

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



**Helical aromatic oligoamide foldamers for peptide
macrocycle and protein surface recognition**

Lingfei Wang

aus

Wuhan, Hubei, China

2025

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 11.03.2025

Lingfei Wang

Dissertation eingereicht am 11.03.2025

1. Gutachter: Prof. Dr. Ivan Huc

2. Gutachter: Prof. Dr. Franz Bracher

Mündliche Prüfung am 22.04.2025

Table of Content

List of Publication.....	5
1. Abstract.....	6
2. Introduction.....	8
2.1 Target protein surface – approach for drug discovery.....	8
2.2 Protein-protein interaction (PPI).....	9
2.3 Biological display technologies and RaPID system.....	12
2.4 Foldamers as templates to target protein surface.....	15
2.5 Approaches to obtaining structural information of a model protein-foldamