association can be further tuned by charges borne by the helices and by pH. Selective heterodimerization can also be implemented. This interface may serve as a tool for the programmed assembly of various entities in water, including in combination with aromatic helix bundling, the initial and unmet objective of the work reported here, about which progress will be reported in due course.

Acknowledgements: This work was supported by the DFG (Excellence Cluster 114, CIPSM). D. Gill is gratefully acknowledged for contributing synthetic precursors, and C. Glas for assistance with NMR measurements. Synchrotron data were collected at beamline P14 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany). We thank Dr. S. Panneerselvam for his assistance in using the beamline. Open Access funding enabled and organized by Projekt DEAL.

Conflict of interest: The authors declare no conflict of interest.

** Handedness reversal of long helices can be kinetically blocked in water, unless some more flexible units are introduced.

Supplementary Information

for:

Discrete Stacked Dimers of Aromatic Oligoamide Helices

Daniel Bindl, Pradeep K. Mandal, Lars Allmendinger, and Ivan Huc

List of Abbreviations

AcOH	acetic acid
CD	circular dichroism
CyHex	cyclohexane
DCM	dichloromethane
DIPEA	*N*,*N*-diisopropylethylamine
DMF	*N*,*N*-dimethylformamide
DMSO	dimethyl sulfoxide
EI	electron ionization
ESI	electrospray ionization
EtOAc	ethyl acetate
Fmoc	fluorenylmethoxycarbonyl
HMBC	heteronuclear multiple bond correlation
HMQC	heteronuclear multiple quantum correlation
HSQC	heteronuclear single quantum correlation
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
MeOH	methanol
MW	molecular weight
NMR	nuclear magnetic resonance
RP	reversed phase
RT	room temperature
SPFS	solid phase foldamer synthesis
TEA	triethylamine
THF	tetrahydrofuran
TLC	thin layer chromatography
TMSP	3-(trimethylsilyl)propionic-2,2,3,3-d_4 acid sodium salt
UV/Vis	ultraviolet-visible

6.1 Supplementary figures

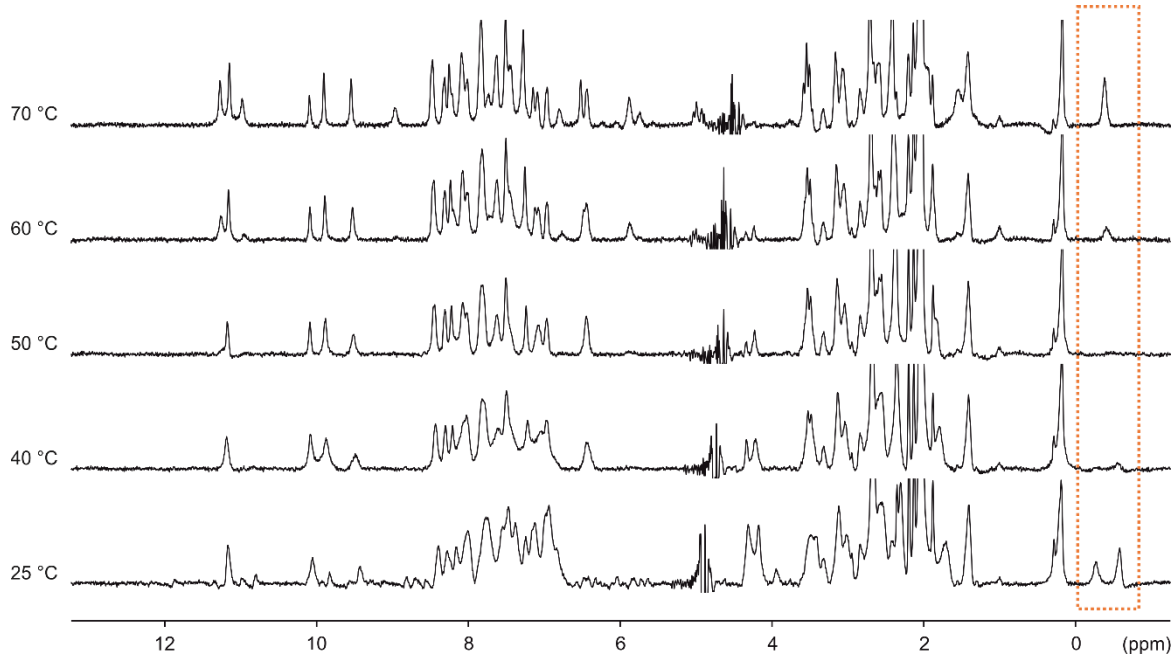


Fig. 77 Variable temperature ^{1}H NMR of compound **1** (400 MHz, 0.35 mM in 12 mM ammonium acetate buffer pH 8.5). Coalescence is observed at 50 °C for the chiral methyl group of the BRme unit (orange box).

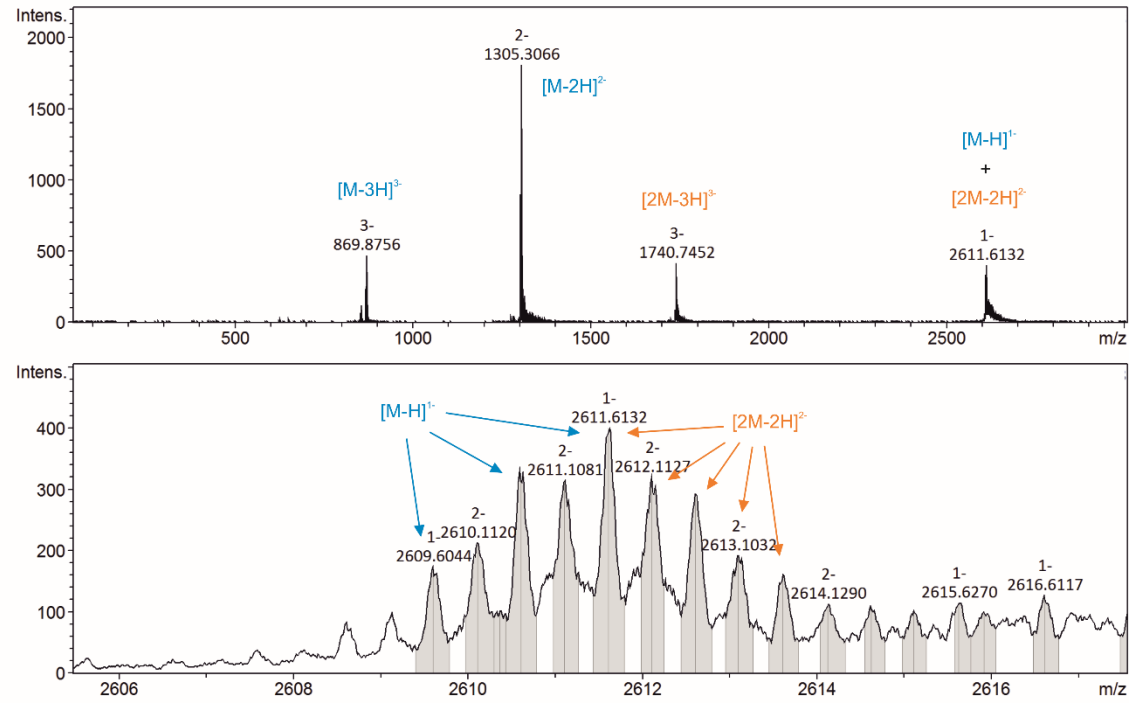


Fig. 78 ESI-MS spectrum of compound **1** (direct infusion of an aqueous sample). Peaks originating from monomeric and dimeric species are highlighted in blue and orange, respectively. The $[M-H]^{1-}$ peak is superimposed by the $[2M-2H]^{2-}$ peak, while the mass envelope suggests a major population of the $[2M-2H]^{2-}$ species.

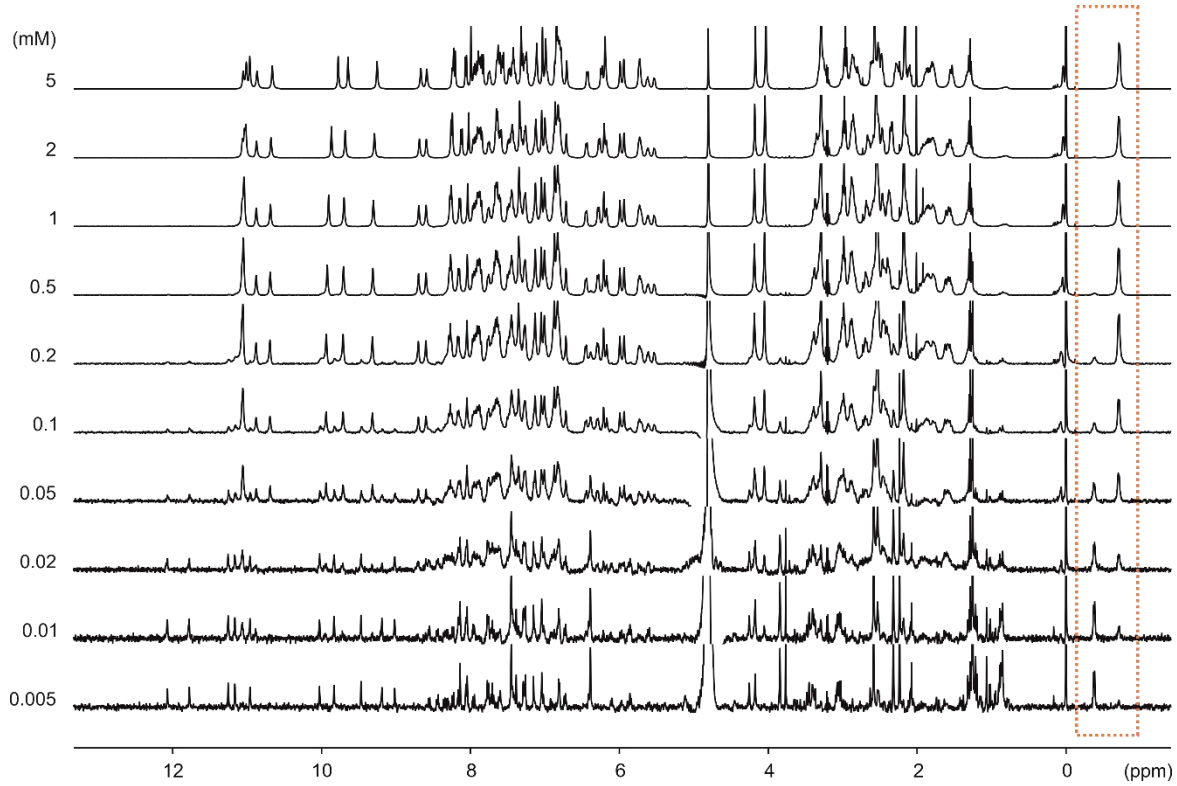


Fig. 79 Variable concentration ^{1}H NMR spectra of compound **1** (500 MHz, 27 mM sodium phosphate buffer pH 7.0). Concentrations of **1** are shown next to the respective spectrum, buffer concentrations were kept constant. Integrals used for binding curve plotting are marked in orange.

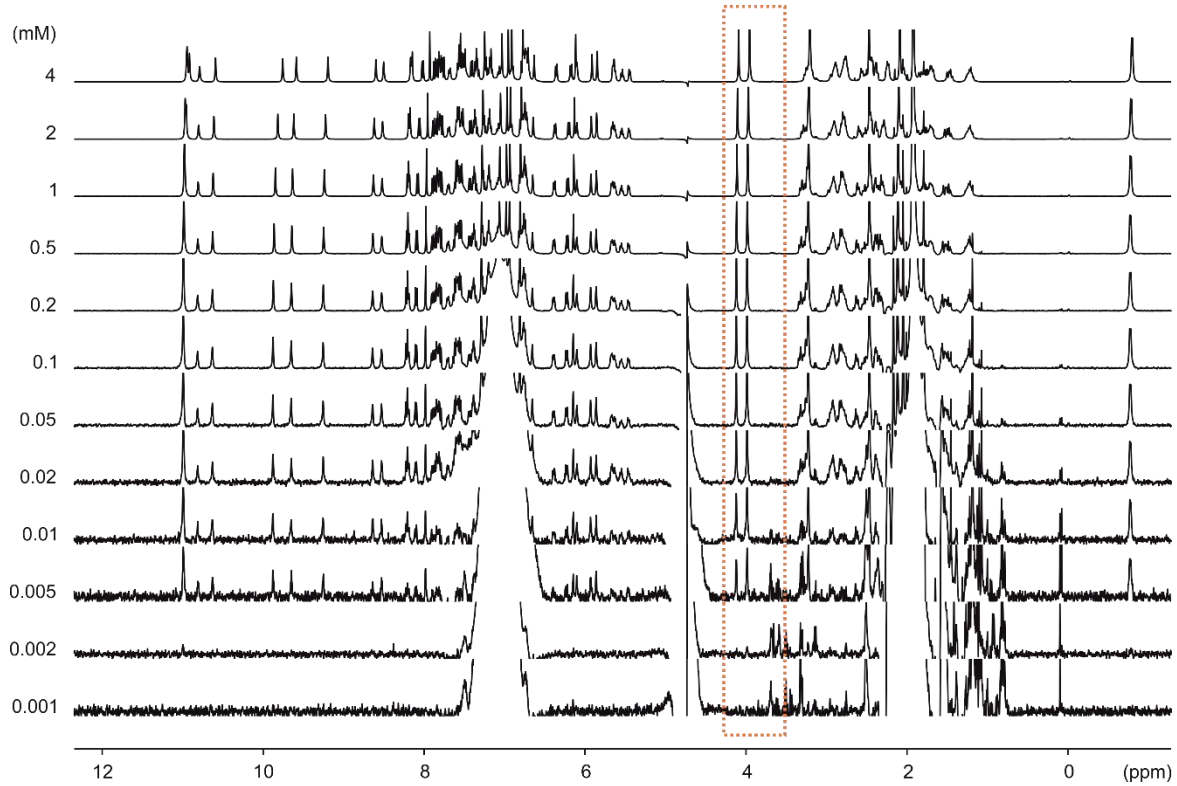
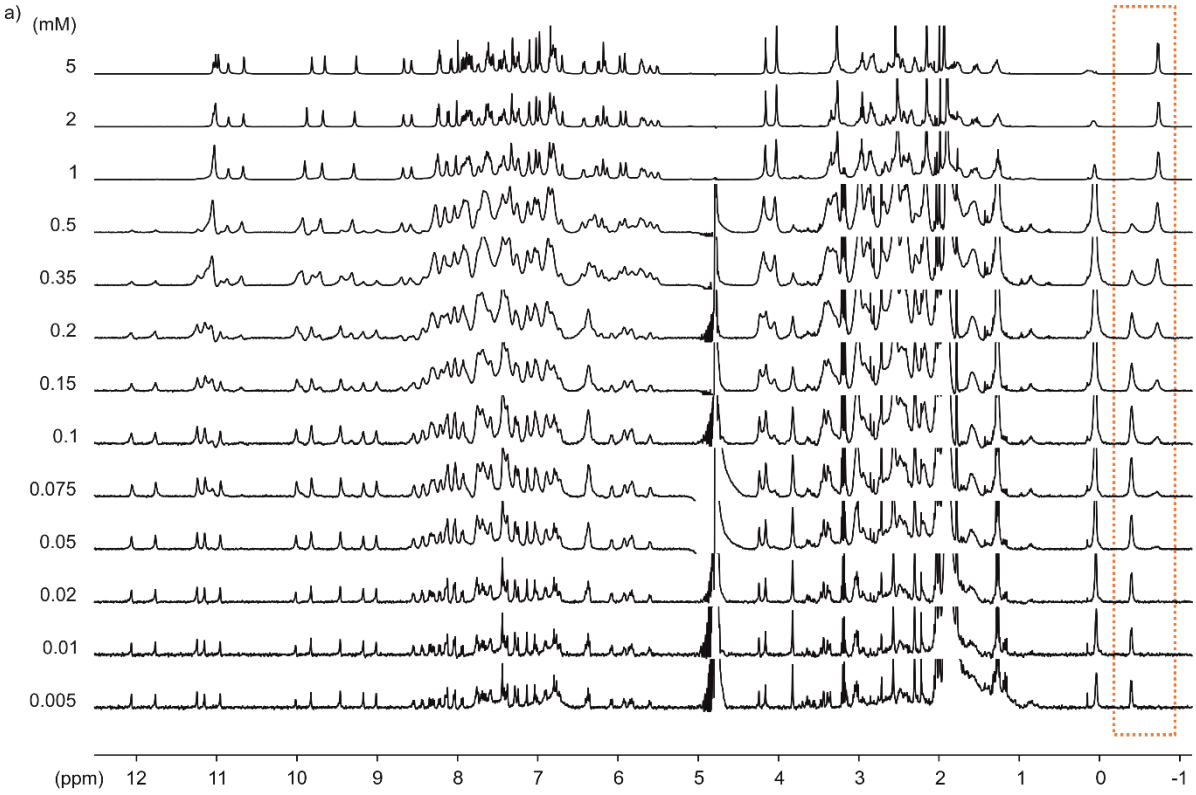
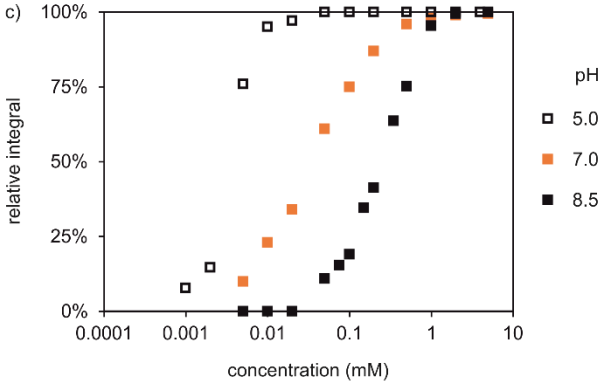


Fig. 80 Variable concentration ^{1}H NMR spectra of compound **1** (500 MHz, 14 mM ammonium acetate buffer pH 5.0). Concentrations of **1** are shown next to the respective spectrum, buffer concentrations were kept constant. Integrals used for binding curve plotting are marked in orange.



b) pH

pH 7.0		pH 5.0		pH 8.5	
c (mM)	Integral	c (mM)	Integral	c (mM)	Integral
5	100%	4	100%	5	100%
2	99%	2	100%	2	100%
1	98%	1	100%	1	95%
0.5	96%	0.5	100%	0.5	75%
0.2	87%	0.2	100%	0.35	64%
0.1	75%	0.1	100%	0.2	41%
0.05	61%	0.05	100%	0.15	35%
0.02	34%	0.02	97%	0.1	19%
0.01	23%	0.01	95%	0.075	15%
0.005	10%	0.005	76%	0.05	11%
		0.002	15%	0.02	0%
		0.001	8%	0.01	0%
				0.005	0%



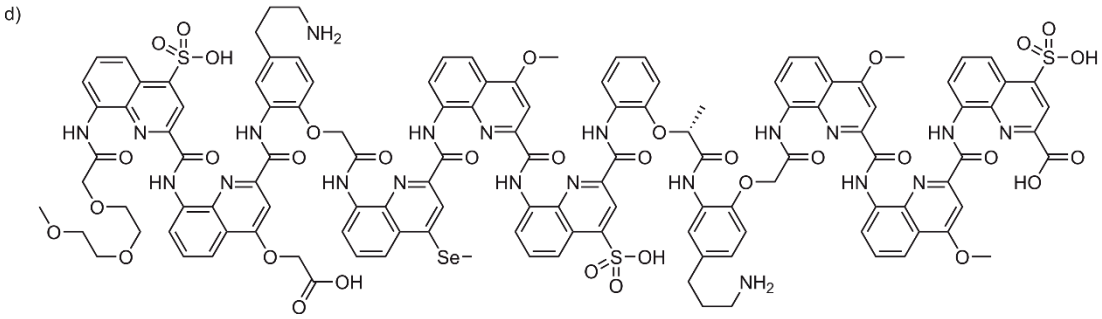


Fig. 81 a) Variable concentration ^{1}H NMR spectra of compound **1** (500 MHz, 12 mM ammonium acetate buffer pH 8.5). Concentrations of **1** are shown next to the respective spectrum, buffer concentrations were kept constant. Integrals used for binding curve plotting are marked in orange. Data (b) and plot (c) of the relative integrals of the aggregated compared to the monomeric species. d) Chemical structure of compound **1**.

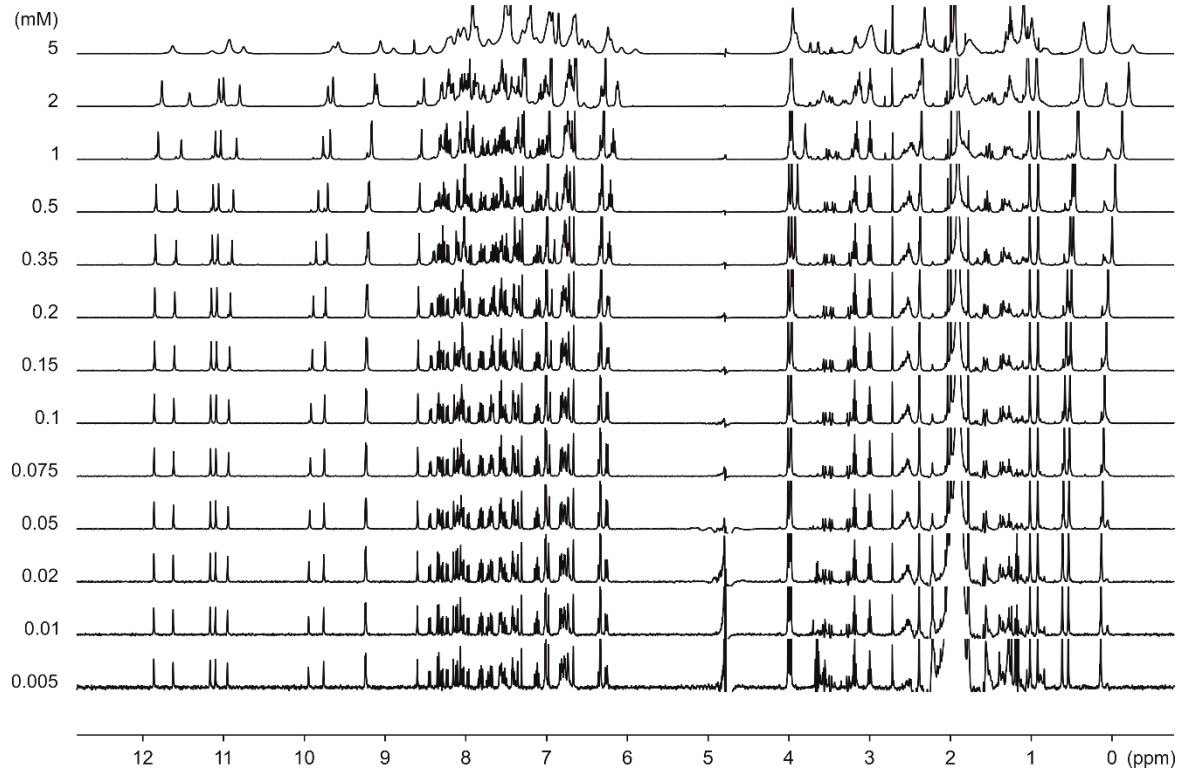


Fig. 82 Variable concentration ^{1}H NMR spectra of compound **2** (500 MHz, 12 mM ammonium acetate buffer pH 8.5). Concentrations of **2** are shown next to the respective spectrum, buffer concentrations were kept constant.

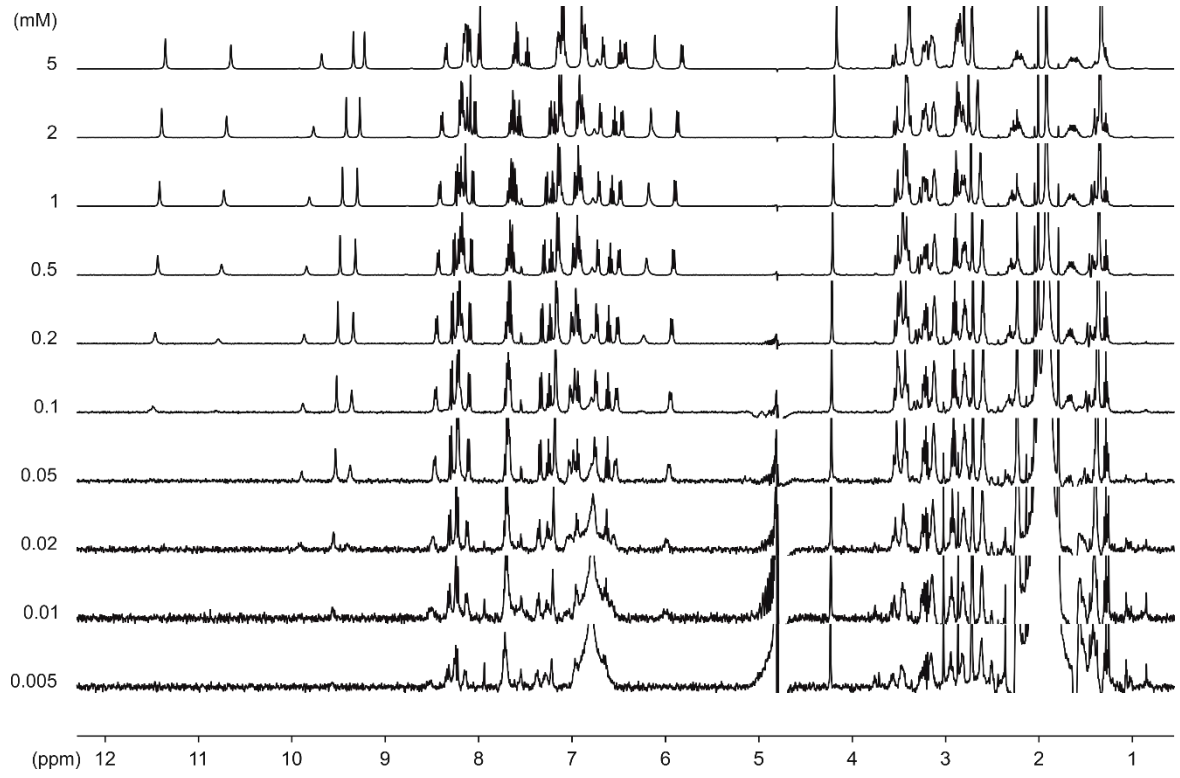
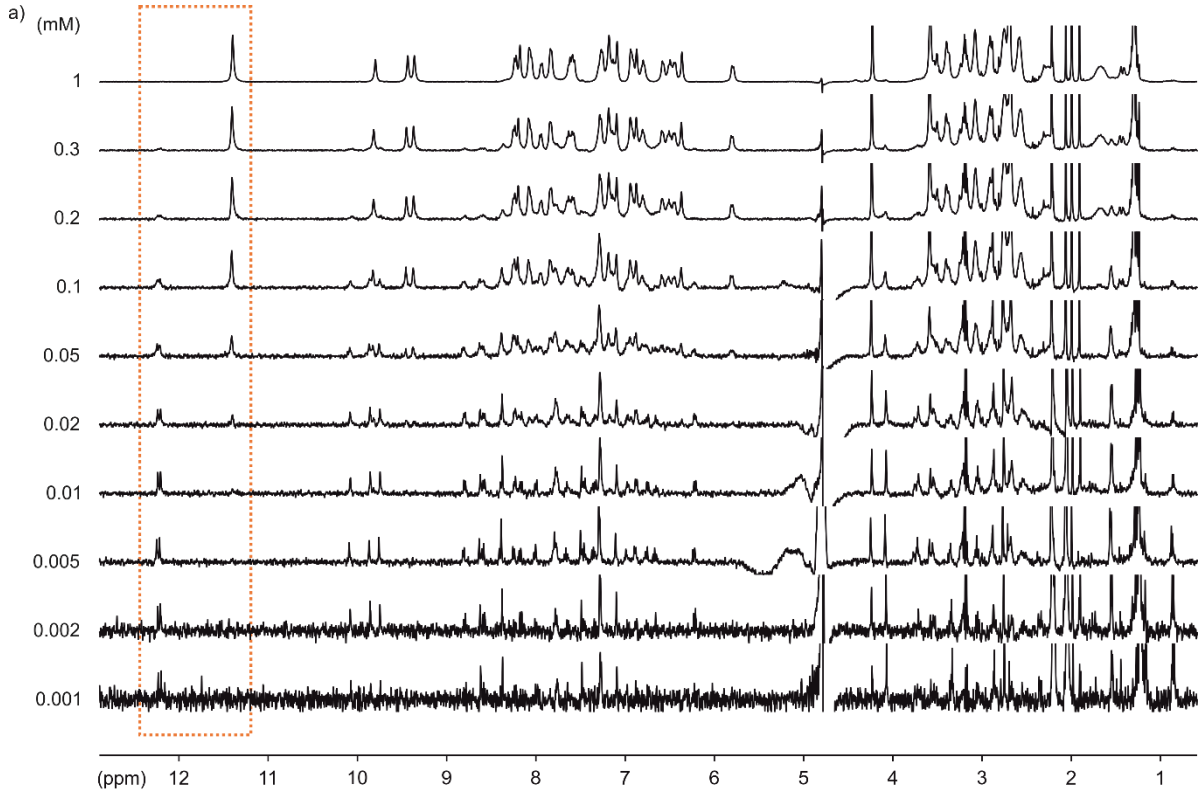
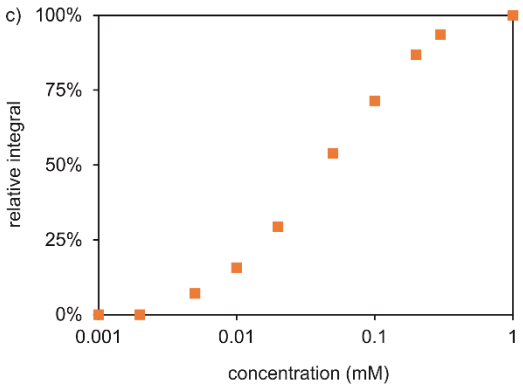


Fig. 83 Variable concentration ^{1}H NMR spectra of compound **5** (500 MHz, 12 mM ammonium acetate buffer pH 8.5). Concentrations of **5** are shown next to the respective spectrum, buffer concentrations were kept constant.



b)

c (mM)	Integral
1	100%
0.3	94%
0.2	87%
0.1	71%
0.05	54%
0.02	29%
0.01	16%
0.005	7%
0.002	0%
0.001	0%



d)

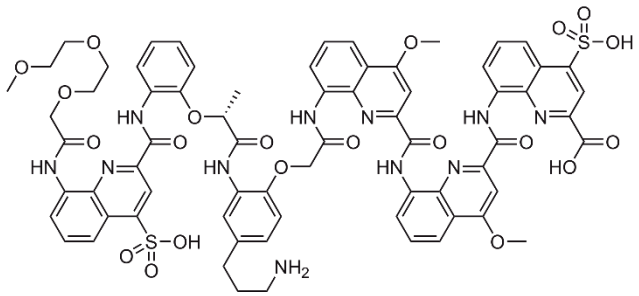
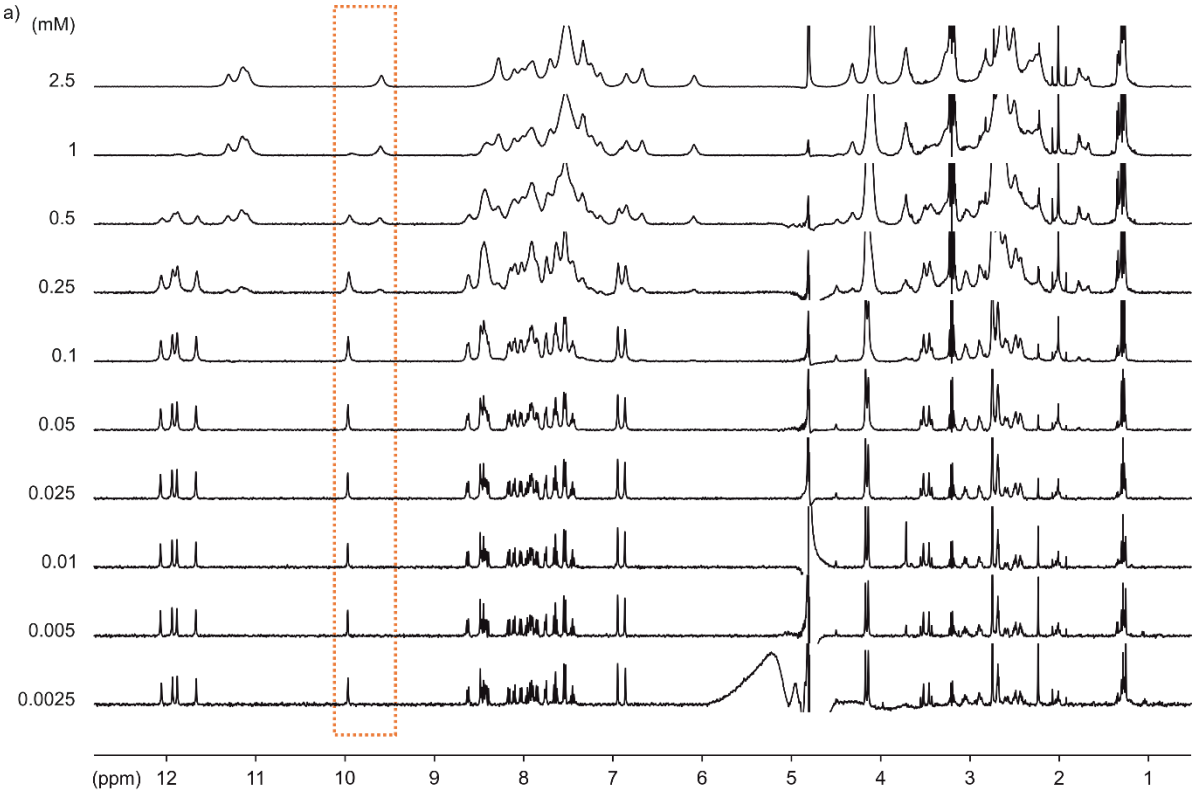
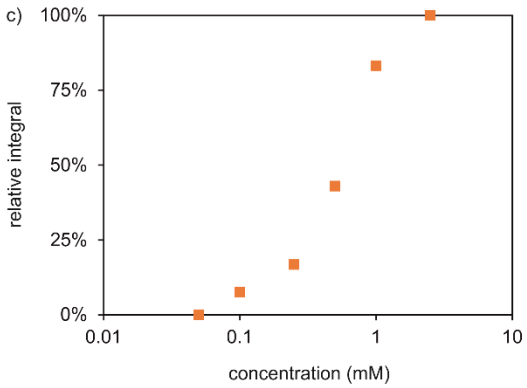


Fig. 84 a) Variable concentration ^{1}H NMR spectra of compound **4** (500 MHz, 27 mM sodium phosphate buffer pH 7.0). Concentrations of **4** are shown next to the respective spectrum, buffer concentrations were kept constant. Integrals used for binding curve plotting are marked in orange. Data (b) and plot (c) of the relative integrals of the aggregated compared to the monomeric species. d) Chemical structure of compound **4**.



b)

c (mM)	Integral
2.5	100%
1	83%
0.5	43%
0.25	17%
0.1	8%
0.05	0%
0.025	0%
0.01	0%
0.005	0%
0.0025	0%



d)

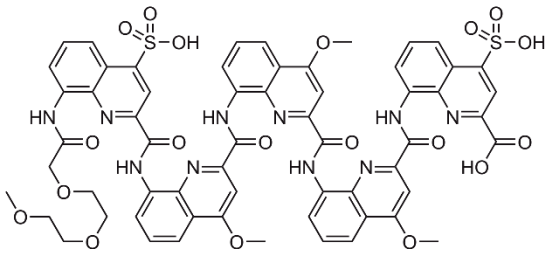
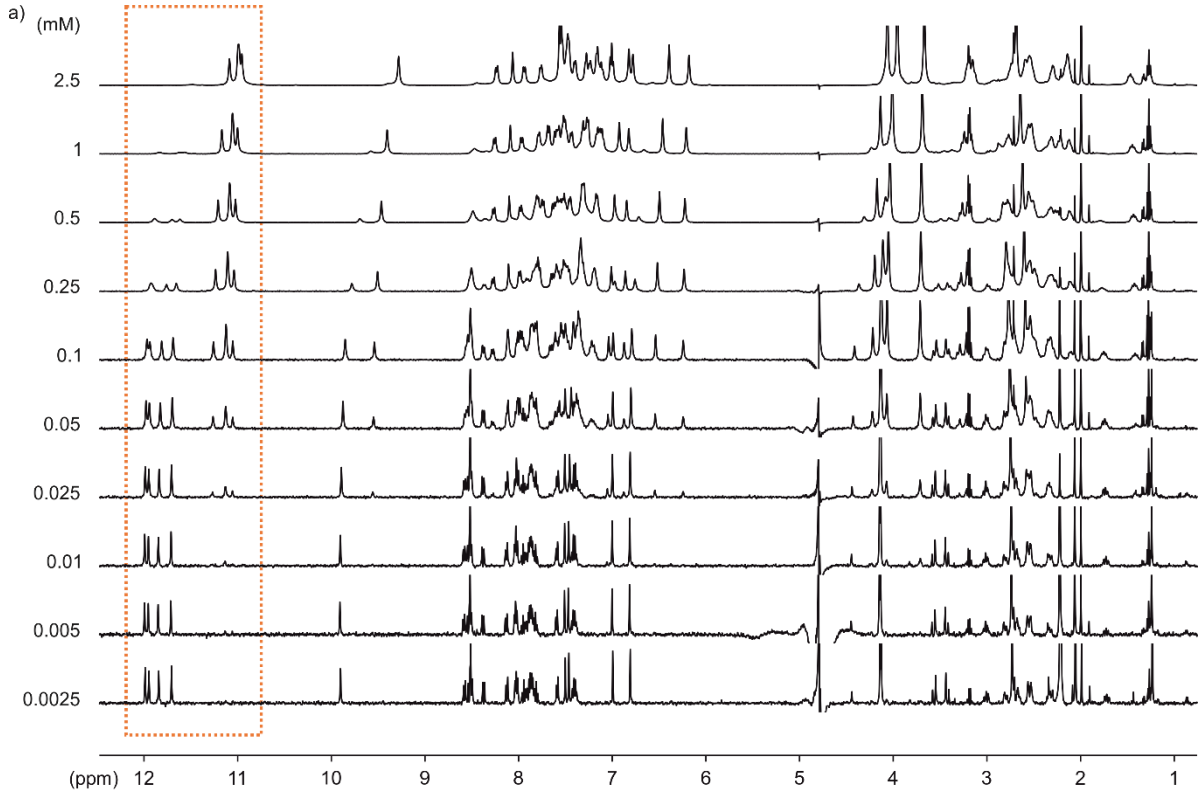
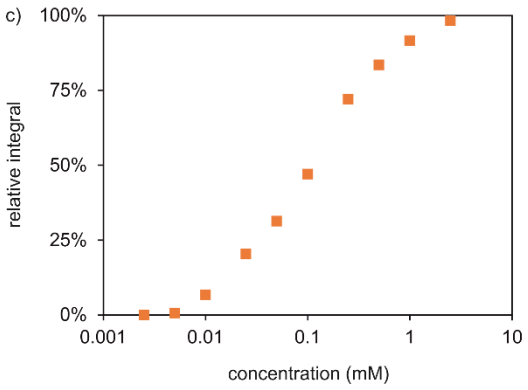


Fig. 85 a) Variable concentration ^{1}H NMR spectra of compound **6** (500 MHz, 27 mM sodium phosphate buffer pH 7.0). Concentrations of *P*-**6**/*M*-**6** (total concentration has been divided by two, since only homomeric dimerization is possible) are shown next to the respective spectrum, buffer concentrations were kept constant. Integrals used for binding curve plotting are marked in orange. Data (b) and plot (c) of the relative integrals of the aggregated compared to the monomeric species. d) Chemical structure of compound **6**.



b)

c (mM)	Integral
2.5	98%
1	92%
0.5	84%
0.25	72%
0.1	47%
0.05	31%
0.025	20%
0.01	7%
0.005	1%
0.0025	0%



d)

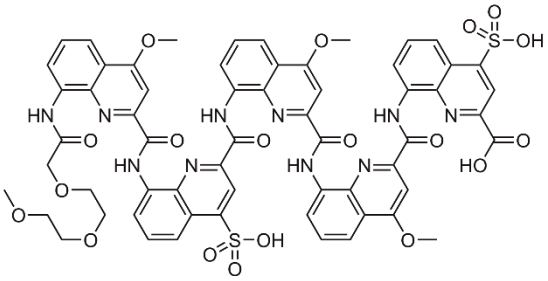
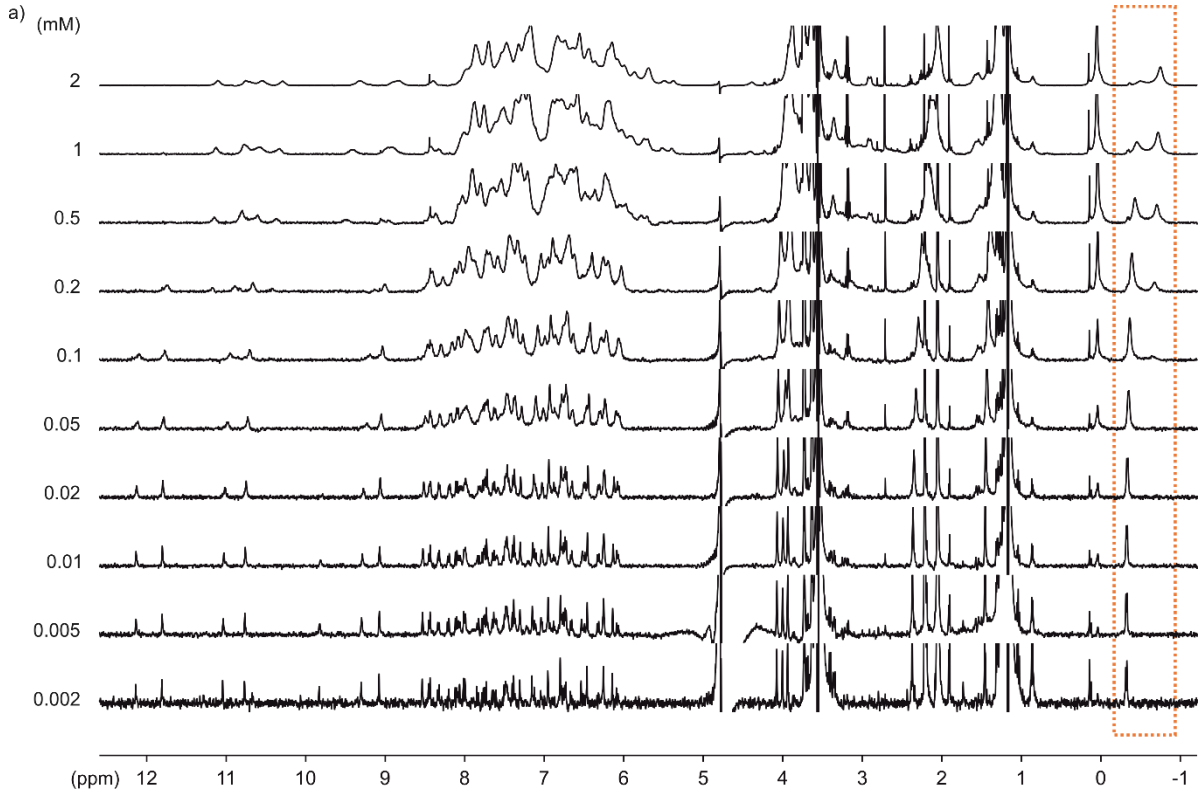
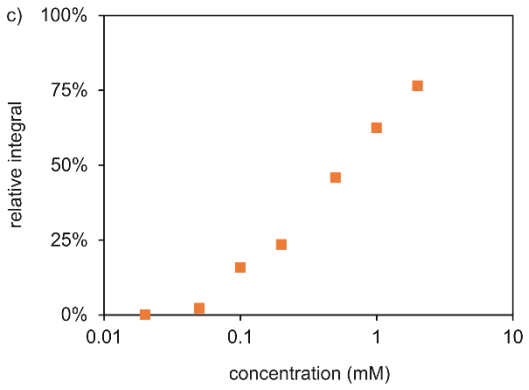


Fig. 86 a) Variable concentration ^{1}H NMR spectra of compound **7** (500 MHz, 27 mM sodium phosphate buffer pH 7.0). Concentrations of *P*-**7**/*M*-**7** (total concentration has been divided by two, since only homomeric dimerization is possible) are shown next to the respective spectrum, buffer concentrations were kept constant. Integrals used for binding curve plotting are marked in orange. Data (b) and plot (c) of the relative integrals of the aggregated compared to the monomeric species. d) Chemical structure of compound **7**.



b)

c (mM)	Integral
2	77%
1	63%
0.5	46%
0.2	24%
0.1	16%
0.05	2%
0.02	0%
0.01	0%
0.005	0%
0.002	0%



d)

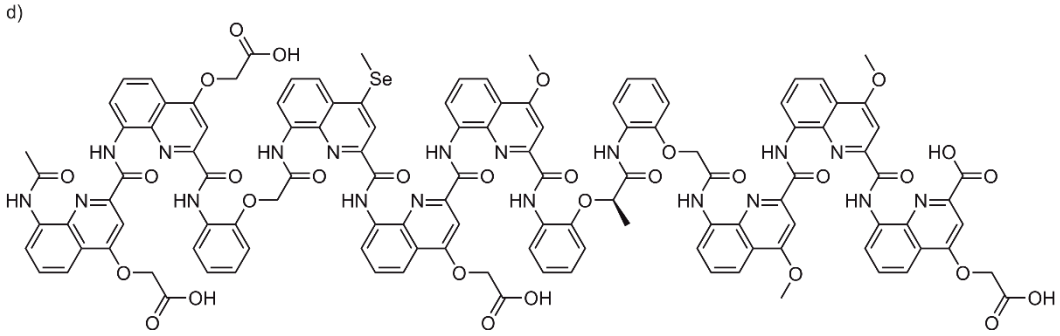


Fig. 87 a) Variable concentration ^{1}H NMR spectra of compound **8** (500 MHz, 27 mM sodium phosphate buffer pH 7.0). Concentrations of **8** are shown next to the respective spectrum, buffer concentrations were kept constant. Integrals used for binding curve plotting are marked in orange. Data (b) and plot (c) of the relative integrals of the aggregated compared to the monomeric species. d) Chemical structure of compound **8**.

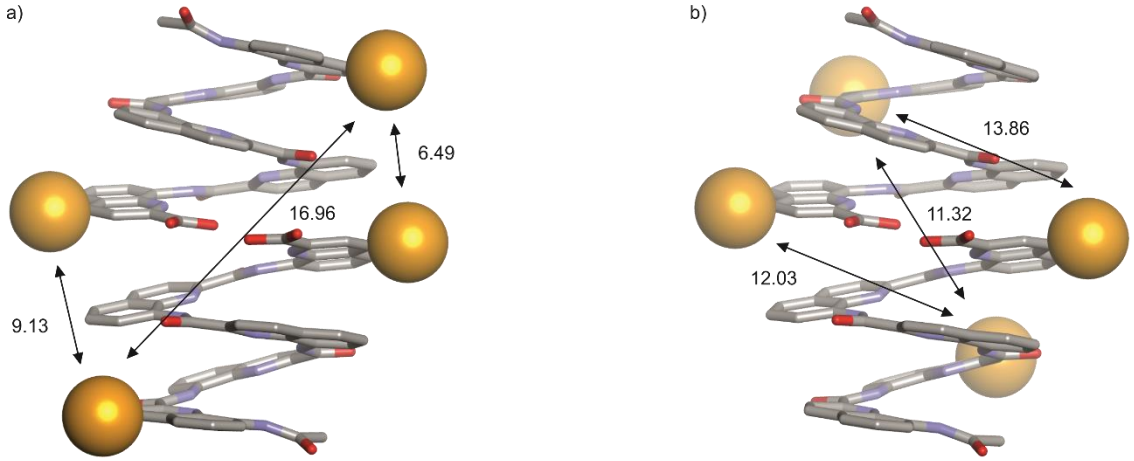


Fig. 88 Models of compounds **6** (a) and **7** (b) (Maestro, Method: TNCG, Forcefield: OPLS3, Solvent: water). Hydrogen atoms and side chains are omitted for clarity. The sulfur atoms of the Q^{Sul} units are shown as yellow balls. Intermolecular distances between sulfur atoms (except between the C-terminal Q^{Sul} units) is depicted in Angstrom (Å). Compound **6** shows shorter distances (>10 Å) compared to **7**.

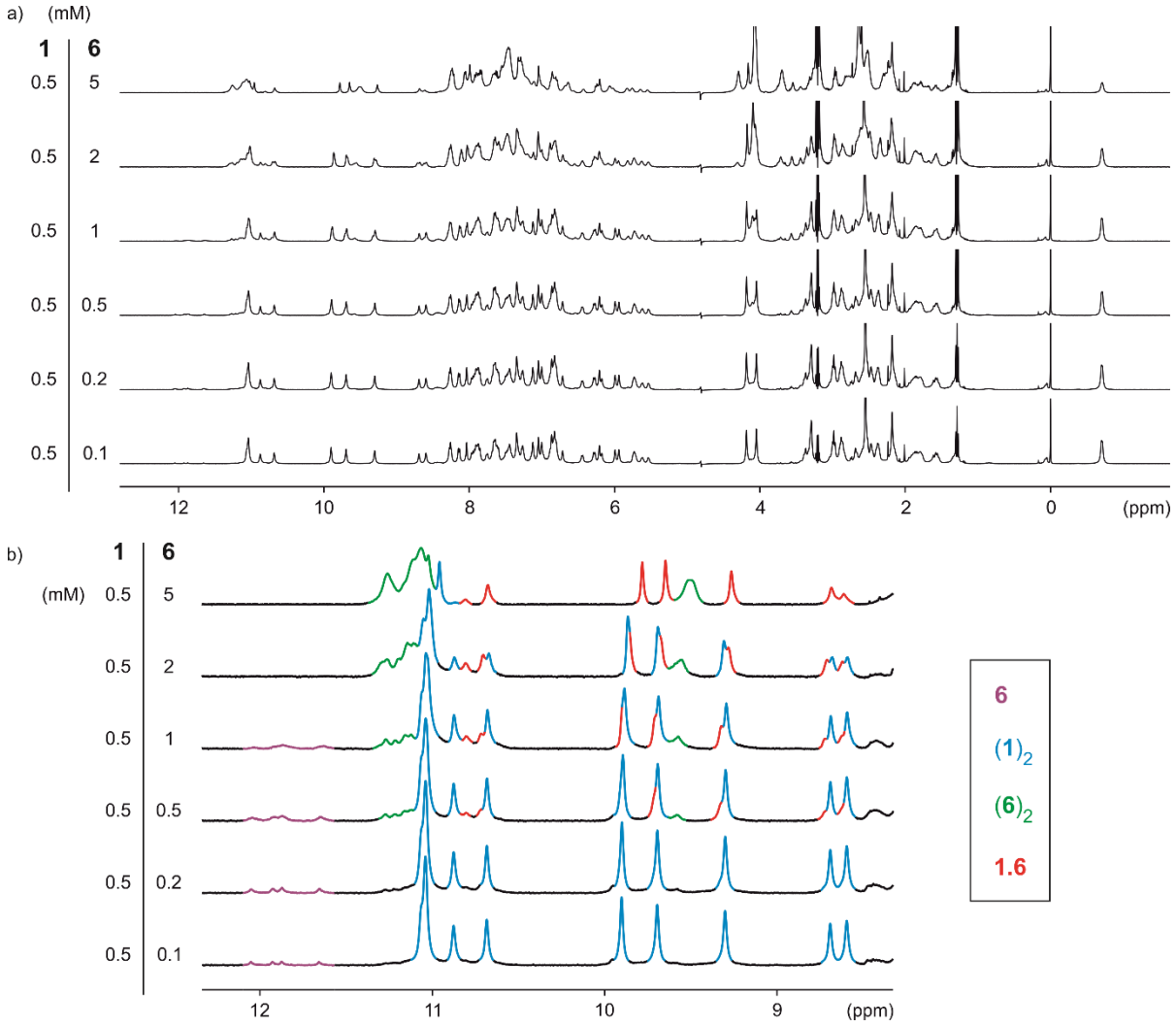
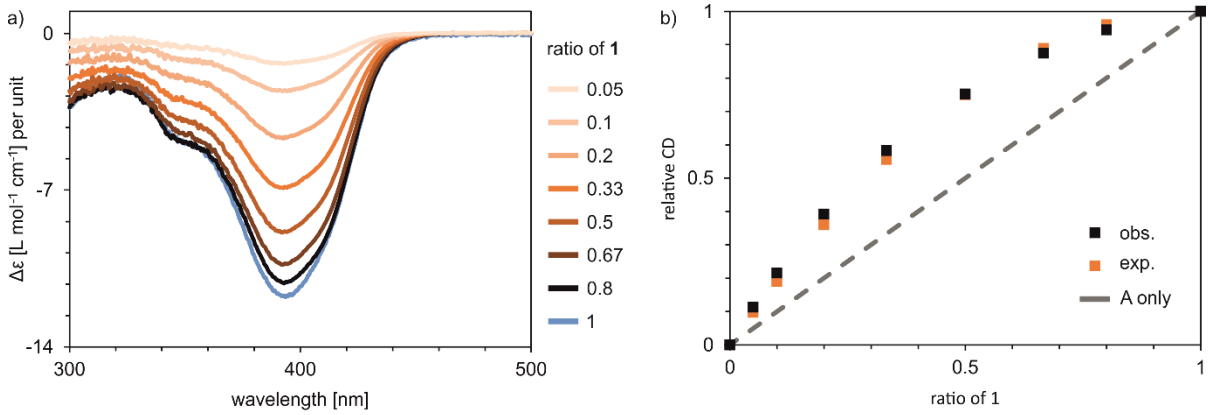


Fig. 89 Variable concentration ^{1}H NMR spectra of mixtures of compound **1** and **6** (500 MHz, 27 mM sodium phosphate buffer pH 7.0). Concentrations of **1** and **6** are shown next to the respective spectrum, buffer concentrations

were kept constant. a) Full spectra without coloration. b) Zoom on the amide region of the spectra. Signals corresponding to different compounds and/or aggregates are highlighted in different colors.



c)

ratio		complexes			CD	relative CD		
1	3	1-1	1-3	3-3	obs.	exp.	obs.	only A
1	0	1	0	0	-11.8	1.00	1.00	1.00
0.80	0.20	0.64	0.32	0.04	-11.1	0.96	0.94	0.80
0.67	0.33	0.44	0.44	0.11	-10.3	0.89	0.87	0.67
0.50	0.50	0.25	0.50	0.25	-8.85	0.75	0.75	0.50
0.33	0.67	0.11	0.44	0.44	-6.85	0.56	0.58	0.33
0.20	0.80	0.04	0.32	0.64	-4.62	0.36	0.39	0.20
0.10	0.90	0.01	0.18	0.81	-2.54	0.19	0.22	0.10
0.05	0.95	0.00	0.10	0.90	-1.32	0.10	0.11	0.05
0	1	0	0	1	0.00	0.00	0.00	0.00

Fig. 90 a) CD spectra of mixtures of compounds **1** and **3** at total concentrations of 0.5 mM in H_2O. b) Observed (obs.) relative CD at the maxima of these mixtures plotted with the expected (exp.) values for stochastic heterodimer formation (orange) and no interaction (grey line). c) Table showing the ratios of stochastic dimer formation and the expected relative CD at different mixing ratios of compounds **1** and **3**.

6.2 Materials and Methods

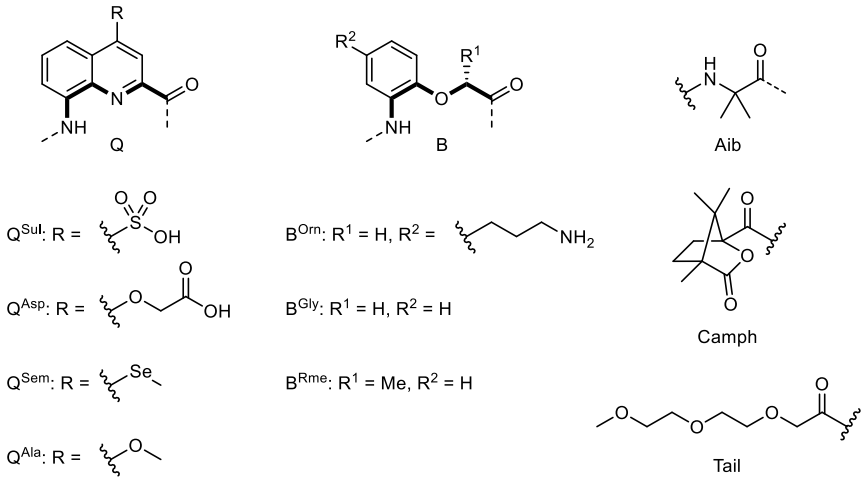


Fig. 91 Fmoc-acid and other building blocks used in this study. Fmoc-Q^{Sul}-OH,[64] Fmoc-Q^{Asp}-OH,[26a] Fmoc-Q^{Ala}-OH,[96] Fmoc-B^{Gly}-OH[93] and Fmoc-B^{Rme}-OH[93] have been described previously. Fmoc-Q^{Sem}-OH will be described elsewhere. For a detailed procedure to Fmoc-B^{Orn}(Boc)-OH (**13**), see section 6.2.3.

6.2.1 General

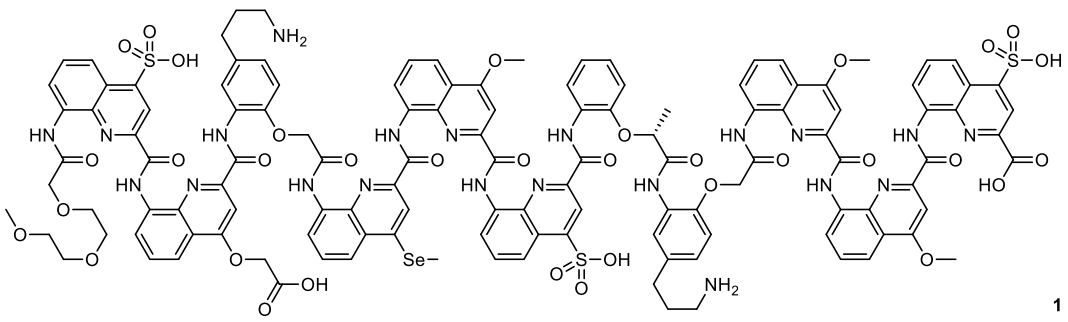
Commercial reagents (Suppliers: Abcr, Fisher Scientific, Merck, Sigma-Aldrich, TCI or VWR) were used without further purification unless otherwise stated. Cl-MPA ProTide™ resin LL and Rink Amide MBHA resin were purchased from CEM. Peptide grade *N*,*N*-dimethylformamide (DMF) was purchased from Carlo Erba. Anhydrous chloroform, triethylamine (TEA) and *N*,*N*-diisopropylethylamine (DIPEA) 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 CryProbe™ Prodigy. Measurements were performed at 25 °C unless stated otherwise. In case of aqueous buffers and pure water solutions, mixtures of H_2O/D_2O 9/1 (*v*/*v*) were used, and water suppression was performed with excitation sculpting unless stated otherwise. 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.[97] 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, ^{1}H-^{13}C HSQC and ^{1}H-^{13}C HMBC spectra using standard pulse sequences from the Bruker pulse program library applying standard processing parameters. 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). UV absorbance was monitored at 300 nm if not stated otherwise. Simple ultraviolet–visible (UV/Vis) absorbance measurements were done with a Thermo Fisher Scientific Nanodrop One instrument using a 1 cm quartz cuvette. Circular dichroism (CD) spectra were measured on a Jasco J-810 spectrometer. Measurements were performed at 20 °C if not stated otherwise. Manual microwave-assisted solid-phase foldamer synthesis (SPFS) was performed via a CEM® Discover Bio microwave peptide synthesizer. The temperature within the reactor vessel was monitored with an optical fiber probe. Automated SPFS was done via a Gyros Protein Technologies PurePep Chorus synthesizer with induction heating.

6.2.2 Solid phase synthesis procedures

Oligomers were synthesized according to previously reported SPFS protocols,[26] hereafter referred to as standard method. Fmoc acid building blocks were activated *in situ* by generating the respective acid chlorides prior to coupling.

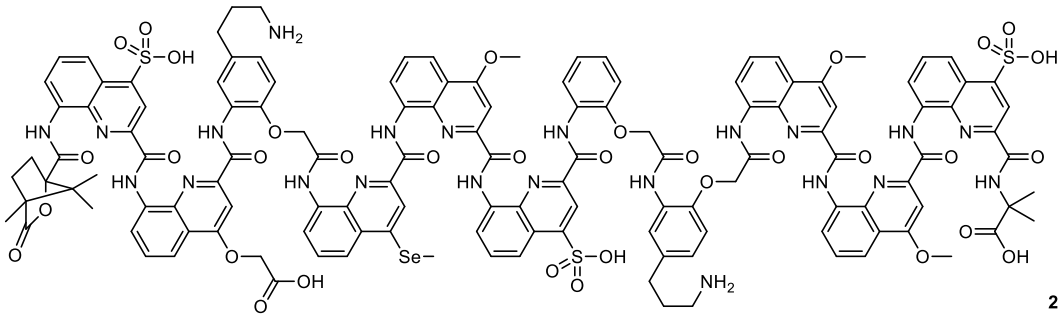
General acylation method: In the microwave vessel: after the resin (1.0 equiv.) was washed with anhydrous THF (4×), DIPEA (10.0 equiv.) and acyl chloride (5.0 equiv.) in anhydrous THF (1 mL per 100 mg resin; not less than 2 mL) were added and the suspension was heated to 50 °C for 15 min (25 W, ramp to 50 °C over 5 min, hold at 50 °C for 15 min). The resin was washed with anhydrous THF (3×) and the coupling step was repeated once. Then, the resin was washed again with anhydrous THF (1×) and DMF (5×), and kept suspended in DMF (if stored longer than 24 h, it was kept at 4 °C).



1

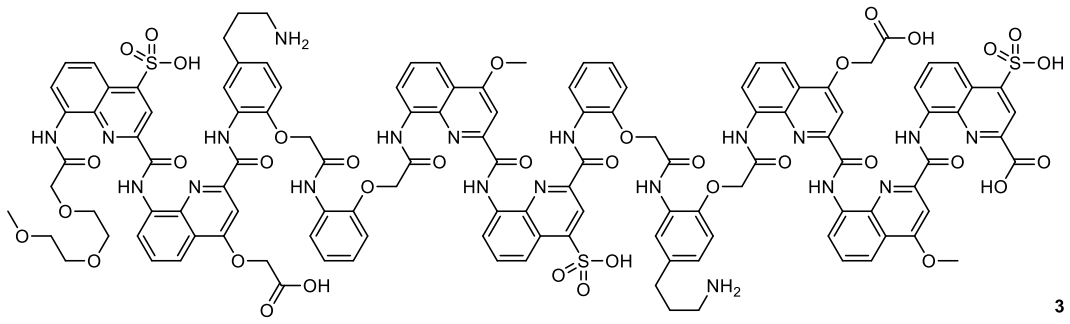
Compound 1: Oligomer **1** was synthesized on ProTide™ resin (0.16 mmol g^{-1}, 45 µmol scale) according to the standard method (manually). Loading of the first monomer: 0.11 mmol g^{-1} (71%).

The final Tail group was installed via the standard method (using Tail-OH). After purification by semi-prep HPLC (C18, 0–25B, 50 °C; A: 13 mmol ammonium acetate buffer pH 8.5, B: acetonitrile), the title compound was obtained as a yellow solid (33.0 mg, 12.6 µmol, 28%; HPLC-purity: 98%). **^{1}H NMR** (500 MHz, 1 mM in 27 mM sodium phosphate buffer pH 7.0): δ = 11.04 (s, 3H), 10.88 (s, 1H), 10.69 (s, 1H), 9.90 (s, 1H), 9.70 (s, 1H), 9.30 (s, 1H), 8.69 (s, 1H), 8.60 (s, 1H), 8.27 (d, *J* = 9.5 Hz, 2H), 8.14 (d, *J* = 8.6 Hz, 1H), 8.04 (s, 1H), 7.96 (d, *J* = 9.3 Hz, 1H), 7.94–7.82 (m, 2H), 7.80–7.72 (m, 1H), 7.72–7.63 (m, 2H), 7.61 (d, *J* = 8.8 Hz, 1H), 7.49 (d, *J* = 9.2 Hz, 1H), 7.47–7.39 (m, 1H), 7.39–7.30 (m, 2H), 7.30–7.22 (m, 1H), 7.13 (s, 2H), 7.05 (s, 1H), 7.00 (s, 1H), 6.92–6.76 (m, 5H), 6.72 (s, 1H), 6.45 (d, *J* = 8.3 Hz, 1H), 6.28 (d, *J* = 9.6 Hz, 1H), 6.21 (s, 1H), 6.17 (s, 1H), 5.99 (s, 1H), 5.94 (s, 1H), 5.79–5.67 (m, 1H), 5.61 (s, 1H), 5.53 (s, 1H), 4.18 (s, 2H), 4.05 (s, 2H), 3.38 (d, *J* = 12.6 Hz, 1H), 3.29 (s, 4H), 2.98 (t, *J* = 8.6 Hz, 2H), 2.94–2.80 (m, 1H), 2.68 (s, 1H), 2.54 (s, 5H), 2.50–2.43 (m, 1H), 2.43–2.33 (m, 2H), 2.18 (s, 4H), 1.90–1.82 (m, 1H), 1.82–1.72 (m, 1H), 1.65–1.49 (m, 2H), 1.37–1.24 (m, 1H), -0.71 (d, *J* = 7.4 Hz, 3H). **HRMS** (ESI$^-$) *m/z* calcd. for $C_{124}H_{108}N_{21}O_{34}S_3Se$: 2610.5700 $(M-H)^-$; found: 2610.5909.



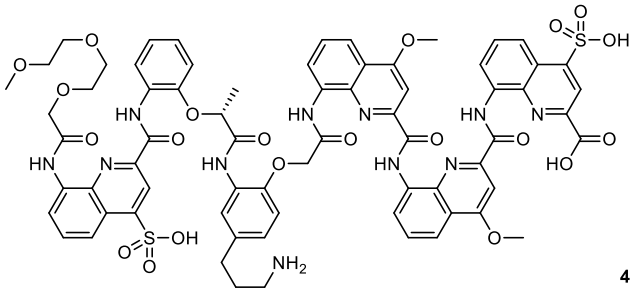
2

Compound 2: Oligomer **2** was synthesized on ProTide™ resin (0.16 mmol g^{-1}, 20 µmol scale) according to the standard method (manually). Loading of the first monomer: 0.16 mmol g^{-1} (quant.). The final **Camph** group was installed via the general acylation method. After purification by semi-prep HPLC (C18, 0–30B, 50 °C; A: 13 mmol ammonium acetate buffer pH 8.5, B: acetonitrile), the title compound was obtained as a yellow solid (18.0 mg, 6.67 µmol, 33%; HPLC-purity: 90%). **^{1}H NMR** (500 MHz, 1 mM in H_2O/D_2O 9:1): δ = 11.84 (s, 1H), 11.60 (s, 1H), 11.13 (s, 1H), 11.06 (s, 1H), 10.87 (s, 1H), 9.81 (s, 1H), 9.70 (s, 1H), 9.22 (s, 1H), 9.19 (s, 1H), 8.57 (s, 1H), 8.42 (d, *J* = 8.3 Hz, 1H), 8.33 (d, *J* = 8.3 Hz, 1H), 8.27 (d, *J* = 9.5 Hz, 2H), 8.23 (d, *J* = 8.5 Hz, 1H), 8.13 (s, 1H), 8.09 (d, *J* = 8.1 Hz, 1H), 8.04 (d, *J* = 9.0 Hz, 1H), 8.03–7.99 (m, 2H), 7.97–7.90 (m, 2H), 7.81 (d, *J* = 8.4 Hz, 1H), 7.77 (d, *J* = 9.3 Hz, 1H), 7.71–7.64 (m, 1H), 7.62 (d, *J* = 8.4 Hz, 1H), 7.60–7.46 (m, 5H), 7.44–7.37 (m, 2H), 7.35 (d, *J* = 8.5 Hz, 1H), 7.33–7.26 (m, 3H), 7.13 (d, *J* = 8.5 Hz, 1H), 7.11–7.05 (m, 1H), 7.04–7.00 (m, 2H), 6.99–6.94 (m, 2H), 6.93 (s, 1H), 6.82–6.68 (m, 8H), 6.66 (s, 2H), 6.34 (s, 1H), 6.32 (s, 2H), 6.30 (s, 1H), 6.23 (d, *J* = 9.3 Hz, 1H), 6.19 (d, *J* = 9.1 Hz, 1H), 4.00 (s, 3H), 3.96 (s, 3H), 3.94 (s, 3H), 3.54 (d, *J* = 15.6 Hz, 1H), 3.47 (d, *J* = 15.2 Hz, 1H), 3.23 (d, *J* = 16.0 Hz, 2H), 3.18 (t, *J* = 8.3 Hz, 3H), 3.00 (t, *J* = 8.4 Hz, 3H), 2.65–2.44 (m, 5H), 2.37 (s, 3H), 2.03–1.72 (m, 5H), 1.54 (dd, *J* = 19.0, 15.1 Hz, 3H), 1.37 (d, *J* = 15.0 Hz, 1H), 1.01 (s, 5H), 0.91 (s, 3H), 0.46 (s, 3H), 0.44 (s, 3H), -0.09 (s, 3H). **HRMS** (ESI$^-$) *m/z* calcd. for $C_{130}H_{113}N_{22}O_{34}S_3Se$: 2701.6122 $(M-H)^-$; found: 2701.7377.



3

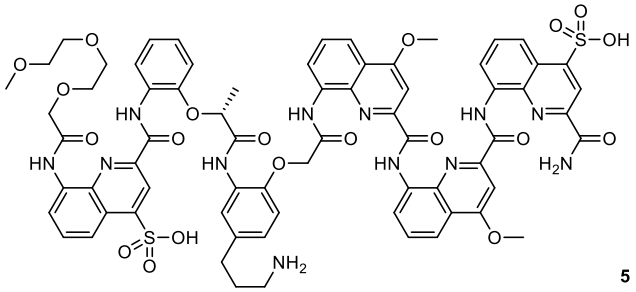
Compound 3: Oligomer **3** was synthesized on ProTide™ resin (0.23 mmol g^{-1}, 23 µmol scale) according to the standard method (manually). Loading of the first monomer: 0.23 mmol g^{-1} (98%). The final Tail group was installed via the standard method (using Tail-OH). After purification by semi-prep HPLC (C18, 0–20B, 25 °C; A: 13 mM ammonium acetate buffer pH 8.5, B: acetonitrile), the title compound was obtained as a yellow solid (10.4 mg, 4.11 µmol, 18%; HPLC-purity: >99%). **^{1}H NMR** (500 MHz, 0.5 mM in H_2O/D_2O 9:1): δ = 11.32 (s, 1H), 10.93 (s, 1H), 10.83 (s, 1H), 9.90 (s, 1H), 9.41 (s, 1H), 9.22 (s, 1H), 8.90 (s, 1H), 8.80 (s, 1H), 8.22 (d, *J* = 8.8 Hz, 1H), 8.16 (d, *J* = 7.9 Hz, 1H), 8.09 (d, *J* = 8.0 Hz, 1H), 8.07–8.01 (m, 2H), 7.97 (s, 1H), 7.83 (s, 1H), 7.81 (s, 2H), 7.73 (d, *J* = 9.1 Hz, 1H), 7.70 (s, 1H), 7.69 (s, 1H), 7.66–7.51 (m, 4H), 7.47 (t, *J* = 8.3 Hz, 1H), 7.36 (s, 1H), 7.31 (s, 1H), 7.25–7.15 (m, 3H), 7.14–6.97 (m, 5H), 6.92 (s, 1H), 6.84 (t, *J* = 7.8 Hz, 1H), 6.77 (t, *J* = 7.9 Hz, 1H), 6.70 (d, *J* = 8.3 Hz, 1H), 6.67 (d, *J* = 3.5 Hz, 1H), 6.61 (d, *J* = 7.7 Hz, 1H), 6.57–6.50 (m, 1H), 6.48 (d, *J* = 7.7 Hz, 1H), 6.42 (s, 1H), 6.35 (s, 1H), 6.19 (s, 1H), 6.08 (s, 1H), 6.04 (d, *J* = 8.9 Hz, 1H), 6.00 (d, *J* = 8.8 Hz, 1H), 5.95 (d, *J* = 8.6 Hz, 1H), 5.77 (d, *J* = 6.8 Hz, 1H), 3.79 (s, 3H), 3.45 (s, 3H), 3.36 (d, *J* = 13.9 Hz, 1H), 3.23 (d, *J* = 14.4 Hz, 1H), 3.10–2.83 (m, 8H), 2.67–2.60 (m, 2H), 2.58 (s, 3H), 2.57–2.42 (m, 6H), 2.32–2.19 (m, 3H), 2.06 (s, 2H), 1.80 (qt, *J* = 14.4, 7.5 Hz, 1H), 1.64 (t, *J* = 10.3 Hz, 1H), 1.43 (d, *J* = 14.6 Hz, 1H), 1.36–1.21 (m, 1H), 0.97–0.85 (m, 2H). **HRMS** (ESI^-) *m/z* calcd. for $C_{121}H_{105}N_{20}O_{37}S_3$: 2525.6117 $(M-H)^-$; found: 2525.5560.



4

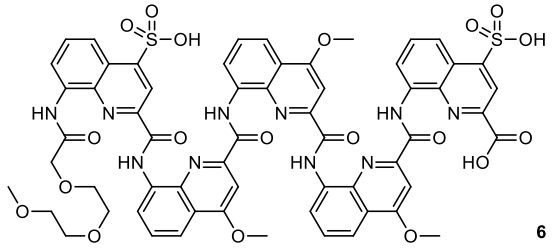
Compound 4: Oligomer **4** was synthesized on ProTide™ resin (0.15 mmol g^{-1}, 25 µmol scale) according to the standard method (automated). Loading of the first monomer: 0.13 mmol g^{-1} (88%). The final Tail group was installed via the standard method (using Tail-OH). After purification by semi-prep HPLC (C18, 0–30B, 50 °C; A: 13 mM ammonium acetate buffer pH 8.5, B: acetonitrile), the title compound was obtained as a yellow solid (1.24 mg, 0.859 µmol, 14%; HPLC-purity: >99%). **^{1}H NMR** (500 MHz, 1 mM in 27 mmol sodium phosphate buffer pH 7.0): δ = 11.39 (s, 2H), 9.79 (s, 1H), 9.43 (s, 1H), 9.35 (s, 1H), 8.22 (d, *J* = 10.3 Hz, 1H), 8.17 (s, 1H), 8.12–8.00

(m, 1H), 7.93 (d, *J* = 9.1 Hz, 1H), 7.82 (d, *J* = 6.0 Hz, 4H), 7.68–7.52 (m, 3H), 7.36–7.21 (m, 2H), 7.21–7.03 (m, 3H), 6.98–6.88 (m, 1H), 6.86 (s, 1H), 6.84–6.73 (m, 1H), 6.57 (d, *J* = 7.3 Hz, 1H), 6.49 (t, *J* = 8.3 Hz, 2H), 6.44 (d, *J* = 8.6 Hz, 1H), 6.36 (s, 1H), 5.79 (d, *J* = 10.3 Hz, 1H), 4.22 (s, 2H), 3.57 (s, 4H), 3.54–3.45 (m, 1H), 3.45–3.31 (m, 1H), 3.27–3.12 (m, 3H), 3.07 (s, 3H), 2.89 (d, *J* = 13.7 Hz, 3H), 2.74 (s, 4H), 2.68 (s, 5H), 2.63–2.51 (m, 4H), 2.40–2.14 (m, 4H), 1.99 (s, 4H), 1.79–1.56 (m, 3H), 1.42 (d, *J* = 17.2 Hz, 1H), 1.36–1.20 (m, 5H). **HRMS** (ESI⁻) *m/z* calcd. for $C_{69}H_{64}N_{11}O_{21}S_2$: 1446.3725 $(M-H)^-$; found: 1446.4677.



5

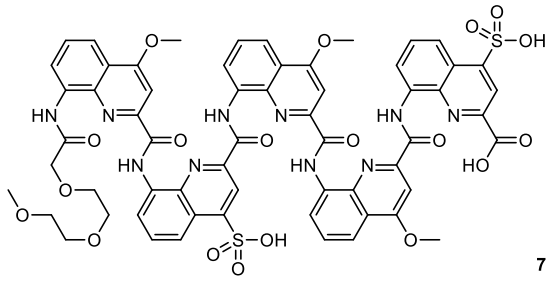
Compound 5: Oligomer **5** was synthesized on Rink Amide MBHA resin (0.33 mmol g^{-1}, 25 µmol scale) according to the standard method (automated). Loading of the first monomer: 0.33 mmol g^{-1} (quant.). The final Tail group was installed via the standard method (using Tail-OH). After purification by semi-prep HPLC (C18, 0–40B, 50 °C; A: 13 mmol ammonium acetate buffer pH 8.5, B: acetonitrile), the title compound was obtained as a yellow solid (18.5 mg, 12.8 µmol, 51%; HPLC-purity: 96%). **^{1}H NMR** (500 MHz, 1 mM in 12 mM ammonium acetate buffer pH 8.5): δ = 11.40 (s, 1H), 10.71 (s, 1H), 9.79 (s, 1H), 9.44 (s, 1H), 9.28 (s, 1H), 8.40 (d, *J* = 8.6 Hz, 1H), 8.22 (d, *J* = 9.2 Hz, 1H), 8.18 (d, *J* = 9.5 Hz, 1H), 8.14 (d, *J* = 15.4 Hz, 3H), 8.04 (d, *J* = 8.4 Hz, 1H), 7.66 (d, *J* = 8.9 Hz, 1H), 7.64–7.60 (m, 2H), 7.58 (d, *J* = 8.6 Hz, 1H), 7.25 (d, *J* = 9.1 Hz, 1H), 7.19 (t, *J* = 8.5 Hz, 1H), 7.15–7.06 (m, 4H), 6.95 (d, *J* = 8.2 Hz, 1H), 6.93–6.84 (m, 3H), 6.76 (s, 1H), 6.69 (d, *J* = 9.0 Hz, 1H), 6.55 (t, *J* = 8.5 Hz, 1H), 6.46 (d, *J* = 9.6 Hz, 1H), 6.16 (s, 1H), 5.88 (d, *J* = 9.3 Hz, 1H), 3.51 (d, *J* = 15.5 Hz, 1H), 3.47–3.34 (m, 6H), 3.28–3.16 (m, 3H), 3.15–3.05 (m, 3H), 2.87 (t, *J* = 8.6 Hz, 2H), 2.79 (dt, *J* = 13.0, 4.8 Hz, 1H), 2.71 (s, 3H), 2.66–2.54 (m, 3H), 2.32–2.13 (m, 3H), 1.74–1.53 (m, 2H), 1.40 (d, *J* = 15.9 Hz, 1H), 1.33 (d, *J* = 6.7 Hz, 4H), 1.27 (t, *J* = 7.9 Hz, 1H). **HRMS** (ESI⁻) *m/z* calcd. for $C_{69}H_{65}N_{12}O_{20}S_2$: 1445.3885 $(M-H)^-$; found: 1445.3797.



6

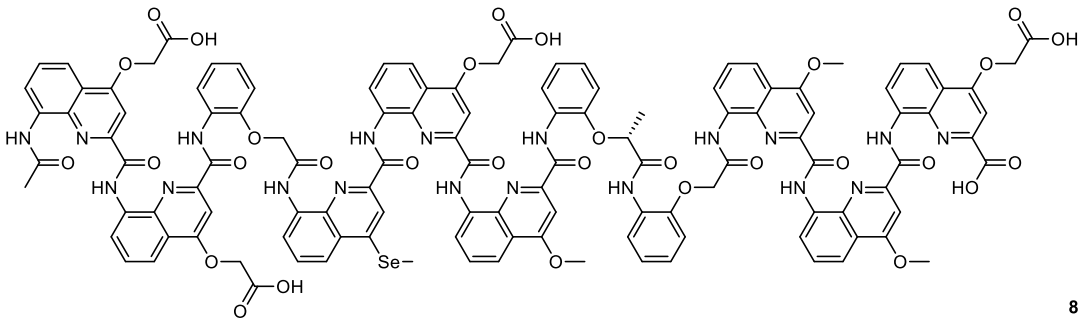
Compound 6: Oligomer **6** was synthesized on ProTide™ resin (0.15 mmol g^{-1}, 25 µmol scale) according to the standard method (manually). Loading of the first monomer: 0.13 mmol g^{-1} (88%). The final Tail group was installed via the standard method (using Tail-OH). After purification by semi-prep HPLC (C18, 0–50B, 50 °C; A: 13 mmol ammonium acetate buffer pH 8.5, B:

acetonitrile), the title compound was obtained as a yellow solid (11.3 mg, 8.82 µmol, 35%; HPLC-purity: 98%). **^{1}H NMR** (500 MHz, 1 mM in 12 mM ammonium acetate buffer pH 8.5): δ = 12.05 (s, 1H), 11.92 (s, 1H), 11.86 (s, 1H), 11.65 (s, 1H), 9.94 (s, 1H), 8.61 (d, *J* = 9.3 Hz, 1H), 8.53–8.30 (m, 4H), 8.14 (d, *J* = 10.2 Hz, 1H), 8.09 (d, *J* = 8.8 Hz, 1H), 8.01 (d, *J* = 9.0 Hz, 1H), 7.98–7.78 (m, 2H), 7.73 (d, *J* = 6.2 Hz, 1H), 7.67–7.57 (m, 2H), 7.52 (d, *J* = 8.5 Hz, 2H), 7.44 (t, *J* = 9.4 Hz, 1H), 6.93 (s, 1H), 6.84 (s, 1H), 4.15 (s, 2H), 4.12 (s, 2H), 3.59–3.34 (m, 3H), 3.10–2.97 (m, 1H), 2.94–2.83 (m, 1H), 2.74 (s, 4H), 2.67 (s, 3H), 2.57 (d, *J* = 14.7 Hz, 1H), 2.53–2.33 (m, 2H). **HRMS** (ESI$^-$) *m/z* calcd. for $C_{60}H_{49}N_{10}O_{19}S_2$: 1277.2622 $(M\text{-}H)^-$; found: 1277.3584.



7

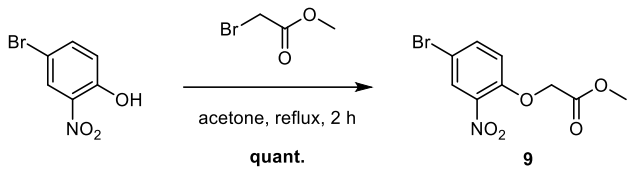
Compound 7: Oligomer **7** was synthesized on ProTide™ resin (0.15 mmol g^{-1}, 25 µmol scale) according to the standard method (automated). Loading of the first monomer: 0.11 mmol g^{-1} (74%). The final Tail group was installed via the standard method (using Tail-OH). After purification by semi-prep HPLC (C18, 0–50B, 50 °C; A: 13 mmol ammonium acetate buffer pH 8.5, B: acetonitrile), the title compound was obtained as a yellow solid (13.2 mg, 10.3 µmol, 41%; HPLC-purity: 96%). **^{1}H NMR** (500 MHz, 5 mM in 27 mM sodium phosphate buffer pH 7.0): δ = 11.04 (s, 1H), 10.95 (s, 1H), 10.93 (s, 1H), 10.90 (s, 1H), 9.23 (s, 1H), 8.18 (d, *J* = 10.0 Hz, 1H), 8.01 (s, 1H), 7.89 (d, *J* = 10.1 Hz, 1H), 7.71 (d, *J* = 8.8 Hz, 1H), 7.59–7.46 (m, 3H), 7.46–7.37 (m, 2H), 7.35 (d, *J* = 8.6 Hz, 1H), 7.23 (d, *J* = 7.4 Hz, 1H), 7.19 (d, *J* = 7.7 Hz, 1H), 7.16–7.03 (m, 2H), 6.95 (t, *J* = 8.5 Hz, 1H), 6.77 (s, 1H), 6.73 (s, 1H), 6.34 (s, 1H), 6.13 (s, 1H), 4.01 (s, 3H), 3.90 (s, 4H), 3.62 (s, 3H), 3.22–3.04 (m, 3H), 2.75–2.60 (m, 4H), 2.60–2.40 (m, 7H), 2.30–2.20 (m, 1H), 2.15–2.04 (m, 2H), 1.41 (t, *J* = 11.7 Hz, 1H). **HRMS** (ESI$^-$) *m/z* calcd. for $C_{60}H_{49}N_{10}O_{19}S_2$: 1277.2622 $(M\text{-}H)^-$; found: 1277.2303.



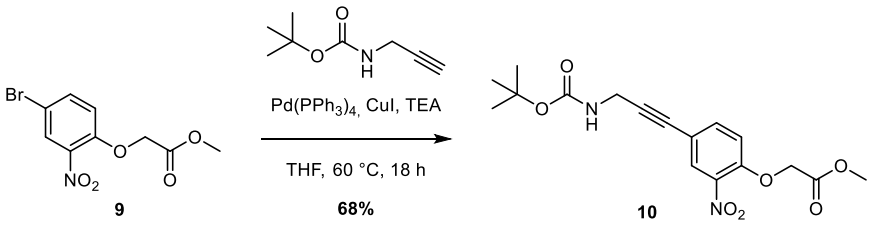
8

Compound 8: The synthesis of this oligomer has already been published previously.[93]

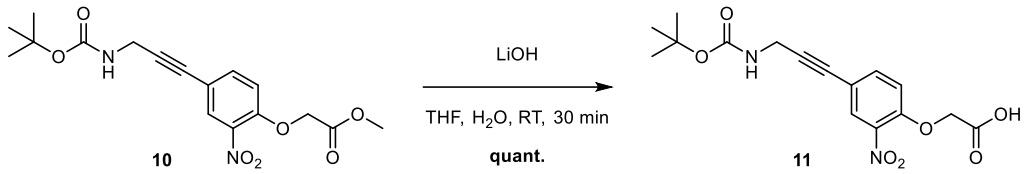
6.2.3 Monomer synthesis procedures



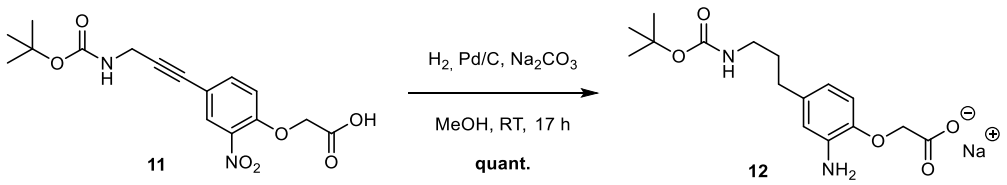
Compound 9: 4-Bromo-2-nitrophenol (10.0 g, 45.9 mmol, 1.0 equiv.), methyl bromoacetate (4.78 ml, 50.5 mmol, 1.1 equiv.) and K_2CO_3 (6.97 g, 50.5 mmol, 1.1 equiv.) were suspended in anhydrous acetone (180 ml). The reaction mixture was refluxed (bath temp.: 75 °C) for 2 h, filtered and solvents were evaporated *in vacuo*. Then, the residue was diluted with H_2O, extracted with DCM (2×) and the combined organic phases were dried over $MgSO_4$. After removing the solvents *in vacuo*, the title compound (13.7 g, 45.9 mmol, quant.) was obtained as a light-yellow solid. ($C_9H_8BrNO_5$, MW = 290.07 g mol^{-1}). $\mathbf{R_f}$ (CyHex/EtOAc 9:1) = 0.11. **^{1}H NMR** (500 MHz, $CDCl_3$): δ = 7.99 (d, *J* = 2.46 Hz, 1H, C3-H), 7.62 (dd, *J* = 8.94 Hz, 2.47 Hz, 1H, C5-H), 6.89 (d, *J* = 8.93 Hz, 1H, C6-H), 4.77 (s, 2H, C7-H), 3.80 (s, 3H, C9-H). **^{13}C NMR** (126 MHz, $CDCl_3$): δ = 168.0 (C8), 150.5 (C1), 140.9 (C2), 136.9 (C5), 128.7 (C3), 117.0 (C6), 113.7 (C4), 66.7 (C7), 52.7 (C9). **HRMS** (ESI$^+$) *m/z* calcd. for $C_9H_9BrNO_5$: 289.9659 (M+H)$^+$; Found: 289.9666.



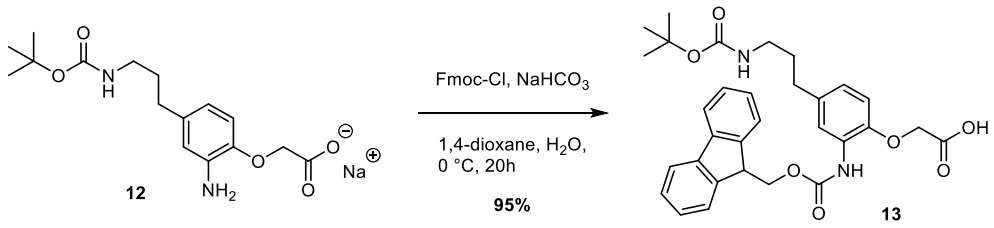
Compound 10: Compound **9** (5.73 g, 19.8 mmol, 1.0 equiv.), *N*-Boc-propargylamine (4.60 g, 29.6 mmol, 1.5 equiv.) and TEA (6.89 ml, 49.4 mmol, 2.5 equiv.) were dissolved in anhydrous THF (80 ml). The solution was degassed by freeze pump cycles (3×) and put under Ar atmosphere. Then, CuI (188 mg, 988 µmol, 5 mol%) and $Pd(PPh_3)_4$ (1.14 g, 988 µmol, 5 mol%) were added and the reaction mixture was stirred at 60 °C for 18 h. After diluting with H_2O, the mixture was extracted with DCM (3×), dried over $MgSO_4$ and solvents were removed *in vacuo*. The residue was triturated in Et_2O/CyHex 9:1, filtered and washed with CyHex. Finally, the precipitate was purified further by column chromatography (SiO_2, CyHex/EtOAc 8:2 → 6:4) yielding the title compound as a white solid (4.86 g, 13.3 mmol, 68%). ($C_{17}H_{20}N_2O_7$, MW = 364.35 g mol^{-1}). $\mathbf{R_f}$ (CyHex/EtOAc 6:4) = 0.39. **^{1}H NMR** (500 MHz, $CDCl_3$): δ = 7.91 (d, *J* = 2.08 Hz, 1H, C3-H), 7.53 (dd, *J* = 8.70 Hz, 2.12 Hz, 1H, C5-H), 6.90 (d, *J* = 8.69 Hz, 1H, C6-H), 4.79 (s, 2H, C7-H), 4.76 (s, 1H, N13-H), 4.14 (d, *J* = 5.69 Hz, 2H, C12-H), 3.81 (s, 3H, C9-H), 1.47 (s, 9H, C16-H, C17-H, C18-H). **^{13}C NMR** (126 MHz, $CDCl_3$): δ = 168.0 (C8), 155.4 (C14), 151.0 (C1), 140.2 (C2), 137.1 (C5), 129.2 (C3), 117.0 (C4), 115.1 (C6), 87.1 (C11), 80.4 (C10, C15), 66.5 (C7), 52.8 (C9), 31.2 (C13), 28.5 (C16, C17, C18). **HRMS** (ESI$^+$) *m/z* calcd. for $C_{17}H_{20}N_2O_7Na$: 387.1163 (M+Na)$^+$; Found: 387.1163.



Compound 11: Compound **10** (2.45 g, 6.72 mmol, 1.0 equiv.) was dissolved in THF (100 ml). After the addition of LiOH (193 mg, 8.07 mmol, 1.2 equiv.) in H_2O (25 ml), the solution was stirred at RT for 30 min. Then, the mixture was acidified using citric acid $_{(aq.)}$ (1 M). The resulting solution was extracted with DCM (3×) and dried over $MgSO_4$. After removing the solvents *in vacuo*, the title compound was obtained as a white solid (2.36 g, 6.72 mmol, quant.). ($C_{16}H_{18}N_2O_7$, MW = 350.33 g mol^{-1}). **^{1}H NMR** (500 MHz, DMSO-d_6): δ = 13.32 (s, 1H, O9-H), 7.90 (d, *J* = 2.15 Hz, 1H, C3-H), 7.63 (dd, *J* = 8.82 Hz, 2.20 Hz, 1H, C5-H), 7.36 (t, *J* = 5.74 Hz, 1H, N13-H), 7.24 (d, *J* = 8.81 Hz, 1H, C6-H), 4.92 (s, 2H, C7-H), 3.97 (d, *J* = 5.71 Hz, 2H, C12-H), 1.40 (s, 9H, C16-H, C17-H, C18-H). **^{13}C NMR** (126 MHz, DMSO-d_6): δ = 169.1 (C8), 155.3 (C14), 150.3 (C1), 139.5 (C2), 136.6 (C5), 127.5 (C3), 115.7 (C6), 114.9 (C4), 88.1 (C11), 79.1 (C10), 78.3 (C15), 65.5 (C7), 30.1 (C12), 28.2 (C16, C17, C18). **HRMS** (ESI$^-$) *m/z* calcd. for $C_{16}H_{17}N_2O_7$: 349.1041 (M-H)$^-$; Found: 349.1040.



Compound 12: Compound **11** (2.36 g, 6.74 mmol, 1.0 equiv.) and Na_2CO_3 (713 mg, 6.74 mmol, 1.0 equiv.) were suspended in MeOH (250 ml). After the solution was quickly degassed by vacuum–N_2 cycles (3×), Pd/C (236 mg, 10 wt. % loading) was added and the N_2 atmosphere was replaced by H_2. The reaction mixture was stirred at RT for 17 h, filtered over Celite© and washed with MeOH. Solvents were evaporated *in vacuo* yielding the title compound as a light-yellow solid (2.33 g, 6.74 mmol, 1.0 equiv.). ($C_{16}H_{23}N_2O_5Na$, MW = 346.36 g mol^{-1}). **^{1}H NMR** (500 MHz, DMSO-d_6): δ = 6.80 (t, *J* = 5.73 Hz, 1H, N14-H), 6.53 (d, *J* = 8.04 Hz, 1H, C6-H), 6.40 (d, *J* = 2.09 Hz, 1H, C3-H), 6.23 (dd, *J* = 8.10 Hz, 2.13 Hz, 1H, C5-H), 4.90 (s, 2H, N10-H), 3.98 (s, 2H, C7-H), 2.89 (q, *J* = 6.67 Hz, 2H, C13-H), 2.33 (t, *J* = 7.72 Hz, 2H, C11-H), 1.57 (p, *J* = 7.38 Hz, 2H, C12-H), 1.37 (s, 9H, C17-H, C18-H, C19-H). **^{13}C NMR** (126 MHz, DMSO-d_6): δ = 171.4 (C8), 155.6 (C15), 145.4 (C1), 139.0 (C2), 134.1 (C4), 115.7 (C5), 114.5 (C6), 114.1 (C3), 77.3 (C16), 70.8 (C7), 39.5 (C13), 32.2 (C11), 31.6 (C12), 28.3 (C17, C18, C19). **HRMS** (ESI$^-$) *m/z* calcd. for $C_{16}H_{22}N_2O_5$: 323.1612 (M-H)$^-$; Found: 323.1610.



Compound 13: Compound **12** (2.33 g, 6.74 mmol, 1.0 equiv.) and $NaHCO_3$ (2.83 g, 33.7 mmol, 5.0 equiv.) were dissolved in H_2O (150 ml). Then, Fmoc-Cl (2.27 g, 8.76 mmol, 1.3 equiv.) in 1,4-dioxane (150 ml) were added at 0 °C over 1 h. The reaction mixture was stirred at 0 °C for 1 h and then at RT for 18 h. After the mixture was acidified using citric acid $_{(aq.)}$ (1 M), it was extracted with DCM (3×), dried over $MgSO_4$ and solvents were removed *in vacuo*. The residue was purified by column chromatography (SiO_2, CyHex/EtOAc 1:1 + 1% DIPEA → CyHex/EtOAc 1:1 → CyHex/EtOAc 1:1 + 1% AcOH → EtOAc + 1% AcOH) to yield the title compound as a white solid (3.51 g, 6.42 mmol, 95%). ($C_{31}H_{34}N_2O_7$, MW = 546.62 g mol^{-1}). $\mathbf{R_f}$ (EtOAc/MeOH 98:2 + 1% AcOH) = 0.36. **^{1}H NMR** (500 MHz, DMSO-d_6): δ = 13.08 (s, 1H, O9-H), 8.65 (s, 1H, N19-H), 7.91 (dt, *J* = 7.57, 0.97 Hz, 2H, C25-H, C27-H), 7.74 (d, *J* = 7.44 Hz, 2H, C23-H, C30-H), 7.52 (s, 1H, C3-H), 7.43 (td, *J* = 7.46, 1.14 Hz, 2H, C25-H, C28-H), 7.34 (td, *J* = 7.49, 1.18 Hz, 2H, C24-H, C29-H), 6.89 (d, *J* = 8.39 Hz, 1H, C6-H), 6.87–6.82 (m, 2H, C5-H, N13-H), 4.69 (s, 2H, C7-H), 4.41 (d, *J* = 7.11 Hz, 2H, C21-H), 4.31 (t, *J* = 7.06 Hz, 1H, C22-H), 2.92 (q, *J* = 6.65 Hz, 2H, C12-H), 2.48–2.41 (m, 2H, C10-H), 1.61 (quint., *J* = 7.30 Hz, 2H, C11-H), 1.37 (s, 9H, C16-H, C17-H, C18-H). **^{13}C NMR** (126 MHz, DMSO-d_6): δ = 170.7 (C8), 155.6 (C14), 153.5 (C20), 146.7 (C1), 143.8 (C22a, C30a), 140.7 (C26a, C26b), 135.0 (C4), 127.7 (C25, C28), 127.5 (C2), 127.1 (C24, C29), 125.3 (C23, C30), 123.7 (C5), 120.2 (C26, C27), 113.6 (C6), 77.4 (C15), 66.4 (C7), 66.1 (C21), 46.6 (C22), 39.4 (C12), 32.0 (C10), 31.5 (C11), 28.3 (C16, C17, C18). **HRMS** (ESI$^-$) *m/z* calcd. for $C_{31}H_{33}N_2O_7$: 545.2293 (M-H)$^-$; Found: 545.2289.

6.3 Advanced NMR measurements

6.3.1 Assignment of oligomer 1

The 2D NOESY spectrum was 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 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 sine-square 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.

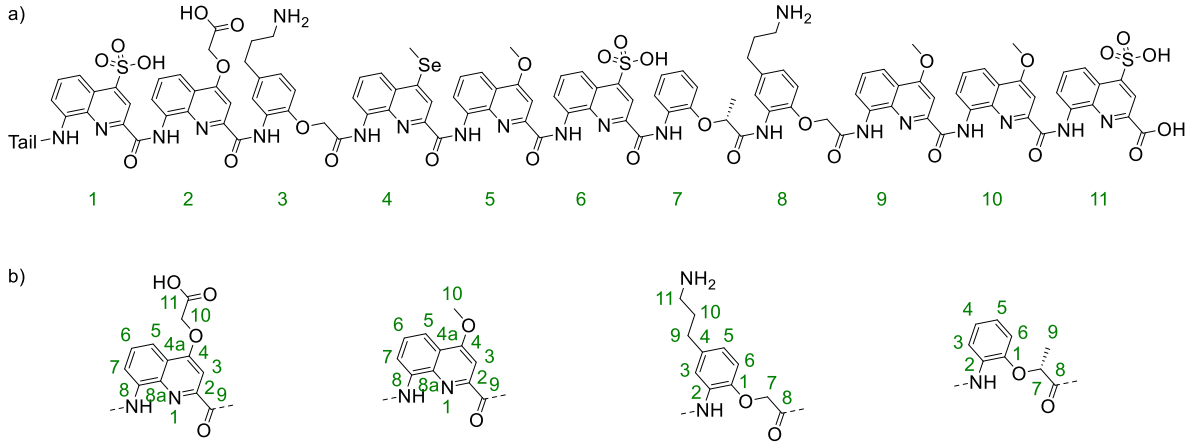


Fig. 92 a) Numbering of the units of compound **1** used in NMR assignment. b) Representative numbering of the carbon atoms in Q- and B-units.

Table 5 Assignment of the ^{1}H chemical shifts of compound **1** in 27 mM sodium phosphate buffer pH 7.0 H2O/D2O 9:1 at 25 °C (500 MHz).

Monomer	Atom	^{1}H (ppm)	Monomer	Atom	^{1}H (ppm)
Q1 (Sul)	H3	8.05	B7 (Rme)	H3	6.46
	H5	8.29		H4	5.61
	H6	7.68		H5	5.67
	H7	8.16		H6	5.51
	NH	9.93		H7	2.82
				H9	-0.71
				NH	8.71
Q2 (Asp)	H3	7.04	B8 (Orn)	H3	5.92
	H5	7.86		H5	5.73
	H6	7.34		H6	5.12
	H7	7.62		H7	3.05; 2.02
	H10	4.87; 4.75		H9	1.75
	NH	11.60		H10	1.32
				H11	2.88
				NH	6.15
B3 (Orn)	H3	7.35	Q9 (Ala)	H3	7.14
	H5	6.82		H5	7.50
	H6	6.30		H6	7.00
	H7	3.34; 1.61		H7	6.86
	H9	3.34; 1.61		H10	4.19
	H10	1.84		NH	8.60
	H11	2.99			
	NH	9.72			
Q4 (Sem)	H3	6.21	Q10 (Ala)	H3	5.99
	H5	7.46		H5	6.84
	H6	7.27		H6	6.80
	H7	6.88		H7	7.27
	H10	2.18		H10	3.29
	NH	9.32		NH	10.88
Q5 (Ala)	H3	6.88	Q11 (Sul)	H3	6.72
	H5	7.89		H5	7.65
	H6	7.64		H6	7.14
	H7	7.91		H7	7.77
	H10	4.05		NH	11.07
	NH	11.6			
Q6 (Sul)	H3	7.00			
	H5	7.96			
	H6	7.42			
	H7	8.27			
	NH	10.7			

Table 6 Assignment of the ^{13}C chemical shifts of compound **1** in 27 mM sodium phosphate buffer pH 7.0 H_2O/D_2O 9:1 at 25 °C (500 MHz).

Monomer	Atom	^{13}C (ppm)	Monomer	Atom	^{13}C (ppm)
Q1 (Sul)	C3	120.0	B7 (Rme)	C1	149.3
	C4	152.4		C3	120.1
	C4a	127.3		C4	124.7
	C5	123.8		C5	127.0
	C6	133.3		C6	117.0
	C7	120.4		C7	81.8
	C8	136.4		C8	168.3
	C8a	139.8		C9	16.4
Q2 (Asp)	C3	102.4	B8 (Orn)	C1	142.4
	C4	166.2		C2	135.0
	C4a	124.4		C3	118.4
	C5	119.8		C4	126.0
	C6	130.8		C5	125.6
	C7	119.6		C6	111.4
	C8	133.7		C7	68.3
	C8a	139.9		C9	32.7
	C10	70.6		C10	29.1
				C11	42.5
B3 (Orn)	C1	145.0	Q9 (Ala)	C3	102.0
	C2	137.4		C4	167.0
	C3	119.8		C4a	123.5
	C4	127.7		C5	119.2
	C5	127.3		C6	130.4
	C6	113.5		C7	119.4
	C7	68.2		C8	132.7
	C9	34.0		C8a	138.5
	C10	31.4		C10	59.5
	C11	42.0			
Q4 (Sem)	C3	117.3	Q10 (Ala)	C3	100.3
	C4	153.7		C4	164.6
	C4a	130.5		C4a	122.8
	C5	123.1		C5	119.1
	C6	131.7		C6	130.7
	C7	119.6		C7	117.8
	C8	133.4		C8	134.3
	C8a	136.4		C8a	134.4
	C10	8.2		C10	58.6
Q5 (Ala)	C3	101.1	Q11 (Sul)	C3	120.3
	C4	166.9		C4	152.9
	C4a	125.1		C4a	125.3
	C5	119.8		C5	123.5
	C6	131.0		C6	131.8
	C7	119.8		C7	118.1
	C8	134.8		C8	135.9
	C8a	140.7		C8a	139.9
	C10	59.3			
Q6 (Sul)	C3	117.9			
	C4	152.4			
	C4a	126.7			
	C5	123.9			
	C6	132.9			
	C7	119.9			
	C8	135.0			
	C8a	139.2			

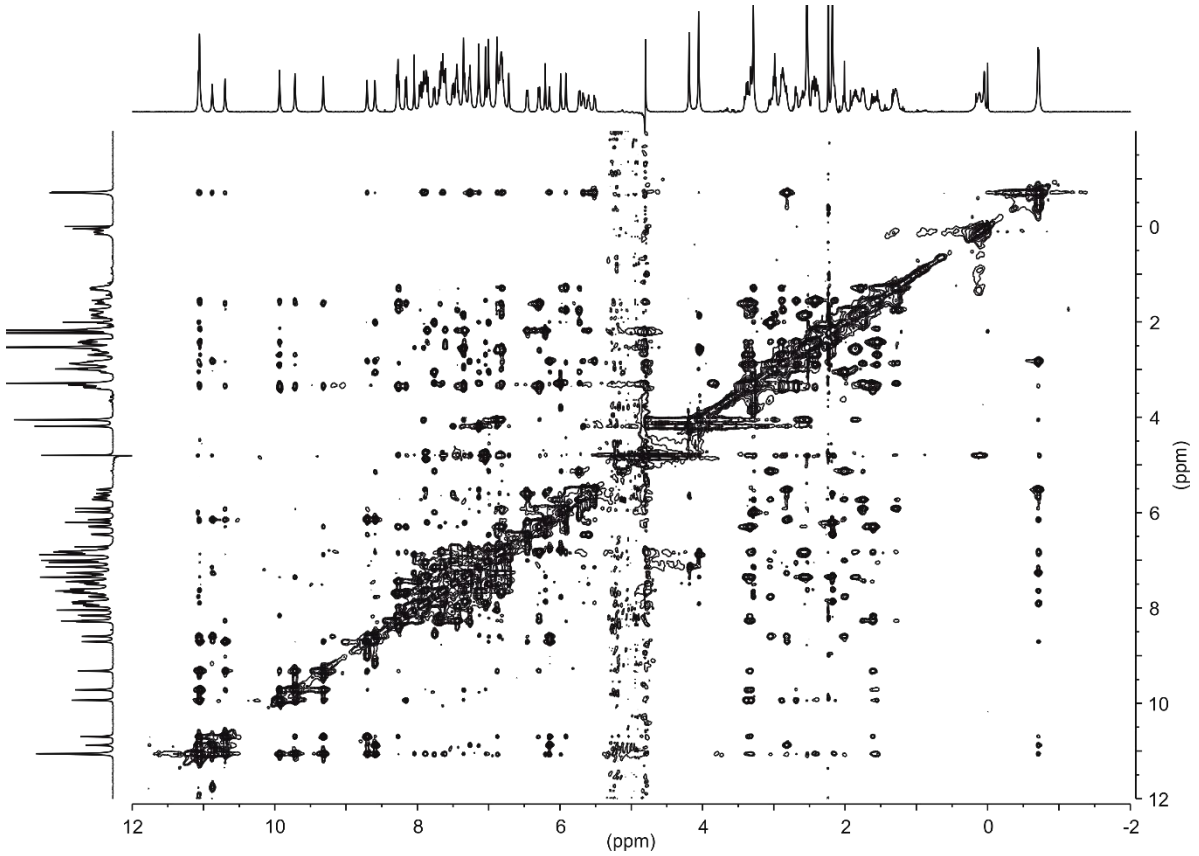


Fig. 93 ^{1}H-^{1}H NOESY spectrum of **1** (500 MHz, 27 mM sodium phosphate buffer pH 7.0).

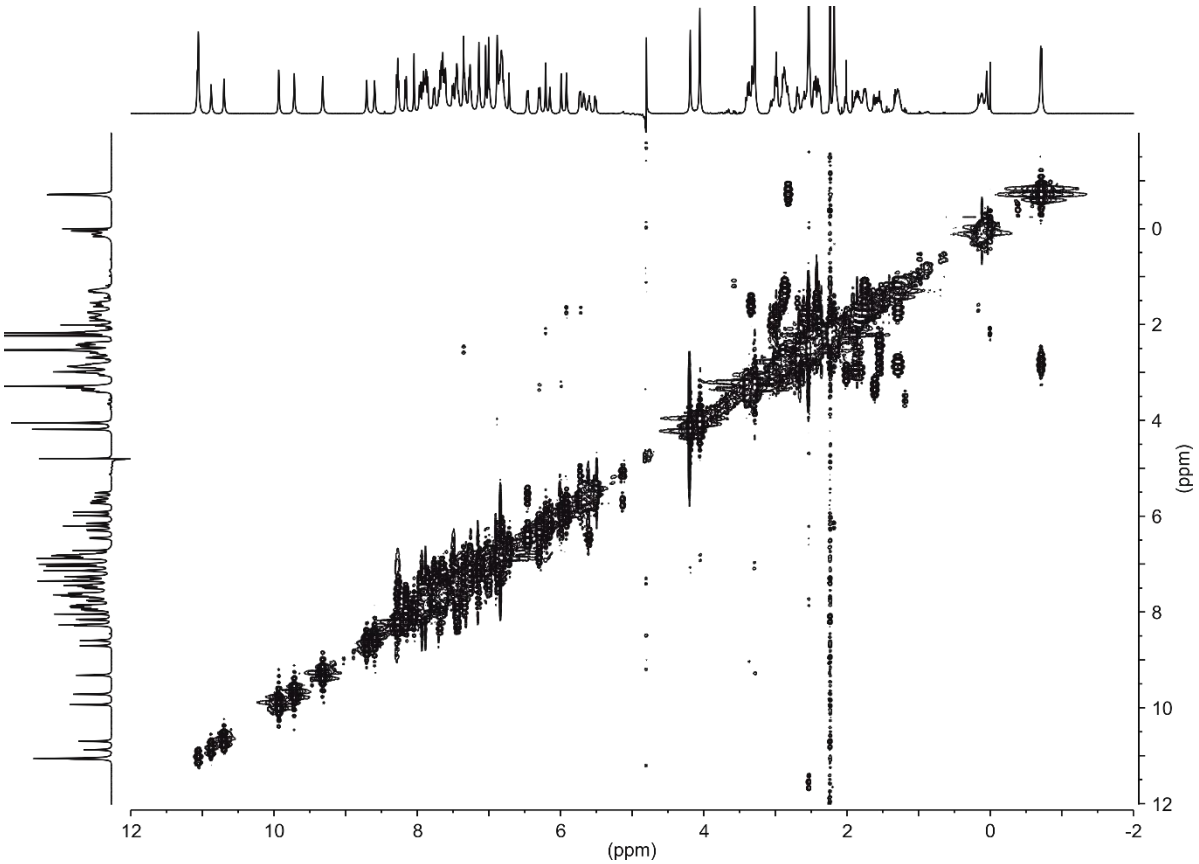


Fig. 94 ^{1}H-^{1}H COSY spectrum of **1** (500 MHz, 27 mM sodium phosphate buffer pH 7.0).

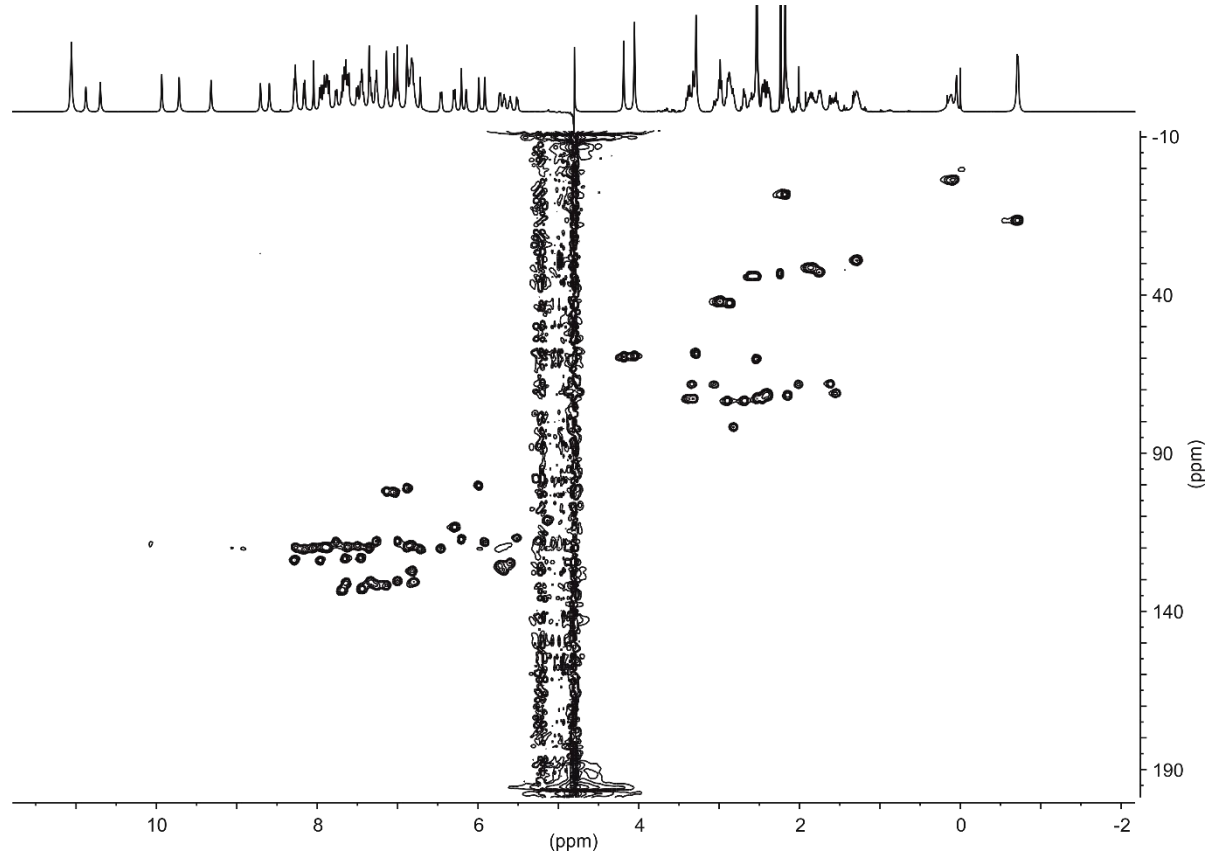


Fig. 95 ^{1}H-^{13}C HSQC spectrum of **1** (500 MHz, 27 mM sodium phosphate buffer pH 7.).

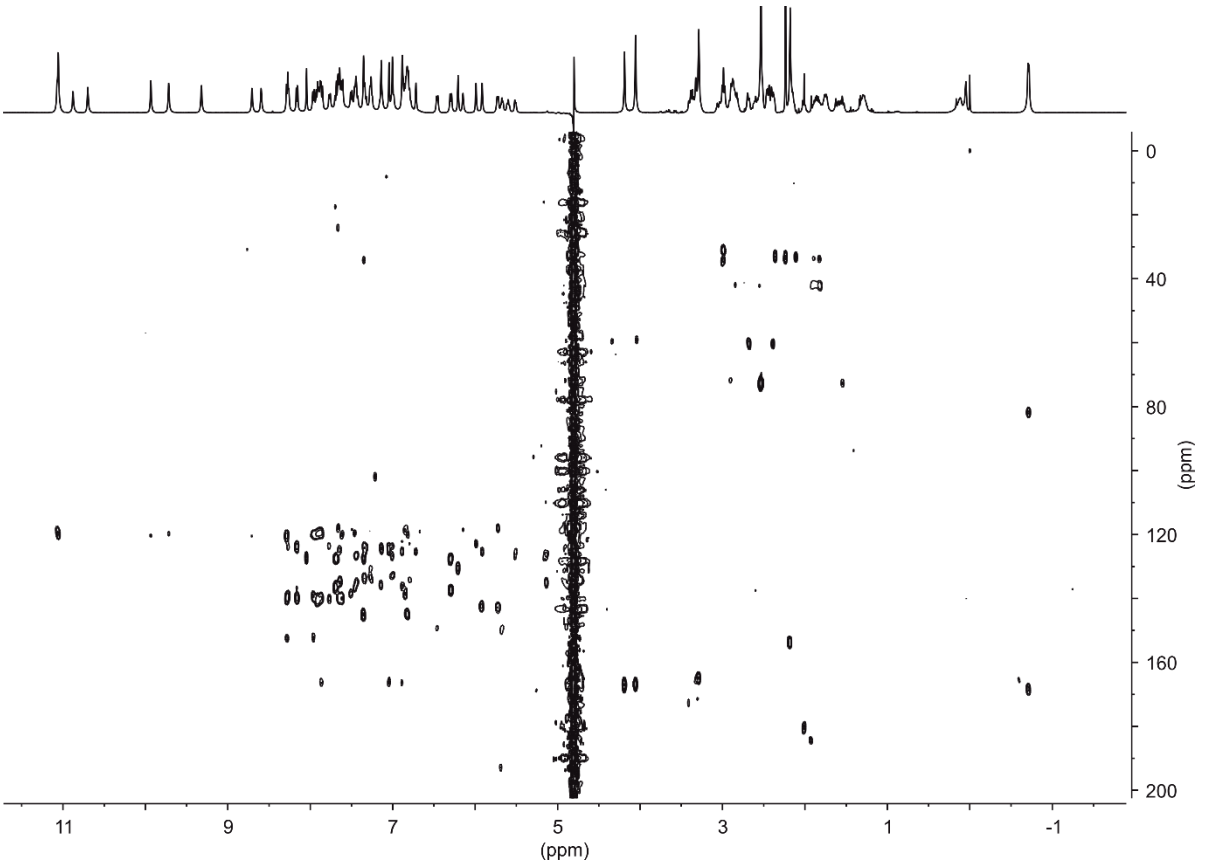


Fig. 96 ^{1}H-^{13}C HMBC spectrum of **1** (500 MHz, 27 mM sodium phosphate buffer pH 7.0).

6.3.2 DOSY NMR of oligomers 1 and 2

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 65k complex data points were acquired. DOSY processing was performed with the DOSY processing tool from MestReNova (v.12.0.0-20080) employing the Peak Heights Fit algorithm including the options "use existing peaks", "autocorrect peak positions", and "overlapped peaks analysis" with 32 points in diffusion dimension and a window of $1.00*10^{-4}$ to $1.00*10^{-8}$ cm^2 s^{-1}.

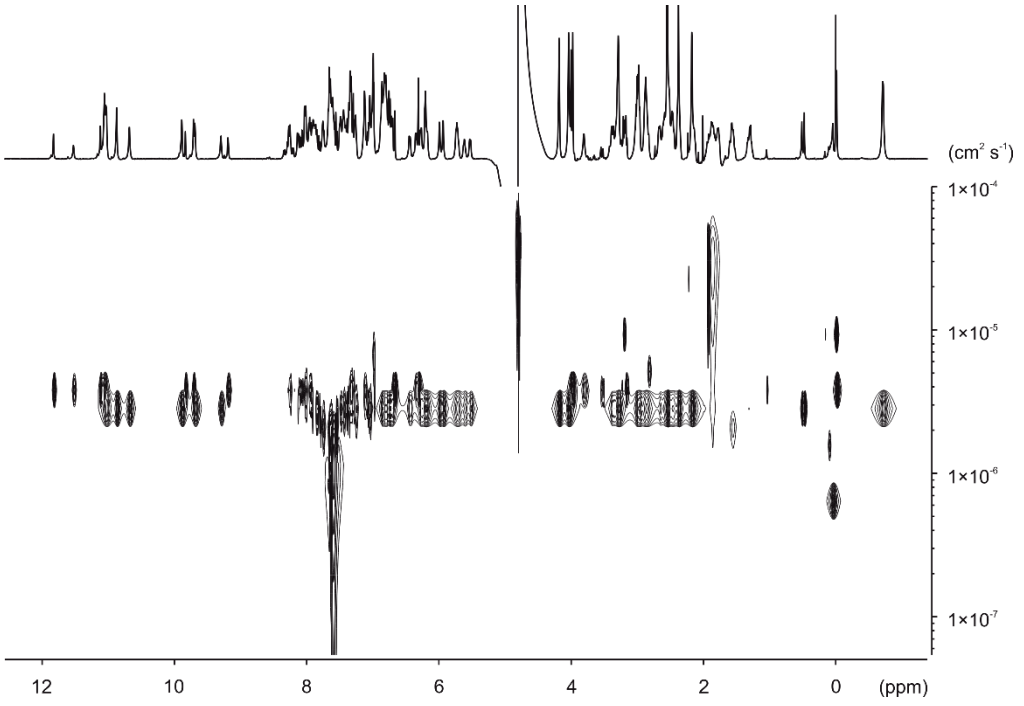


Fig. 97 DOSY spectrum of a mixture of **1** (1 mM) and **2** (1 mM) in 27 mM sodium phosphate buffer pH 7.0.

6.4 X-ray Crystallography

After lyophilization, compound **1** was dissolved in pure water to a final concentration of 10 mM. Crystallization trials were performed at 20 °C by standard aqueous sitting drop vapor diffusion method and subsequently optimized by hanging drop vapor diffusion method at 20 °C. Hexagonal prisms (Fig. 98) were obtained by mixing 0.5 µL of **1** and 0.5 µL of crystallization reagent comprising 20% w/v polyethylene glycol 8,000, 10 mM TRIS buffer, pH 7.5, 10 mM calcium chloride; equilibrated against 500 µL of crystallization reagent in the reservoir. For data collection, a single crystal was fished using a MiTeGen microloop, cryo-protected in a solution composed of 30% v/v polyethylene glycol 400 and 20% w/v polyethylene glycol 8,000 and flash frozen in liquid nitrogen.

Synchrotron data was collected on beam line P14 operated by EMBL Hamburg at the Petra III storage ring (DESY, Hamburg, Germany) using 0.9762 Å wavelength. During data collection, the crystal was cooled to 100 K. The crystal was exposed for 0.008 s and 0.2° oscillation per frame and a rotation pass of 360° was measured using an EIGER 16M detector. Data were processed using *autoPROC* pipeline.[104, 122, 138] The crystal belonged to the monoclinic space group $P2_1$ with unit cell parameters a = 21.400 (9) Å, b = 35.063 (1) Å, c = 23.820 (9) Å and β = 100.84 (4)°, V = 17,554 (10) Å^3 and two independent molecules in the asymmetric unit.

The structure was solved using dual space method with the program *SHELXD*[105] and refined by the full-matrix least-squares method on F^2 with *SHELXL*-2014[106] within *Olex2*.[107] After each refinement step, visual inspection of the model and the electron-density maps were carried out using *Olex2* and *Coot*.[108] AFIX, DFIX, EADP and FLAT instructions were used to improve the geometry of molecules and temperature parameters. One CH_3 group was severely disordered in one diethylene glycol tail and omitted. All non-H atoms were refined with anisotropic displacement parameters. Hydrogen atoms were placed at idealized positions. Restraints on anisotropic displacement parameters were implemented with DELU, SIMU, RIGU and ISOR instructions. In the final stage of refinement *SQUEEZE*[109] procedure from Platon suite was introduced to remove unmodeled electron density. Calculated total potential solvent accessible void volume and electron counts per unit cell were 7519.3 Å^3 and 2307, respectively.

Statistics of data collection and refinement are described in Table S1. The final cif file was checked using IUCr's checkcif algorithm. Due to large volume fractions of disordered solvent molecules, weak diffraction intensity and poor resolution, a number of A- and B-level remain in the checkcif file. These alerts are inherent to the data and refinement procedures and illustrate the limited practicality of the checkcif tool for medium- size molecule crystallography. They are listed below and have been divided into two groups. The first group illustrates weak quality of the data and refinement statistics if compared to that expected for small molecule structures from highly diffracting crystals. The second group is connected to decisions made during refinement and explained below. Atomic coordinates and structure factors for $(\mathbf{1})_2$ was deposited in the Cambridge Crystallographic Data Centre (CCDC) with accession code 2122518. The data is available free of charge upon request (www.ccdc.cam.ac.uk/).

Group 1 alerts:

THETM01_ALERT_3_A The value of sine(theta_max)/wavelength is less than 0.550
Calculated sin(theta_max)/wavelength = 0.4673
PLAT029_ALERT_3_A _diffrn_measured_fraction_theta_full value Low . 0.915 Why?
PLAT341_ALERT_3_B Low Bond Precision on C-C Bonds 0.0197 Ang.
PLAT911_ALERT_3_B Missing FCF Refl Between Thmin & STh/L= 0.467
PLAT241_ALERT_2_B High 'MainMol' Ueq as Compared to Neighbors of
PLAT242_ALERT_2_B Low 'MainMol' Ueq as Compared to Neighbors of

Group 2 alerts:

SHFSU01_ALERT_2_A The absolute value of parameter shift to su ratio > 0.20
Absolute value of the parameter shift to su ratio given 2.141
Additional refinement cycles did not improve this.
PLAT080_ALERT_2_A Maximum Shift/Error 2.14 Why?
Additional refinement cycles did not improve this.
PLAT031_ALERT_4_A Refined Extinction Parameter within Range 0.000 Sigma
Extinction parameter was removed.

Table 7 Crystallographic data and refinement details for $(\mathbf{1})_2$.

Chemical formula	$C_{123.5}\ H_{103.75}\ N_{20.75}\ O_{34}\ S_3\ Se_1$
Formula weight	2597.66
Temperature	100 (2) K
Wavelength	0.9762 Å
Crystal system	Monoclinic
Space group	$P2_1$
Unit cell dimensions	a = 21.400 (9) Å, b = 35.063 (10) Å, c = 23.820 (9) Å α = 90°, β = 100.848 (4)°, γ = 90°
Volume	17554 (10) $Å^3$
Z	4
Density (calculated)	0.983 g/cm^3
Absorption coefficient	0.719 μ/mm^{-1}
Color and shape	Yellow, prism
Crystal size	0.05 x 0.05 x 0.18 mm
Index ranges	$-19 \leq h \leq 19$, $-32 \leq k \leq 32$, $-22 \leq l \leq 22$
Completeness to θ = 27.14°	91.2 %
Reflections collected	91128
Independent reflections	27368 [R_{int} = 0.0745, R_{sigma} = 0.0655]
Data/restraints/parameters	27368/1066/2079
Goodness-of-fit on F^2	1.086
Final R indexes [$I > 2\sigma(I)$]	R_1 = 0.0899, wR_2 = 0.2555
Final R indexes [all data]	R_1 = 0.1235, wR_2 = 0.2950
Largest diff. peak and hole	0.57/-0.29 e $Å^{-3}$
CCDC #	2122518

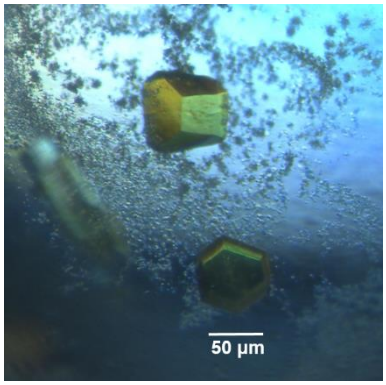


Fig. 98 Crystals of $(\mathbf{1})_2$ observed under crossed polarizing microscope.

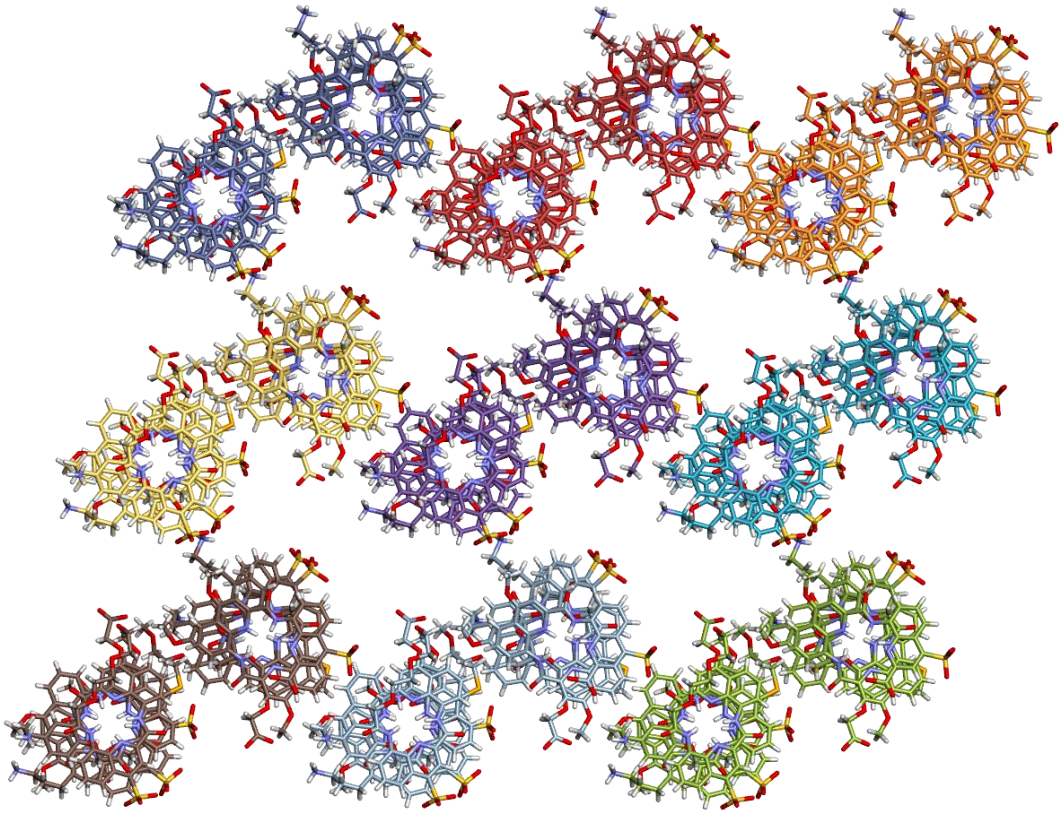


Fig. 99 Packing of **1** in the crystal viewed down the b-axis. Helices in the same unit cell are shown in the same color.

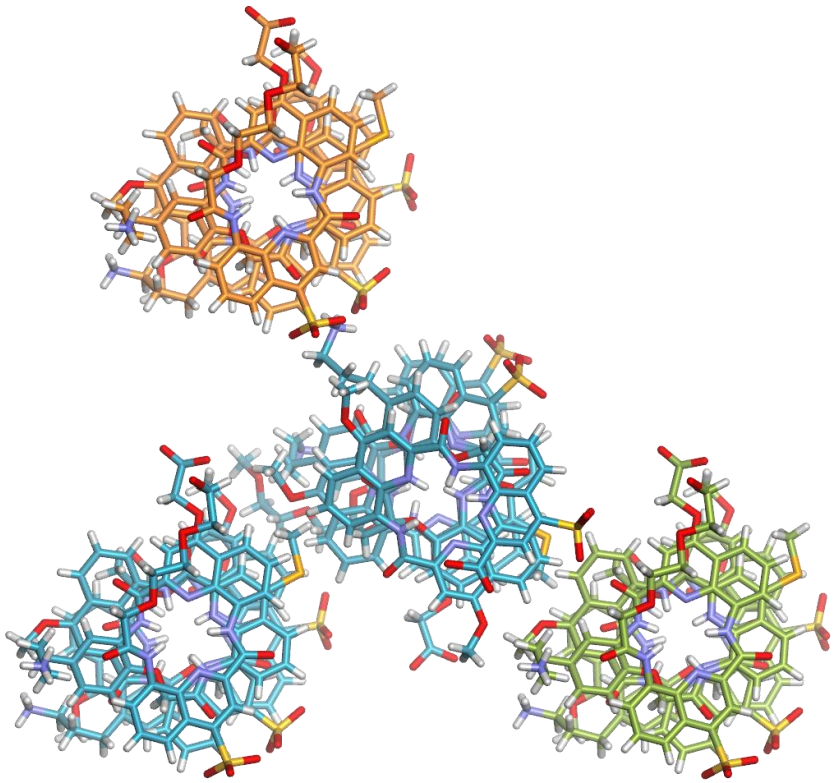


Fig. 100 Coplanar neighbors in the crystal structure of **1**. Helices in the same unit cell are shown in the same color. Each helix (illustrated by the middle helix in this figure) has three different coplanar neighbors: two are mediated by hydrophobic contacts (blue and green neighbor), one is mediated by a salt bridge (orange neighbor).

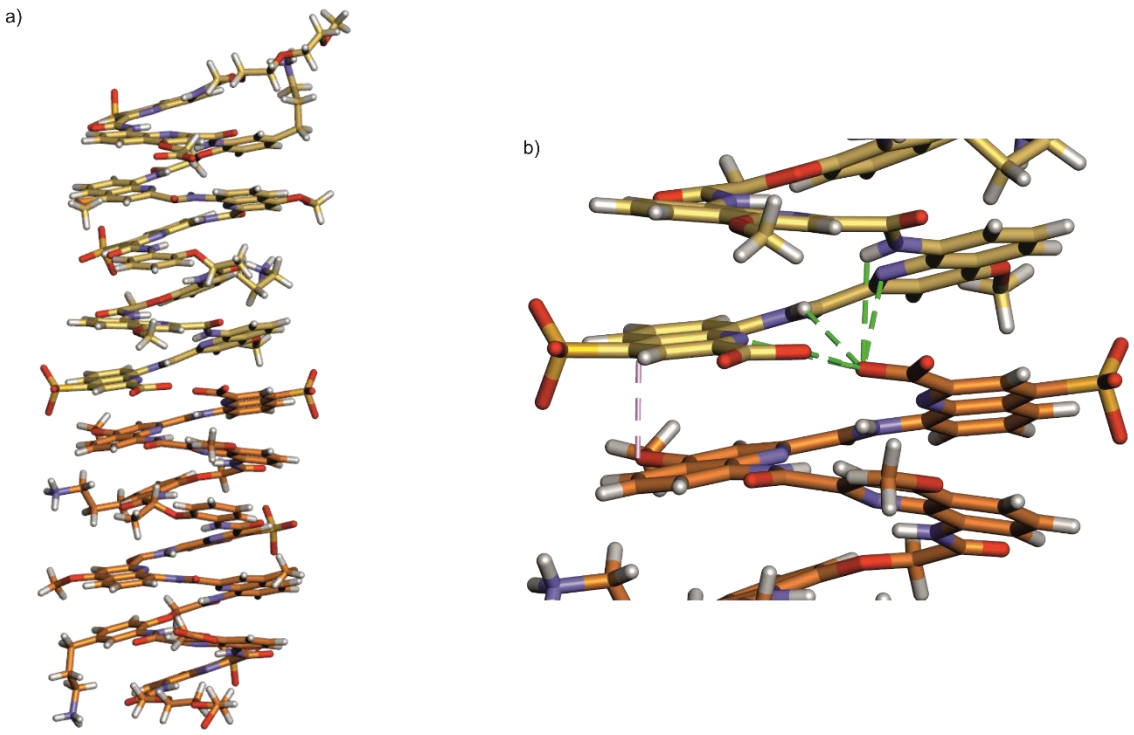


Fig. 101 Stacking of helices via their C-termini in the crystal structure of **1** showing two whole helices (a) and a zoom on the intermolecular interface (b). Hydrogen bonding interactions and hydrophobic contacts are shown by green and pink dashed lines, respectively.

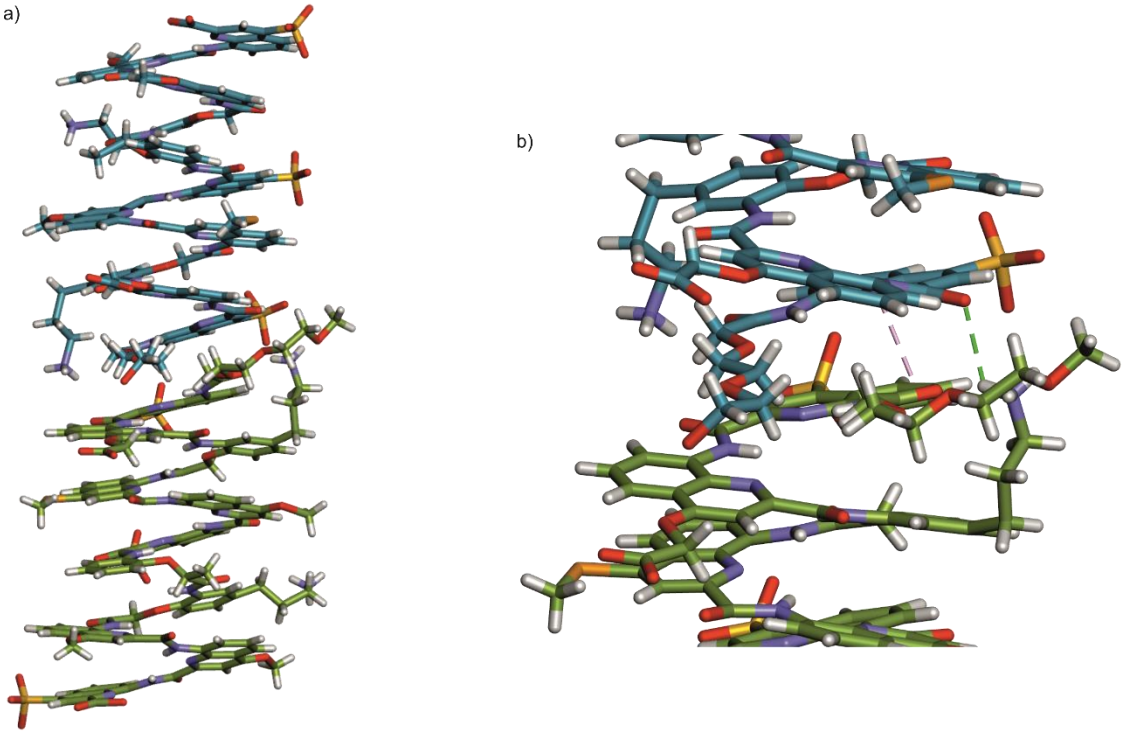


Fig. 102 Stacking of helices via their N-termini in the crystal structure of **1** showing two whole helices (a) and a zoom on the intermolecular interface (b). Hydrogen bonding interactions and hydrophobic contacts are shown by green and pink dashed lines, respectively.

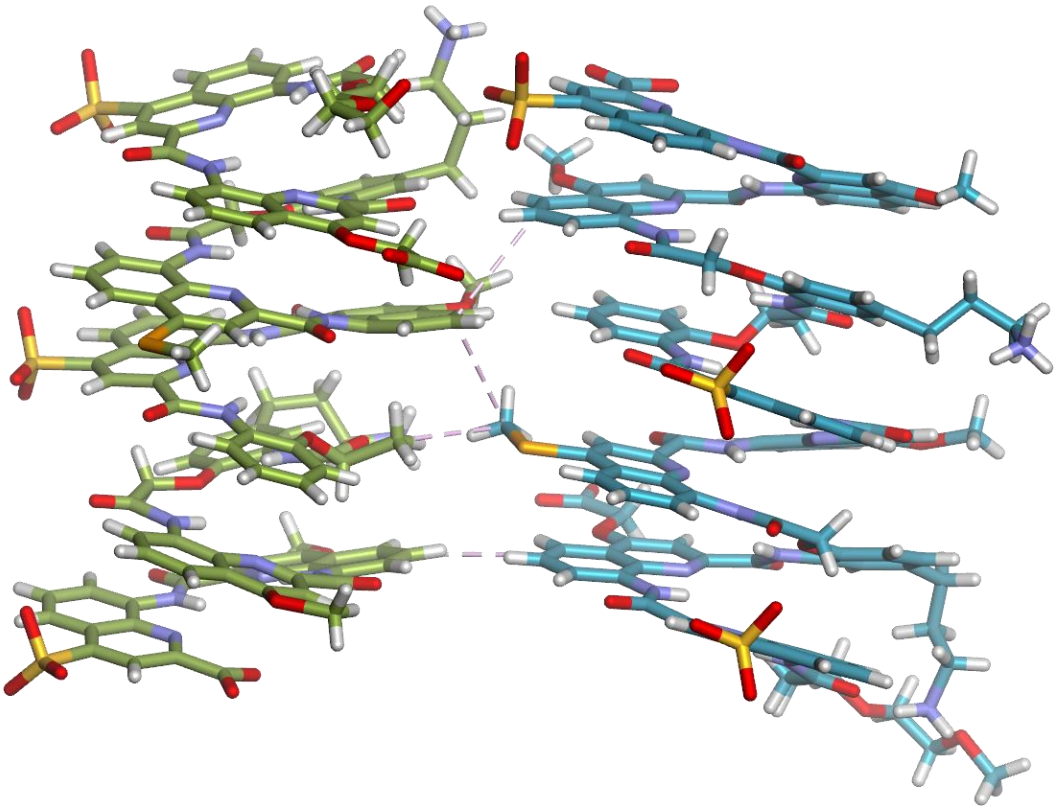


Fig. 103 Tightest side-to-side helix-helix interface mediated by hydrophobic contacts in the crystal structure of **1**. Close hydrophobic contacts are highlighted by dashed lines.

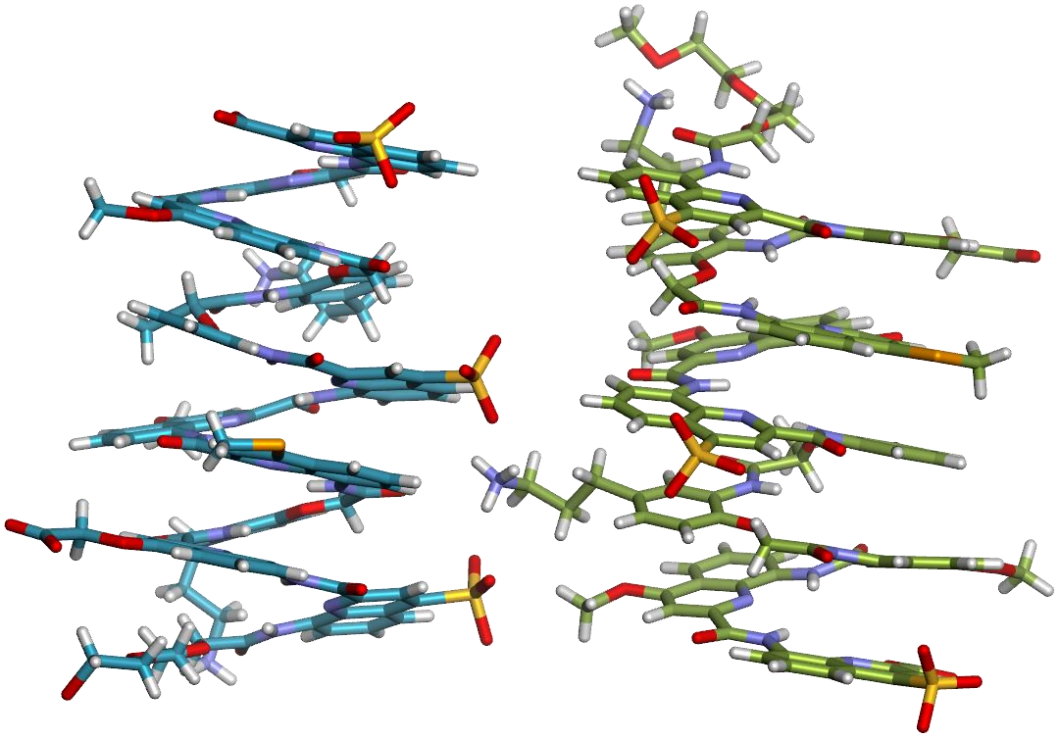


Fig. 104 Intermolecular salt-bridge observed between the side chains of a B^{Orn} and two Q^{Sul} units in the crystal structure of **1**.

6.5 Spectra and Chromatograms

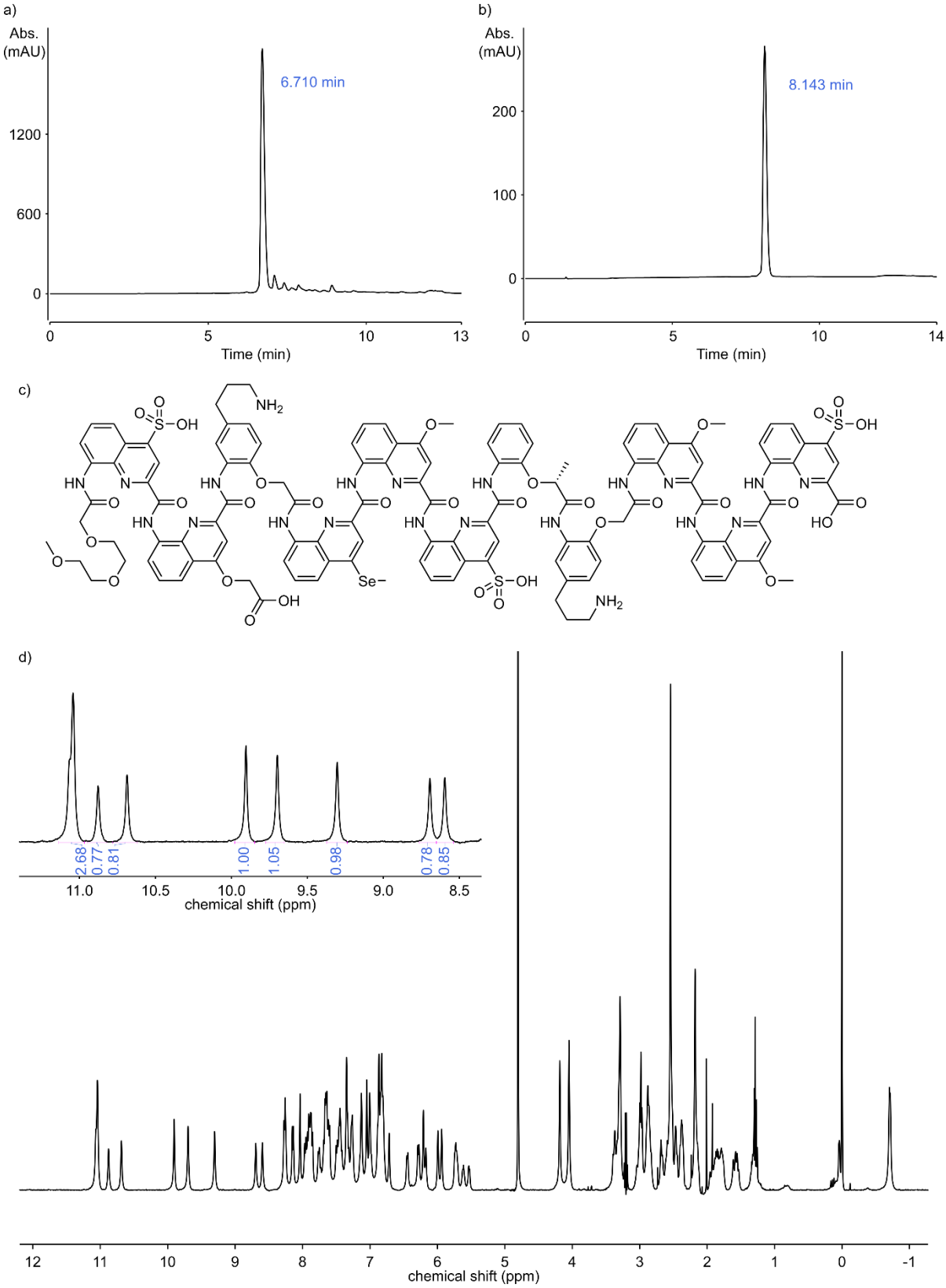


Fig. 105 Analytical data of compound **1**. HPLC chromatograms after cleavage from the resin (a) and after purification (b) (C8, 0–40B, 50 °C; A: 13mM ammonium acetate buffer pH 8.5, B: acetonitrile). c) Chemical structure of compound **1**. d) ^{1}H NMR spectrum (500 MHz, 1.0 mM in 27 mM sodium phosphate buffer pH 7.0).

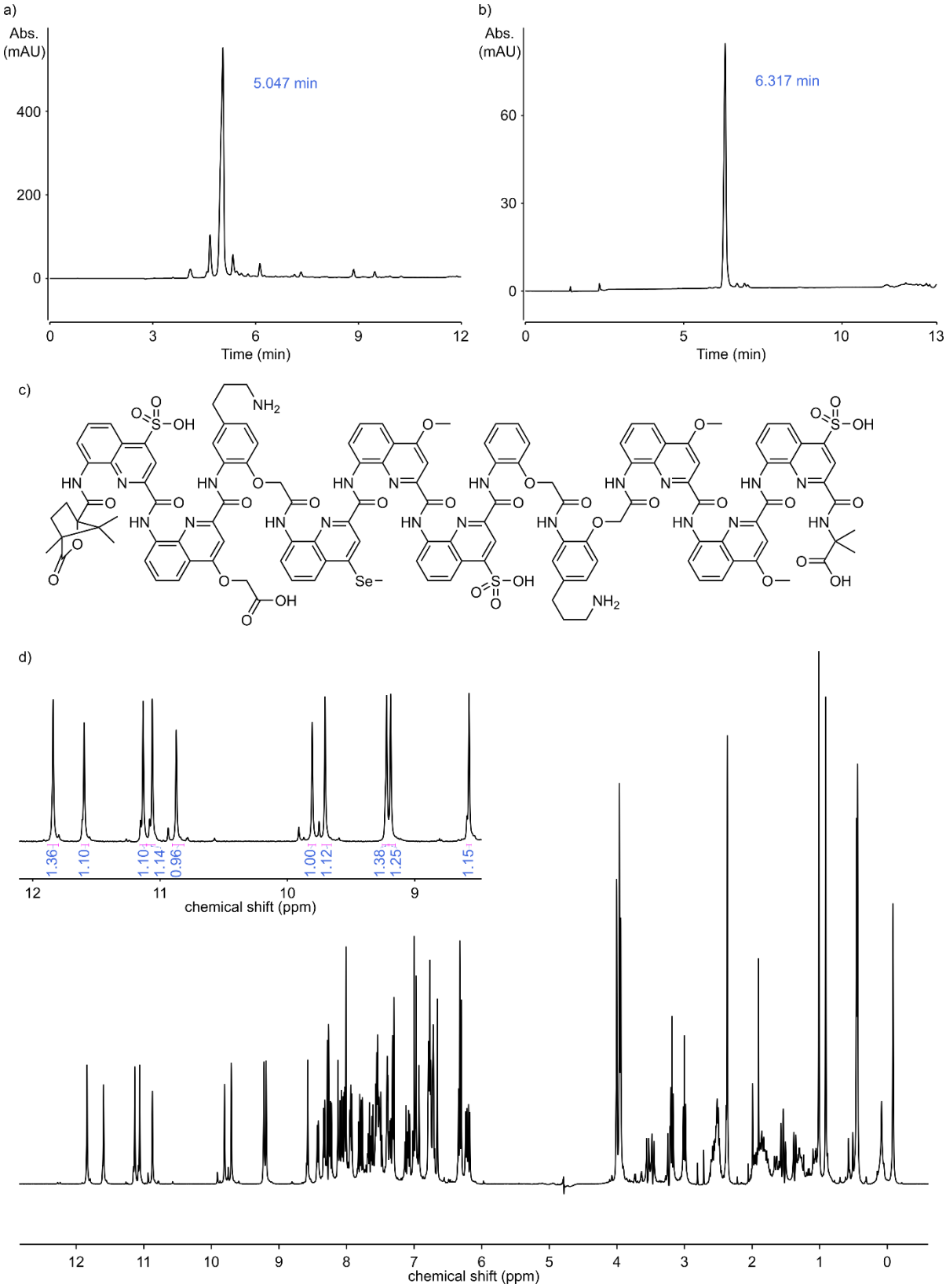


Fig. 106 Analytical data of compound **2**. HPLC chromatograms after cleavage from the resin (a) (C8, 0–60B, 50 °C; A: 13mM ammonium acetate buffer pH 8.5, B: acetonitrile) and after purification (b) (C8, 0–35B, 50 °C; A: 13mM ammonium acetate buffer pH 8.5, B: acetonitrile). c) Chemical structure of compound **2**. d) ^{1}H NMR spectrum (500 MHz, 1.0 mM in H_2O/D_2O 9:1).

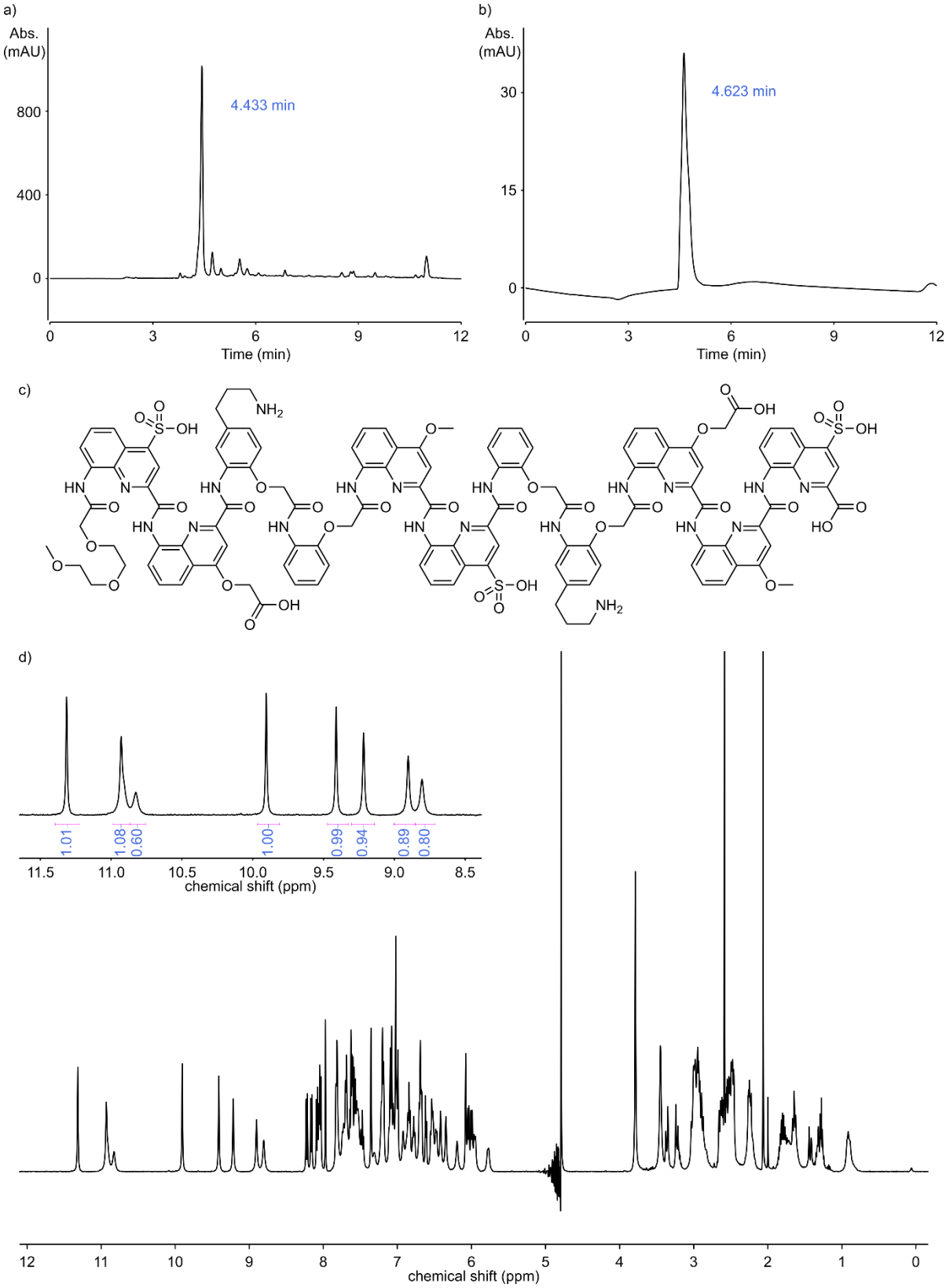


Fig. 107 Analytical data of compound **3**. HPLC chromatograms after cleavage from the resin (a) (C8, 0–40B, 50 °C; A: 13mM ammonium acetate buffer pH 8.5, B: acetonitrile) and after purification (b) (C8, 0–60B, 25 °C; A: 13mM ammonium acetate buffer pH 8.5, B: acetonitrile). c) Chemical structure of compound **3**. d) ^{1}H NMR spectrum (500 MHz, 0.5 mM in H_2O/D_2O 9:1).

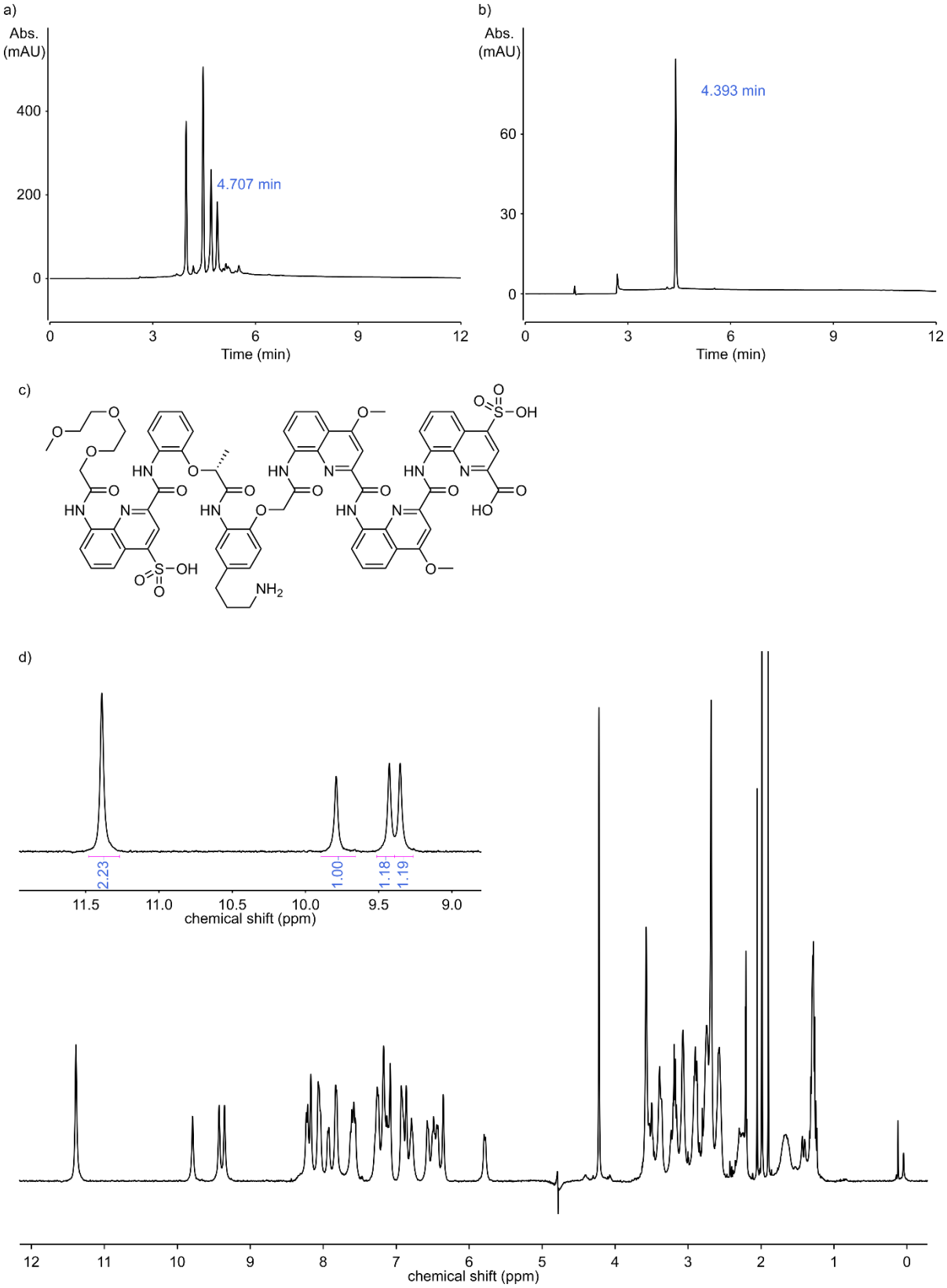


Fig. 108 Analytical data of compound **4**. HPLC chromatograms after cleavage from the resin (a) and after purification (b) (C18, 0–100B, 50 °C; A: 13mM ammonium acetate buffer pH 8.5, B: acetonitrile). c) Chemical structure of compound **4**. d) ^{1}H NMR spectrum (500 MHz, 1.0 mM in 27 mM sodium phosphate buffer pH 7.0).

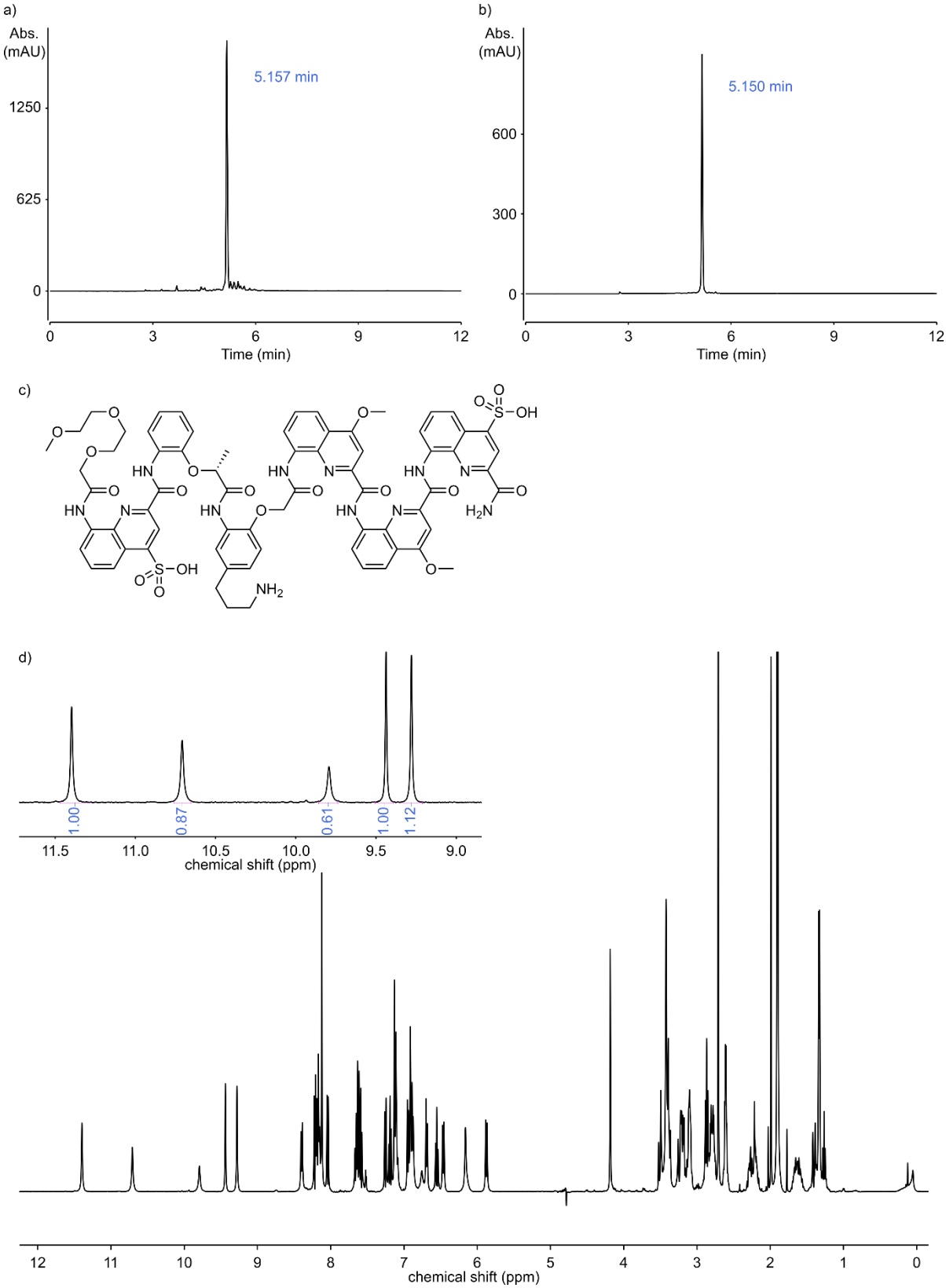


Fig. 109 Analytical data of compound **5**. HPLC chromatograms after cleavage from the resin (a) and after purification (b) (C18, 0–100B, 50 °C; A: 13mM ammonium acetate buffer pH 8.5, B: acetonitrile). c) Chemical structure of compound **5**. d) ^{1}H NMR spectrum (500 MHz, 1.0 mM in 12 mM ammonium acetate buffer pH 7.0).

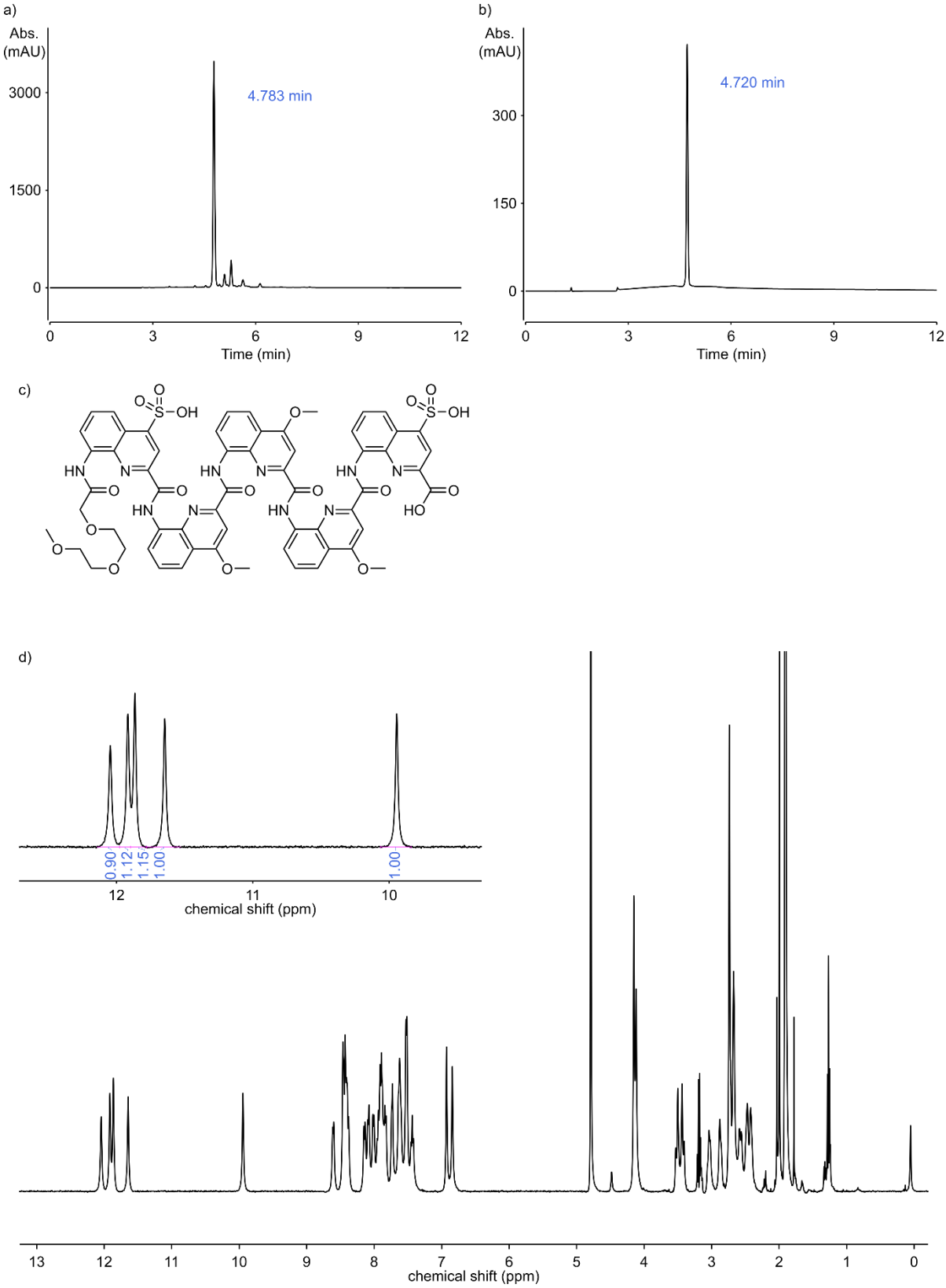


Fig. 110 Analytical data of compound **6**. HPLC chromatograms after cleavage from the resin (a) and after purification (b) (C18, 0–100B, 50 °C; A: 13mM ammonium acetate buffer pH 8.5, B: acetonitrile). c) Chemical structure of compound **6**. d) ^{1}H NMR spectrum (500 MHz, 1.0 mM in 12 mM ammonium acetate buffer pH 7.0).

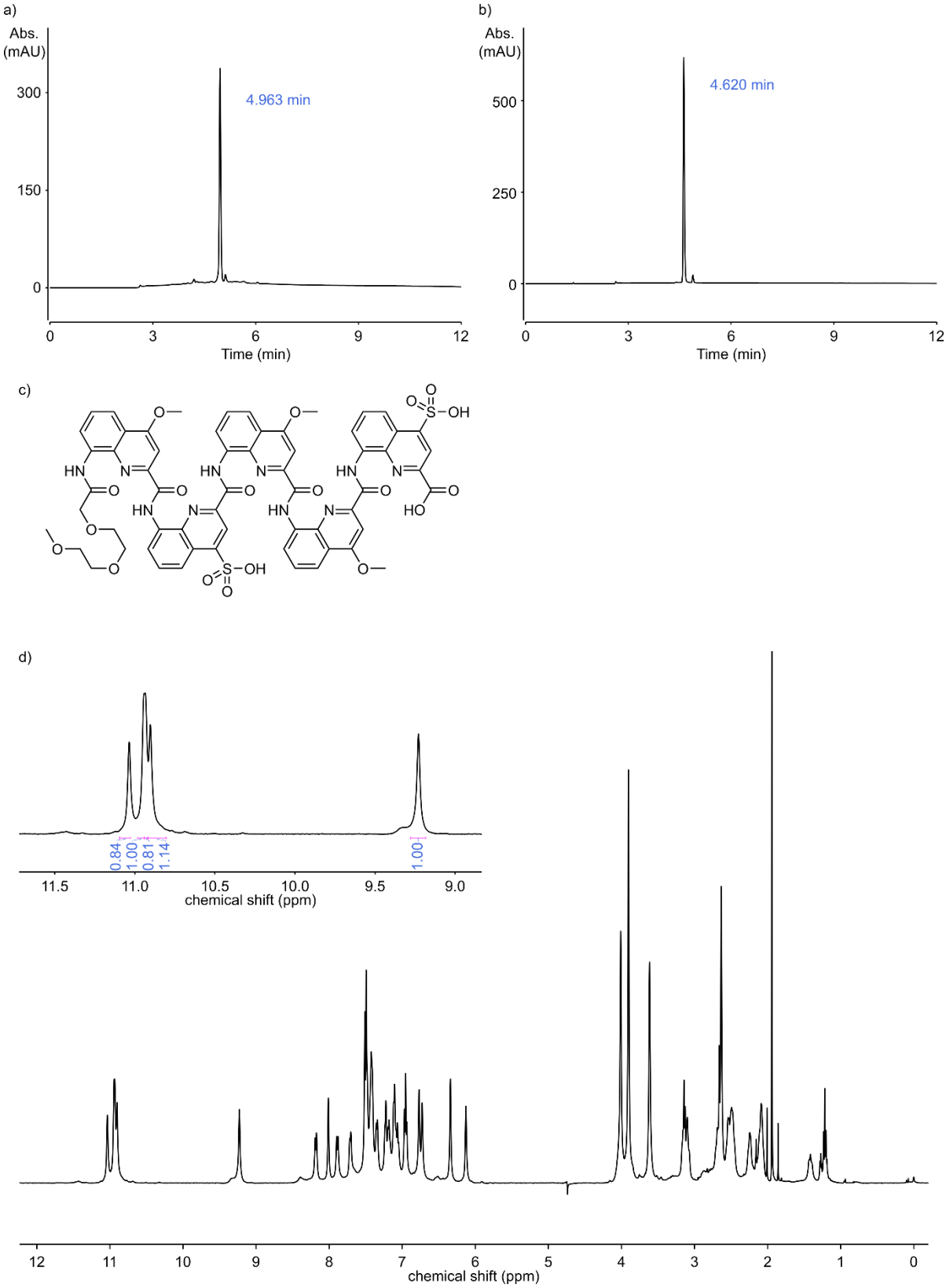


Fig. 111 Analytical data of compound **7**. HPLC chromatograms after cleavage from the resin (a) and after purification (b) (C18, 0–30B, 50 °C; A: 13mM ammonium acetate buffer pH 8.5, B: acetonitrile). c) Chemical structure of compound **7**. d) ^{1}H NMR spectrum (500 MHz, 5.0 mM in 27 mM sodium phosphate buffer pH 7.0).

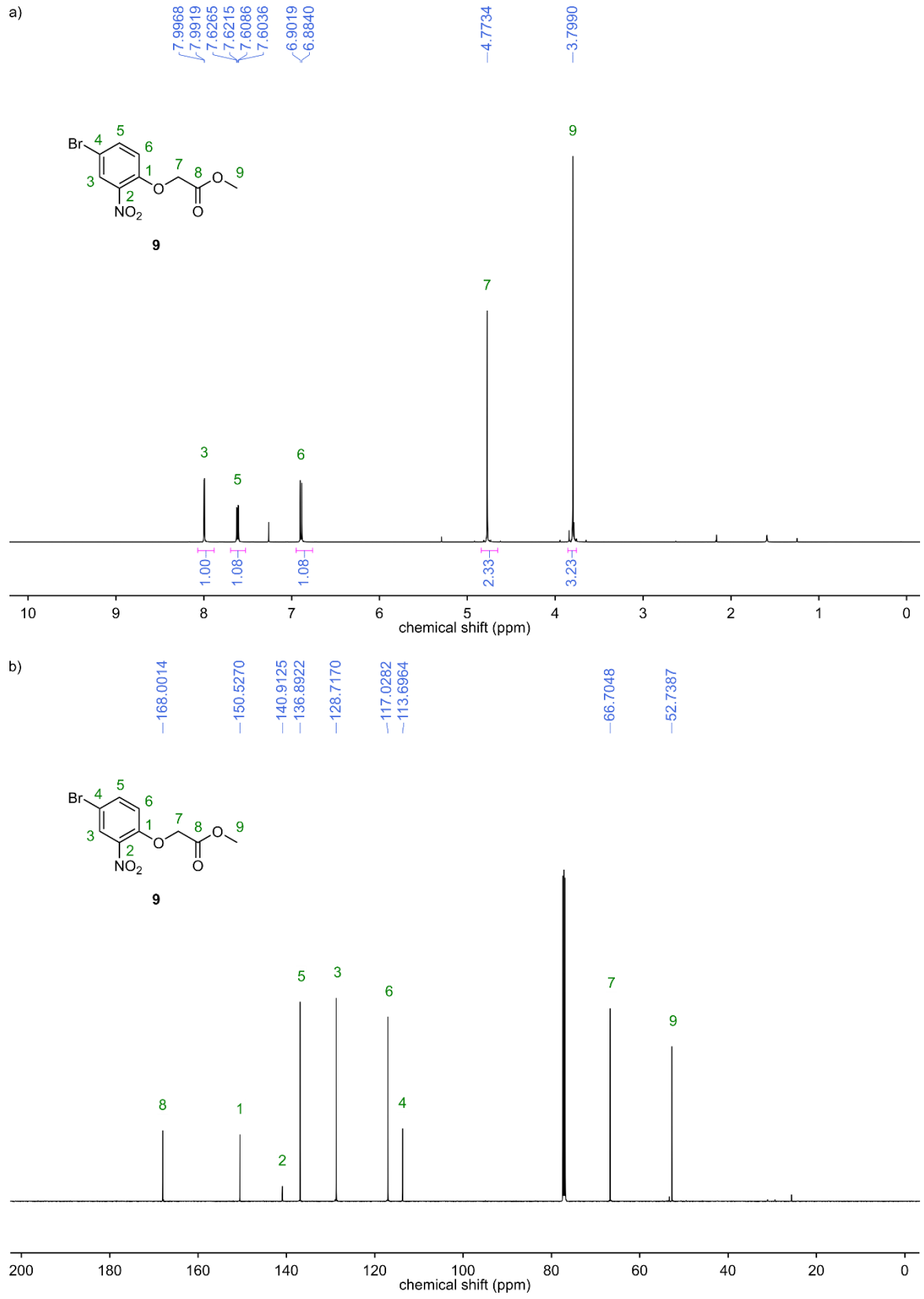


Fig. 112 NMR spectra of compound **9**. a) ^{1}H NMR (500 MHz, $CDCl_3$). b) ^{13}C NMR (126 MHz, $CDCl_3$).

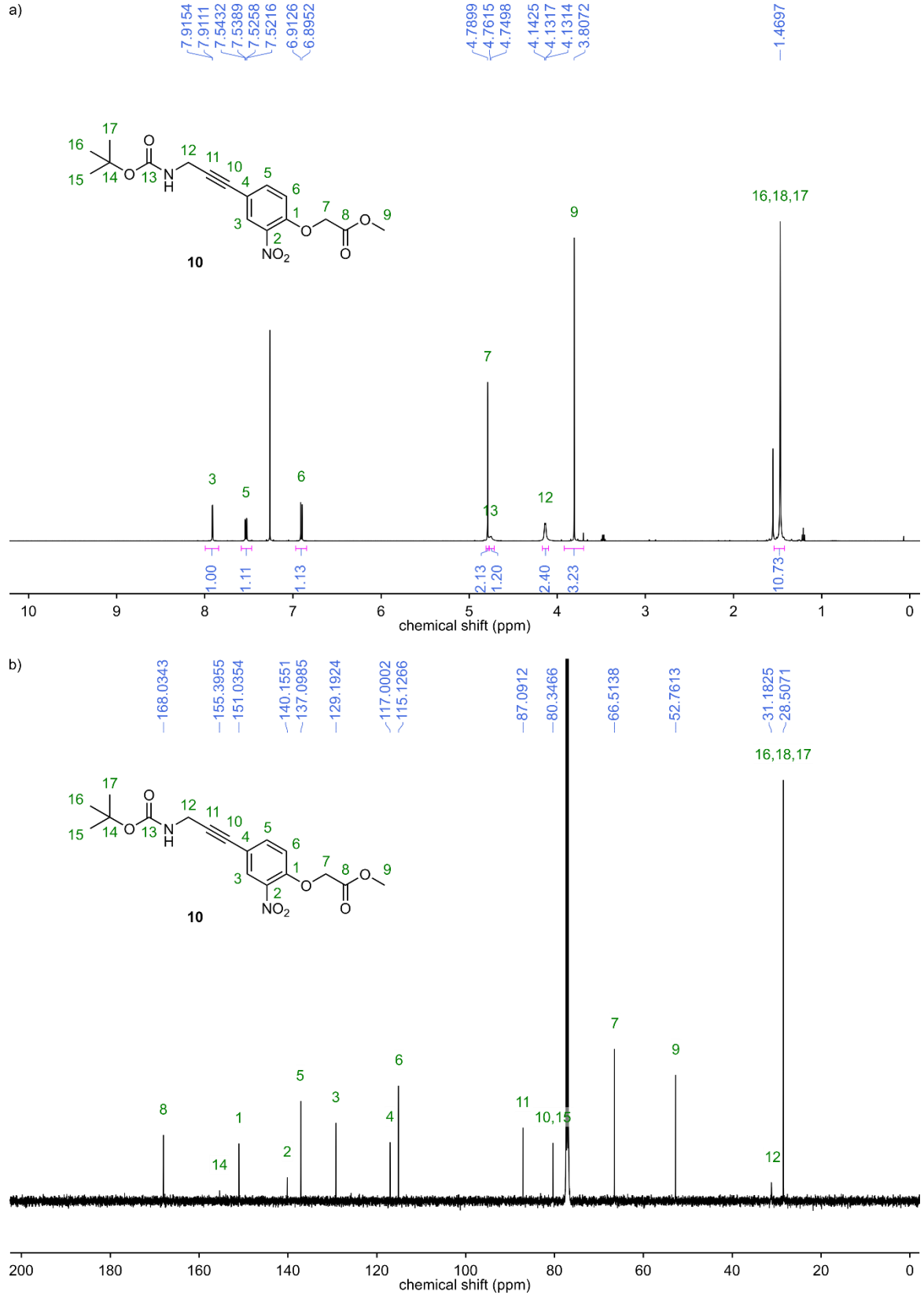


Fig. 113 NMR spectra of compound **10**. a) ^{1}H NMR (500 MHz, $CDCl_3$). b) ^{13}C NMR (126 MHz, $CDCl_3$).

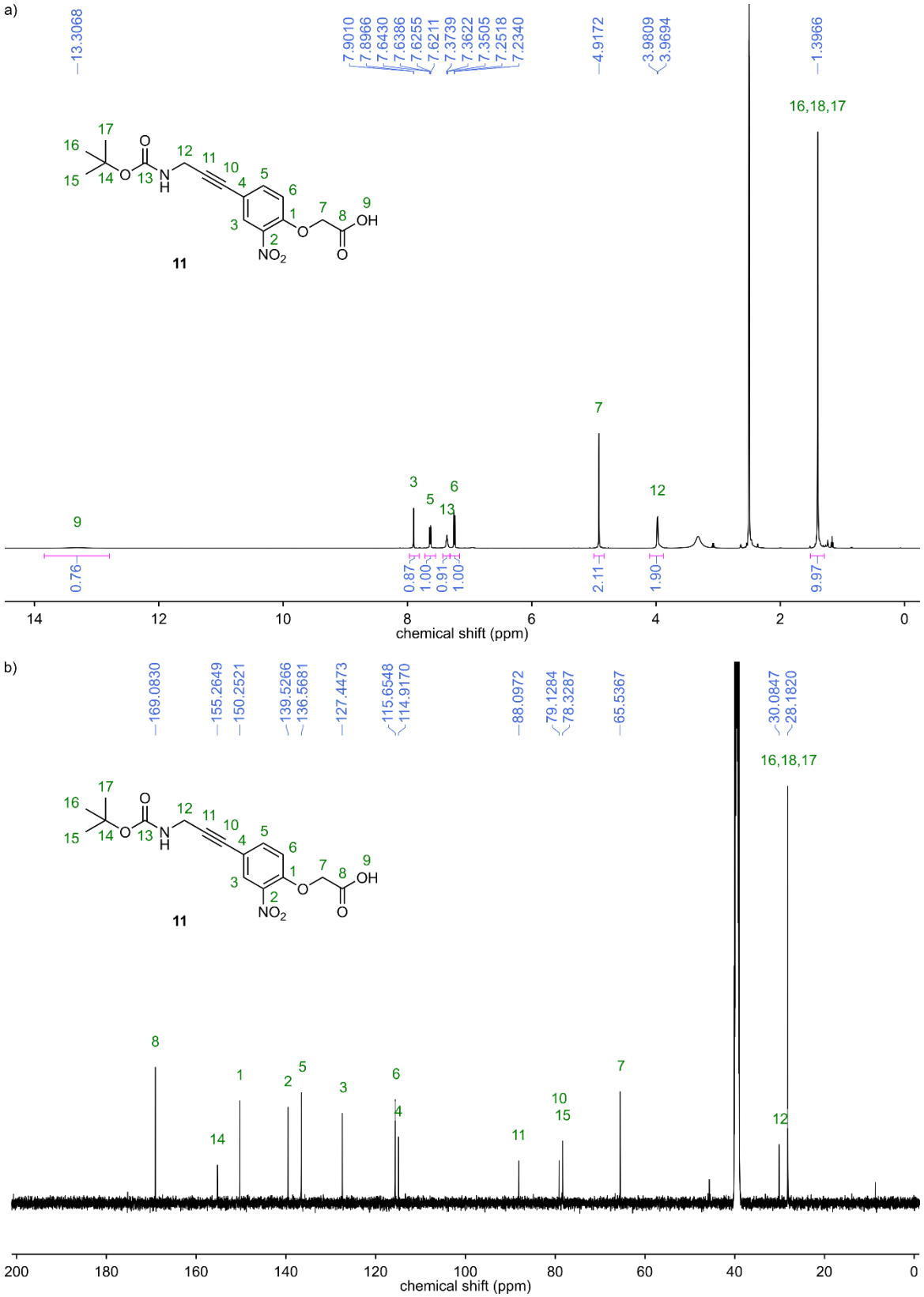


Fig. 114 NMR spectra of compound **11**. a) ^{1}H NMR (500 MHz, DMSO-d_6). b) ^{13}C NMR (126 MHz, DMSO-d_6).

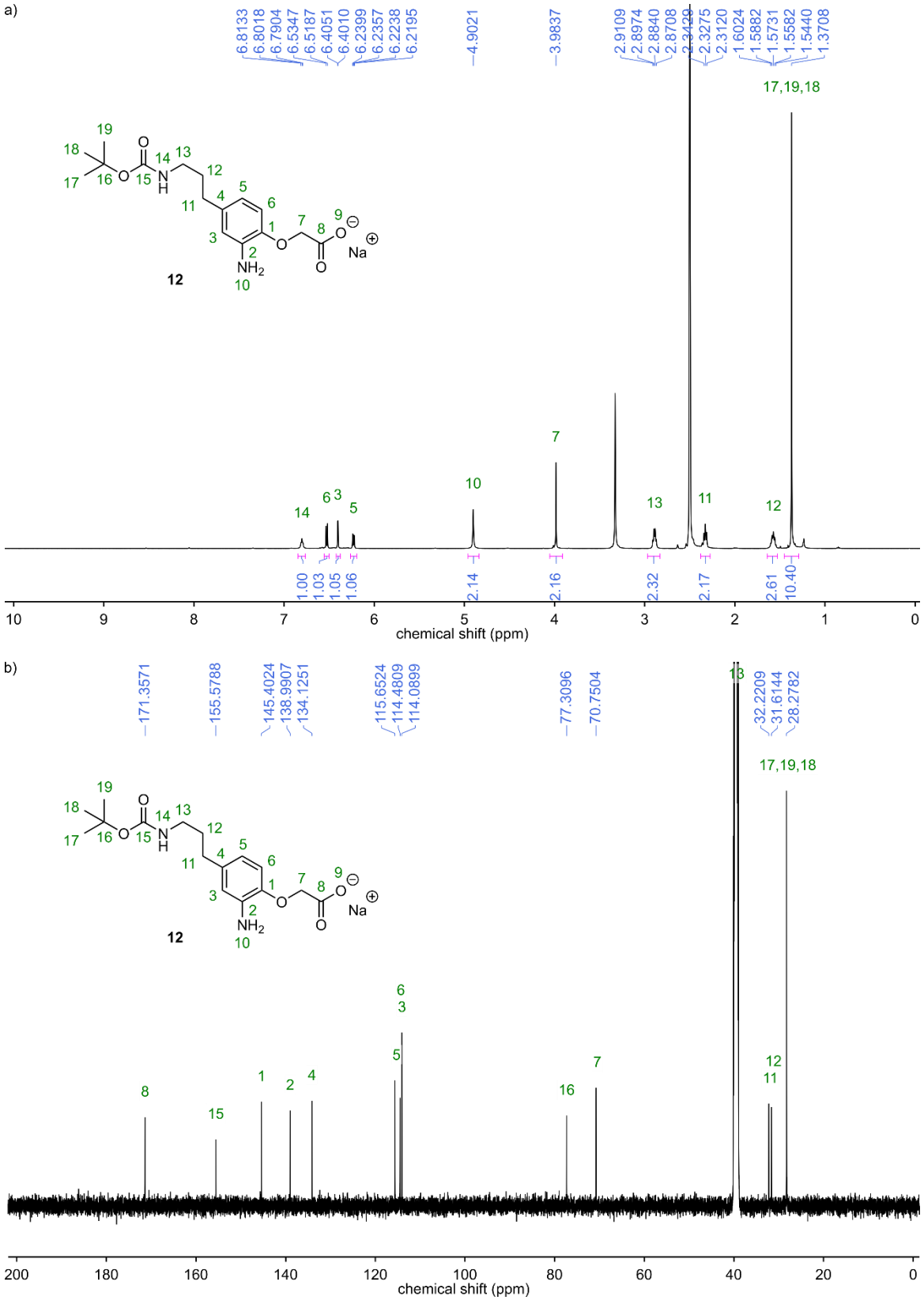


Fig. 115 NMR spectra of compound **12**. a) ^{1}H NMR (500 MHz, DMSO-d$_6$). b) ^{13}C NMR (126 MHz, DMSO-d$_6$).

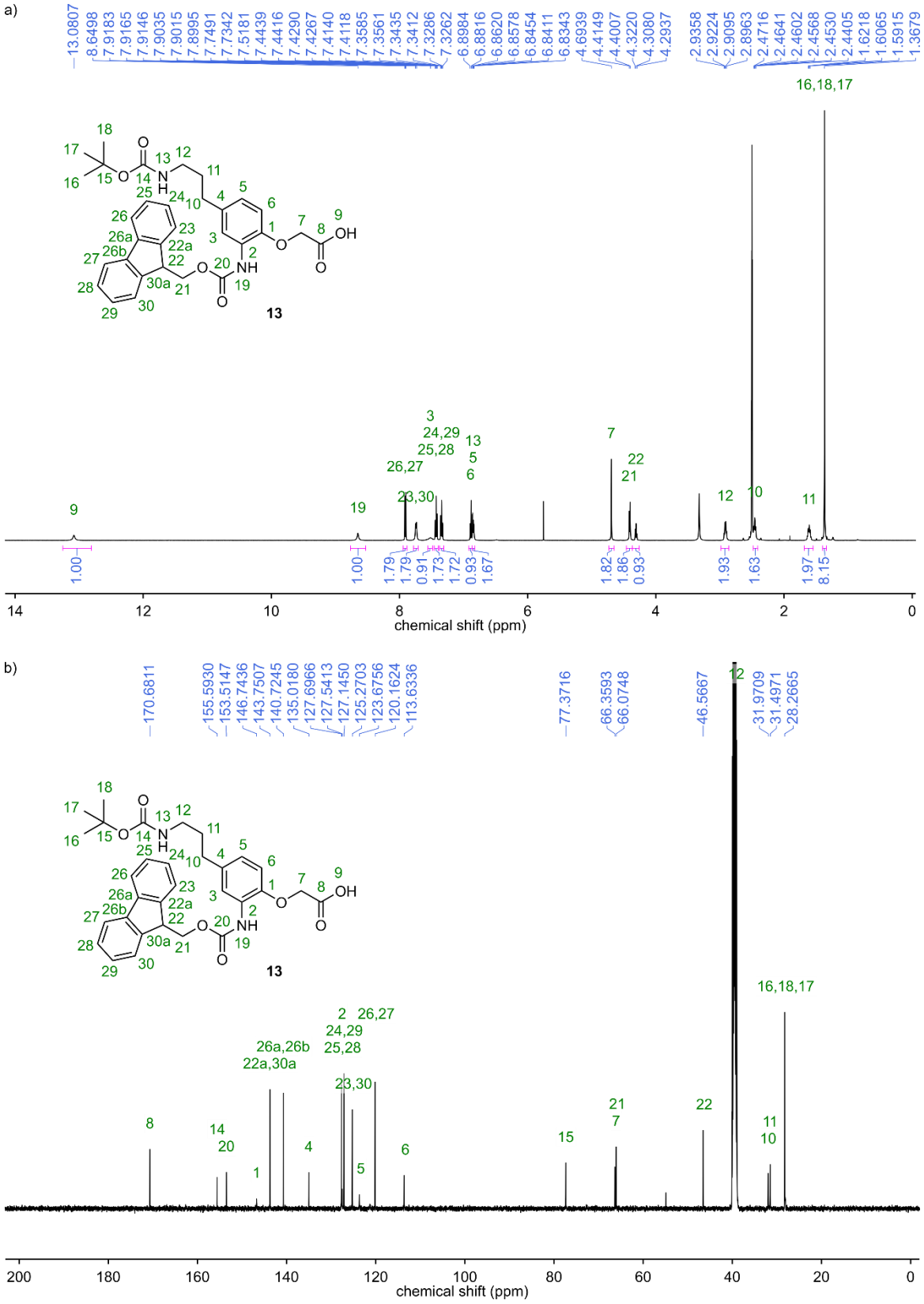


Fig. 116 NMR spectra of compound **13**. a) ^{1}H NMR (500 MHz, DMSO-d$_6$). b) ^{13}C NMR (126 MHz, DMSO-d$_6$).

7 Conclusion and Perspectives

7.1 Conclusions from published/submitted work

Through this work, a new monomer family based on 2-(2-aminophenoxy)acetic acid (B) was developed. The synthesis of B-units is not only quick, cheap, and easy, it also facilitates the introduction of a variety of side chains, especially in position 4 of the aromatic ring. This side chain position is equivalent to position 6 on quinoline units, providing an alternative for these building blocks that are usually more difficult to synthesize. B-monomers can be combined with Q-units without significantly altering the quinoline-helix structure. The strong folding propensity of Q-monomers was illustrated by the finding that a 12-mer consisting of two Q-units and ten B-units still adopts the canonical aromatic helix fold. Conversely, an oligomer made purely of B-monomers does not fold into a well-defined structure in aqueous conditions. Most surprisingly, it was demonstrated that $(PB)_n$-oligomers possess the canonical fold as well, despite lacking any Q-units. These findings show that the aromatic δ-amino acid units B, P and Q are interchangeable in the design of helical aromatic foldamers. The ratio and order of the building blocks dictate the stability and side chain positioning of the resulting oligomer, allowing a fine tuning of helix shape and inversion kinetics. Thus, these new building blocks are especially interesting in foldamer-protein binding applications.

A chiral B-monomer (B^{Rme}) that bears a methyl group and a proton on the asymmetric carbon was also synthesized. It was demonstrated that inserting this moiety in the middle of a quinoline helix fully biases (ratio of >99/1) handedness towards *M*-helicity in the case of (*R*)-chirality. Interestingly, the difference of methyl vs. H is enough to create the necessary energy difference for this bias, illustrating the compactness and strong folding propensity of Q-helices. The bias is even stronger than that of the previously established camphanic acid moiety, which also has the disadvantage of blocking the N-terminus for further functionalization. Additionally, multiple B^{Rme} may be combined in the same helix to further enhance the bias. When placing B^{Rme} at the penultimate position to the N-terminus, full handedness bias was observed as well, while inserting it penultimate to the C-terminus still led to *M*-helicity, albeit only with a ratio of 87/13. This new monomer brings critical advantages compared to the previously used groups for handedness control:[46] it does not block N- or C-terminus for other functionalizations; it does not significantly alter the physical properties such as solubility and shape of the molecules; and it provides increased handedness inversion kinetics by itself, which is crucial for reaching equilibrium of *P*-*M*-conversion in aqueous conditions. Ultimately, B^{Rme} enables designs that rely on both controlled helicity and free N- and C-termini.

In the third work, discrete binding was discovered between the C-terminal cross sections of Q-helices leading to homochiral head-to-head dimers in aqueous media. Slow exchange between monomeric and dimeric species on the NMR timescale was observed, despite the aggregation being mediated only by one hydrogen bond between C-terminal carboxy-functions and aromatic stacking. Under neutral pH, dissociation constants of up to $K_d = 30 \pm 8$ µmol were achieved. Binding strength

depended on the nature of the side chains near the interface and the pH. Aggregation may also be completely prevented by modifying the C-terminus with for example a bulky AiB-residue. Exclusively heteromeric dimers can be produced by using a primary amide terminus on one binding partner and the right concentration- and pH-window. Additionally, handedness bias may be transferred through binding from a chiral helix to an achiral binding partner. This new binding interface may be used to aid crystal growth through dimerization, or ultimately, in the design of bigger self-assembled structures in combination with other binding motives.

7.2 Continuing challenges for foldamer self-assembly

The production of aromatic δ-amino acid foldamers with a significant number of monomers that promote aggregation (e.g., isobutoxy side chains) necessitate long and difficult oligomer syntheses imposing a limit on the number of groups that can be utilized in a given design. Current solid phase synthesis protocols can lead to mixtures with deletions that may be difficult to purify even when capping is applied, resulting in low overall yields. Additionally, monomer building blocks are usually not commercially available and must be synthesized individually beforehand. This labor- and material-intensive process impedes the screening of a large number of designs and side chain compositions, thus one must rely on molecular modeling to “screen” for promising candidates. Especially poor coupling yields were encountered when coupling the Q^{Sul}-unit, the only monomer used without a protecting group (PG) on its side chain. However, this unit also proved to be very useful in providing water solubility and crystal growth ability, features that are especially valuable in designs that involve many floppy hydrophobic side chains.

The synthesis of protected Q^{Sul} esters is not feasible since sulfonyl chlorides of electron poor N-heterocycles are inherently instable.[139] To circumvent this, a new sulfonate monomer based on the B-unit was developed (B^{Sul}). Sulfonates cannot be protected with conventional PGs because they are themselves exceptionally good leaving groups. Thus, a Boc-protected neopentyl derived alcohol (Bnp) based on the literature was introduced as a PG.[140] Neopentyl esters of sulfonic acids are stable towards nucleophiles like piperidine or hydroxide ions as well as acids. To effect deprotection under mild conditions, a self-elimination mechanism is utilized. It is triggered by Boc deprotection of the amine, which, after being exposed to basic conditions, substitutes the sulfonate group via an intramolecular S_N2-reaction generating a pyrrolidine ring and the unprotected sulfonate (Fig. 117). This monomer might serve as a replacement to Q^{Sul} in the future, simplifying the synthesis of water soluble and well-crystallizable oligomers (see section 8.2) for more information).

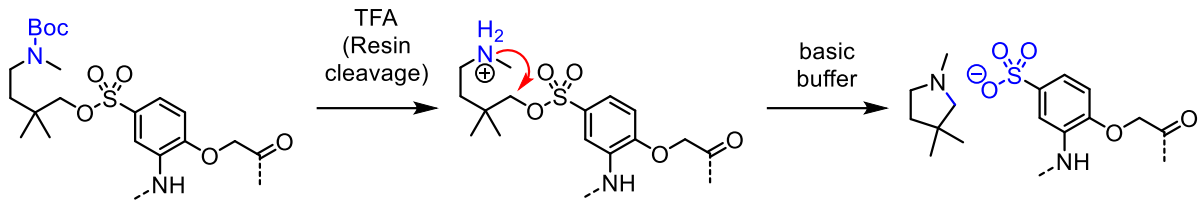


Fig. 117 – Mechanism for the deprotection of the B^{Sul}-monomer. Moieties involved in reactions are highlighted in blue. The intramolecular nucleophilic attack is indicated with a red arrow.

This is very important because X-ray crystallography is often the only way to obtain precise structural information on binding interfaces. Obtaining well-diffracting crystals depends on the intrinsic properties of the oligomer such as the nature the side chains and the overall shape of the molecule. Solid proof of an aggregation phenomenon in solution requires thorough analysis that may be difficult to obtain. Multimeric species observed in mass spectrometry can provide hints on the multiplicity of the aggregates; however, there is no guarantee it reflects the state present in solution. Dilution experiments monitored by NMR spectroscopy are a prime tool to get insights on aggregation behavior. But fast exchange often is observed on the NMR timescale, limiting one's ability to extract detailed information on the binding phenomena. Furthermore, the nature of the buffer, salt concentration and pH may have a significant influence on binding strength or may influence NMR chemical shifts, making it difficult to draw conclusions from simple dilution studies.

Gaining information on the exact binding mode of a given design through X-ray crystallography presents a bottleneck in the workflow of finding and/or optimizing discrete foldamer aggregates. Over the course of this work, the following observations regarding crystal growth and crystal packing have been made: The aromatic cross sections of Q-helices which present big hydrophobic surfaces may lead to unintended aggregation (see the discovery presented in chapter 6). Furthermore, packing forces resulting from stacking of helix cross sections during crystallization may perturb binding interactions present in solution, potentially yielding wrong information on the aggregation behavior. This can be prevented by end-groups that disrupt stacking such as a C-terminal AiB or an N-terminal Camph. However, these groups also often significantly hamper crystal growth (empirical observation), and, unlike for proteins, low resolution is a serious problem for solving X-ray data sets of foldamer molecules because the software tools developed to solve low resolution protein data are not yet applicable here.

7.3 Lessons and ideas for future designs

The basis of most designs intended to form helix bundles in water is a hydrophobic core that should drive aggregation through hydrophobic effects. However, natural binding motifs are not easily transferred to aromatic oligoamide helices. Shape complementarity is the major principle for controlling selectivity in α-peptide coiled coils. For the still closely related β-peptides and β-ureas the structural outcome of aggregates is already challenging to predict just based on their sequence.[7, 73] Additionally, most aromatic oligomers synthesized and investigated so far had a length of ~10–15 units. A 15-mer, for example, possesses three side chains in each of the five distinct side chain arrays at the side of the helix. Three hydrophobic groups such as isobutyl side chains might not be enough to promote aggregation as strongly as anticipated. Therefore, longer sequences are necessary to achieve tight binding. Alternatively, side chains in position 6 of the quinoline ring can be utilized to increase hydrophobicity. Fig. 118 shows models of possible oligomers utilizing this approach. Depending on which 6-position side chains are introduced, different arrays of hydrophobic and salt bridging groups can be generated.

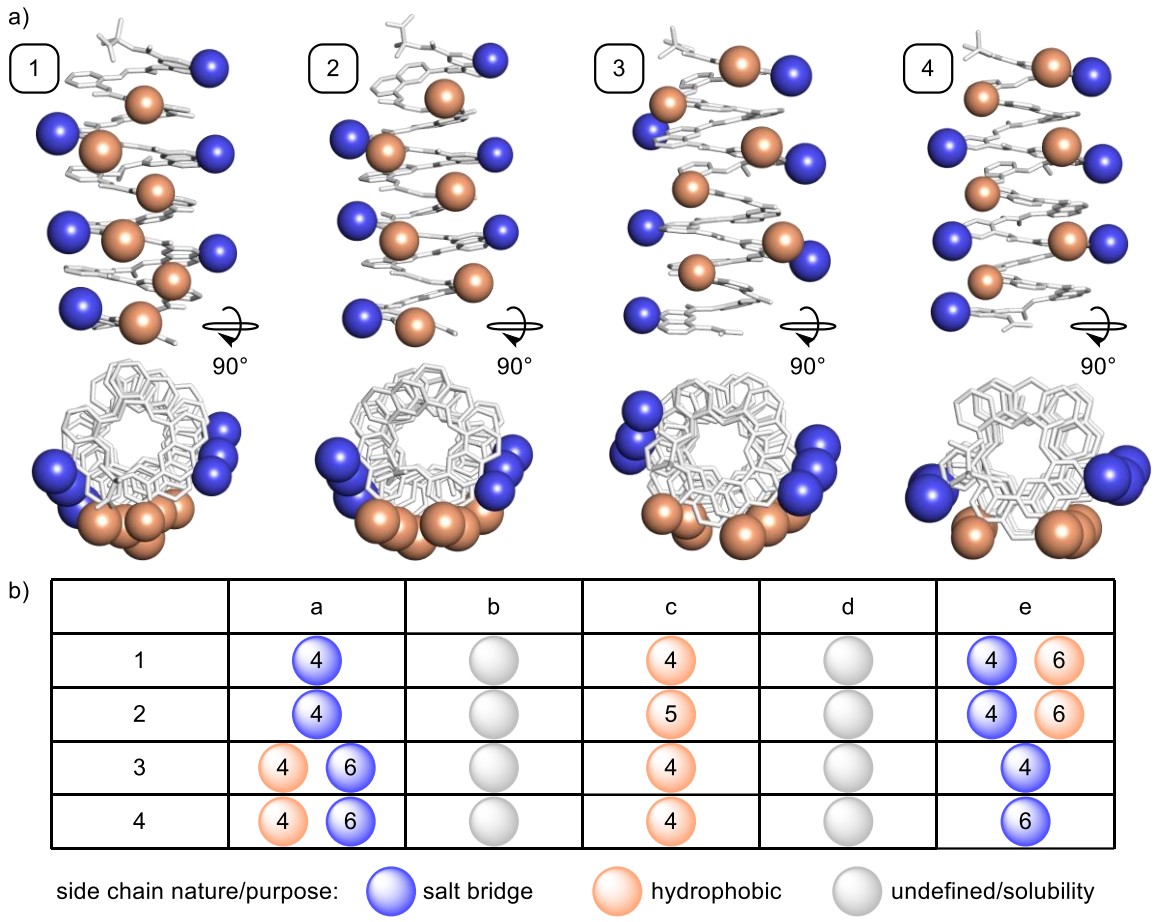


Fig. 118 – Design ideas including 4,6-disubstituted Q-units. a) Molecular models (Maestro, forcefield: OPLS3, method: TNCG, solvent: water) of designs 1–4 possessing two arrays of salt bridging side chains (blue) and two arrays of hydrophobic side chains (orange) utilizing one 4,6-disubstituted Q-monomer per turn. b) Substitution patterns and sequence order of designs 1–4. The numbers in the circles indicate the position of the side chain on a Q-unit.

Another promising candidate for optimization emerged from the “flat” surface designs (Fig. 119). Originally intended to have B-monomers at the binding interface, racemic oligomer **1** was shown to have a rather low dissociation constant of K_d = 65 ± 6 µmol in neutral aqueous phosphate buffer (Fig. 120). Although the crystal structure could confirm the formation of an antiparallel dimer (at least in the solid state) mediated by two kinds of salt bridges, the B-monomers were not at the interface, but on the sides of the bundle. Only one set of signals in the ^{1}H NMR spectra was observed over the measured concentration range meaning exchange between monomeric and aggregated species is fast on the NMR timescale in solution. Therefore, it is not possible to conclude if the solution data represent the bundle observed in the solid state or the average of one or more other aggregation phenomena. The structure is nevertheless a good starting point for further investigations and design improvements. Possible ideas in this direction include elongations of the sequence, an optimization of side chains to obtain better shape complementarity and/or more or stronger salt bridges, and tests whether the B-monomers play a crucial role in the binding process, for example, by increasing the flexibility of the helices.

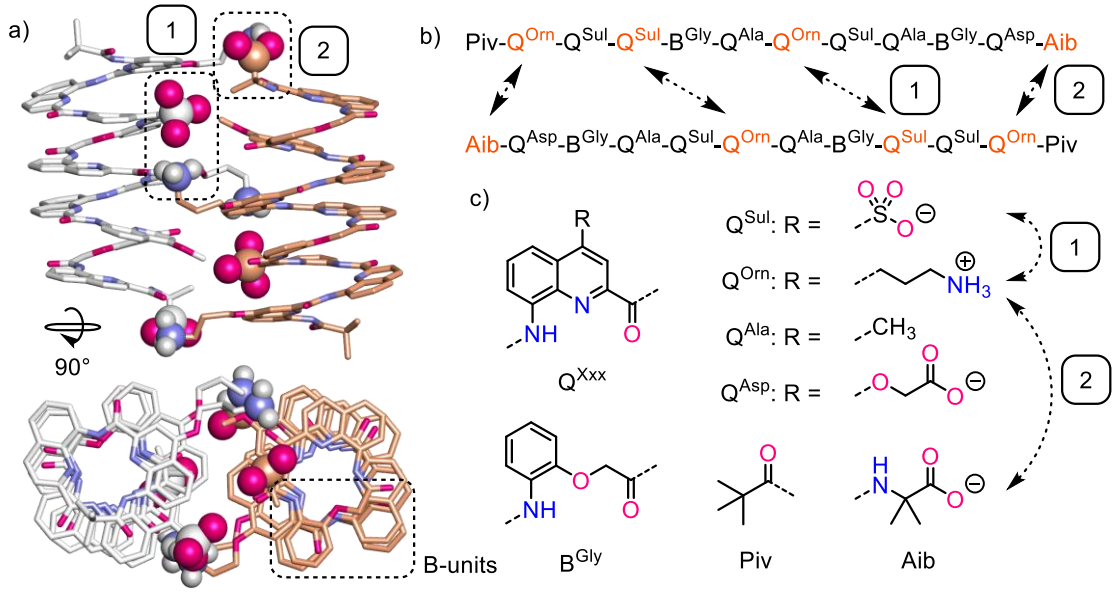
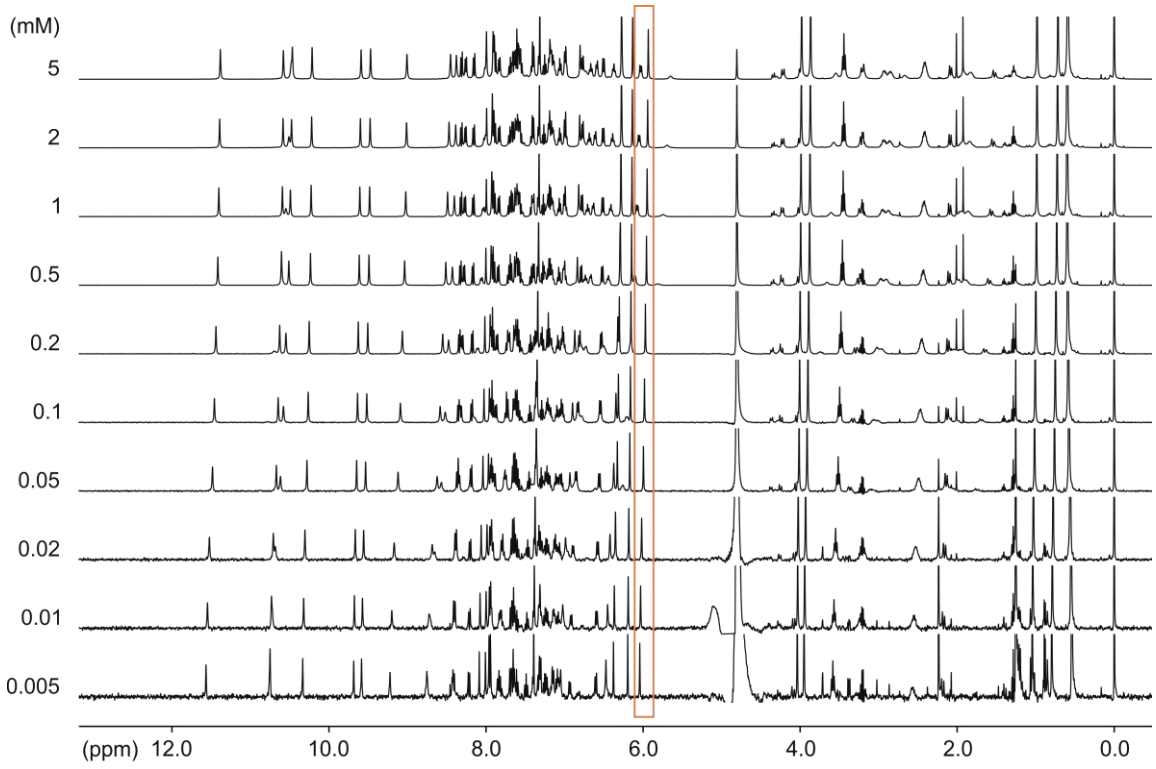


Fig. 119 – Salt bridge mediated dimer with flat interface design. a) Crystal structure of **1**, an aromatic oligoamide helix dimer mediated by two kinds of salt bridges (1) and (2) marked by dashed boxes (only in the side view). The position of B-units is shown with a dashed box in the top view. Individual molecules are shown in different colors and groups involved in hydrogen bonding are shown as balls. b) Sequence of oligomer **1**. Units involved in salt bridges are highlighted in orange and salt bridged are indicated by dashed arrows. c) Formulas of the building blocks utilized in this design (with protonation states as in neutral aqueous buffer). The salt bridges observed in the crystal are indicated by dashed arrows. Oxygen and nitrogen atoms are shown in purple and blue, respectively.

8 Experimental (unpublished)

For general materials and methods see published supplementary information (sections 4.2.1, 5.3.1, 6.2.1).

8.1 Dilution study of oligomer 1



c (mM)	(ppm)
5	5.9335
2	5.9414
1	5.9484
0.5	5.9565
0.2	5.9705
0.1	5.9822
0.05	5.9956
0.02	6.0182
0.01	6.032
0.005	6.0422

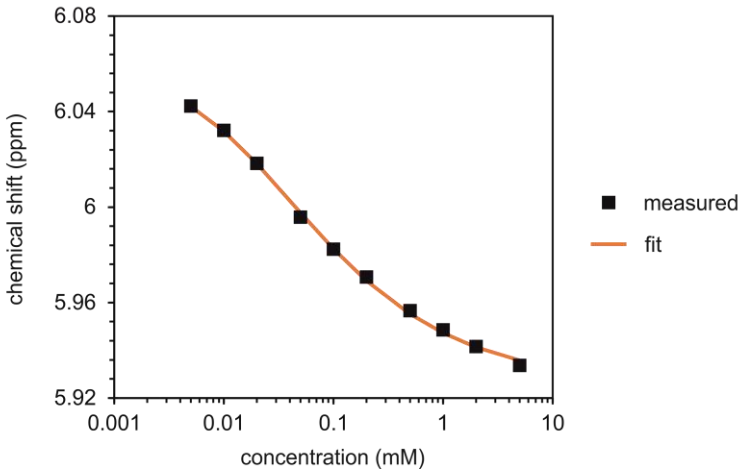
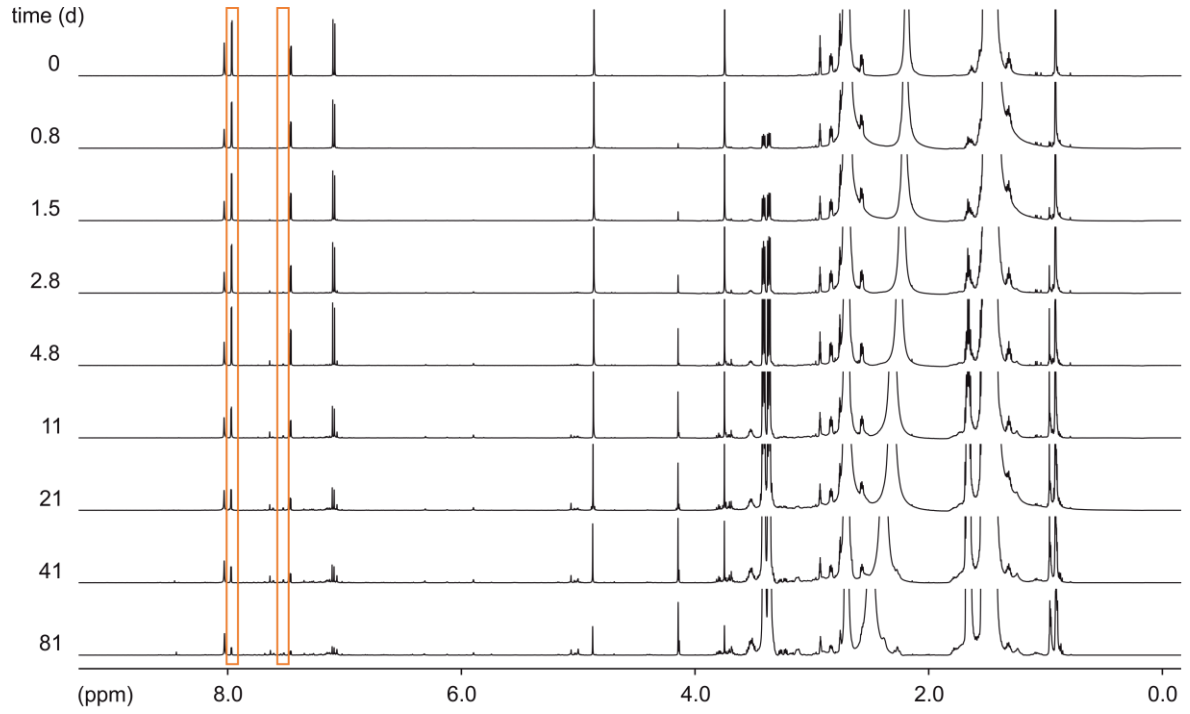
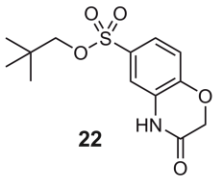


Fig. 120 Dilution study of oligomer **1** monitored by ^{1}H NMR (500 MHz, 27 mM sodium phosphate buffer pH 7.0). The chemical shifts of one signal (indicated by the orange box) were fitted with a one-to-one binding model (software: HypNMR 2008).

8.2 Sulfonate monomer stability and deprotection

The Fmoc deprotection step, which usually involves a 20% piperidine solution in DMF, is crucial in solid phase foldamer synthesis and is repeated many times throughout a given synthesis. Thus, side chain protecting groups must be considerably stable against these conditions. Otherwise, the resin-bound oligomers would participate in a variety of side reactions resulting in poor overall yields. To assess the stability of B^{Sul}-units protected with neopentyl alcohol derived esters, compound **22** was subjected to 20% piperidine in DMF and monitored by ^{1}H NMR over time (Fig. 121). A model compound lacking the additional Boc group was chosen for this study to avoid side reactions and only assess the stability of the sulfonate ester. The observed half-life of 97.3 days should be enough to allow long oligomer syntheses without significant degradation of the protecting group. For example, for 20 deprotection steps a cumulative exposure to piperidine of 200 min (0.14 days) is necessary, meaning that 99.9% of the B^{Sul}-units would still be intact.





time (d)	integral
0	0
0.8	1%
1.5	2%
2.8	2%
4.8	4%
11	9%
21	15%
41	26%
81	43%

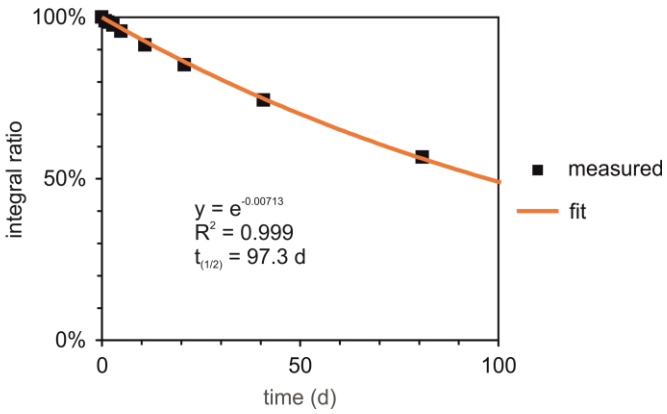


Fig. 121 Stability study of the B^{Sul} model compound **22** against 20% piperidine in DMF monitored by ^{1}H NMR (500 MHz, DMF-d_7). The ratios of selected integrals (indicated by orange boxes) were fitted with an exponential fit.

Next, the deprotection of B^{Sul} was investigated. For this, model compound **23** was subjected to the typical resin cleavage conditions (2.5% H_2O and 2.5% TIPS in TFA) for 2 h. After evaporating most of the solvents, the residue was dissolved in H_2O and the resulting solution was basified with $NaOH_{(aq.)}$ (1 M) and incubated for 30 min. A clean conversion was observed by RP-HPLC, and mass analysis of the reaction solution confirmed the presence of the deprotected sulfonate **24** (Fig. 122). (HRMS $(ESI)^-$ *m/z* calcd. for $C_8H_6NO_5S$: 227.9972 $(M-H)^-$; Found: 228.0820). These results validate the design of the protecting group developed for B^{Sul}. It is considerably stable against piperidine and allows mild cleavage under standard global deprotection conditions. For intramolecular self-elimination, incubation in a neutral or slightly basic aqueous buffer should be sufficient.

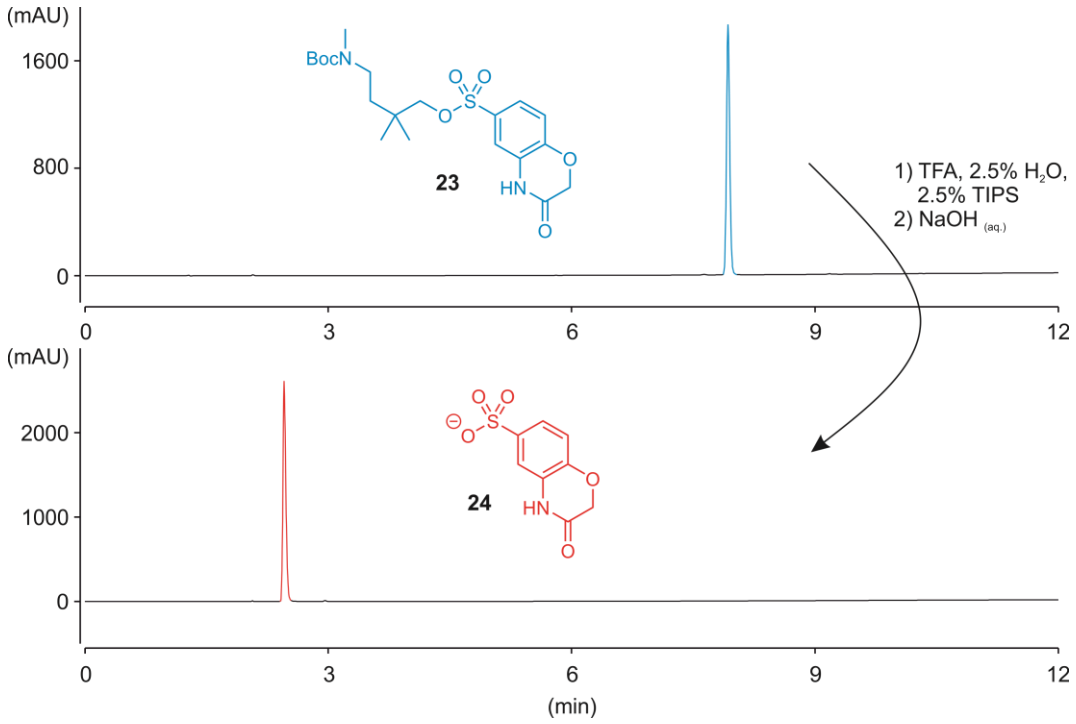
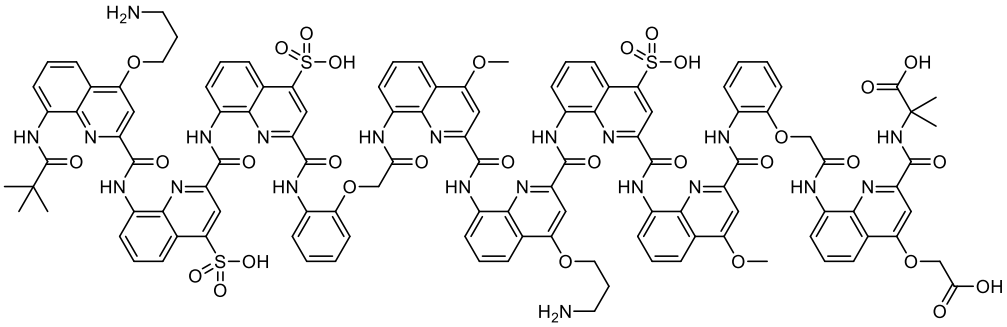


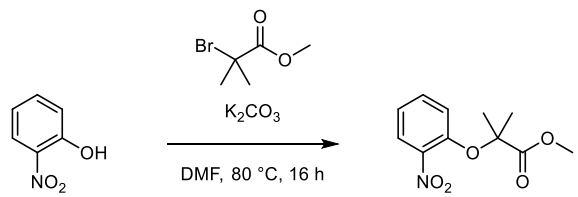
Fig. 122 RP-HPLC chromatograms of the deprotection test of compound **23** (C18, 25 °C, 0–100B, A: H_2O + 0.1% TFA, B: acetonitrile + 0.1% TFA). The peaks of the starting material and product are highlighted in blue and red, respectively.

8.3 Synthetic procedures



1

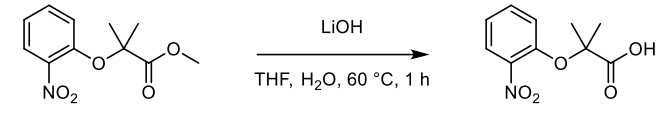
Compound 1: Oligomer **1** was synthesized on ProTide™ resin according to established protocols (0.16 mmol g^{-1}, 30 µmol scale).[26] Initial loading of the resin: 96% (0.15 mmol g^{-1}). The crude product was purified by semi-prep. HPLC (C8, 0–40B, 50 °C, A: 13 mM NH_4OAc buffer pH 8.5, B: acetonitrile) yielding 23.6 mg (9.96 µmol, 33%, HPLC purity: 94.4%) of the title compound as a yellow solid. ($C_{115}H_{99}N_{21}O_{31}S_3$, MW = 2367.35 g mol^{-1}). **^{1}H NMR** (500 MHz, 1 mM in 27 mM sodium phosphate buffer pH 7.0): δ = 11.40 (s, 1H), 10.59 (s, 1H), 10.55 (s, 1H), 10.49 (s, 1H), 10.23 (s, 1H), 9.61 (s, 1H), 9.48 (s, 1H), 9.02 (s, 1H), 8.49 (s, 1H), 8.40 (s, 1H), 8.32 (d, *J* = 9.1 Hz, 1H), 8.27 (d, *J* = 8.2 Hz, 1H), 8.16 (d, *J* = 9.0 Hz, 1H), 8.03 (d, *J* = 9.9 Hz, 1H), 8.00 (s, 1H), 7.95–7.87 (m, 3H), 7.84 (d, *J* = 8.9 Hz, 1H), 7.71 (d, *J* = 8.5 Hz, 1H), 7.69–7.65 (m, 1H), 7.64–7.60 (m, 1H), 7.59 (d, *J* = 8.0 Hz, 1H), 7.56 (d, *J* = 8.6 Hz, 1H), 7.53 (d, *J* = 4.5 Hz, 1H), 7.45–7.37 (m, 2H), 7.34 (s, 1H), 7.32 (s, 1H), 7.27 (t, *J* = 8.3 Hz, 1H), 7.24–7.15 (m, 2H), 7.14 (d, *J* = 8.0 Hz, 1H), 7.07 (t, *J* = 7.9 Hz, 1H), 7.03–6.95 (m, 2H), 6.82 (s, 1H), 6.78 (d, *J* = 8.3 Hz, 1H), 6.71 (t, *J* = 8.4 Hz, 1H), 6.64 (d, *J* = 8.6 Hz, 1H), 6.51 (d, *J* = 9.0 Hz, 1H), 6.41 (t, *J* = 8.2 Hz, 1H), 6.28 (s, 2H), 6.14 (s, 1H), 6.07 (d, *J* = 9.5 Hz, 1H), 5.95 (s, 1H), 5.75 (s, 1H), 4.23 (d, *J* = 14.2 Hz, 1H), 4.02 (s, 1H), 3.99 (s, 3H), 3.87 (s, 3H), 3.61 (s, 1H), 3.45 (t, *J* = 8.4 Hz, 2H), 3.30–3.14 (m, 1H), 3.05–2.80 (m, 2H), 2.53–2.33 (m, 2H), 2.10 (d, *J* = 12.7 Hz, 1H), 2.04–1.77 (m, 3H), 1.56 (d, *J* = 15.6 Hz, 1H), 1.29 (t, *J* = 7.8 Hz, 1H), 0.99 (s, 3H), 0.73 (s, 3H), 0.60 (s, 9H). **HRMS** $(ESI)^-$ *m/z* calcd. for $C_{115}H_{98}N_{21}O_{31}S_3$: 2364.5905 $(M-H)^-$; Found: 2364.6137.



2

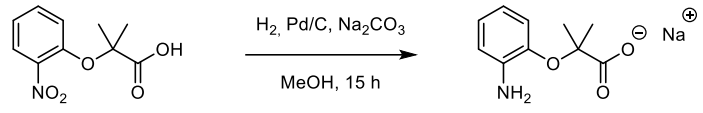
Compound 2: 2-Nitrophenol (5.00 g, 35.9 mmol, 1.0 equiv.), methyl 2-bromo-2-methylpropanoate (9.30 ml, 71.9 mmol, 2.0 equiv.) and K_2CO_3 (5.46 g, 39.5 mmol, 1.1 equiv.) were suspended in anhydrous DMF (50 ml). The reaction mixture was stirred at 80 °C for 16 h. After diluting the mixture with EtOAc, it was washed with H_2O (2×), with brine (1×), dried over $MgSO_4$ and solvents

were evaporated *in vacuo*. Purification by column chromatography (SiO_2, CyHex/EtOAc 9:1 → 8:2) yielded the title compound (7.04 g, 29.4 mmol, 85%) as a white solid. ($C_{11}H_{13}NO_5$, MW = 239.23 g mol^{-1}). $\mathbf{R_f}$ (CyHex/EtOAc 8:2) = 0.36. **^{1}H NMR** (500 MHz, $CDCl_3$): δ = 7.75 (dd, *J* =8.1 Hz, 1.7 Hz, 1H, C3-H), 7.43 (ddd, *J* = 8.5 Hz, 7.4 Hz, 1.7 Hz, 1H, C5-H), 7.08 (ddd, *J* = 8.4 Hz, 7.5 Hz, 1.2 Hz, 1H, C4-H), 6.94 (dd, *J* = 8.5 Hz, 1.2 Hz, 1H, C6-H), 3.79 (s, 3H, C11-H), 1.65 (s, 6H, C8-H, C9-H). **^{13}C NMR** (500 MHz, $CDCl_3$): δ = 174.2 (C10), 148.9 (C1), 143.3 (C2), 133.1 (C5), 125.4 (C3), 122.2 (C4), 120.1 (C6), 81.6 (C7), 52.9 (C11), 25.2 (C8, C9). **HRMS** (ESI)$^+$ *m/z* calcd. for $C_{11}H_{14}NO_5$: 240.0866 $(M+H)^+$; Found: 240.0870. (Modified procedure, analytical data is in line with literature).[141]



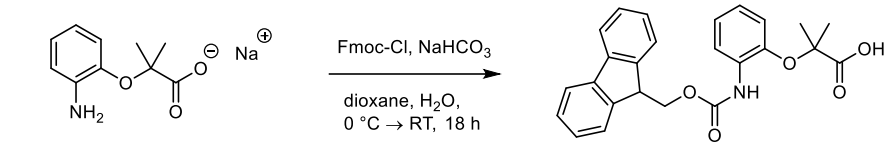
2 **3**

Compound 3: Compound **2** (7.04 g, 29.4 mmol, 1.0 equiv.) was dissolved in THF (80 ml). After the addition of LiOH (1.06 g, 44.1 mmol, 1.5 equiv.) in H_2O (20 ml), the solution was stirred at 60 °C for 1 h. The mixture was acidified to pH ~2 using HCl $_{(aq.)}$ (1 M), extracted with DCM (3×) and dried over $MgSO_4$. Evaporation of the solvents *in vacuo* yielded the title compound (6.60 g, 29.3 mmol, quant.) as a white solid. ($C_{10}H_{11}NO_5$, MW 225.20 g mol^{-1}). $\mathbf{R_f}$ (DCM/MeOH 97/3 + 0.5% AcOH) = 0.43. **^{1}H NMR** (500 MHz, DMSO-d_6): δ = 13.37 (s, 1H, O11-H), 7.83 (dd, *J* = 8.1 Hz, 1.7 Hz, 1H, C3-H), 7.59 (ddd, *J* = 8.5 Hz, 7.4 Hz, 1.7 Hz, 1H, C5-H), 7.16 (ddd, *J* = 8.3 Hz, 7.4 Hz, 1.1 Hz, 1H, C4-H), 7.05 (dd, *J* = 8.5 Hz, 1.1 Hz, 1H, C6-H), 1.55 (s, 6H, C-H, C-H). **^{13}C NMR** (500 MHz, DMSO-d_6): δ = 174.1 (C10), 147.8 (C1), 142.2 (C2), 133.3 (C5), 124.7 (C3), 121.8 (C4), 119.0 (C6), 80.6 (C7), 24.8 (C8, C9). **HRMS** (ESI)$^-$ *m/z* calcd. for $C_{10}H_{10}NO_5$: 224.0564 $(M-H)^-$; Found: 224.0563.



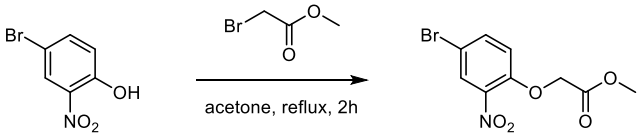
3 **4**

Compound 4: Compound **3** (6.60 g, 29.3 mmol, 1.0 equiv.) and Na_2CO_3 (3.11 g, 29.3 mmol, 1.0 equiv.) were dissolved in methanol (600 ml). The solution was quickly degassed by vacuum-N_2 cycles (3×), then, Pd/C (660 mg, 10 wt. % loading) was added and the N_2 atmosphere was replaced by H_2. After stirring for 15 h the reaction mixture was filtered over celite©, and solvents were removed *in vacuo*. The crude product (6.37 g, 29.3 mmol, quant.) was obtained as a slightly brown solid and was used in the next step without further purification. ($C_{10}H_{12}NO_3Na$, MW = 217.20 g mol^{-1}). **^{1}H NMR** (500 MHz, DMSO-d_6): δ = 6.73 (d, *J* = 7.95 Hz, 1H, C6-H), 6.61 (t, *J* = 7.47 Hz, 1H, C4-H), 6.52 (d, *J* = 7.68 Hz, 1H, C3-H), 6.31 (t, *J* = 7.53 Hz, 1H, C5-H), 5.37 (s, 2H, N12-H), 1.28 (s, 6H, C8-H, C9-H). **^{13}C NMR** (500 MHz, DMSO-d_6): δ = 177.3 (C10), 143.0 (C1), 142.3 (C2), 121.8 (C4), 120.7 (C6), 114.8 (C5), 113.8 (C3), 81.3 (C7), 26.4 (C8, C9). **HRMS** (EI)$^+$ *m/z* calcd. for $C_{10}H_{12}NO_3$: 195.0890 $(M)^+$; Found: 195.0869.



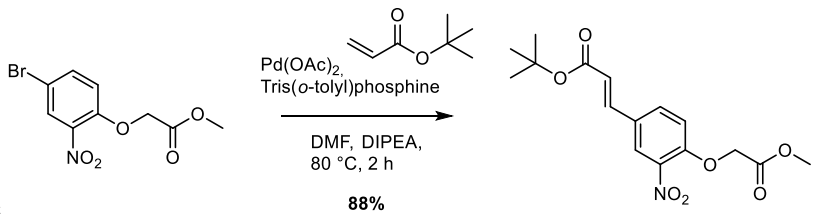
4 **84%** **5**

Compound 5: Compound **4** (6.37 g, 29.3 mmol, 1.0 equiv.) and $NaHCO_3$ (12.3 g, 147 mmol, 5.0 equiv.) were dissolved in H_2O (300 ml). After the suspension was cooled to 0 °C, Fmoc-Cl (9.86 g, 38.1 mmol, 1.3 equiv.) in 1,4-dioxane (300 ml) were added over 1 h. Then, the reaction mixture was stirred at 0 °C for 1 h and at RT for 16 h. The mixture was acidified to pH ~2 using $HCl_{(aq.)}$ (1 M), extracted with DCM (3×), dried over $MgSO_4$ and solvents were removed *in vacuo*. Purification by column chromatography (SiO_2, CyHex/EtOAc 7:3 + 1% DIPEA → CyHex/EtOAc 7:3 → CyHex/EtOAc 7:3 + 1% AcOH → CyHex/EtOAc 1:1 + 1% AcOH) yielded the title compound (10.3 g, 24.6 mmol, 84%) as a white solid. ($C_{25}H_{23}NO_5$, MW = 417.46 g mol^{-1}). $\mathbf{R_f}$ (CyHex/EtOAc 7:3 + 1% AcOH) = 0.29. **^{1}H NMR** (500 MHz, DMSO-d_6): δ = 13.33 (s, 1H, O11-H), 8.95 (s, 1H C12-H), 7.91 (d, *J* = 7.6 Hz, 2H C19-H, C20-H), 7.75 (d, *J* = 7.5 Hz, 2H C16-H, C23-H), 7.69 (s, 1H C3-H), 7.43 (td, *J* = 7.5 Hz, 1.1 Hz, 2H C18-H, C21-H), 7.34 (td, *J* = 7.5 Hz, 1.2 Hz, 2H C17-H, C22-H), 7.03–6.95 (m, 2H, C4-H, C5-H), 6.91 (dd, *J* = 7.4 Hz, 2.2 Hz, 1H, C6-H), 4.39 (d, *J* = 7.4 Hz, 2H, C14-H), 4.31 (t, *J* = 7.3 Hz, 1H, C15-H), 1.50 (s, 6H, C8-H, C9-H). **^{13}C NMR** (500 MHz, DMSO-d_6): δ = 175.9 (C10), 153.5 (C13), 145.2 (C1), 143.7 (C15a, C23a), 140.7 (C19a, C19b), 130.7 (C2), 127.7 (C18, C21), 127.1 (C17, C22), 125.3 (C16, C23), 123.4 (C5), 122.6 (C4), 121.4 (C3), 120.2 (C19, C20), 119.3 (C6), 79.8 (C7), 66.2 (C14), 46.5 (C15), 24.8 (C8, C9). **HRMS** (EI)$^+$ *m/z* calcd. for $C_{25}H_{23}NO_5$: 417.1571 $(M)^+$; Found: 417.1576.



quant. **6**

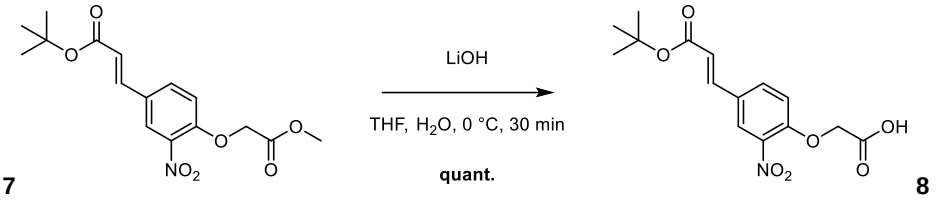
For the synthesis of compound **6** see section 6.2.3).



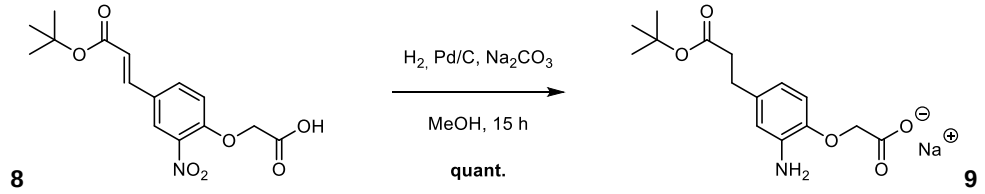
6 **88%** **7**

Compound 7: Compound **6** (2 g, 6.89 mmol, 1.0 equiv.), $Pd(OAc)_2$ (39.0 mg, 172 µmol, 2.5 mol%) and Tris(*o*-tolyl)phosphine (210 mg, 689 µmol, 0.1 equiv.) were dissolved in anhydrous DMF (7 ml) and DIPEA (7 ml) under N_2 atmosphere. Then, *tert*-butyl acrylate (1.52 ml, 10.3 mmol, 1.5 equiv.) was added and the reaction mixture was stirred at 80 °C for 2 h. After letting it cool to RT, H_2O was added and the mixture was extracted with DCM (3×). The combined organic phases were dried over $MgSO_4$ and solvents were removed *in vacuo*. Purification by column

chromatography (SiO_2, CyHex/EtOAc 8:2) yielded the title compound (2.57 g, 7.62 mmol, 88%) as a white solid. ($C_{16}H_{19}NO_7$, MW = 337.33 g mol^{-1}). $\mathbf{R_f}$ (CyHex/EtOAc 7:3) = 0.24. **^{1}H NMR** (400 MHz, $CDCl_3$): δ = 8.02 (d, *J* = 2.6 Hz, 1H, C3-H), 7.63 (dd, *J* = 9.0 Hz, 2.1 Hz, 1H, C5-H), 7.49 (d, *J* = 16.0 Hz, 1H, C10-H), 6.97 (d, *J* = 8.8 Hz, 1H, C6-H), 6.33 (d, *J* = 16.0 Hz, 1H, C11-H), 4.82 (s, 2H, C7-H), 3.81 (s, 3H, C9-H), 1.53 (s, 9H, C14-H, C15-H, C16-H). **^{13}C NMR** (101 MHz, $CDCl_3$): δ = 168.0 (C8), 165.7 (C12), 152.1 (C1), 140.6 (C2), 140.3 (C10), 133.2 (C5), 129.0 (C4), 125.1 (C3), 121.8 (C11), 115.3 (C6), 81.1 (C13), 66.4 (C7), 52.8 (C9), 28.3 (C14, C15, C16). **HRMS** (ESI)$^+$ *m/z* calcd. for $C_{16}H_{19}NO_7Na$: 360.1054 $(M+Na)^+$; Found: 360.1055.

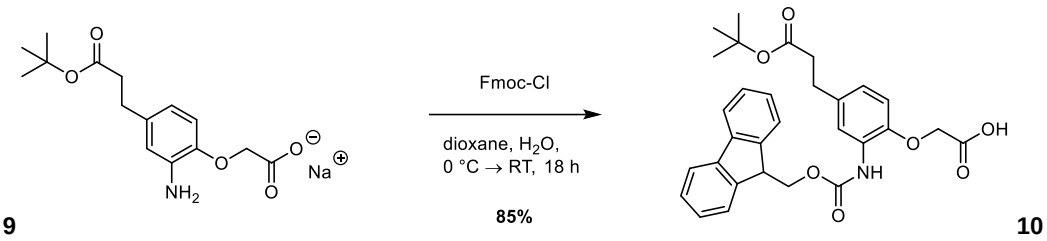


Compound 8: Compound **7** (2.57 g, 7.62 mmol, 1.0 equiv.) was dissolved in THF (200 ml). LiOH (182 mg, 7.62 mmol, 1.0 equiv.) in H_2O (50 ml) was added at 0 °C and the resulting solution was stirred at 0 °C for 30 min. Then, citric acid $_{(aq.)}$ (1 M, 7.62 ml, 7.62 mmol, 1.0 equiv.) were added and the mixture was extracted with DCM (3×). The combined organic phases were dried over $MgSO_4$ and evaporated *in vacuo* yielding the title compound (2.46 g, 7.62 mmol, quant.) as a white solid. ($C_{15}H_{17}NO_7$, MW = 323.30 g mol^{-1}). **^{1}H NMR** (500 MHz, DMSO-d_6): δ =13.26 (s, 1H, O9-H), 8.25 (d, *J* = 2.2 Hz, 1H, C3-H), 7.97 (dd, *J* = 8.8 Hz, 2.3 Hz, 1H, C5-H), 7.55 (d, *J* = 16.0 Hz, 1H, C10-H), 7.29 (d, *J* = 8.9 Hz, 1H, C6-H), 6.57 (d, *J* = 16.0 Hz, 1H, C11-H), 4.97 (s, 2H, C7-H), 1.48 (s, 9H, C14-H, C15-H, C16-H). **^{13}C NMR** (500 MHz, DMSO-d_6): δ =169.1 (C8), 165.4 (C12), 151.2 (C1), 141.0 (C10), 140.0 (C2), 133.2 (C5), 127.5 (C4), 124.5 (C3), 120.5 (C11), 115.4 (C6), 80.1 (C13), 65.4 (C7), 27.8 (C14, C15, C16). **HRMS** (ESI)$^-$ *m/z* calcd. for $C_{15}H_{16}NO_7$: 322.0932 $(M-H)^-$; Found: 322.0932.

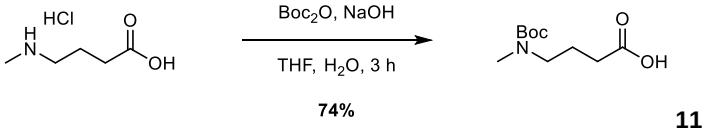


Compound 9: Compound **8** (2.46 g, 7.62 mmol, 1.0 equiv.) and Na_2CO_3 (808 mg, 7.62 mmol, 1.0 equiv.) were suspended in methanol (250 ml). The solution was quickly degassed by vacuum-N_2 cycles (3×), then, Pd/C (246 mg 10 wt. % loading) was added and the N_2 atmosphere was replaced by H_2. After stirring for 15 h the reaction mixture was filtered over celite©, and solvents were removed *in vacuo*. The crude product (2.42 g, 7.62 mmol, quant.) was obtained as a slightly brown solid and was used in the next step without further purification. ($C_{15}H_{20}NO_5Na$, MW = 317.32 g mol^{-1}). **^{1}H NMR** (500 MHz, DMSO-d_6): δ = 6.52 (d, *J* = 8.07 Hz, 1H, C6-H), 6.42 (d, *J* = 2.10 Hz, 1H, C3-H), 6.25 (dd, *J* = 8.14 Hz, 2.14 Hz, 1H, C5-H), 4.91 (s, 2H, N17-H), 3.98 (s, 2H, C7-H), 2.59 (t, *J* = 7.65 Hz, 2H, C10-H), 2.38 (t, *J* = 7.68 Hz, 2H, C11-H), 1.37 (s, 9H, C14-H, C15-

H, C16-H). **^{13}C NMR** (500 MHz, DMSO-d_6): δ =172.2 (C12), 171.7 (C8), 146.1 (C1), 139.5 (C2), 133.4 (C4), 116.0 (C5), 114.9 (C6), 114.5 (C3), 80.0 (C13), 71.2 (C7), 37.3 (C11), 30.7 (C10), 28.2 (C14, C15, C16). **HRMS** (ESI)$^-$ *m/z* calcd. for $C_{15}H_{20}NO_5$: 294.1347 (M-H)$^-$; Found: 294.1346.

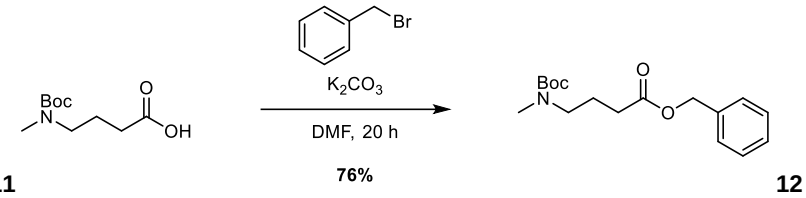


Compound 10: Compound **9** (2.52 g, 7.62 mmol, 1.0 equiv.) and $NaHCO_3$ (3.20 g, 38.1 mmol, 5.0 equiv.) were dissolved in H_2O (150 ml). After cooling the suspension to 0 °C, Fmoc-Cl (2.56 g, 9.90 mmol, 1.3 equiv.) in 1,4-dioxane (150 ml) was added over 1 h. The resulting mixture was stirred at 0 °C for 1 h, then at RT for 16 h. After adjusting to pH ~2 using citric acid $_{(aq.)}$ (1 M, ~50 ml), the suspension was extracted with DCM (3×). The combined organic phases were dried over $MgSO_4$ and solvents were removed *in vacuo*. Purification of the residue by column chromatography (SiO_2, CyHex/EtOAc 1:1 + 1% DIPEA → CyHex/EtOAc 1:1 → CyHex/EtOAc 1:1 + 1% AcOH → EtOAc + 1% AcOH) yielded the title compound (3.37 g, 6.51 mmol, 85%) as a white solid. ($C_{30}H_{31}NO_7$, MW = 517.58 g mol^{-1}). $\mathbf{R_f}$ (EtOAc + 2% AcOH) = 0.23. **^{1}H NMR** (500 MHz, DMSO-d_6): δ = 13.08 (s, 1H, O9-H), 8.65 (s, 1H, N17-H), 7.91 (d, *J* = 7.5 Hz, 2H, C24-H, C25-H), 7.74 (d, *J* = 7.5 Hz, 2H, C21-H, C28-H), 7.53 (s, 1H, C3-H), 7.43 (t, *J* = 7.4 Hz, 2H, C23-H, C26-H), 7.34 (t, *J* = 7.4 Hz, 2H, C22-H, C27-H), 6.89 (d, *J* = 8.4 Hz, 1H, C6-H), 6.87 (dd, *J* = 8.4 Hz, 1.89 Hz, 1H, C5-H), 4.69 (s, 2H, C7-H), 4.41 (d, *J* = 7.1 Hz, 2H, C19-H), 4.31 (t, *J* = 7.0 Hz, 1H, C20-H), 2.71 (t, *J* = 7.5 Hz, 2H, C10-H), 2.44 (t, *J* = 7.6 Hz, 2H, C11-H), 1.36 (s, 9H, C14-H, C15-H, C16-H). **^{13}C NMR** (500 MHz, DMSO-d_6): δ = 171.5 (C12), 170.6 (C8), 153.5 (C18), 146.9 (C1), 143.7 (C20a, C28a), 140.7 (C24a, C24b), 133.8 (C4), 128.2 (C3), 127.7 (C23, C26), 127.5 (C2), 127.1 (C22, C27), 125.3 (C21, C28), 123.6 (C5), 120.2 (C24, C25), 113.6 (C6), 79.7 (C13), 66.3 (C7), 66.1 (C19), 46.6 (C20), 36.5 (C11), 29.9 (C10), 27.7 (C14, C15, C16). **HRMS** (ESI)$^-$ *m/z* calcd. for $C_{30}H_{30}NO_7$: 516.2028 (M-H)$^-$; Found: 516.2024.



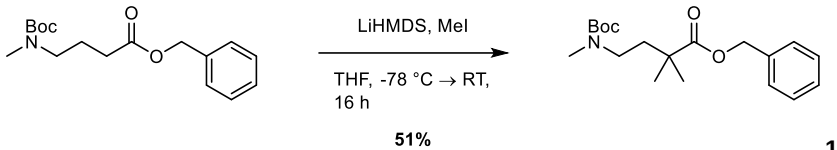
Compound 11 (Bnp): 4-(Methylamino)butyric acid hydrochloride (25.0 g, 163 mmol, 1.0 equiv.) was dissolved in THF (250 ml). NaOH $_{(aq.)}$ (2 M, 163 ml, 326 mmol, 2.0 equiv.) and Boc_2O (46.2 g, 212 mmol, 1.3 equiv.) were added at 0 °C, and the reaction mixture was stirred at RT for 3 h. Organic solvents were evaporated *in vacuo*, H_2O was added, and the mixture was washed with Et_2O (2×). Then, the aqueous phase was acidified using citric acid (aq.) (1 M) and extracted with DCM (3×). After the combined organic phases were dried over $MgSO_4$, solvents were evaporated *in vacuo* yielding the title compound (26.2 g, 120 mmol, 74%) as a colorless oil.

($C_{10}H_{19}NO_4$, MW = 217.27 g mol^{-1}). **^{1}H NMR** (400 MHz, DMSO-d_6): δ = 12.05 (s, 1H, O11-H), 3.16 (t, *J* = 7.0 Hz, 2H, C4-H), 2.74 (s, 3H, C5-H), 2.17 (t, *J* = 7.2 Hz, 2H, C2-H), 1.67 (quint., *J* = 7.2 Hz, 2H, C3-H), 1.38 (s, 9H, C8-H, C9-H, C10-H). **^{13}C NMR** (101 MHz, DMSO-d_6): δ = 174.0 (C1), 154.8 (C6), 78.3 (C7), 47.0 (C4), 33.6 (C5), 30.5 (C2), 28.0 (C8, C9, C10), 22.6 (C3). **HRMS** (ESI)$^-$ *m/z* calcd. for $C_{10}H_{18}NO_4$: 216.1241 (M-H)$^-$; Found: 216.1240.



11 **12**

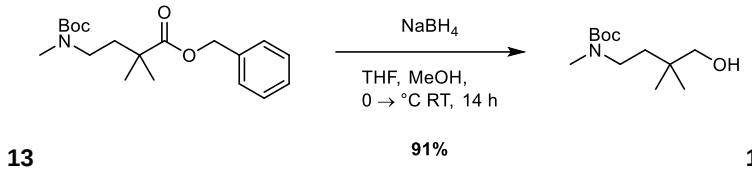
Compound 12: Compound **11** (20.0 g, 92.1 mmol, 1.0 equiv.), K_2CO_3 (38.2 g, 276 mmol, 3.0 equiv.) and benzyl bromide (13.1 ml, 120 mmol, 1.3 equiv.) were suspended in anhydrous DMF (100 ml). The reaction mixture was stirred for 20 h, then, H_2O was added, and the mixture was extracted with DCM (3×). The combined organic phases were washed with brine, dried over $MgSO_4$, and solvents were evaporated *in vacuo*. Purification of the residue by column chromatography (SiO_2, CyHex/EtOAc 9:1) yielded the title compound (21.5 g, 69.8 mmol, 76%) as a colorless oil. ($C_{17}H_{25}NO_4$, MW = 307.39 g mol^{-1}). **R_f** (CyHex/EtOAc 9:1) = 0.18. **^{1}H NMR** (500 MHz, $CDCl_3$): δ = 7.44–7.29 (m, 5H, C13-H, C14-H, C15-H, C16-H, C17-H), 5.12 (s, 2H, C11-H), 3.25 (t, *J* = 7.1 Hz, 2H, C4-H), 2.82 (s, 3H, C5-H), 2.36 (t, *J* = 7.4 Hz, 2H, C2-H), 1.85 (quint., *J* = 7.3 Hz, 2H, C3-H), 1.44 (s, 9H, C8-H, C9-H, C10-H). **^{13}C NMR** (126 MHz, $CDCl_3$): δ = 173.2 (C1), 155.9 (C6), 136.1 (C12), 128.7 (C13, C17), 128.4 (C14, C15, C16), 66.4 (C11), 48.2 (C4), 34.2 (C5), 31.5 (C2), 28.6 (C8, C9, C10), 23.3 (C3). **HRMS** (ESI)$^+$ *m/z* calcd. for $C_{17}H_{26}NO_4$: 308.1856 (M+H)$^+$; Found: 308.1860.



12 **13**

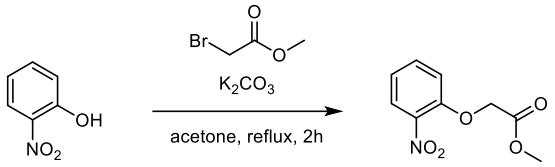
Compound 13: Compound **12** (21.5 g, 70.0 mmol, 1.0 equiv.) was dissolved in THF (55 ml) and the solution was cooled to −78 °C. LiHMDS (1 M in THF, 70.0 ml, 70.0 mmol, 1.0 equiv.) was added dropwise, the mixture was stirred for 15 min, then, iodomethane (4.35 ml, 70.0 mmol, 1.0 equiv.) was added dropwise. After the mixture was stirred for 2 h, LiHMDS (1 M in THF, 140.0 ml, 140.0 mmol, 2.0 equiv.) was added dropwise, the mixture was stirred for 15 min, then, iodomethane (8.71 ml, 140.0 mmol, 2.0 equiv.) was added dropwise. The reaction mixture was allowed to warm to RT and stirred for another 13 h. The solution was acidified with $NH_4Cl_{(aq.)}$ (saturated) and extracted with DCM (3×). After the combined organic phases were dried over $MgSO_4$, solvents were evaporated *in vacuo*, and the residue was purified by column chromatography (SiO_2, CyHex/EtOAc 95:5 → 9:1). The title compound (11.9 g, 35.3 mmol, 51%) was obtained as a colorless oil. ($C_{19}H_{29}NO_4$, MW = 335.21 g mol^{-1}). **R_f** (CyHex/EtOAc 9:1) = 0.28. **^{1}H NMR** (400 MHz, $CDCl_3$): δ = 7.44–7.29 (m, 5H, C15-H, C16-H, C17-H, C18-H, C19-H), 5.12 (s, 2H, C13-H), 3.13

(s, 2H, C4-H), 2.73 (s, 3H, C7-H), 1.75 (t, *J* = 7.8 Hz, 2H, C3-H), 1.43 (s, 9H, C10-H, C11-H, C12-H), 1.23 (s, 6H, C5-H, C6-H). **^{13}C NMR** (101 MHz, $CDCl_3$): δ = 177.3 (C1), 155.7 (C8), 136.3 (C14), 128.7 (C16, C18), 128.3 (C17), 128.2 (C15, C19), 79.4 (C9), 66.4 (C13), 45.5 (C4), 41.1 (C2), 38.1 (C3), 34.1 (C7), 28.6 (C10, C11, C12), 25.3 (C5, C6). **HRMS** (ESI)$^-$ *m/z* calcd. for $C_{19}H_{30}NO_4$: 336.2169 $(M+H)^+$; Found: 336.2174.



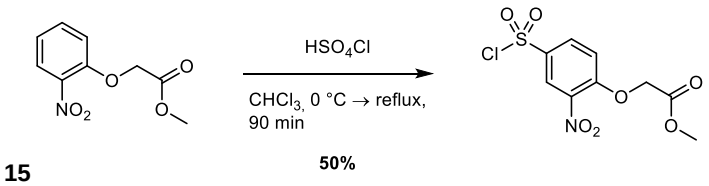
13 **14**

Compound 14: Compound **13** (11.9 g, 35.3 mmol, 1.0 equiv.) was dissolved in THF (60 ml) and methanol (30 ml). Sodium borohydride (4.01 g, 106 mmol, 3.0 equiv.) was added at 0 °C and the reaction mixture was stirred at RT for 14 h. After acidifying with citric acid $_{(aq.)}$ (1 M), H_2O was added, and the mixture was extracted with DCM (3×). The combined organic phases were dried over $MgSO_4$, solvents were evaporated *in vacuo*, and the residue was purified by column chromatography (SiO_2, CyHex/EtOAc 8:2 → 7:3 → 6:2) yielding the title compound (7.43 g, 32.1 mmol, 91%) as a colorless oil. ($C_{12}H_{25}NO_3$, 231.34 g mol^{-1}). **R_f** (CyHex/EtOAc 7:3) = 0.25. **^{1}H NMR** (400 MHz, DMSO-d_6): δ = 4.49 (t, *J* = 5.4 Hz, 1H, O13-H), 3.17–3.10 (m, 2H, C4-H), 3.08 (d, *J* = 5.4 Hz, 2H, C1-H), 2.73 (s, 3H, C7-H), 1.38 (s, 9H, C10-H, C11-H, C12-H), 1.37–1.29 (m, 2H, C3-H), 0.80 (s, 6H, C5-H, C6-H). **^{13}C NMR** (101 MHz, DMSO-d_6): δ = 154.6 (C8), 78.1 (C9), 70.0 (C1), 44.6 (C4), 36.0 (C3), 34.1 (C2), 33.6 (C7), 28.1 (C10, C11, C12), 23.8 (C5, C6). **HRMS** (ESI)$^-$ *m/z* calcd. for $C_{12}H_{24}NO_3$: 230.1762 $(M-H)^-$; Found: 230.1397.



15

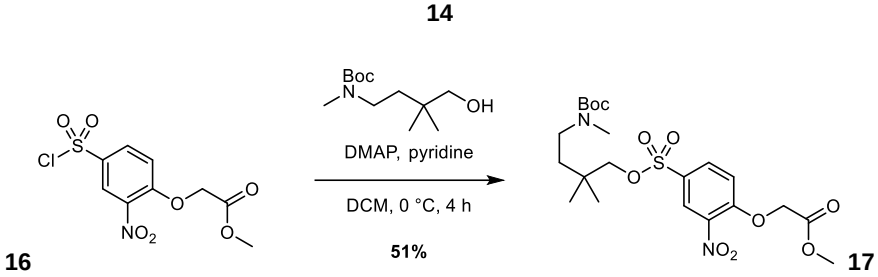
Compound **15** is described in section 5.3.3.



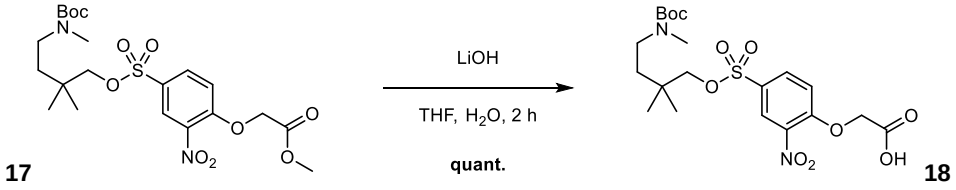
15 **16**

Compound 16: Compound **15** (15.0 g, 71.0 mmol, 1.0 equiv.) was dissolved in chloroform (30 ml). After the addition of chlorosulfuric acid (11.8 ml, 178 mmol, 2.5 equiv.) in chloroform (45 ml) at 0 °C, the mixture was refluxed (bath temperature: 80 °C) for 90 min. The reaction was quenched by the addition of H_2O and the mixture was extracted with DCM (3×), combined organic phases were dried over $MgSO_4$ and solvents were evaporated *in vacuo*. After purification of the residue by column chromatography (SiO_2, CyHex/EtOAc 8:2 → 7:3), the title compound (11.0 g,

35.4 mmol, 50%) was obtained as a brown oil (that became a solid after a while in the fridge). ($C_9H_8NO_7SCl$, MW = 309.67 g mol^{-1}). **R_f** (CyHex/EtOAc 7:3) = 0.19. **^{1}H NMR** (400 MHz, $CDCl_3$): δ = 8.54 (d, *J* = 2.5 Hz, 1H, C3-H), 8.17 (dd, *J* = 9.0, 2.5 Hz, 1H, C5-H), 7.14 (d, *J* = 9.0 Hz, 1H, C6-H), 4.94 (s, 2H, C7-H), 3.85 (s, 3H, C9-H). **^{13}C NMR** (101 MHz, $CDCl_3$): δ = 167.0 (C8), 155.9 (C1), 139.7 (C2), 137.0 (C4), 132.6 (C5), 125.8 (C3), 115.4 (C6), 66.3 (C7), 53.1 (C9). **HRMS** (ESI)$^-$ *m/z* calcd. for $C_9H_8NO_7SCl_2$: 343.9404 (M+Cl)$^-$; Found: 343.9409.

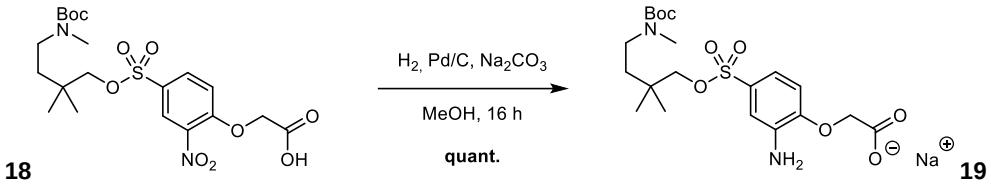


Compound 17: Compound **14** (7.43 g, 32.1 mmol, 1.0 equiv.), 4-dimethylaminopyridine (3.92 g, 32.1 mmol, 1.0 equiv.), and pyridine (5.17 ml, 64.2 mmol, 2.0 equiv.) were dissolved in anhydrous DCM (80 ml). Compound **16** (9.95 g, 32.1 mmol, 1.0 equiv.) in anhydrous DCM (20 ml) was added dropwise at 0 °C and the reaction mixture was stirred at 0 °C for 4 h. Then, the mixture was acidified with citric acid $_{(aq.)}$ (1 M) and extracted with DCM (3×). The combined organic phases were dried over $MgSO_4$, and solvents were evaporated *in vacuo*. After column purification (SiO_2, CyHex/EtOAc 8:2 → 7:3 → 1:1), the title compound (8.33 g, 16.5 mmol, 51%) was obtained as a brown oil. ($C_{21}H_{32}N_2O_{10}S$, MW = 504.18 g mol^{-1}). **R_f** (CyHex/EtOAc 6:4) = 0.23. **^{1}H NMR** (400 MHz, $CDCl_3$): δ = 8.40 (d, *J* = 2.2 Hz, 1H, C3-H), 8.04 (d, *J* = 8.8 Hz, 1H, C5-H), 7.10 (d, *J* = 8.9 Hz, 1H, C6-H), 4.90 (s, 2H, C7-H), 3.84 (s, 3H, C9-H), 3.78 (s, 2H, C10-H), 3.17 (t, *J* = 8.2 Hz, 2H, C13-H), 2.80 (s, 3H, C16-H), 1.48 (d, *J* = 8.3 Hz, 2H, C12-H), 1.44 (s, 9H, C19-H, C20-H, C21-H), 0.94 (s, 6H, C14-H, C15-H). **^{13}C NMR** (101 MHz, $CDCl_3$): δ = 167.3 (C8), 156.0 (C17), 154.9 (C1), 139.9 (C2), 133.0 (C5), 129.5 (C4), 126.3 (C3), 115.3 (C6), 79.6 (C18), 78.9 (C10), 66.3 (C7), 53.0 (C9), 44.2 (C13), 34.2 (C12), 33.8 (C11, C16), 28.6 (C19, C20, C21), 23.8 (C14, C15). **HRMS** (ESI)$^+$ *m/z* calcd. for $C_{21}H_{33}N_2O_{10}S$: 505.1850 (M+H)$^+$; Found: 505.1850.



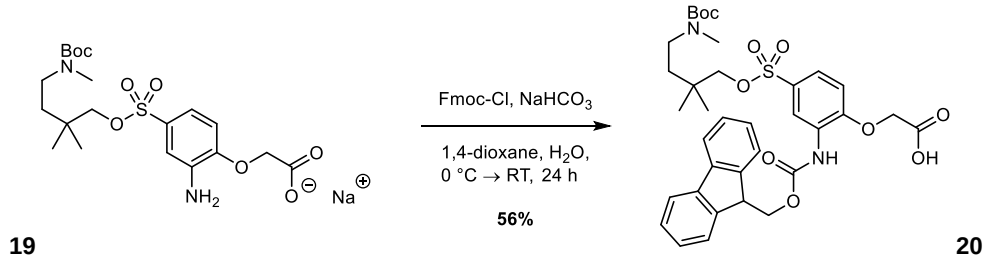
Compound 18: Compound **17** (8.33 g, 16.5 mmol, 1.0 equiv.) was dissolved in THF (160 ml). LiOH (474 mg, 19.8 mmol, 1.2 equiv.) in H_2O (40 ml) was added at 0 °C, and the reaction mixture was stirred at 0 °C for 2 h. After acidifying the mixture with citric acid $_{(aq.)}$ (1 M), it was extracted with DCM (3×), and the combined organic phases were dried over $MgSO_4$. Evaporation of the solvents *in vacuo* yielded the title compound (8.48 g, 16.5 mmol, quant.) as a brown oil. ($C_{20}H_{30}N_2O_{10}S$, MW = 490.52 g mol^{-1}). **R_f** (DCM/MeOH 95:5 + 1% AcOH) = 0.18. **^{1}H NMR** (400 MHz, DMSO-d_6): δ = 13.40 (s, 1H, O9-H), 8.42 (d, *J* = 2.3 Hz, 1H, C3-H), 8.14 (dd, *J* = 9.0,

2.4 Hz, 1H, C5-H), 7.54 (d, *J* = 9.0 Hz, 1H, C6-H), 5.07 (s, 2H, C7-H), 3.83 (s, 2H, C10-H), 3.16–3.02 (m, 2H, C13-H), 2.70 (s, 3H, C16-H), 1.38 (s, 2H, C12-H), 1.36 (s, 9H, C19-H, C20-H, C21-H), 0.84 (s, 6H, C14-H, C15-H). **^{13}C NMR** (101 MHz, DMSO-d_6): δ = 168.7 (C8), 154.5 (C17), 154.3 (C1), 139.3 (C2), 133.0 (C5), 127.2 (C4), 125.1 (C3), 116.6 (C6), 78.6 (C10), 78.4 (C18), 65.9 (C7), 43.6 (C13), 35.6 (C11), 33.6 (C16), 33.1 (C12), 28.0 (C19, C20, C21), 23.1 (C14, C15). **HRMS** (ESI)$^-$ *m/z* calcd. for $C_{20}H_{29}N_2O_{10}S$: 489.1548 $(M-H)^-$; Found: 489.1548.



18 **19**

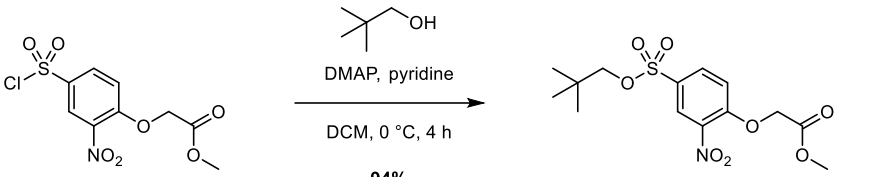
Compound 19: Compound **18** (8.09 g, 16.5 mmol, 1.0 equiv.) and Na_2CO_3 (1.75 g, 16.5 mmol, 1.0 equiv.) were suspended in MeOH (350 ml). The solution was quickly degassed by vacuum-N_2 cycles (3×), then, Pd/C (809 mg 10 wt. % loading) was added and the N_2 atmosphere was replaced by H_2. After stirring for 16 h the reaction mixture was filtered over celite©, washed with methanol and solvents were removed *in vacuo*. The crude product (8.93 g, 16.5 mmol, quant.) was obtained as a brown solid and was used in the next step without further purification. ($C_{20}H_{31}N_2O_8SNa$, MW = 482.52 g mol^{-1}). **^{1}H NMR** (500 MHz, DMSO-d_6): δ = 7.06 (d, *J* = 2.4 Hz, 1H, C3-H), 6.96 (dd, *J* = 8.4, 2.4 Hz, 1H, C5-H), 6.72 (d, *J* = 8.5 Hz, 1H, C6-H), 5.30 (s, 2H, N10-H), 4.17 (s, 2H, C7-H), 3.61 (s, 2H, C11-H), 3.08 (d, *J* = 10.5 Hz, 2H, C14-H), 2.70 (s, 3H, C17-H), 1.38 (s, 2H, C13-H), 1.37 (s, 9H, C20-H, C21-H, C22-H), 0.83 (s, 6H, C15-H, C16-H). **^{13}C NMR** (126 MHz, DMSO-d_6): δ = 169.4 (C8), 154.5 (C18), 150.7 (C1), 139.2 (C2), 125.7 (C4), 116.2 (C5), 111.9 (C6), 110.9 (C3), 78.3 (C19), 77.1 (C11), 69.1 (C7), 44.1 (C14), 33.6 (C17), 33.0 (C12), 28.1 (C20, C21, C22), 23.3 (C15, C16). **HRMS** (ESI)$^-$ *m/z* calcd. for $C_{20}H_{31}N_2O_8S$: 459.1807 $(M-H)^-$; Found: 459.1808.



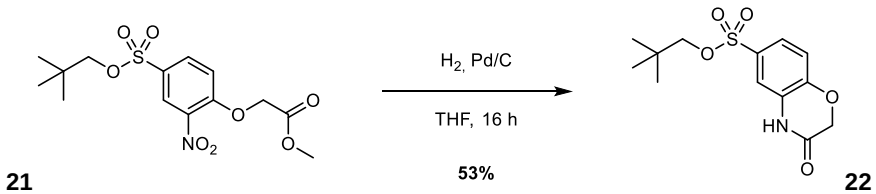
19 **20**

Compound 20: Compound **19** (7.96 g, 16.5 mmol, 1.0 equiv.) and $NaHCO_3$ (6.93 g, 82.5 mmol, 5.0 equiv.) were dissolved in H_2O (300 ml). Fmoc-Cl (5.55 g, 21.5 mmol, 1.3 equiv.) in 1,4-dioxane (300 ml) was added dropwise at 0 °C over 1 h. The reaction mixture was stirred at 0 °C for another 1 h, then, it was stirred at RT for 22 h. After acidifying with citric acid $_{(aq.)}$ (1 M), the mixture was extracted with DCM (3×), the combined organic phases were dried over $MgSO_4$, and solvents were evaporated *in vacuo*. Column chromatography (SiO_2, CyHex/EtOAc 1:1 + 1% DIPEA → 1:1 → EtOAc + 1% AcOH → EtOAc/MeOH 9:1 + 1% AcOH) of the residue yielded the title compound (6.43 g, 9.23 mmol, 56%) as a white solid. ($C_{35}H_{42}N_2O_{10}S$, MW = 682.79 g mol^{-1}). **R_f** (EtOAc/MeOH 9:1 + 1% AcOH) = 0.17. **^{1}H NMR** (500 MHz, DMSO-d_6): δ = 13.27 (s, 1H, O9-H),

9.17 (s, 1H, N22-H), 8.34 (d, *J* = 2.3 Hz, 1H, C3-H), 7.91 (d, *J* = 7.5 Hz, 2H, C29-H, C30-H), 7.78 (d, *J* = 7.6 Hz, 2H, C26-H, C33-H), 7.59 (d, *J* = 7.6 Hz, 1H, C5-H), 7.43 (t, *J* = 7.5 Hz, 2H, C28-H, C31-H), 7.34 (td, *J* = 7.4, 1.2 Hz, 2H, C27-H, C32-H), 7.22 (d, *J* = 8.8 Hz, 1H, C6-H), 4.91 (s, 2H, C7-H), 4.44 (d, *J* = 7.2 Hz, 2H, C24-H), 4.32 (t, *J* = 7.1 Hz, 1H, C25-H), 3.70 (s, 2H, C10-H), 3.05 (t, *J* = 7.9 Hz, 2H, C13-H), 2.67 (s, 3H, C16-H), 1.35 (s, 2H, C12-H), 1.34 (s, 9H, C19-H, C20-H, C21-H), 0.82 (s, 6H, C14-H, C15-H). **^{13}C NMR** (126 MHz, DMSO-d_6): δ = 169.5 (C8), 154.5 (C17), 153.6 (C23), 152.1 (C1), 143.6 (C25a, C33a), 140.7 (C29a, C29b), 128.5 (C4), 127.7 (C28, C31), 127.2 (C2), 127.1 (C27, C32), 125.3 (C26, C33), 123.8 (C5), 120.2 (C29, C30), 119.2 (C3), 113.1 (C6), 78.3 (C18), 77.7 (C10), 66.5 (C24), 65.6 (C7), 46.5 (C25), 44.1 (C13), 35.4 (C11), 33.6 (C16), 33.0 (C12), 28.0 (C19, C20, C21), 23.2 (C14, C15). **HRMS** (ESI)$^-$ *m/z* calcd. for $C_{35}H_{41}N_2O_{10}S$: 681.2487 $(M-H)^-$; Found: 681.2486.

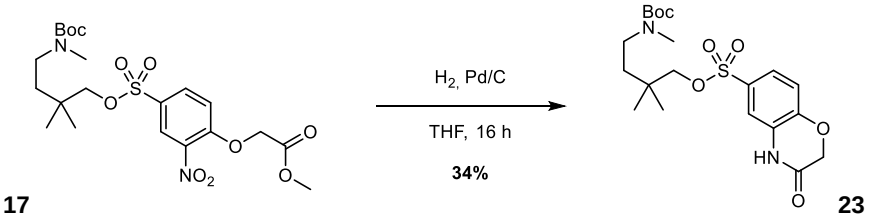


Compound 21: Neopentyl alcohol (313 mg, 3.55 mmol, 1.1 equiv.), 4-dimethylaminopyridine (395 mg, 3.23 mmol, 1.0 equiv.), and pyridine (520 µl, 6.46 mmol, 2.0 equiv.) were dissolved in anhydrous DCM (8 ml). Compound **16** (1.00 g, 3.23 mmol, 1.0 equiv.) in anhydrous DCM (2 ml) was added dropwise at 0 °C and the reaction mixture was stirred at 0 °C for 4 h. Then, the mixture was acidified with citric acid $_{(aq.)}$ (1 M) and extracted with DCM (3×). The combined organic phases were dried over $MgSO_4$, and solvents were evaporated *in vacuo* yielding the title compound (1.10 g, 3.04 mmol, 94%) as a brown solid. ($C_{14}H_{19}NO_8S$, 361.37 g mol^{-1}). **R_f** (CyHex/EtOAc 6:4) = 0.34. **^{1}H NMR** (400 MHz, $CDCl_3$): δ = 8.40 (d, *J* = 2.3 Hz, 1H, C3-H), 8.03 (dd, *J* = 8.9, 2.3 Hz, 1H, C5-H), 7.09 (d, *J* = 8.9 Hz, 1H, C6-H), 4.90 (s, 2H, C7-H), 3.84 (s, 3H, C9-H), 3.75 (s, 2H, C10-H), 0.93 (s, 9H, C12-H, C13-H, C14-H). **^{13}C NMR** (101 MHz, $CDCl_3$): δ = 167.2 (C8), 154.7 (C1), 139.7 (C2), 133.2 (C5), 129.5 (C4), 126.1 (C3), 115.1 (C6), 80.4 (C10), 66.1 (C7), 52.9 (C9), 31.8 (C11), 26.0 (C12, C13, C14). **HRMS** (ESI)$^-$ *m/z* calcd. for $C_{14}H_{19}NO_8SCl$: 396.0525 $(M+Cl)^-$; Found: 396.0533.



Compound 22: Compound **21** (1.10 g, 3.04 mmol, 1.0 equiv.) and acetic acid (174 µl, 3.04 mmol, 1.0 equiv.) were dissolved in THF (100 ml). The solution was quickly degassed by vacuum-N_2 cycles (3×), then, Pd/C (110 mg 10 wt. % loading) was added and the N_2 atmosphere was replaced by H_2. After stirring for 16 h the reaction mixture was filtered over celite©, and solvents were removed *in vacuo*. The crude product was recrystallized by dissolving it in a small amount of Et_2O and adding

n-hexane. The solution was kept at −20 °C to aid crystallization, then, the precipitate was filtered and washed with *n*-hexane to yield the title compound (482 mg, 1.61 mmol, 53%) as a white solid. ($C_{13}H_{17}NO_5S$, MW = 299.34 g mol^{-1}). $\mathbf{R_f}$ (CyHex/EtOAc 1:1) = 0.33. **^{1}H NMR** (400 MHz, DMSO-d_6): δ = 11.13 (s, 1H, N9-H), 7.57 (d, *J* = 2.2 Hz, 1H, C3-H), 7.54 (dd, *J* = 8.4, 2.3 Hz, 1H, C5-H), 7.22 (d, *J* = 8.4 Hz, 1H, C6-H), 4.96 (s, 2H, C7-H), 3.70 (s, 2H, C10-H), 0.85 (s, 9H, C12-H, C13-H, C14-H). **^{13}C NMR** (101 MHz, DMSO-d_6): δ = 159.6 (C8), 147.8 (C1), 130.1 (C4), 128.6 (C2), 123.9 (C5), 116.9 (C6), 111.7 (C3), 79.4 (C10), 67.9 (C7), 31.3 (C11), 25.6 (C12, C13, C14). **HRMS** (ESI)$^-$ *m/z* calcd. for $C_{13}H_{16}NO_5S$: 298.0755 (M-H)$^-$; Found: 298.0757.



17 **23**

Compound 23: Compound **17** (280 mg, 555 µmol, 1.0 equiv.) was dissolved in THF (30 ml). The solution was quickly degassed by vacuum-N_2 cycles (3×), then, Pd/C (28 mg 10 wt. % loading) was added and the N_2 atmosphere was replaced by H_2. After stirring for 16 h the reaction mixture was filtered over celite©, and solvents were removed *in vacuo*. The residue was subjected to column chromatography (SiO_2, CyHex/EtOAc 6:4 → 4:6), and the crude product was further purified by trituration in acetonitrile. Centrifugation and decantation of the supernatant yielded the title compound (83 mg, 188 µmol, 34%) as a white solid. ($C_{20}H_{30}N_2O_7S$, MW = 442.53 g mol^{-1}). $\mathbf{R_f}$ (CyHex/EtOAc 4:6) = 0.45. **^{1}H NMR** (500 MHz, $CDCl_3$): δ = 9.48 (s, 1H, N9-H), 7.72 (s, 1H, C3-H), 7.52 (dd, *J* = 8.5, 2.2 Hz, 1H, C5-H), 7.06 (d, *J* = 8.5 Hz, 1H, C6-H), 4.68 (s, 2H, C7-H), 3.67 (s, 2H, C10-H), 3.24–3.09 (m, 2H, C13-H), 2.82 (s, 3H, C16-H), 1.52 (s, 9H, C19-H, C20-H, C21-H), 1.44 (d, *J* = 7.9 Hz, 2H, C12-H), 0.98 (s, 6H, C14-H, C15-H). **^{13}C NMR** (126 MHz, $CDCl_3$): δ = 163.8 (C8), 156.1 (C17), 147.9 (C1), 129.5 (C4), 128.0 (C2), 124.2 (C5), 116.9 (C6), 115.6 (C3), 80.5 (C18), 76.8 (C10), 67.4 (C7), 44.6 (C13), 35.5 (C12), 34.3 (C16), 33.6 (C11), 28.8 (C19, C20, C21), 25.1 (C14, C15). **HRMS** (ESI)$^-$ *m/z* calcd. for $C_{20}H_{29}N_2O_7S$: 441.1701 (M-H)$^-$; Found: 441.1711.

8.4 X-ray crystallography

The stock solution of **1** was prepared from a lyophilized sample by dissolving in H_2O to a final concentration of 8 mM. Diffraction-quality crystals of **1** (Fig. 123) were obtained by sitting drop method by adding 1 µL of **1** and 1 µL of the reservoir solution containing 15% *w/v* polyethylene glycol 8000, 50 mM BIS-TRIS propane buffer (pH 6.8), 100 mM ammonium sulfate and 10% glycerol. Volume of reservoir solution was 500 µL. Single crystals of **1** were fished using MiTeGen micro loops, quickly soaked in a cryo-protectant solution of 30% w/v polyethylene glycol 8000 and 20% v/v polyethylene glycol 400 and flash frozen in liquid nitrogen.

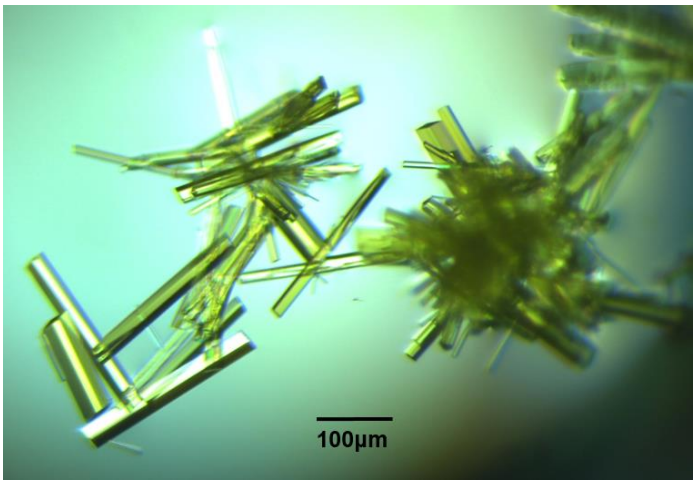


Fig. 123 Crystals of **1** observed under cross polarizing microscope.

Synchrotron data for **1** was collected at the ID30b beamline in European Synchrotron Radiation Facility (Grenoble, France) using PILATUS3 6M 1000µm Si sensor (Dectris) detector. Diffraction data for **1** was processed using *autoPROC*.[104, 122, 138] The structure was solved using intrinsic phasing method with the program *ShelxT*[142] and refined by a full-matrix least squares method on F^2 with *ShelxL*-2014[106] within *Olex2* suite.[107] The initial structure solution revealed most of the main-chain atoms. After each refinement step, visual inspection of the model and the electron-density maps were carried out using *Olex2* and *Coot*.[108] AFIX, DFIX, EADP, SADI and FLAT instructions were used to improve the geometry of the molecule and temperature parameters. All non-H atoms were refined with anisotropic displacement parameters. Hydrogen atoms were placed at idealized positions. Restraints on anisotropic displacement parameters were implemented with DELU, SIMU, RIGU and ISOR instructions. Statistics of data collection and refinement are described in Table 8.

Crystallization experiments and data collection was done with the help of Pradeep Mandal.

Table 8 Crystallographic data and refinement details for **1**.

Compound	**1**
Chemical formula	$C_{115}H_{93}N_{21.5}O_{44}S_3$
Formula weight	2575.78
Temperature	100 K
Wavelength	0.9762 Å
Crystal system	Monoclinic
Space group	$P2_1/n$
Unit cell dimensions	a = 15.566 (3) Å b = 34.166 (7) Å c = 63.600 (12) Å α = 90° β = 94.52 (3)° γ = 90°
Volume	33719 (12) Å^3
Z	8
Density (calculated)	1.015 g/cm^3
Absorption coefficient	0.286 μ/mm^{-1}
Colour and shape	Pale yellow, needles
Crystal size	0.25 x 0.05 x 0.05 mm
Index ranges	-15 ≤ h ≤ 15 -32 ≤ k ≤ 31 -63 ≤ l ≤63
Reflections collected	100786
R_{int}	0.2326
Data/restraints/parameters	31940/95/2864
Goodness-of-fit on F^2	1.524
Final R indexes [$I > 2\sigma(I)$]	R_1 = 0.1774 wR_2 = 0.4208
Final R indexes [all data]	R_1 = 0.2267 wR_2 = 0.4497
Largest diff. peak and hole	1.17/-0.70 e Å^{-3}
Total potential solvent accessible void volume from SQUEEZE	11693.6 Å^3
Electron count/cell	4135

8.5 Spectra and Chromatograms

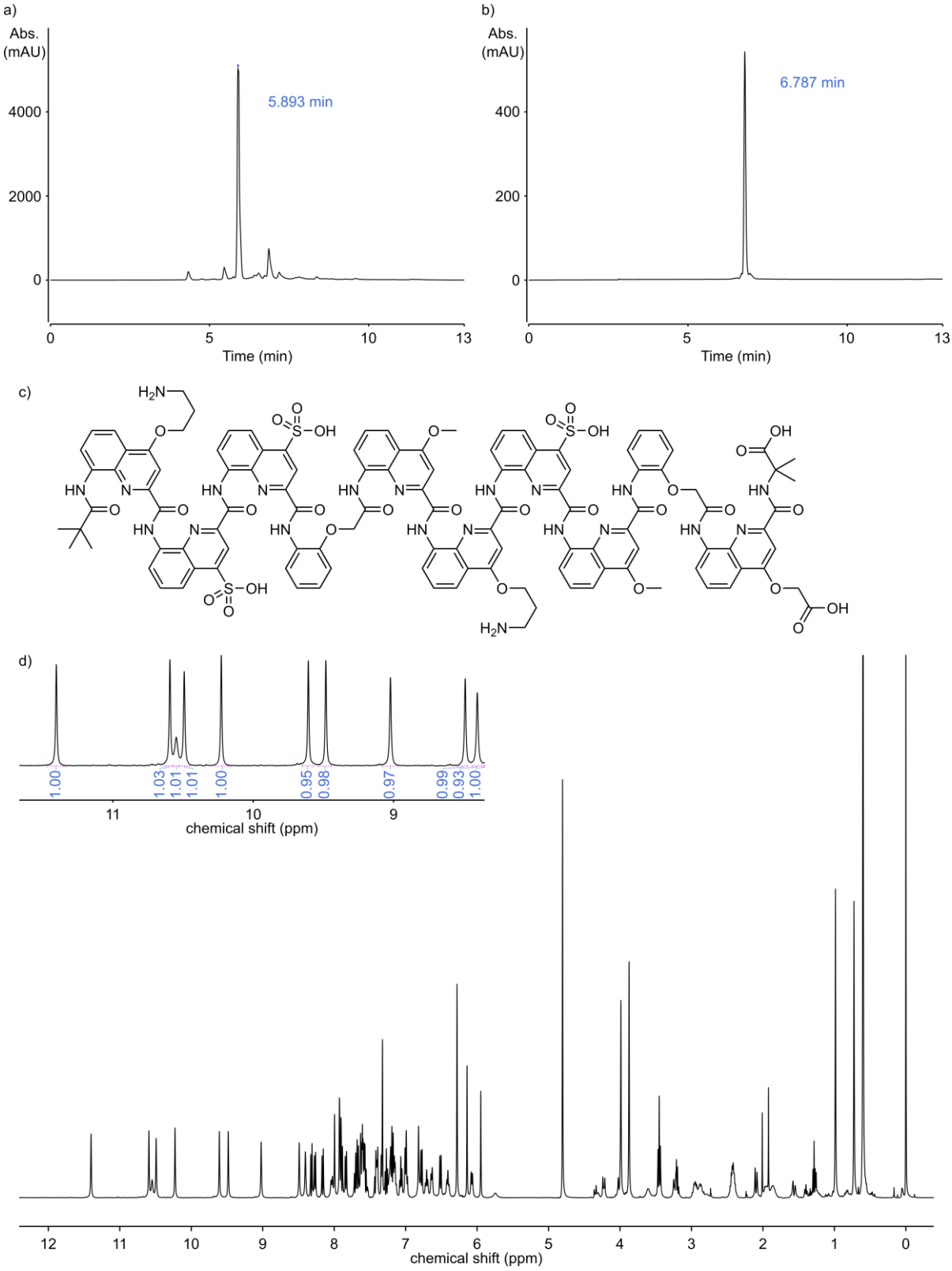


Fig. 124 Analytical data of compound **1**. HPLC chromatograms after cleavage from the resin (a) (C18, 0–45B, 50 °C; A: 13 mM ammonium acetate buffer pH 8.5, B: acetonitrile) and after purification (b) (C8, 0–40B, 50 °C; A: 13 mM ammonium acetate buffer pH 8.5, B: acetonitrile). c) Chemical structure of compound **1**. d) ^{1}H NMR spectrum (500 MHz, 1.0 mM in 27 mM sodium phosphate buffer pH 7.0).

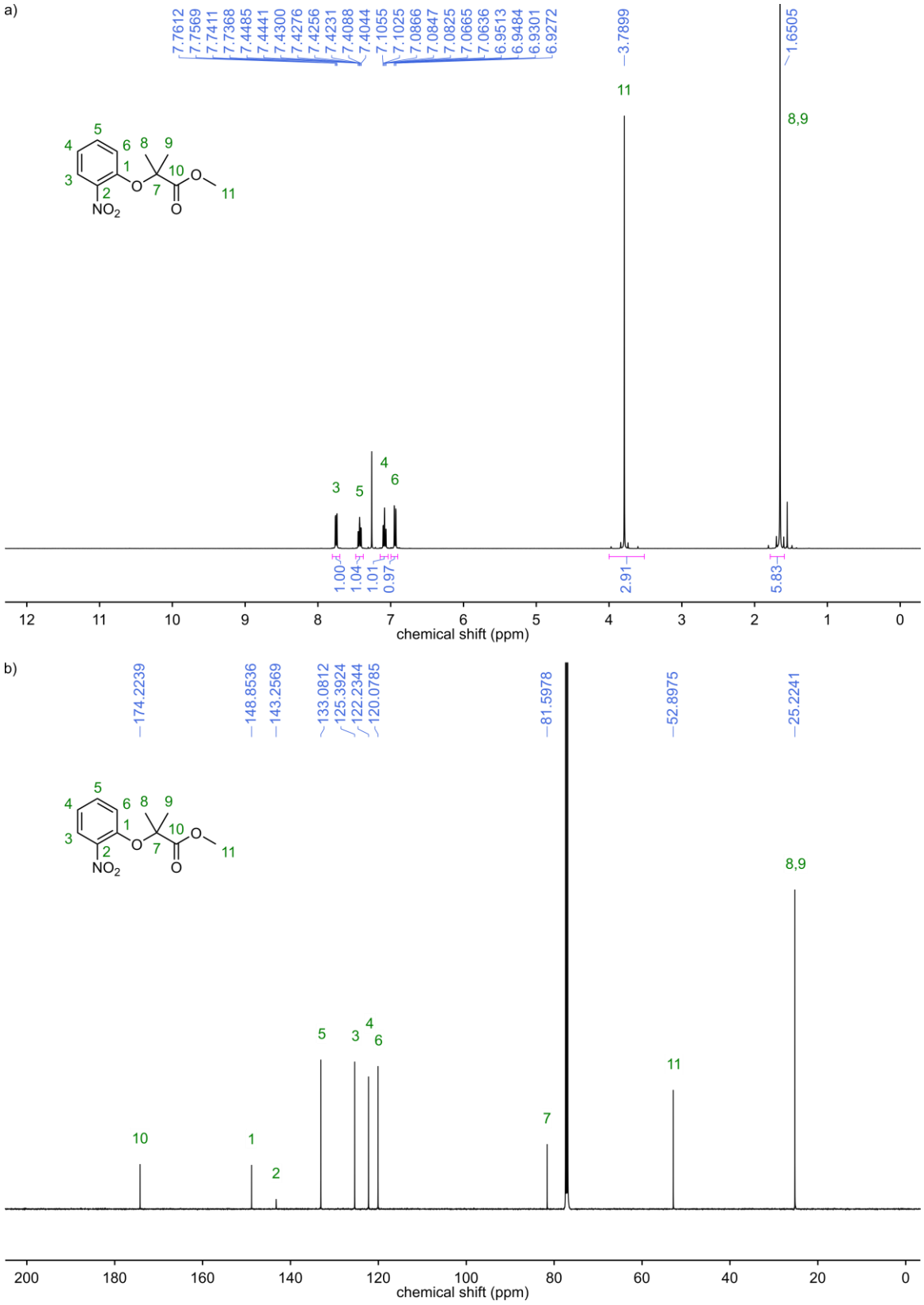


Fig. 125 NMR spectra of compound **2**. a) ^{1}H NMR (500 MHz, $CDCl_3$). b) ^{13}C NMR (126 MHz, $CDCl_3$).

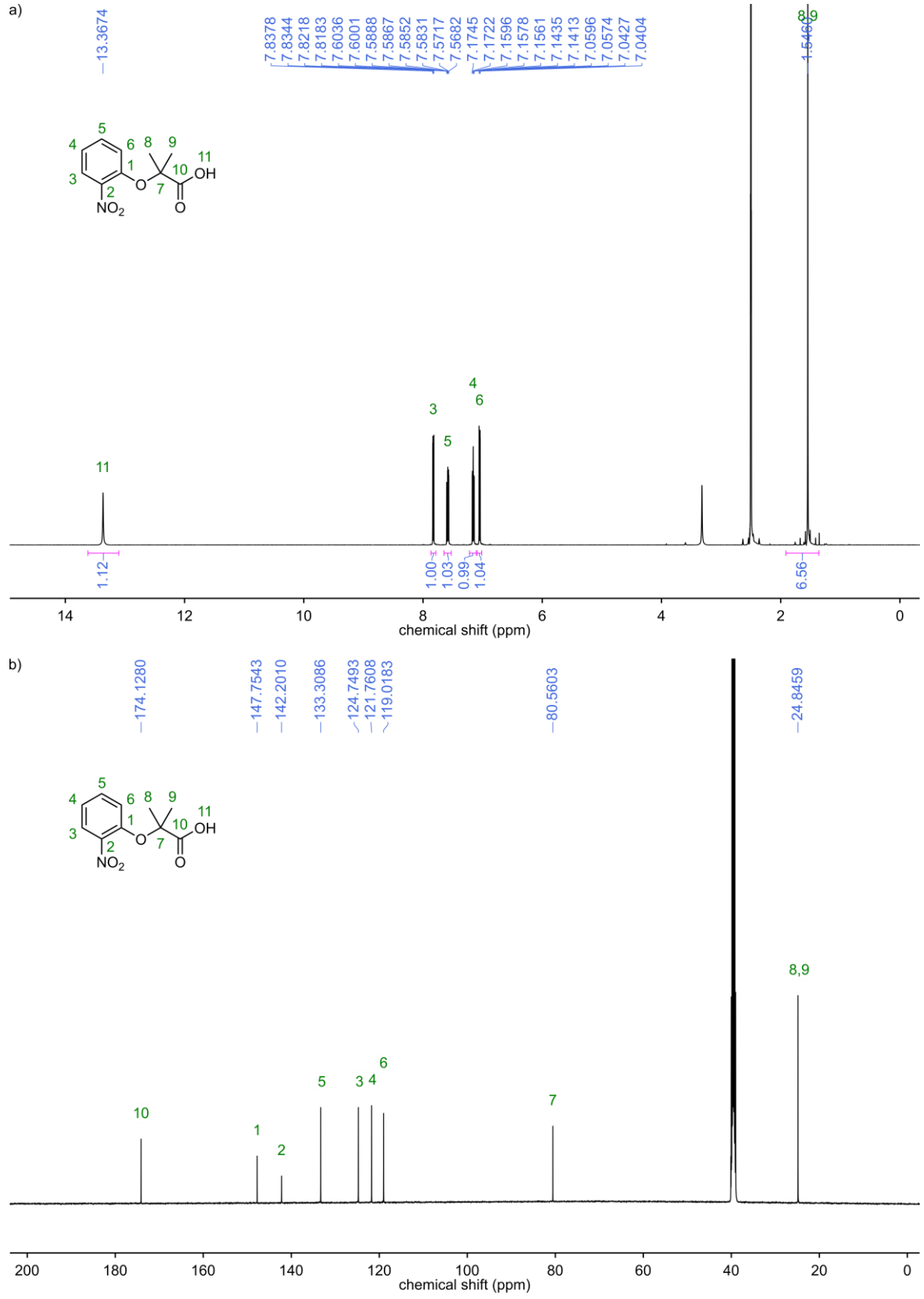


Fig. 126 NMR spectra of compound **3**. a) ^{1}H NMR (500 MHz, DMSO-d$_6$). b) ^{13}C NMR (126 MHz, DMSO-d$_6$).

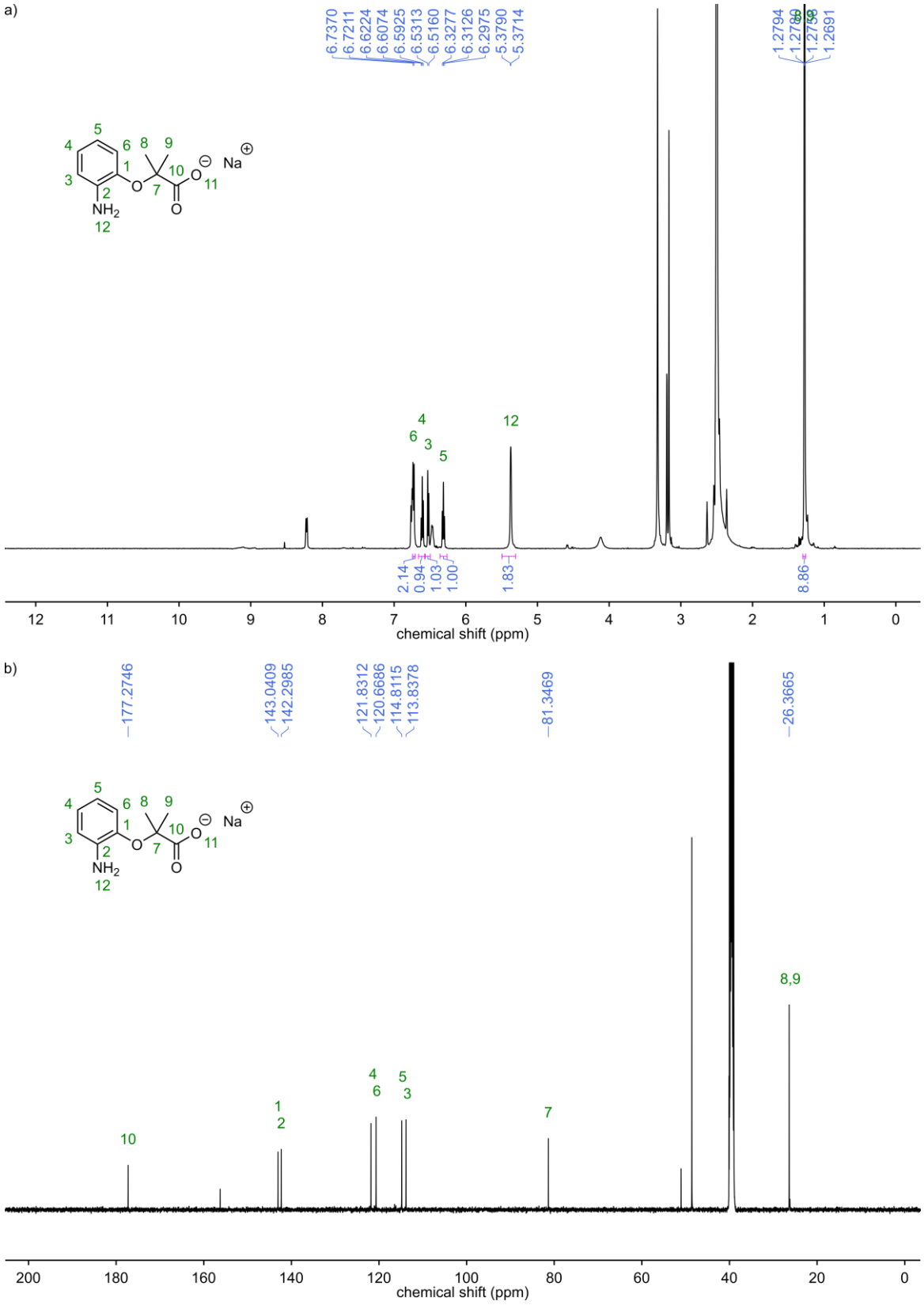


Fig. 127 NMR spectra of compound **4**. a) ^{1}H NMR (500 MHz, DMSO-d$_6$). b) ^{13}C NMR (126 MHz, DMSO-d$_6$).

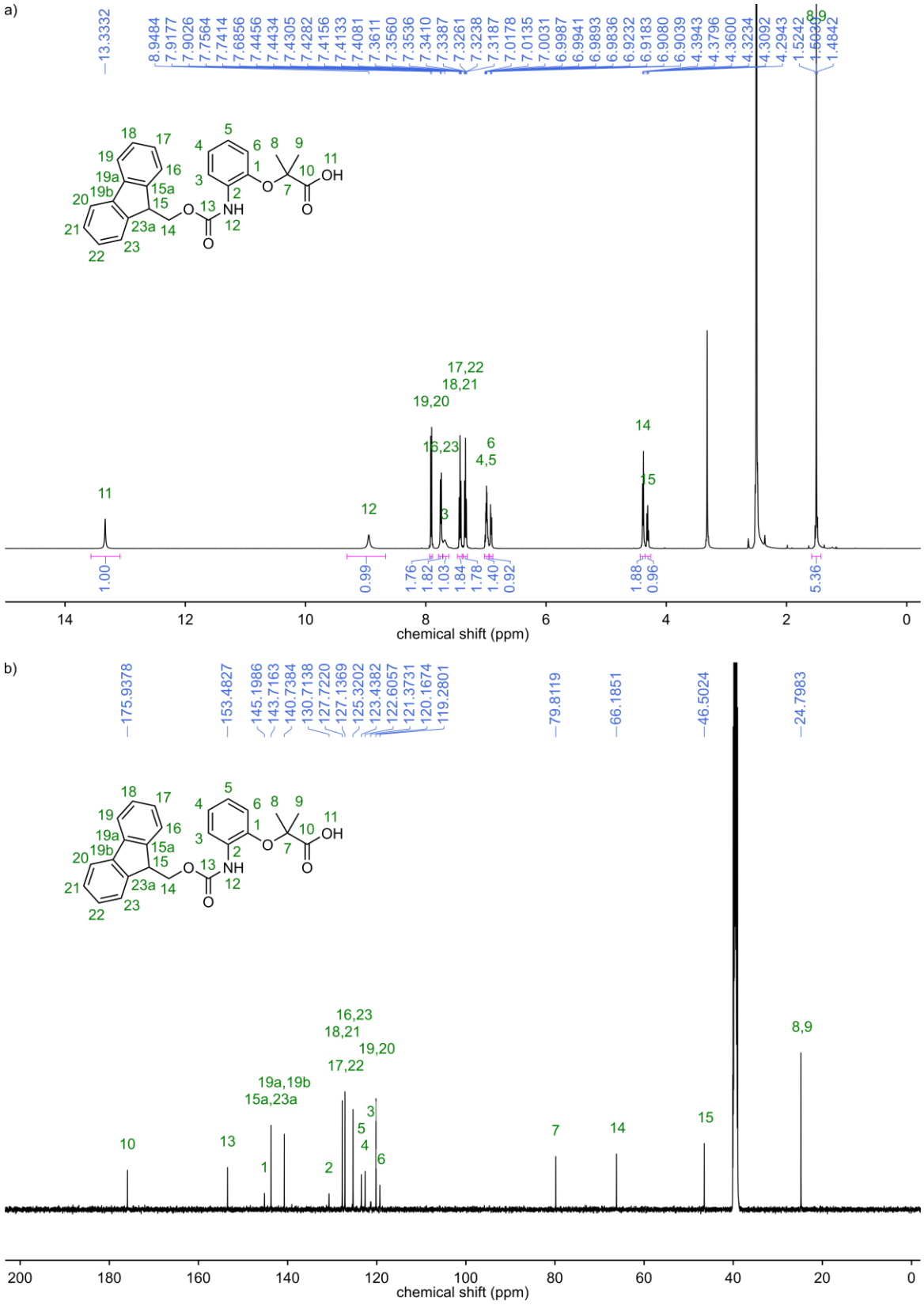


Fig. 128 NMR spectra of compound **5**. a) ^{1}H NMR (500 MHz, DMSO-d$_6$). b) ^{13}C NMR (126 MHz, DMSO-d$_6$).

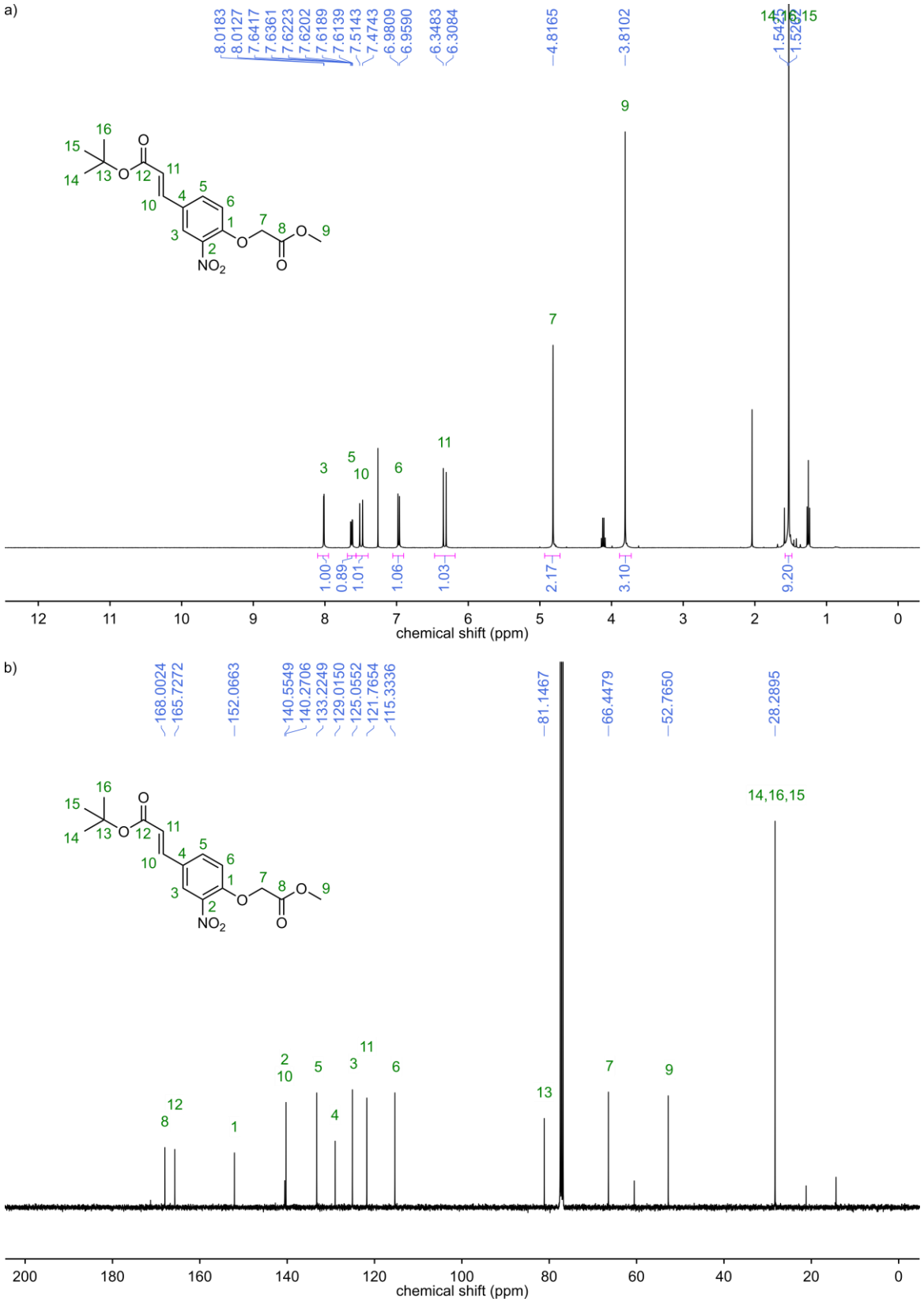


Fig. 129 NMR spectra of compound **7**. a) 1H NMR (400 MHz, $CDCl_3$). b) ^{13}C NMR (101 MHz, $CDCl_3$).

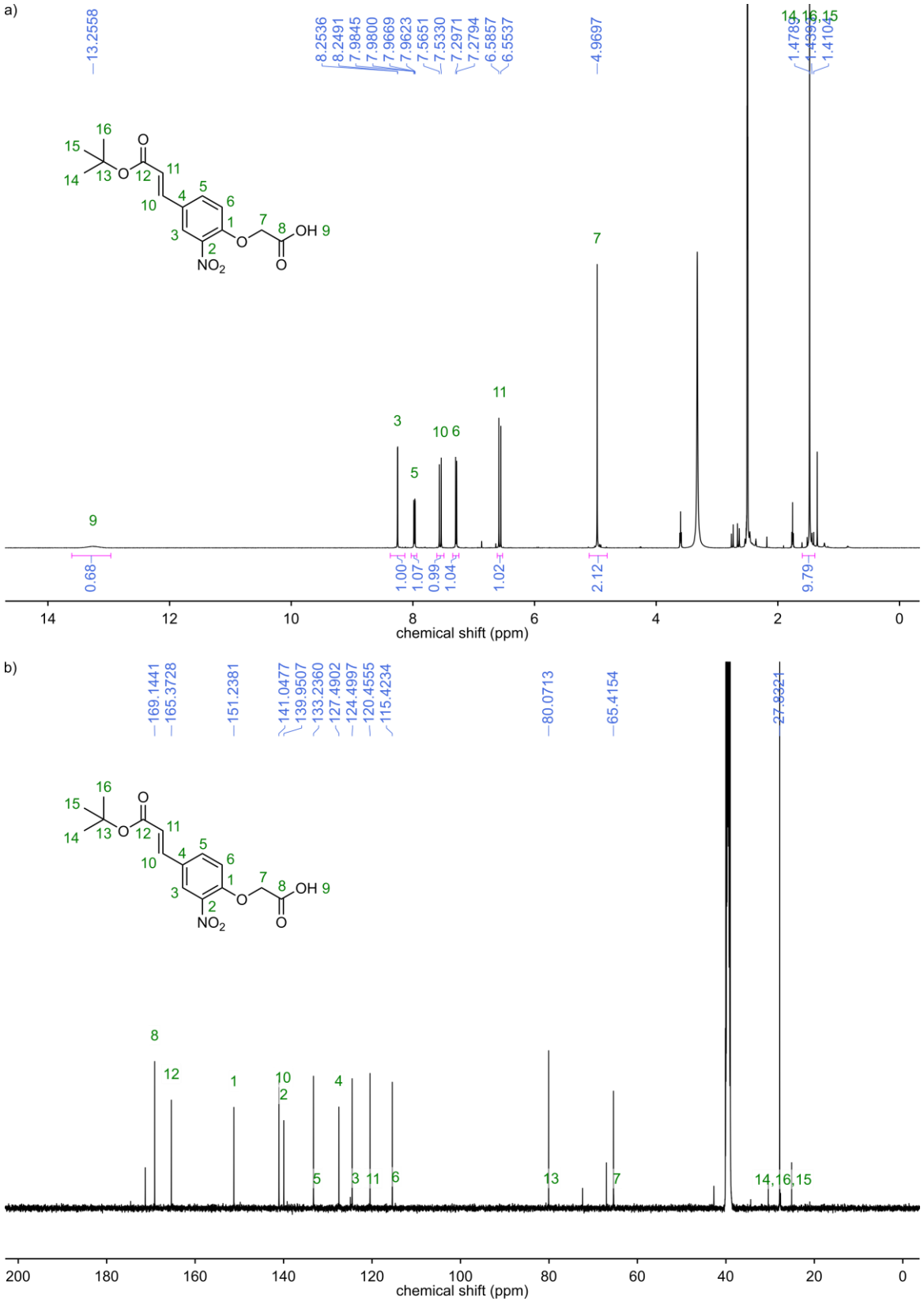


Fig. 130 NMR spectra of compound **8**. a) ^{1}H NMR (500 MHz, DMSO-d$_6$). b) ^{13}C NMR (126 MHz, DMSO-d$_6$).

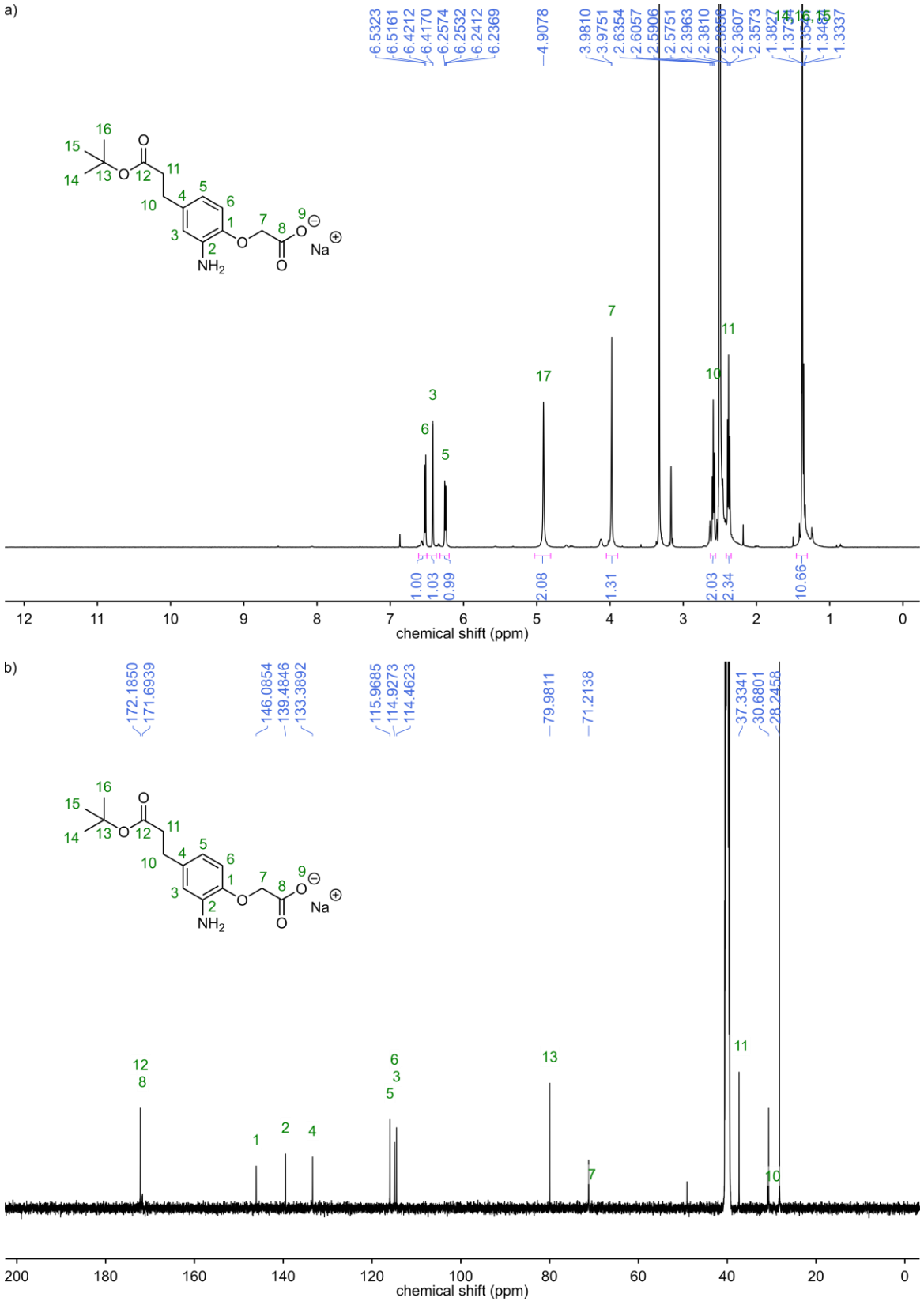


Fig. 131 NMR spectra of compound **9**. a) ^{1}H NMR (500 MHz, DMSO-d$_6$). b) ^{13}C NMR (126 MHz, DMSO-d$_6$).

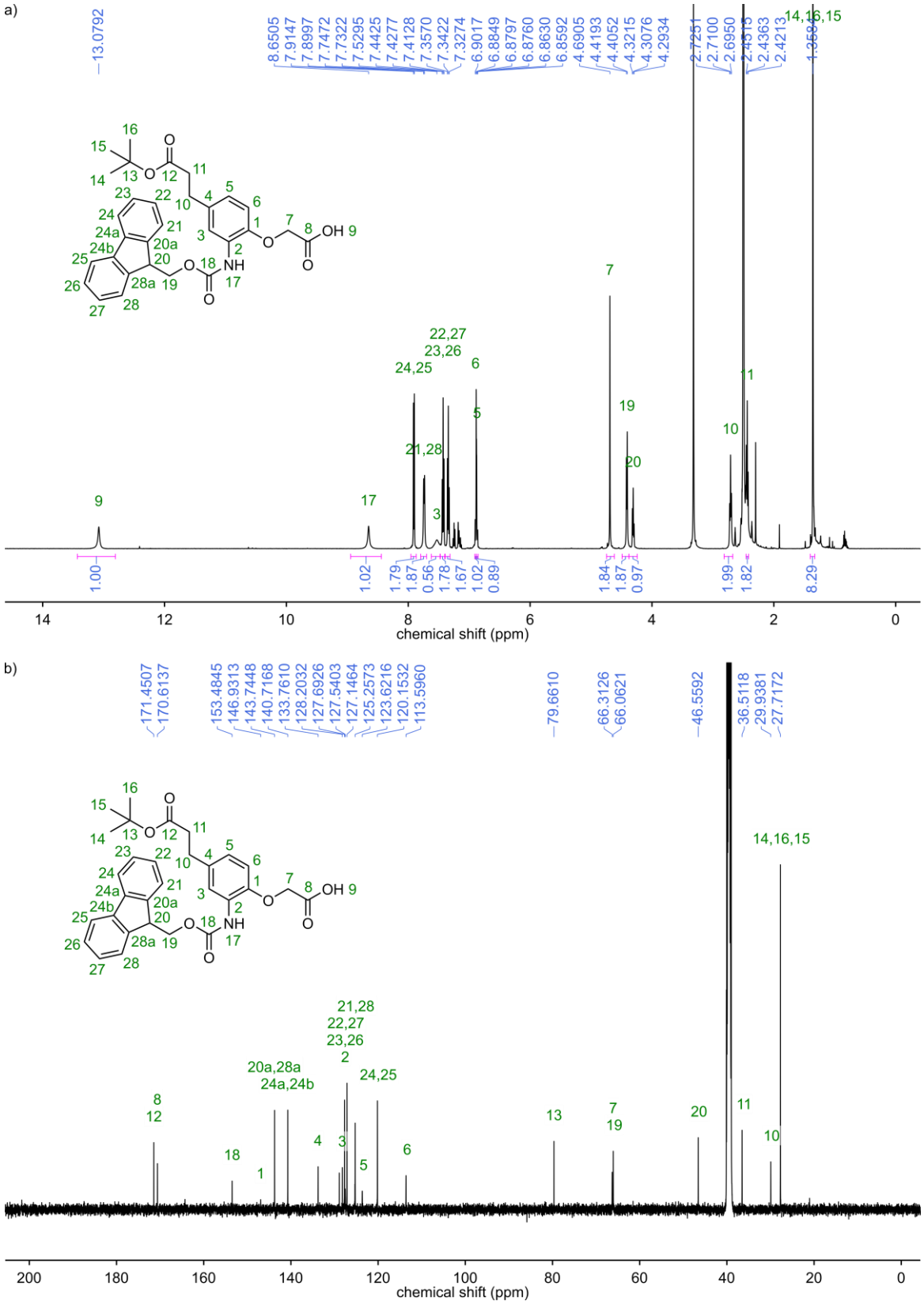


Fig. 132 NMR spectra of compound **10**. a) ^{1}H NMR (500 MHz, DMSO-d$_6$). b) ^{13}C NMR (126 MHz, DMSO-d$_6$).

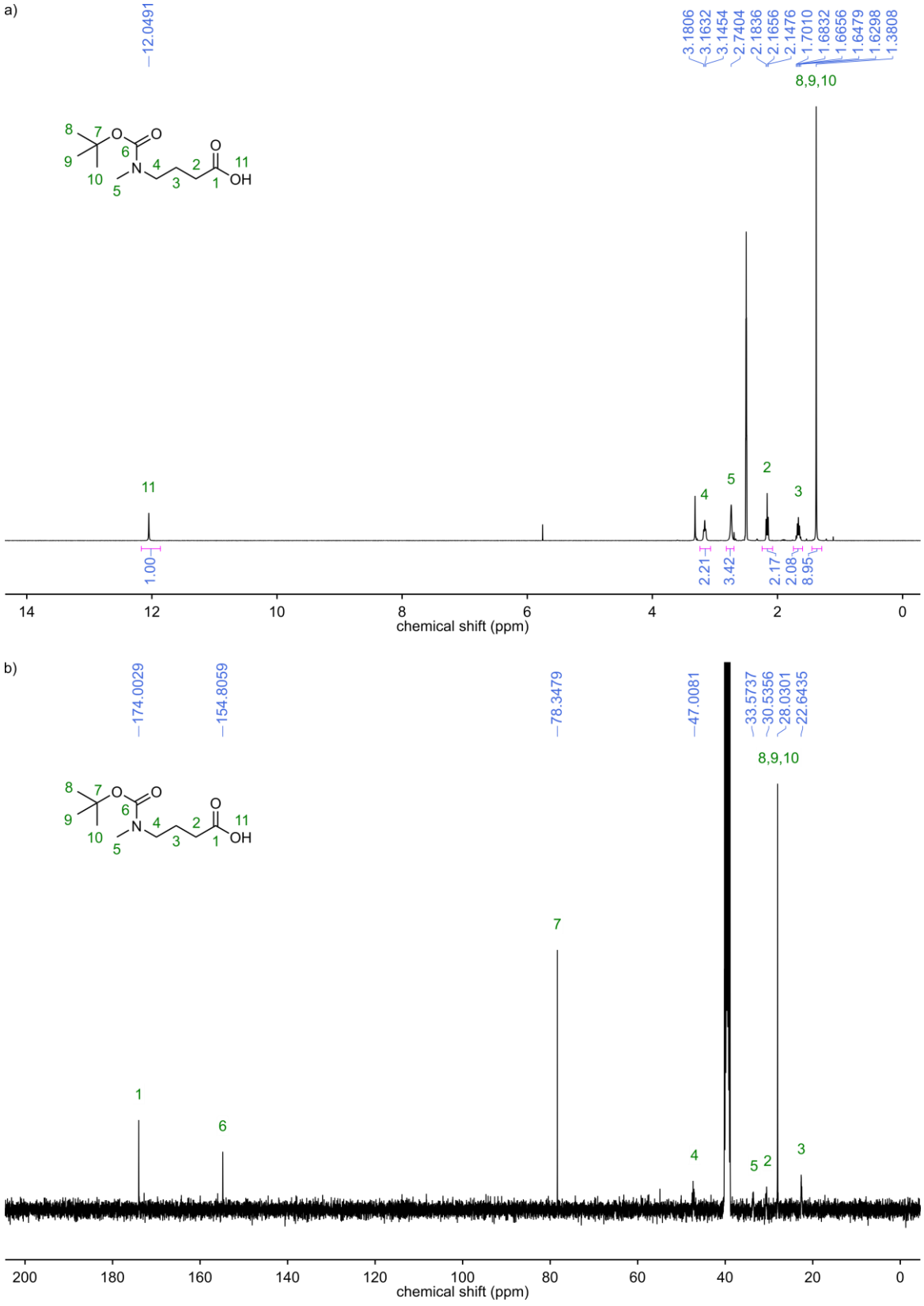


Fig. 133 NMR spectra of compound **11**. a) ^{1}H NMR (400 MHz, DMSO-d$_6$). b) ^{13}C NMR (101 MHz, DMSO-d$_6$).

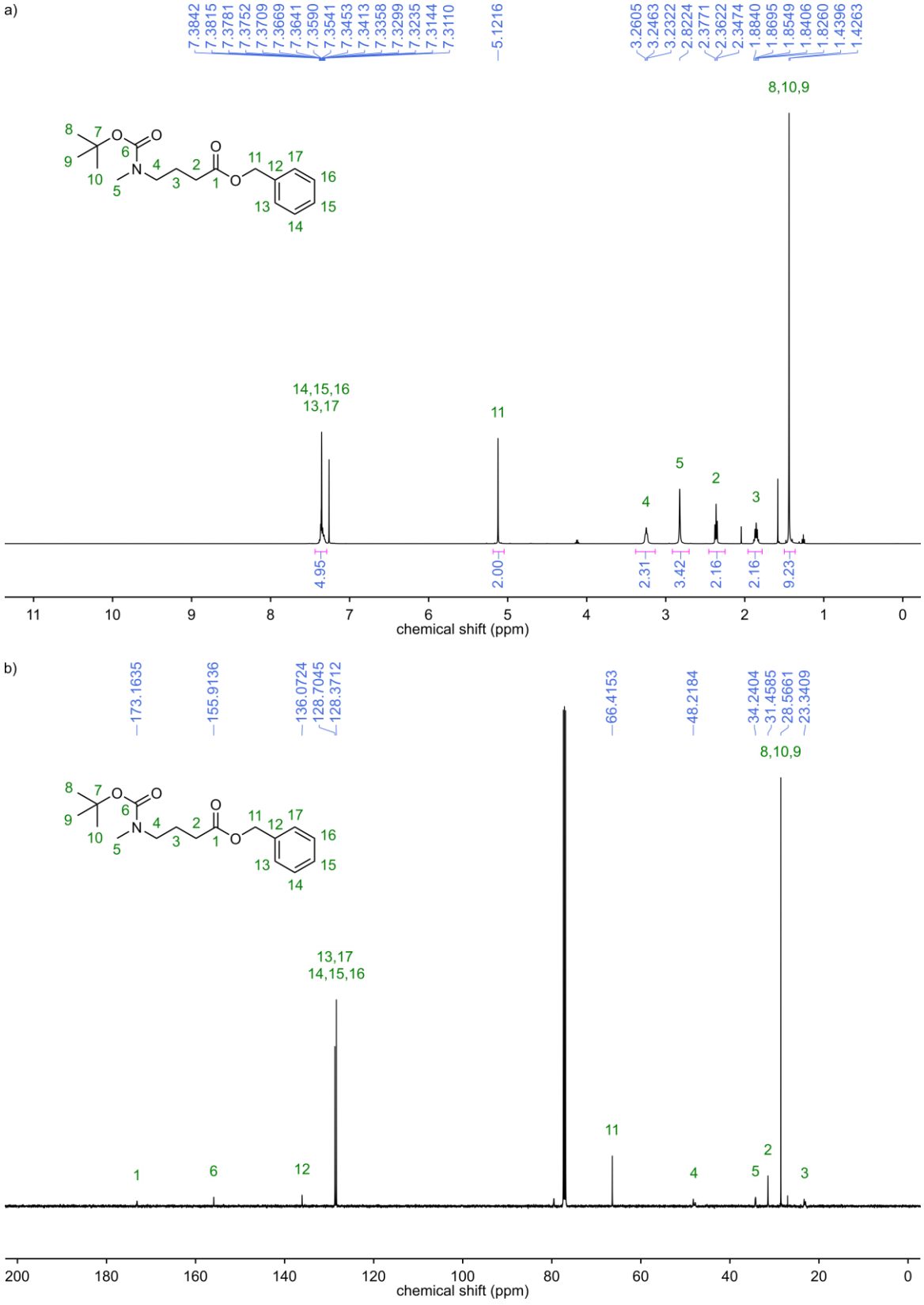


Fig. 134 NMR spectra of compound **12**. a) ^{1}H NMR (500 MHz, $CDCl_3$). b) ^{13}C NMR (126 MHz, $CDCl_3$).

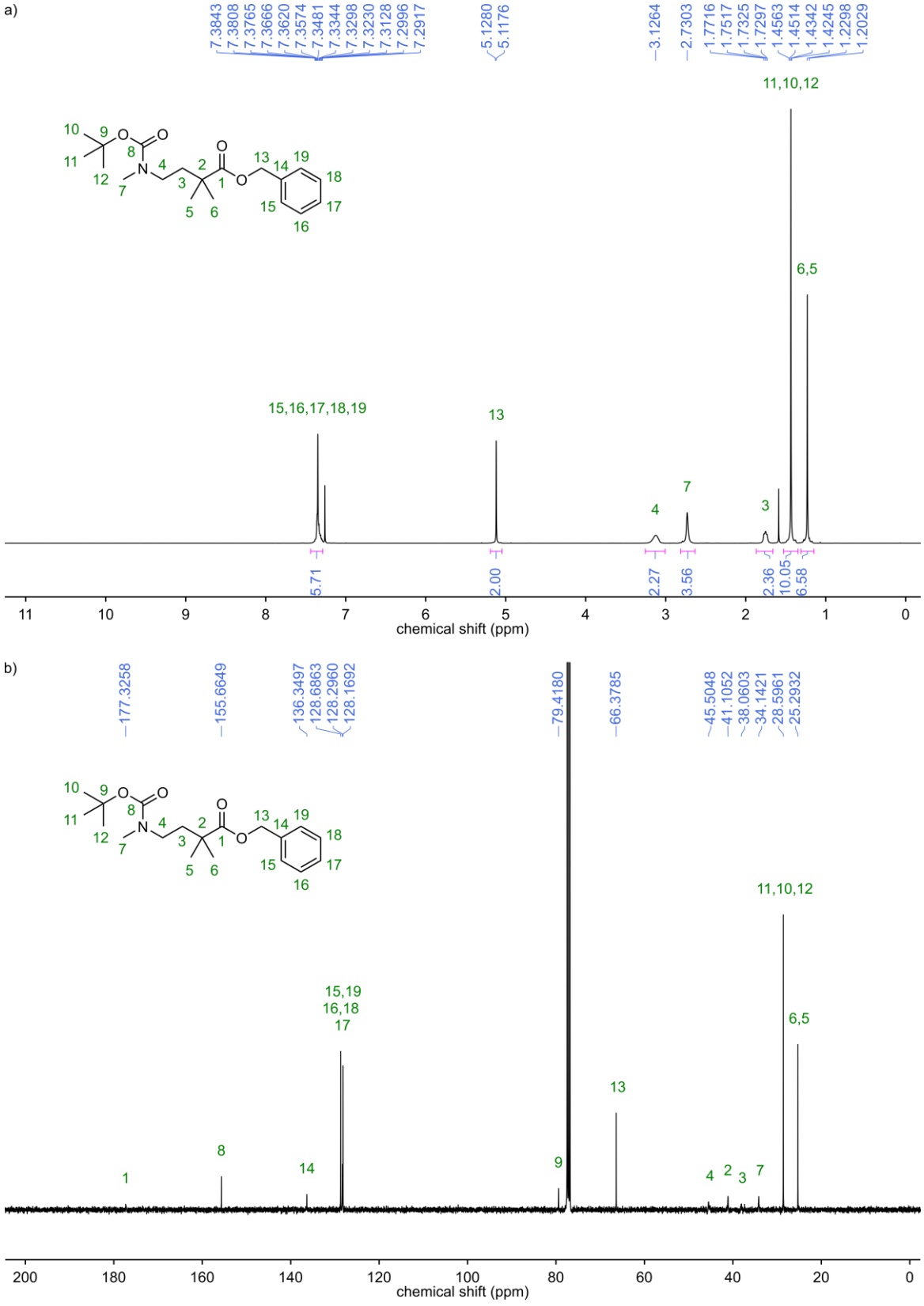


Fig. 135 NMR spectra of compound **13**. a) ^{1}H NMR (400 MHz, $CDCl_3$). b) ^{13}C NMR (101 MHz, $CDCl_3$).

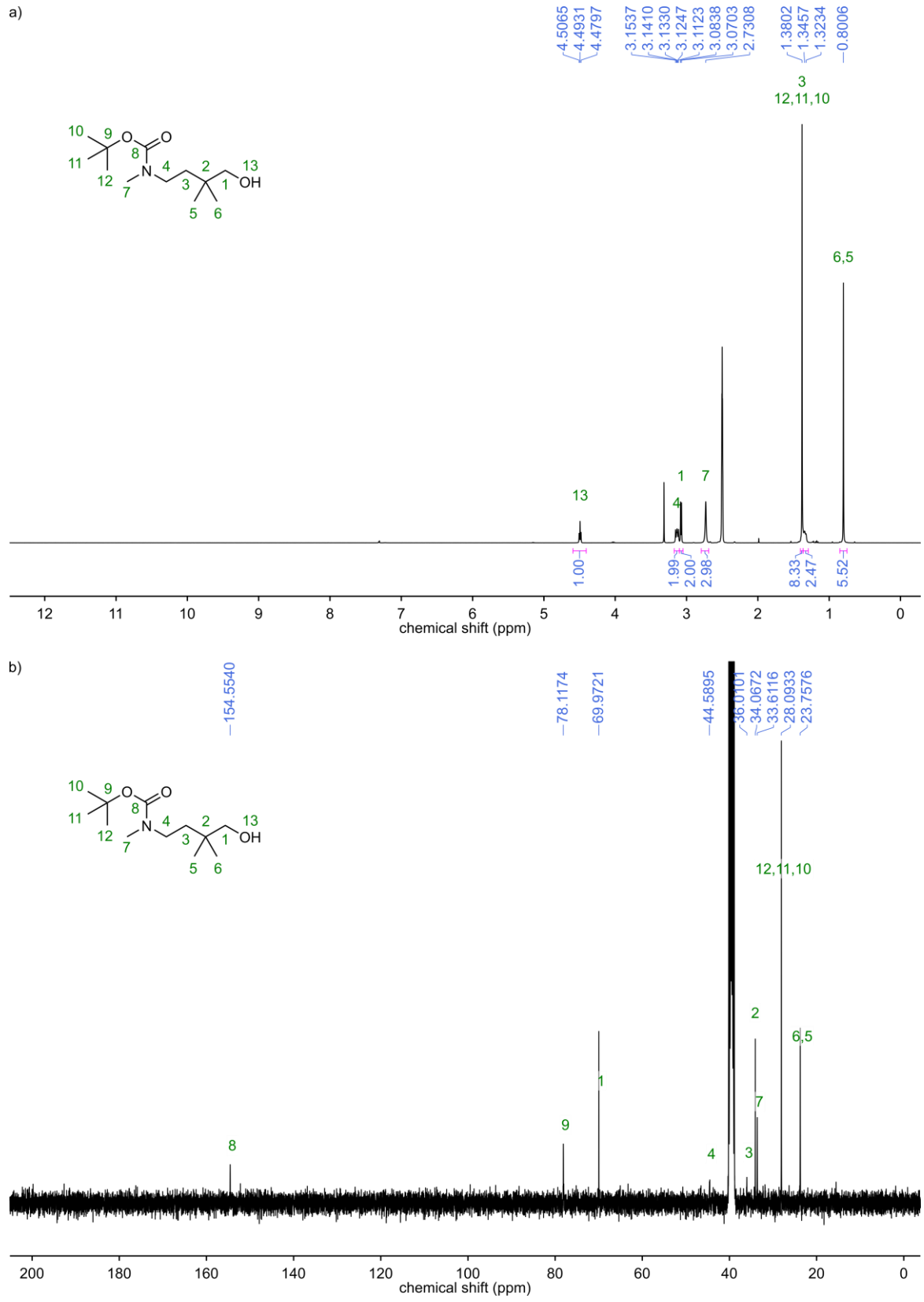


Fig. 136 NMR spectra of compound **14**. a) ^{1}H NMR (400 MHz, DMSO-d$_6$). b) ^{13}C NMR (101 MHz, DMSO-d$_6$).

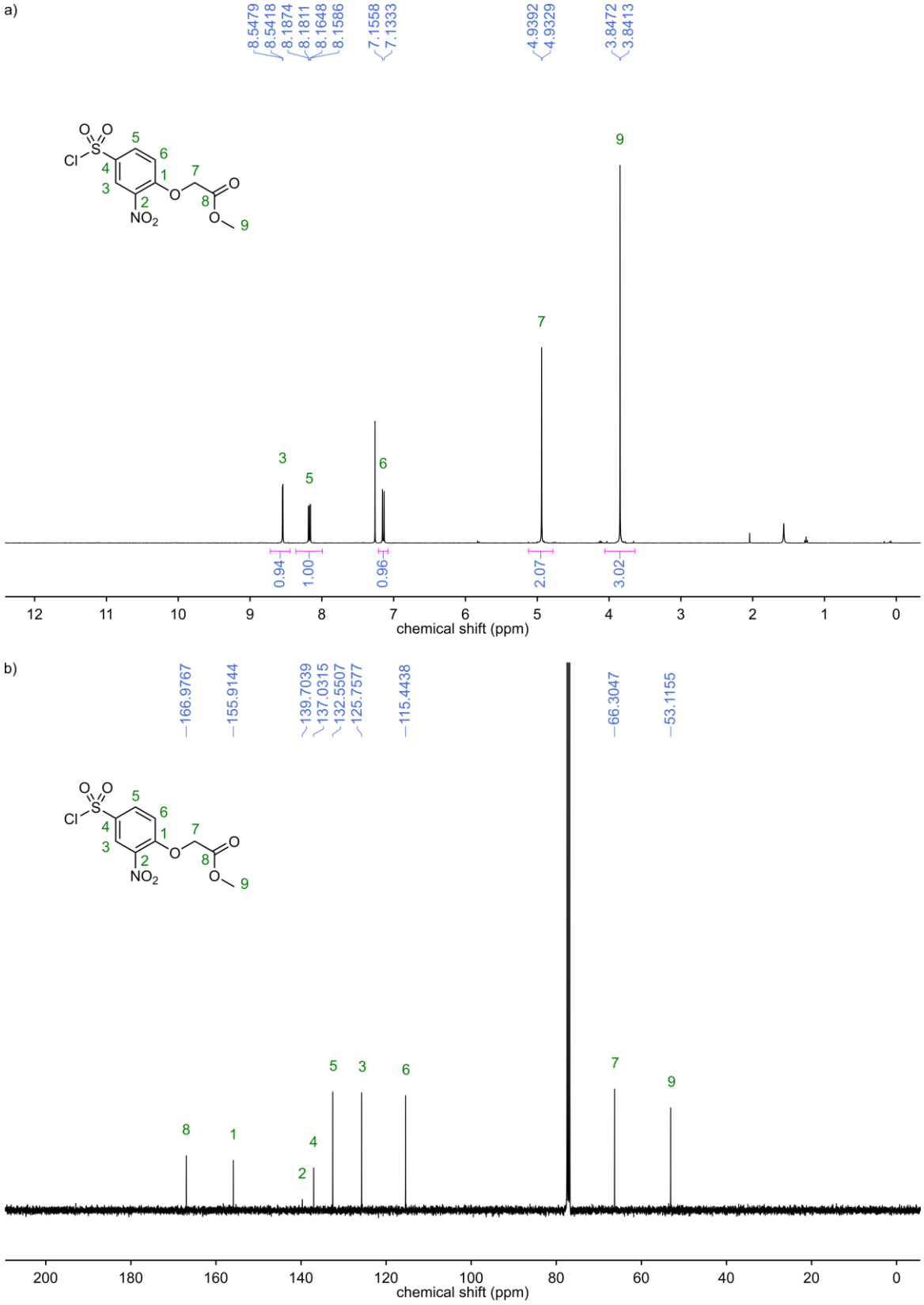


Fig. 137 NMR spectra of compound **16**. a) ^{1}H NMR (400 MHz, $CDCl_3$). b) ^{13}C NMR (101 MHz, $CDCl_3$).

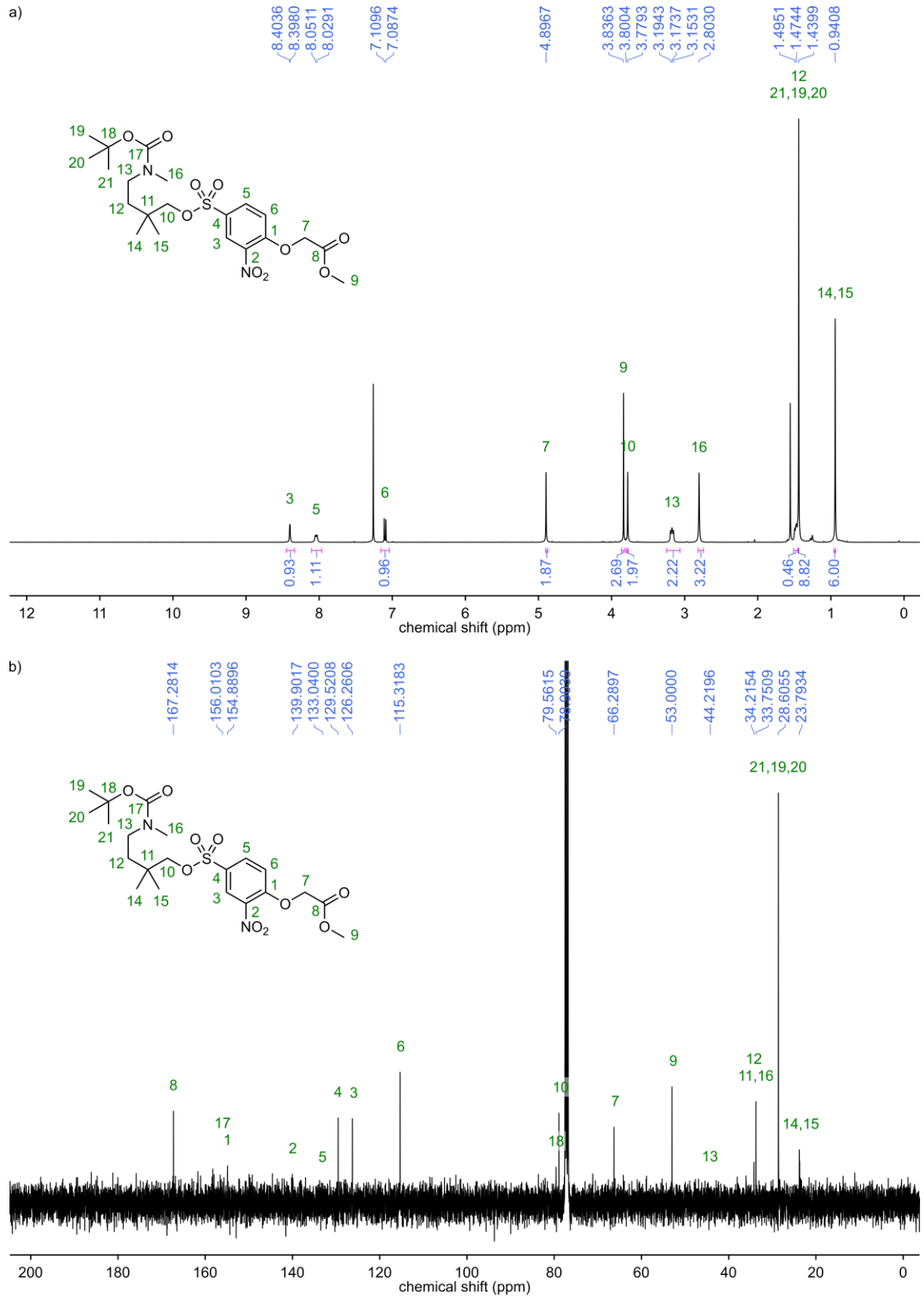


Fig. 138 NMR spectra of compound **17**. a) ^{1}H NMR (400 MHz, $CDCl_3$). b) ^{13}C NMR (101 MHz, $CDCl_3$).

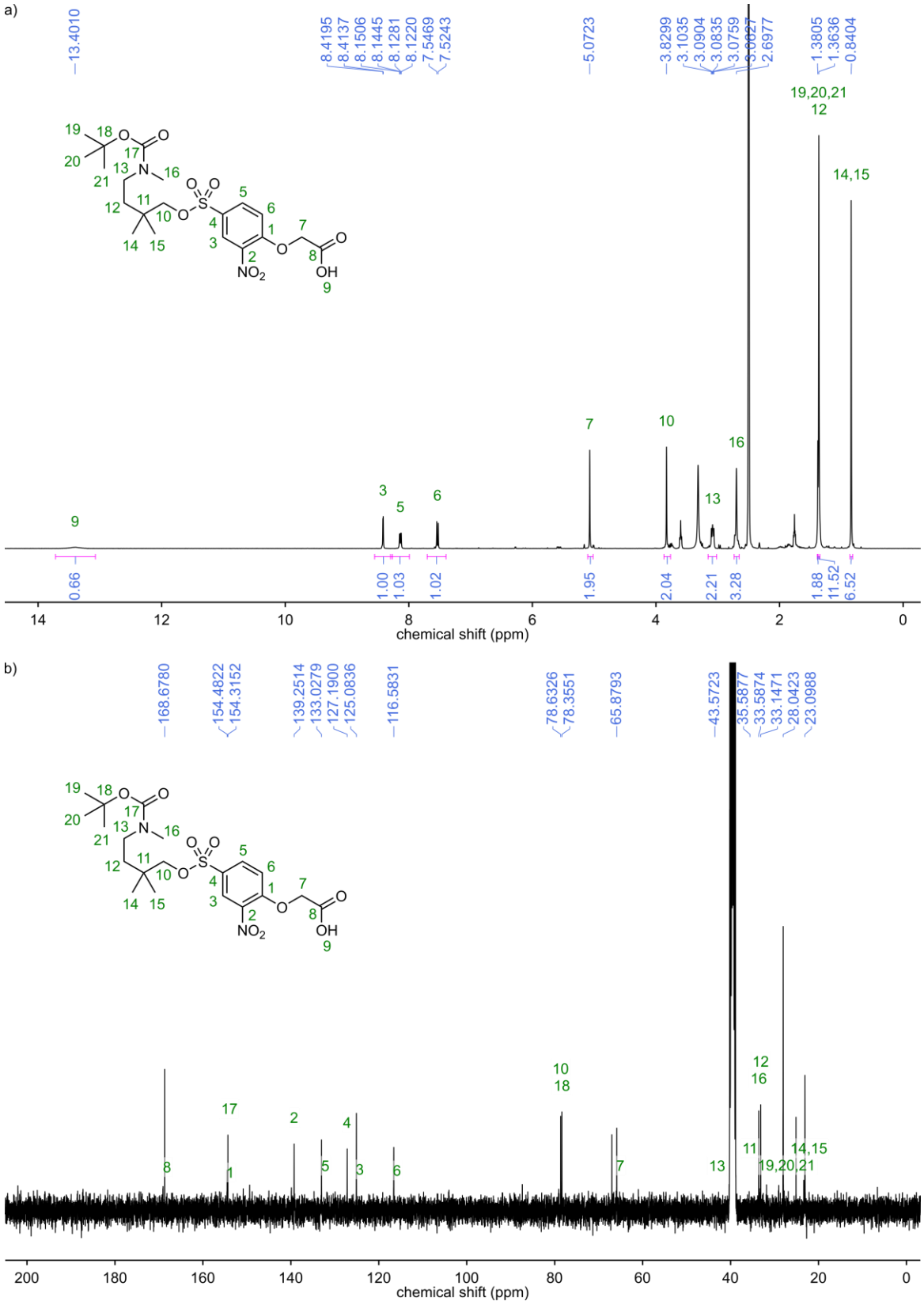


Fig. 139 NMR spectra of compound **18**. a) ^{1}H NMR (400 MHz, DMSO-d$_6$). b) ^{13}C NMR (101 MHz, DMSO-d$_6$).

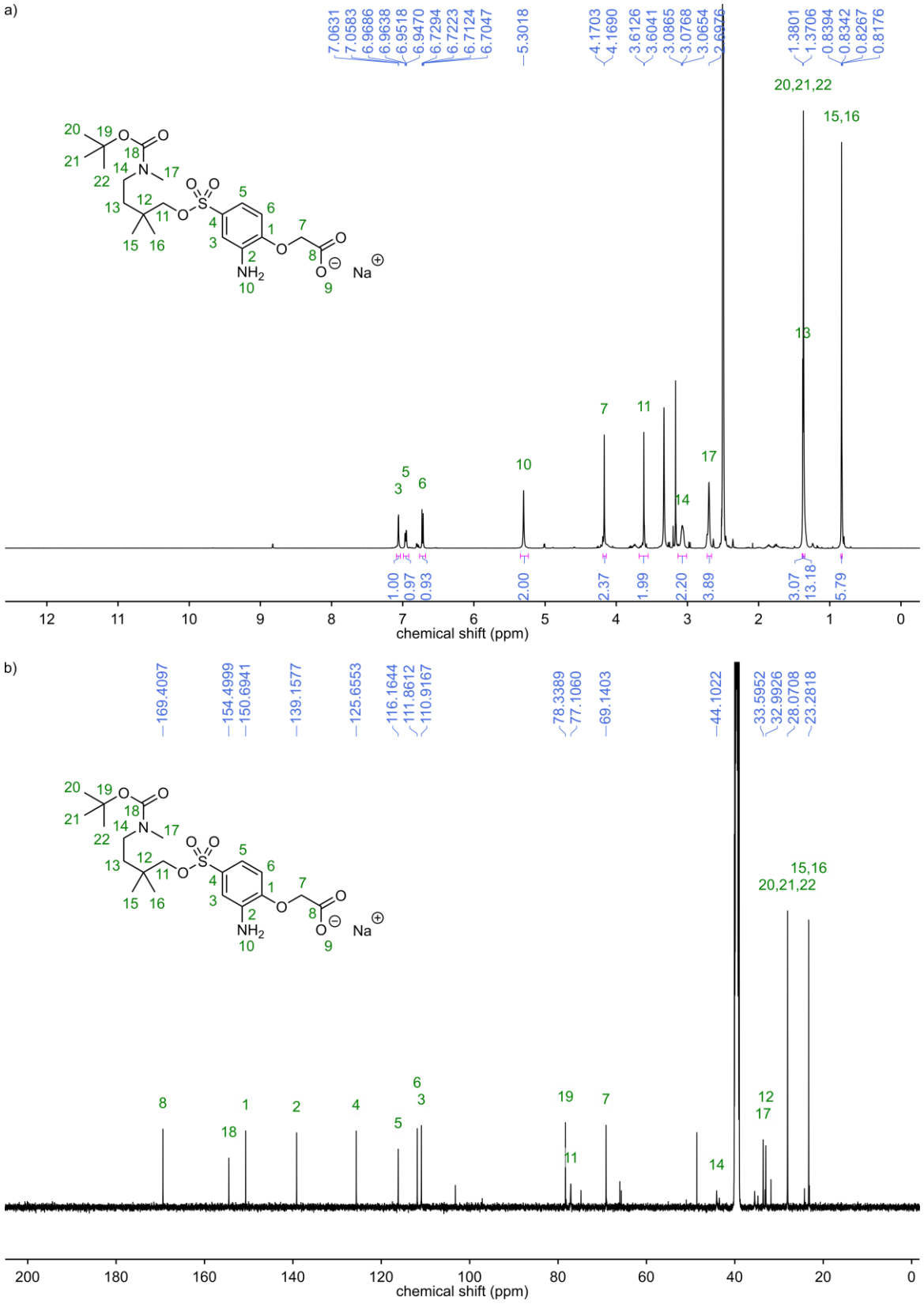


Fig. 140 NMR spectra of compound **19**. a) ^{1}H NMR (500 MHz, DMSO-d$_6$). b) ^{13}C NMR (126 MHz, DMSO-d$_6$).

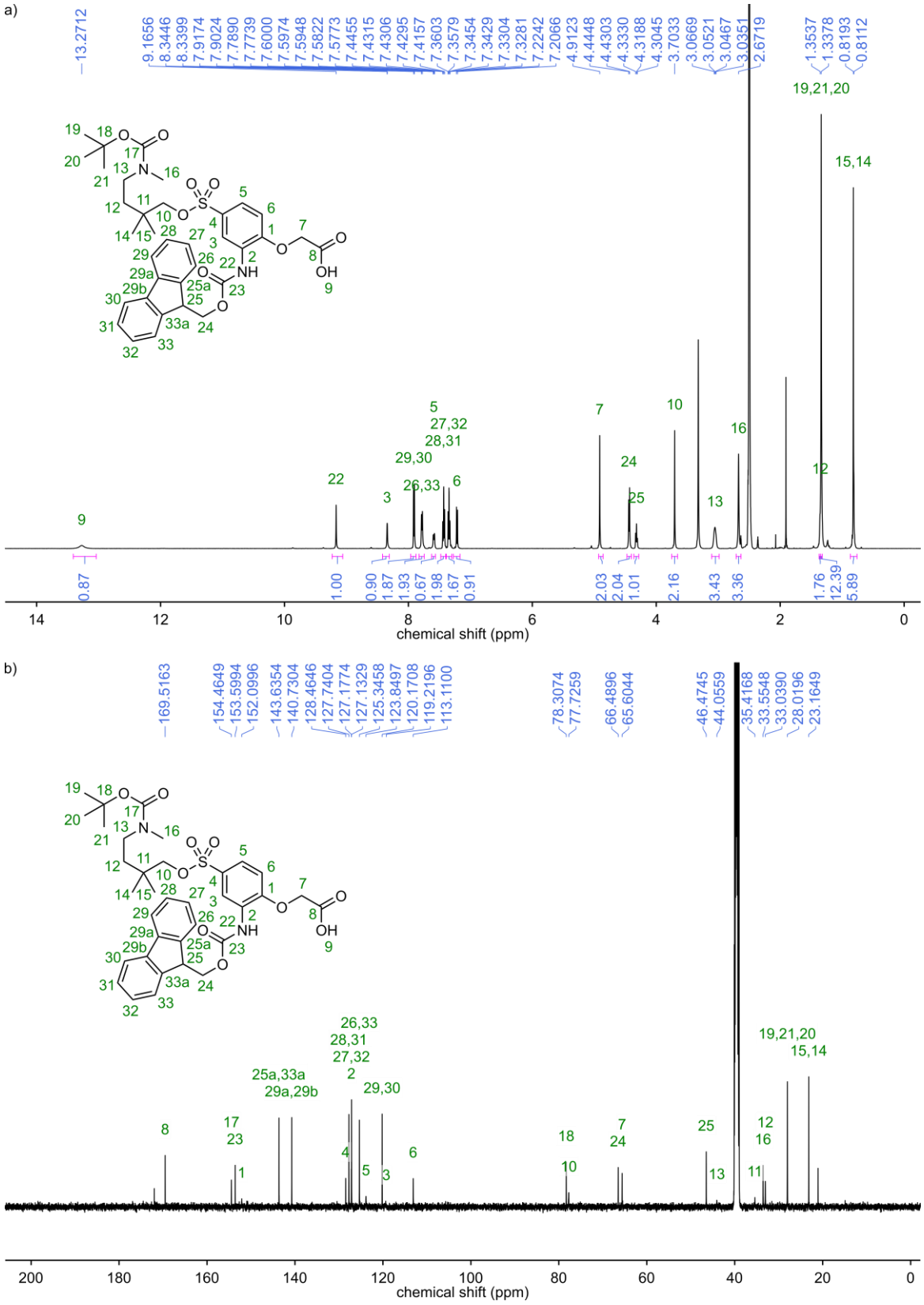


Fig. 141 NMR spectra of compound **20**. a) ^{1}H NMR (500 MHz, DMSO-d$_6$). b) ^{13}C NMR (126 MHz, DMSO-d$_6$).

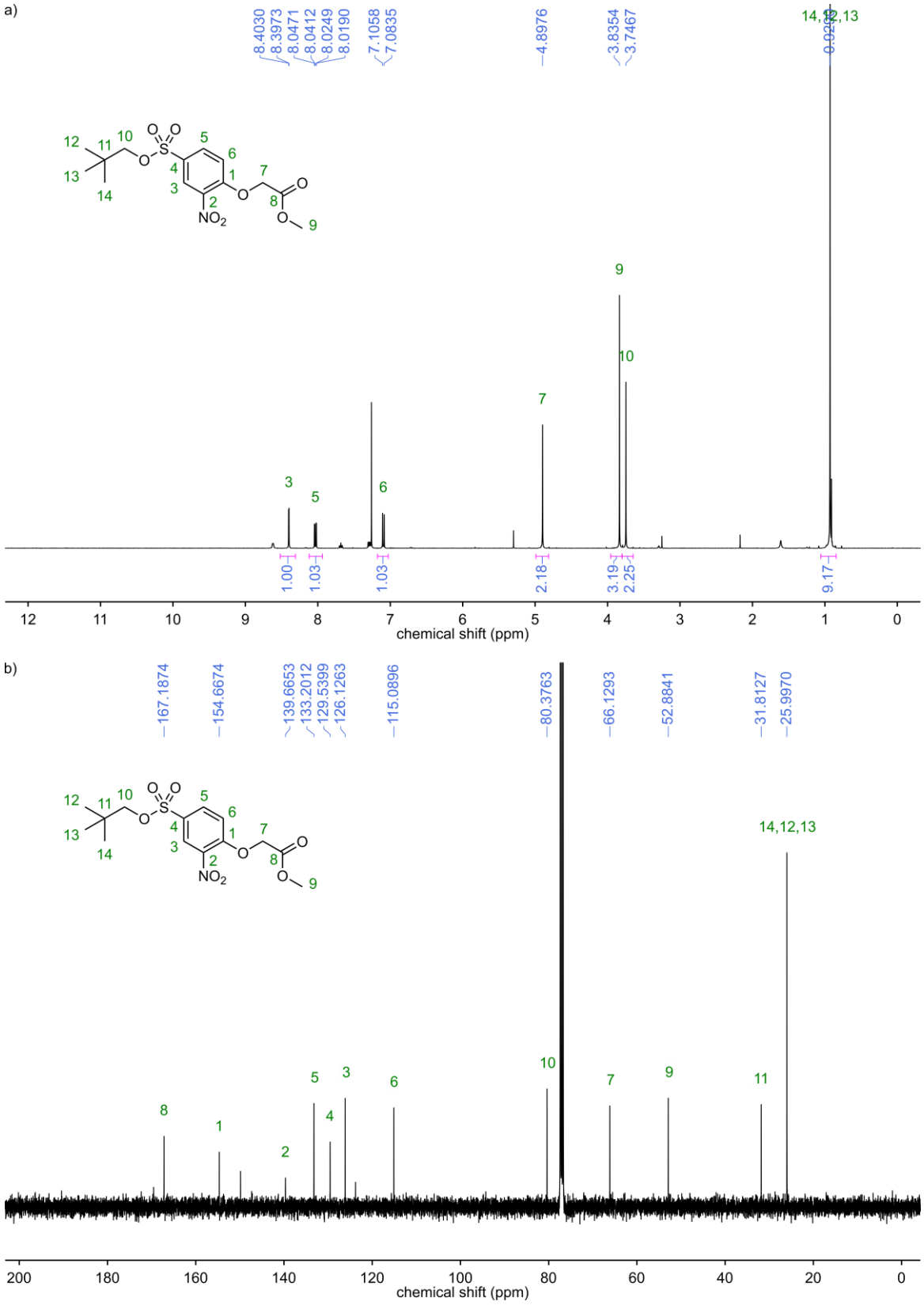


Fig. 142 NMR spectra of compound **21**. a) ^{1}H NMR (400 MHz, $CDCl_3$). b) ^{13}C NMR (101 MHz, $CDCl_3$).

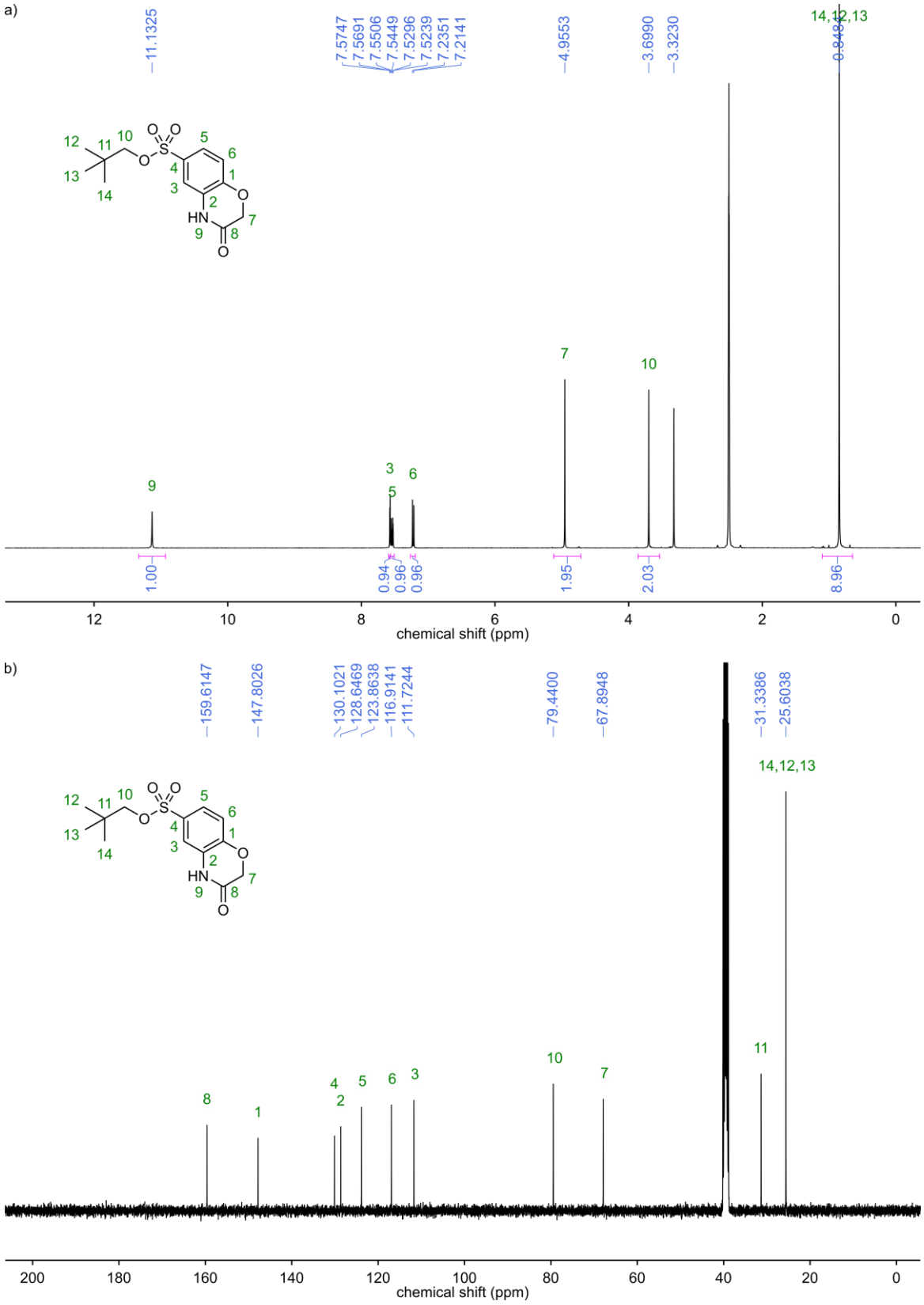


Fig. 143 NMR spectra of compound **22**. a) ^{1}H NMR (400 MHz, DMSO-d$_6$). b) ^{13}C NMR (101 MHz, DMSO-d$_6$).

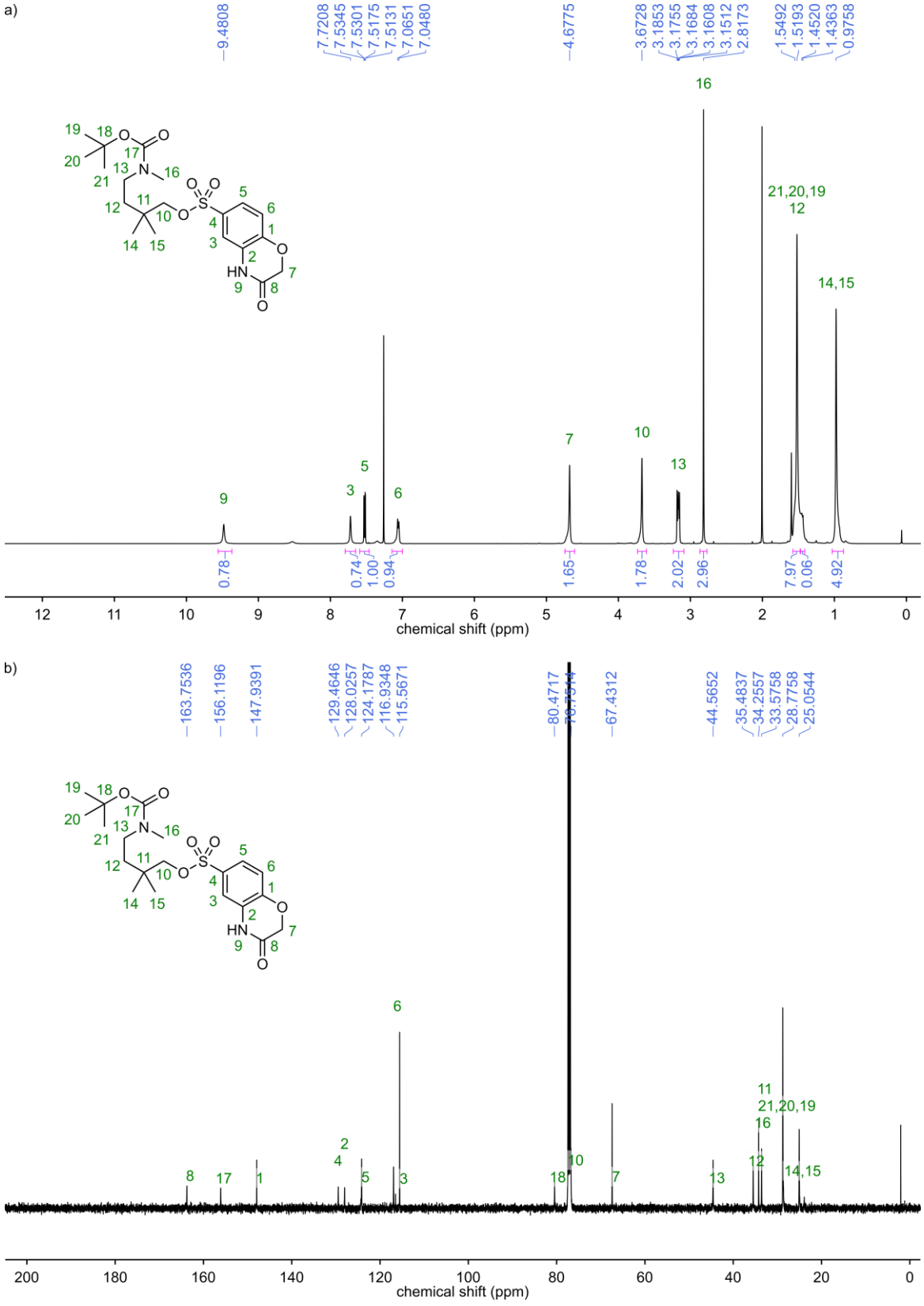


Fig. 144 NMR spectra of compound **23**. a) ^{1}H NMR (500 MHz, $CDCl_3$). b) ^{13}C NMR (126 MHz, $CDCl_3$).

9 References

[1] B. Frolund, B. Ebert, U. Kristiansen, T. Liljefors, P. Krogsgaard-Larsen, *Curr. Top. Med. Chem.* **2002**, *2*, 817–832.

[2] a) P. E. Nielsen, M. Egholm, R. H. Berg, O. Buchardt, *Science* **1991**, *254*, 1497–1500; b) R. J. Simon, R. S. Kania, R. N. Zuckermann, V. D. Huebner, D. A. Jewell, S. Banville, S. Ng, L. Wang, S. Rosenberg, C. K. Marlowe, *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9367–9371.

[3] A. S. Culf, *Biopolymers* **2019**, *110*, e23285.

[4] S. Shakeel, S. Karim, A. Ali, *J. Chem. Technol. Biotechnol.* **2006**, *81*, 892–899.

[5] D. Seebach, M. Overhand, F. N. M. Kühnle, B. Martinoni, L. Oberer, U. Hommel, H. Widmer, *Helv. Chim. Acta* **1996**, *79*, 913–941.

[6] D. H. Appella, L. A. Christianson, I. L. Karle, D. R. Powell, S. H. Gellman, *J. Am. Chem. Soc.* **1996**, *118*, 13071–13072.

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

[8] D. M. Bassani, J.-M. Lehn, G. Baum, D. Fenske, *Angew. Chem. Int. Ed.* **1997**, *36*, 1845–1847.

[9] L. Sebaoun, V. Maurizot, T. Granier, B. Kauffmann, I. Huc, *J. Am. Chem. Soc.* **2014**, *136*, 2168–2174.

[10] C. G. Pappas, P. K. Mandal, B. Liu, B. Kauffmann, X. Miao, D. Komáromy, W. Hoffmann, C. Manz, R. Chang, K. Liu, K. Pagel, I. Huc, S. Otto, *Nat. Chem.* **2020**, *12*, 1180–1186.

[11] a) D. Seebach, D. F. Hook, A. Glättli, *Biopolymers* **2006**, *84*, 23–37; b) M. D. Smith, T. D. W. Claridge, G. W. J. Fleet, G. E. Tranter, M. S. P. Sansom, *Chem. Commun.* **1998**, 2041–2042; c) H.-D. Arndt, B. Ziemer, U. Koert, *Org. Lett.* **2004**, *6*, 3269–3272; d) V. Semetey, D. Rognan, C. Hemmerlin, R. Graff, J.-P. Briand, M. Marraud, G. Guichard, *Angew. Chem. Int. Ed.* **2002**, *41*, 1893–1895; e) A. Salaün, M. Potel, T. Roisnel, P. Gall, P. Le Grel, *J. Org. Chem.* **2005**, *70*, 6499–6502; f) X. Li, D. Yang, *Chem. Commun.* **2006**, 3367–3379; g) M. Gude, U. Piarulli, D. Potenza, B. Salom, C. Gennari, *Tetrahedron Lett.* **1996**, *37*, 8589–8592.

[12] J. C. Nelson, J. G. Saven, J. S. Moore, P. G. Wolynes, *Science* **1997**, *277*, 1793–1796.

[13] R. Scott Lokey, B. L. Iverson, *Nature* **1995**, *375*, 303–305.

[14] Y. Hamuro, S. J. Geib, A. D. Hamilton, *Angew. Chem. Int. Ed.* **1994**, *33*, 446–448.

[15] X. Hu, P. K. Mandal, B. Kauffmann, I. Huc, *ChemPlusChem* **2020**, *85*, 1580–1586.

[16] J. M. Rogers, S. Kwon, S. J. Dawson, P. K. Mandal, H. Suga, I. Huc, *Nat. Chem.* **2018**, *10*, 405–412.

[17] K. Urushibara, Y. Ferrand, Z. Liu, K. Katagiri, M. Kawahata, E. Morvan, R. D'Elia, V. Pophristic, A. Tanatani, I. Huc, *Chem. Eur. J.* **2021**, *27*, 11205–11215.

[18] B. Huang, M. A. Prantil, T. L. Gustafson, J. R. Parquette, *J. Am. Chem. Soc.* **2003**, *125*, 14518–14530.

[19] E. Yashima, K. Maeda, H. Iida, Y. Furusho, K. Nagai, *Chemical Reviews* **2009**, *109*, 6102–6211.

[20] Y. Hamuro, S. J. Geib, A. D. Hamilton, *J. Am. Chem. Soc.* **1996**, *118*, 7529–7541.

[21] Y. Yan, B. Qin, Y. Shu, X. Chen, Y. K. Yip, D. Zhang, H. Su, H. Zeng, *Org. Lett.* **2009**, *11*, 1201–1204.

[22] J. T. Ernst, J. Becerril, H. S. Park, H. Yin, A. D. Hamilton, *Angew. Chem. Int. Ed.* **2003**, *42*, 535–539.

[23] H. Jiang, J.-M. Léger, I. Huc, *J. Am. Chem. Soc.* **2003**, *125*, 3448–3449.

[24] A. Zhang, J. S. Ferguson, K. Yamato, C. Zheng, B. Gong, *Org. Lett.* **2006**, *8*, 5117–5120.

[25] D. Sánchez-García, B. Kauffmann, T. Kawanami, H. Ihara, M. Takafuji, M.-H. l. n. Delville, I. Huc, *J. Am. Chem. Soc.* **2009**, *131*, 8642–8648.

[26] a) B. t. Baptiste, C. l. Douat-Casassus, K. Laxmi-Reddy, F. d. r. Godde, I. Huc, *J. Org. Chem.* **2010**, *75*, 7175–7185; b) X. Hu, S. J. Dawson, Y. Nagaoka, A. Tanatani, I. Huc, *J. Org. Chem.* **2016**, *81*, 1137–1150.

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

[28] D. H. Appella, L. A. Christianson, D. A. Klein, D. R. Powell, X. Huang, J. J. Barchi, S. H. Gellman, *Nature* **1997**, *387*, 381–384.

[29] H. S. Chan, K. A. Dill, *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 447–490.

[30] K. Yoshida, S.-I. Kawamura, T. Morita, S. Kimura, *J. Am. Chem. Soc.* **2006**, *128*, 8034–8041.

[31] J. Brüggemann, S. Bitter, S. Müller, W. M. Müller, U. Müller, N. M. Maier, W. Lindner, F. Vögtle, *Angew. Chem. Int. Ed.* **2007**, *46*, 254–259.

[32] C. A. Hunter, A. Spitaleri, S. Tomas, *Chem. Commun.* **2005**, 3691–3693.

[33] a) N. Delsuc, F. Godde, B. Kauffmann, J.-M. Léger, I. Huc, *J. Am. Chem. Soc.* **2007**, *129*, 11348–11349; b) D. Seebach, S. Abele, K. Gademann, G. Guichard, T. Hintermann, B. Jaun, J. L.

Matthews, J. V. Schreiber, L. Oberer, U. Hommel, H. Widmer, *Helv. Chim. Acta* **1998**, *81*, 932–982.
[34] a) C. Dolain, V. Maurizot, I. Huc, *Angew. Chem. Int. Ed.* **2003**, *42*, 2738–2740; b) D. Kanamori, T.-A. Okamura, H. Yamamoto, N. Ueyama, *Angew. Chem. Int. Ed.* **2005**, *44*, 969–972.
[35] J. Solà, S. P. Fletcher, A. Castellanos, J. Clayden, *Angew. Chem. Int. Ed.* **2010**, *49*, 6836–6839.
[36] M. Barboiu, J.-M. Lehn, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5201–5206.
[37] J. F. Malone, C. M. Murray, G. M. Dolan, R. Docherty, A. J. Lavery, *Chem. Mater.* **1997**, *9*, 2983–2989.
[38] T. Qi, V. Maurizot, H. Noguchi, T. Charoenraks, B. Kauffmann, M. Takafuji, H. Ihara, I. Huc, *Chem. Commun.* **2012**, *48*, 6337–6339.
[39] M. Jewginski, L. Fischer, C. Colombo, I. Huc, C. D. Mackereth, *ChemBioChem* **2016**, *17*, 727–736.
[40] M. Kudo, D. Carbajo López, V. Maurizot, H. Masu, A. Tanatani, I. Huc, *Eur. J. Org. Chem.* **2016**, *2016*, 2457–2466.
[41] M. Vallade, P. Sai Reddy, L. Fischer, I. Huc, *Eur. J. Org. Chem.* **2018**, *2018*, 5489–5498.
[42] S. De, B. Chi, T. Granier, T. Qi, V. Maurizot, I. Huc, *Nat. Chem.* **2018**, *10*, 51–57.
[43] a) M. Akazome, Y. Ishii, T. Nireki, K. Ogura, *Tetrahedron Lett.* **2008**, *49*, 4430–4433; b) M. Akazome, J. Sukegawa, Y. Goto, S. Matsumoto, *Tetrahedron Lett.* **2009**, *50*, 5382–5385.
[44] Y. Ferrand, A. M. Kendhale, B. Kauffmann, A. GréLard, C. C. Marie, V. Blot, M. Pipelier, D. Dubreuil, I. Huc, *J. Am. Chem. Soc.* **2010**, *132*, 7858–7859.
[45] a) E. Yashima, N. Ousaka, D. Taura, K. Shimomura, T. Ikai, K. Maeda, *Chem. Rev.* **2016**, *116*, 13752–13990; b) Z. Liu, X. Hu, A. M. Abramyan, Á. Mészáros, M. Csekei, A. Kotschy, I. Huc, V. Pophristic, *Chem. Eur. J.* **2017**, *23*, 3605–3615.
[46] a) A. M. Kendhale, L. Poniman, Z. Dong, K. Laxmi-Reddy, B. Kauffmann, Y. Ferrand, I. Huc, *J. Org. Chem.* **2011**, *76*, 195–200; b) L. Yang, C. Ma, B. Kauffmann, D. Li, Q. Gan, *Org. Biomol. Chem.* **2020**, *18*, 6643–6650; c) L. Zheng, Y. Zhan, C. Yu, F. Huang, Y. Wang, H. Jiang, *Org. Lett.* **2017**, *19*, 1482–1485.
[47] W. S. Horne, S. H. Gellman, *Acc. Chem. Res.* **2008**, *41*, 1399–1408.
[48] G. V. M. Sharma, K. R. Reddy, P. R. Krishna, A. R. Sankar, K. Narsimulu, S. K. Kumar, P. Jayaprakash, B. Jagannadh, A. C. Kunwar, *J. Am. Chem. Soc.* **2003**, *125*, 13670–13671.
[49] S. Dengler, P. K. Mandal, L. Allmendinger, C. Douat, I. Huc, *Chem. Sci.* **2021**, *12*, 11004–11012.
[50] C. Bao, B. Kauffmann, Q. Gan, K. Srinivas, H. Jiang, I. Huc, *Angew. Chem. Int. Ed.* **2008**, *47*, 4153–4156.
[51] Z. C. Girvin, M. K. Andrews, X. Liu, S. H. Gellman, *Science* **2019**, *366*, 1528–1531.
[52] a) Z. Zhong, Y. Zhao, *Org. Lett.* **2007**, *9*, 2891–2894; b) P. Mateus, A. Jacquet, A. Méndez-Ardoy, A. Boulloy, B. Kauffmann, G. Pecastaings, T. Buffeteau, Y. Ferrand, D. M. Bassani, I. Huc, *Chem. Sci.* **2021**, *12*, 3743–3750.
[53] V. E. Campbell, X. De Hatten, N. Delsuc, B. Kauffmann, I. Huc, J. R. Nitschke, *Nat. Chem.* **2010**, *2*, 684–687.
[54] Q. Gan, Y. Ferrand, C. Bao, B. Kauffmann, A. Grelard, H. Jiang, I. Huc, *Science* **2011**, *331*, 1172–1175.
[55] K. Ziach, C. Chollet, V. Parissi, P. Prabhakaran, M. Marchivie, V. Corvaglia, P. P. Bose, K. Laxmi-Reddy, F. Godde, J.-M. Schmitter, S. Chaignepain, P. Pourquier, I. Huc, *Nat. Chem.* **2018**, *10*, 511–518.
[56] a) J. Iriondo-Alberdi, K. Laxmi-Reddy, B. Bouguerne, C. Staedel, I. Huc, *ChemBioChem* **2010**, *11*, 1679–1685; b) M. Rueping, Y. Mahajan, M. Sauer, D. Seebach, *ChemBioChem* **2002**, *3*, 257–259; c) T. B. Potocky, A. K. Menon, S. H. Gellman, *J. Am. Chem. Soc.* **2005**, *127*, 3686–3687.
[57] J. Fremaux, L. Mauran, K. Pulka‐Ziach, B. Kauffmann, B. Odaert, G. Guichard, *Angew. Chem. Int. Ed.* **2015**, *54*, 9816–9820.
[58] M. Zwillinger, P. S. Reddy, B. Wicher, P. K. Mandal, M. Csékei, L. Fischer, A. Kotschy, I. Huc, *Chem. Eur. J.* **2020**, *26*, 17366–17370.
[59] M. Vallade, M. Jewginski, L. Fischer, J. Buratto, K. Bathany, J.-M. Schmitter, M. Stupfel, F. Godde, C. D. Mackereth, I. Huc, *Bioconjugate Chem.* **2019**, *30*, 54–62.
[60] J. Zhu, R. D. Parra, H. Zeng, E. Skrzypczak-Jankun, X. C. Zeng, B. Gong, *J. Am. Chem. Soc.* **2000**, *122*, 4219–4220.
[61] F. Devaux, X. Li, D. Sluysmans, V. Maurizot, E. Bakalis, F. Zerbetto, I. Huc, A.-S. Duwez, *Chem* **2021**, *7*, 1333–1346.
[62] D. Zheng, L. Zheng, C. Yu, Y. Zhan, Y. Wang, H. Jiang, *Org. Lett.* **2019**, *21*, 2555–2559.
[63] J. Wang, B. Wicher, A. Méndez‐Ardoy, X. Li, G. Pecastaings, T. Buffeteau, D. M. Bassani, V. Maurizot, I. Huc, *Angew. Chem. Int. Ed.* **2021**, *60*, 18461–18466.

[64] X. Hu, S. J. Dawson, P. K. Mandal, X. De Hatten, B. Baptiste, I. Huc, *Chem. Sci.* **2017**, *8*, 3741–3749.
[65] P. B. Harbury, T. Zhang, P. S. Kim, T. Alber, *Science* **1993**, *262*, 1401–1407.
[66] F. H. C. Crick, *Acta Crystallogr.* **1953**, *6*, 689–697.
[67] L. Pauling, R. B. Corey, *Nature* **1953**, *171*, 59–61.
[68] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, *Nature* **2021**, *596*, 583–589.
[69] V. Koronakis, A. Sharff, E. Koronakis, B. Luisi, C. Hughes, *Nature* **2000**, *405*, 914–919.
[70] a) D. Mazzier, S. De, B. Wicher, V. Maurizot, I. Huc, *Chem. Sci.* **2019**, *10*, 6984–6991; b) D. Mazzier, S. De, B. Wicher, V. Maurizot, I. Huc, *Angew. Chem. Int. Ed.* **2020**, *59*, 1606–1610.
[71] K. L. George, W. S. Horne, *J. Am. Chem. Soc.* **2017**, *139*, 7931–7938.
[72] W. S. Horne, J. L. Price, J. L. Keck, S. H. Gellman, *J. Am. Chem. Soc.* **2007**, *129*, 4178–4180.
[73] G. W. Collie, K. Pulka-Ziach, C. M. Lombardo, J. Fremaux, F. Rosu, M. Decossas, L. Mauran, O. Lambert, V. Gabelica, C. D. Mackereth, G. Guichard, *Nat. Chem.* **2015**, *7*, 871–878.
[74] O. Khakshoor, B. Demeler, J. S. Nowick, *J. Am. Chem. Soc.* **2007**, *129*, 5558–5569.
[75] J. Shang, Q. Gan, S. J. Dawson, F. Rosu, H. Jiang, Y. Ferrand, I. Huc, *Org. Lett.* **2014**, *16*, 4992–4995.
[76] M. Li, K. Yamato, J. S. Ferguson, B. Gong, *J. Am. Chem. Soc.* **2006**, *128*, 12628–12629.
[77] Y. Yan, B. Qin, C. Ren, X. Chen, Y. K. Yip, R. Ye, D. Zhang, H. Su, H. Zeng, *J. Am. Chem. Soc.* **2010**, *132*, 5869–5879.
[78] H. Goto, H. Katagiri, Y. Furusho, E. Yashima, *J. Am. Chem. Soc.* **2006**, *128*, 7176–7178.
[79] a) H. Zhao, W. Q. Ong, F. Zhou, X. Fang, X. Chen, S. F. Y. Li, H. Su, N.-J. Cho, H. Zeng, *Chem. Sci.* **2012**, *3*, 2042–2046; b) H. Zhao, S. Sheng, Y. Hong, H. Zeng, *J. Am. Chem. Soc.* **2014**, *136*, 14270–14276; c) Y. Huo, H. Zeng, *Acc. Chem. Res.* **2016**, *49*, 922–930; d) A. Roy, H. Joshi, R. Ye, J. Shen, F. Chen, A. Aksimentiev, H. Zeng, *Angew. Chem. Int. Ed.* **2020**, *59*, 4806–4813.
[80] C. M. Dobson, *Nature* **2003**, *426*, 884–890.
[81] a) S. H. Gellman, *Acc. Chem. Res.* **1998**, *31*, 173–180; b) D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes, J. S. Moore, *Chem. Rev.* **2001**, *101*, 3893–4012; c) G. Guichard, I. Huc, *Chem. Commun.* **2011**, *47*, 5933–5941.
[82] M. S. Newman, D. Lednicer, *J. Am. Chem. Soc.* **1956**, *78*, 4765–4770.
[83] J.-L. Schmitt, A.-M. Stadler, N. Kyritsakas, J.-M. Lehn, *Helv. Chim. Acta* **2003**, *86*, 1598–1624.
[84] A. Tanatani, H. Kagechika, I. Azumaya, R. Fukutomi, Y. Ito, K. Yamaguchi, K. Shudo, *Tetrahedron Lett.* **1997**, *38*, 4425–4428.
[85] a) V. Berl, I. Huc, R. G. Khoury, M. J. Krische, J.-M. Lehn, *Nature* **2000**, *407*, 720–723; b) A. Tanatani, A. Yokoyama, I. Azumaya, Y. Takakura, C. Mitsui, M. Shiro, M. Uchiyama, A. Muranaka, N. Kobayashi, T. Yokozawa, *J. Am. Chem. Soc.* **2005**, *127*, 8553–8561.
[86] a) U. Koert, *J. Prakt. Chem.* **2000**, *342*, 325–333; b) R. P. Cheng, S. H. Gellman, W. F. Degrado, *Chem. Rev.* **2001**, *101*, 3219–3232.
[87] I. Huc, *Eur. J. Org. Chem.* **2004**, *2004*, 17–29.
[88] a) A. M. Ortuño, P. Reiné, S. Resa, L. Alvarez De Cienfuegos, V. Blanco, J. M. Paredes, A. J. Mota, G. Mazzeo, S. Abbate, J. M. Ugalde, V. Mujica, G. Longhi, D. Miguel, J. M. Cuerva, *Org. Chem. Front.* **2021**, *8*, 5071–5086; b) S. Pramanik, B. Kauffmann, S. Hecht, Y. Ferrand, I. Huc, *Chem. Commun.* **2021**, *57*, 93–96.
[89] a) L. Chen, Y.-H. Wang, B. He, H. Nie, R. Hu, F. Huang, A. Qin, X.-S. Zhou, Z. Zhao, B. Z. Tang, *Angew. Chem. Int. Ed.* **2015**, *127*, 4305–4309; b) A. Méndez-Ardoy, N. Markandeya, X. Li, Y.-T. Tsai, G. Pecastaings, T. Buffeteau, V. Maurizot, L. Muccioli, F. Castet, I. Huc, D. M. Bassani, *Chem. Sci.* **2017**, *8*, 7251–7257; c) J. Nejedlý, M. Šámal, J. Rybáček, M. Tobrmanová, F. Szydlo, C. Coudret, M. Neumeier, J. Vacek, J. Vacek Chocholoušová, M. Buděšínský, D. Šaman, L. Bednárová, L. Sieger, I. G. Stará, I. Starý, *Angew. Chem. Int. Ed.* **2017**, *56*, 5839–5843.
[90] a) R. B. Prince, S. A. Barnes, J. S. Moore, *J. Am. Chem. Soc.* **2000**, *122*, 2758–2762; b) J. Garric, J.-M. Léger, I. Huc, *Chem. Eur. J.* **2007**, *13*, 8454–8462; c) W. Q. Ong, H. Zhao, X. Fang, S. Woen, F. Zhou, W. Yap, H. Su, S. F. Y. Li, H. Zeng, *Org. Lett.* **2011**, *13*, 3194–3197; d) N. Chandramouli, Y. Ferrand, G. Lautrette, B. Kauffmann, C. D. Mackereth, M. Laguerre, D. Dubreuil, I. Huc, *Nat. Chem.* **2015**, *7*, 334–341.
[91] H. Jiang, J.-M. Léger, C. Dolain, P. Guionneau, I. Huc, *Tetrahedron* **2003**, *59*, 8365–8374.
[92] P. S. Reddy, B. Langlois d'Estaintot, T. Granier, C. D. Mackereth, L. Fischer, I. Huc, *Chem. Eur. J.* **2019**, *25*, 11042–11047.
[93] D. Bindl, E. Heinemann, P. K. Mandal, I. Huc, *Chem. Commun.* **2021**, *57*, 5662–5665.

[94] C. Dolain, J. M. Leger, N. Delsuc, H. Gornitzka, I. Huc, *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 16146–16151.
[95] R. Wechsel, J. Raftery, D. Cavagnat, G. Guichard, J. Clayden, *Angew. Chem. Int. Ed.* **2016**, *55*, 9657–9661.
[96] J. Buratto, C. Colombo, M. Stupfel, S. J. Dawson, C. Dolain, B. Langlois d'Estaintot, L. Fischer, T. Granier, M. Laguerre, B. Gallois, I. Huc, *Angew. Chem. Int. Ed.* **2014**, *53*, 883–887.
[97] a) L. Pohl, M. Eckle, *Angew. Chem. Int. Ed.* **1969**, *8*, 381–381; b) H. E. Gottlieb, V. Kotlyar, A. Nudelman, *J. Org. Chem.* **1997**, *62*, 7512–7515.
[98] S. Manabe, Y. Ito, *J. Am. Chem. Soc.* **2002**, *124*, 12638–12639.
[99] J. Jancarik, S.-H. Kim, *J. Appl. Crystallogr.* **1991**, *24*, 409–411.
[100] N. E. Chayen, *Structure* **1997**, *5*, 1269–1274.
[101] M. Cianci, G. Bourenkov, G. Pompidor, I. Karpics, J. Kallio, I. Bento, M. Roessle, F. Cipriani, S. Fiedler, T. R. Schneider, *J. Synchrotron Radiat.* **2017**, *24*, 323–332.
[102] G. Winter, *J. Appl. Crystallogr.* **2010**, *43*, 186–190.
[103] G. Winter, D. G. Waterman, J. M. Parkhurst, A. S. Brewster, R. J. Gildea, M. Gerstel, L. Fuentes-Montero, M. Vollmar, T. Michels-Clark, I. D. Young, N. K. Sauter, G. Evans, *Acta Crystallogr. D* **2018**, *74*, 85–97.
[104] a) P. Evans, *Acta Crystallogr. D* **2006**, *62*, 72–82; b) P. R. Evans, G. N. Murshudov, *Acta Crystallogr. D* **2013**, *69*, 1204–1214.
[105] G. M. Sheldrick, *Acta Crystallogr. A* **2008**, *64*, 112–122.
[106] G. M. Sheldrick, *Acta Crystallogr. C* **2015**, *71*, 3–8.
[107] O. V. Dolomanov, L. J. Bourhis, R. J. Gildea, J. A. K. Howard, H. Puschmann, *J. Appl. Crystallogr.* **2009**, *42*, 339–341.
[108] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. D* **2010**, *66*, 486–501.
[109] A. L. Spek, *Acta Crystallogr. D* **2009**, *65*, 148–155.
[110] E. Merlet, K. Moreno, A. Tron, N. McClenaghan, B. Kauffmann, Y. Ferrand, C. Olivier, *Chem. Commun.* **2019**, *55*, 9825–9828.
[111] L. Zheng, D. Zheng, Y. Wang, C. Yu, K. Zhang, H. Jiang, *Org. Biomol. Chem.* **2019**, *17*, 9573–9577.
[112] a) G. Lautrette, B. Wicher, B. Kauffmann, Y. Ferrand, I. Huc, *J. Am. Chem. Soc.* **2016**, *138*, 10314–10322; b) S. Saha, B. Kauffmann, Y. Ferrand, I. Huc, *Angew. Chem. Int. Ed.* **2018**, *57*, 13542–13546.
[113] J. M. Alex, V. Corvaglia, X. Hu, S. Engilberge, I. Huc, P. B. Crowley, *Chem. Commun.* **2019**, *55*, 11087–11090.
[114] A. Zhang, Y. Han, K. Yamato, X. C. Zeng, B. Gong, *Org. Lett.* **2006**, *8*, 803–806.
[115] a) Y. Inai, Y. Kurokawa, N. Kojima, *J. Chem. Soc., Perkin Trans. 2* **2002**, 1850–1857; b) Y. Inai, Y. Kurokawa, T. Hirabayashi, *Biopolymers* **1999**, *49*, 551–564; c) M. De Poli, M. De Zotti, J. Raftery, J. A. Aguilar, G. A. Morris, J. Clayden, *J. Org. Chem.* **2013**, *78*, 2248–2255; d) R. A. Brown, T. Marcelli, M. De Poli, J. Solà, J. Clayden, *Angew. Chem. Int. Ed.* **2012**, *51*, 1395–1399.
[116] a) Z. Dong, G. P. A. Yap, J. M. Fox, *J. Am. Chem. Soc.* **2007**, *129*, 11850–11853; b) H.-Y. Hu, J.-F. Xiang, Y. Yang, C.-F. Chen, *Org. Lett.* **2008**, *10*, 69–72; c) E. Kolomiets, V. Berl, J.-M. Lehn, *Chem. Eur. J.* **2007**, *13*, 5466–5479; d) M. Kudo, V. Maurizot, B. Kauffmann, A. Tanatani, I. Huc, *J. Am. Chem. Soc.* **2013**, *135*, 9628–9631.
[117] C. Dolain, H. Jiang, J.-M. Léger, P. Guionneau, I. Huc, *J. Am. Chem. Soc.* **2005**, *127*, 12943–12951.
[118] C. C. Mcdonald, W. D. Phillips, *J. Am. Chem. Soc.* **1967**, *89*, 6332–6341.
[119] Z. Xu, K. Li, R. Zhai, T. Liang, X. Gui, R. Zhang, *RSC Adv.* **2017**, *7*, 51972–51977.
[120] O. Tour, S. R. Adams, R. A. Kerr, R. M. Meijer, T. J. Sejnowski, R. W. Tsien, R. Y. Tsien, *Nat. Chem. Biol.* **2007**, *3*, 423–431.
[121] A. J. Mccoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr.* **2007**, *40*, 658–674.
[122] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. Mccoy, S. J. Mcnicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, *Acta Crystallogr. D* **2011**, *67*, 235–242.
[123] A. J. Mccoy, *Acta Crystallogr. D* **2007**, *63*, 32–41.
[124] A. W. Schüttelkopf, D. M. F. Van Aalten, *Acta Crystallogr. D* **2004**, *60*, 1355–1363.
[125] G. N. Murshudov, P. Skubák, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr. D* **2011**, *67*, 355–367.
[126] J. Shen, R. Ye, A. Romanies, A. Roy, F. Chen, C. Ren, Z. Liu, H. Zeng, *J. Am. Chem. Soc.* **2020**, *142*, 10050–10058.

[127] a) J. W. Sadownik, E. Mattia, P. Nowak, S. Otto, *Nat. Chem.* **2016**, *8*, 264–269; b) X. Zhou, G. Liu, K. Yamato, Y. Shen, R. Cheng, X. Wei, W. Bai, Y. Gao, H. Li, Y. Liu, F. Liu, D. M. Czajkowsky, J. Wang, M. J. Dabney, Z. Cai, J. Hu, F. V. Bright, L. He, X. C. Zeng, Z. Shao, B. Gong, *Nat. Commun.* **2012**, *3*, 949.

[128] a) L. L. Ong, N. Hanikel, O. K. Yaghi, C. Grun, M. T. Strauss, P. Bron, J. Lai-Kee-Him, F. Schueder, B. Wang, P. Wang, J. Y. Kishi, C. Myhrvold, A. Zhu, R. Jungmann, G. Bellot, Y. Ke, P. Yin, *Nature* **2017**, *552*, 72–77; b) K. F. Wagenbauer, C. Sigl, H. Dietz, *Nature* **2017**, *552*, 78–83.

[129] a) Y.-T. Lai, D. Cascio, T. O. Yeates, *Science* **2012**, *336*, 1129–1129; b) J. B. Bale, S. Gonen, Y. Liu, W. Sheffler, D. Ellis, C. Thomas, D. Cascio, T. O. Yeates, T. Gonen, N. P. King, D. Baker, *Science* **2016**, *353*, 389–394; c) Y. Hsia, J. B. Bale, S. Gonen, D. Shi, W. Sheffler, K. K. Fong, U. Nattermann, C. Xu, P.-S. Huang, R. Ravichandran, S. Yi, T. N. Davis, T. Gonen, N. P. King, D. Baker, *Nature* **2016**, *535*, 136–139; d) E. Golub, R. H. Subramanian, J. Esselborn, R. G. Alberstein, J. B. Bailey, J. A. Chiong, X. Yan, T. Booth, T. S. Baker, F. A. Tezcan, *Nature* **2020**, *578*, 172–176.

[130] a) A. J. Burton, A. R. Thomson, W. M. Dawson, R. L. Brady, D. N. Woolfson, *Nat. Chem.* **2016**, *8*, 837–844; b) T. Lebar, D. Lainšček, E. Merljak, J. Aupič, R. Jerala, *Nat. Chem. Biol.* **2020**, *16*, 513–519.

[131] W. C. Pomerantz, V. M. Yuwono, C. L. Pizzey, J. D. Hartgerink, N. L. Abbott, S. H. Gellman, *Angew. Chem. Int. Ed.* **2008**, *47*, 1241–1244.

[132] a) C. M. Lombardo, G. W. Collie, K. Pulka-Ziach, F. Rosu, V. Gabelica, C. D. Mackereth, G. Guichard, *J. Am. Chem. Soc.* **2016**, *138*, 10522–10530; b) G. W. Collie, R. Bailly, K. Pulka-Ziach, C. M. Lombardo, L. Mauran, N. Taib-Maamar, J. Dessolin, C. D. Mackereth, G. Guichard, *J. Am. Chem. Soc.* **2017**, *139*, 6128–6137.

[133] a) Y. Tanaka, H. Katagiri, Y. Furusho, E. Yashima, *Angew. Chem. Int. Ed.* **2005**, *44*, 3867–3870; b) B. Baptiste, J. Zhu, D. Haldar, B. Kauffmann, J.-M. Léger, I. Huc, *Chem.: Asian J.* **2010**, *5*, 1364–1375; c) Y. Ferrand, A. M. Kendhale, J. Garric, B. Kauffmann, I. Huc, *Angew. Chem. Int. Ed.* **2010**, *49*, 1778–1781.

[134] a) B. Gong, Y. Yan, H. Zeng, E. Skrzypczak-Jankunn, Y. W. Kim, J. Zhu, H. Ickes, *J. Am. Chem. Soc.* **1999**, *121*, 5607–5608; b) E. A. Archer, A. E. Sochia, M. J. Krische, *Chem. Eur. J.* **2001**, *7*, 2059–2066; c) E. A. Archer, M. J. Krische, *J. Am. Chem. Soc.* **2002**, *124*, 5074–5083; d) X. Yang, S. Martinovic, R. D. Smith, B. Gong, *J. Am. Chem. Soc.* **2003**, *125*, 9932–9933; e) H. Gong, M. J. Krische, *J. Am. Chem. Soc.* **2005**, *127*, 1719–1725; f) J. Zhu, J.-B. Lin, Y.-X. Xu, X.-B. Shao, X.-K. Jiang, Z.-T. Li, *J. Am. Chem. Soc.* **2006**, *128*, 12307–12313; g) B. Gong, *Acc. Chem. Res.* **2012**, *45*, 2077–2087; h) J. Atcher, P. Mateus, B. Kauffmann, F. Rosu, V. Maurizot, I. Huc, *Angew. Chem. Int. Ed.* **2021**, *60*, 2574–2577.

[135] D.-W. Zhang, X. Zhao, J.-L. Hou, Z.-T. Li, *Chem. Rev.* **2012**, *112*, 5271–5316.

[136] N. Delsuc, T. Kawanami, J. Lefeuvre, A. Shundo, H. Ihara, M. Takafuji, I. Huc, *ChemPhysChem* **2008**, *9*, 1882–1890.

[137] a) L. Shu, M. Mayor, *Chem. Commun.* **2006**, 4134–4136; b) X. Wu, R. Liu, B. Sathyamoorthy, K. Yamato, G. Liang, L. Shen, S. Ma, D. K. Sukumaran, T. Szyperski, W. Fang, L. He, X. Chen, B. Gong, *J. Am. Chem. Soc.* **2015**, *137*, 5879–5882.

[138] a) W. Kabsch, *Acta Crystallogr. D* **2010**, *66*, 125–132; b) C. Vonrhein, C. Flensburg, P. Keller, A. Sharff, O. Smart, W. Paciorek, T. Womack, G. Bricogne, *Acta Crystallogr. D* **2011**, *67*, 293–302.

[139] T. B. Johnson, J. M. Sprague, *J. Am. Chem. Soc.* **1936**, *58*, 1348–1352.

[140] J. C. Roberts, H. Gao, A. Gopalsamy, A. Kongsjahju, R. J. Patch, *Tetrahedron Lett.* **1997**, *38*, 355–358.

[141] A. R. Gangloff, J. Brown, R. De Jong, D. R. Dougan, C. E. Grimshaw, M. Hixon, A. Jennings, R. Kamran, A. Kiryanov, S. O'Connell, E. Taylor, P. Vu, *Bioorg. Med. Chem. Lett.* **2013**, *23*, 4501–4505.

[142] G. M. Sheldrick, *Acta Crystallogr. A* **2015**, *71*, 3–8.

10 Acknowledgements

I thank Professor Ivan Huc for his supervision, and I greatly appreciate the opportunities he has provided me. He always gave me helpful advice and the inspirational support needed to excel at my research endeavors and achieve my personal career plans.

I also want to thank all the members of the group for creating a collaborative working environment. I could always count on their support and teamwork. Special thanks goes to Valentina Corvaglia who supervised me in the beginning of my PhD. I thank Céline Douat for her continuous help and advice, especially with HPLC analysis and solid phase synthesis. Pradeep Mandal provided invaluable help with crystallizations and X-ray analysis. I also thank Lars Allmendinger, who showed great commitment in his involvement with NMR measurements and data interpretation. Without them all, this work would not have been possible. Finally, I am grateful to Sebastian Dengler, Niklas Böcher, Manuel Loos, and Tim Arczynski for the countless lunch and coffee break discussions and brainstorming.

I thank my friends and family, who have supported me throughout my whole path up to this point. Especially, I am grateful for my parents who did everything to make achieving this goal possible, and Eva Späh who supported me in many ways, including proofreading of this thesis.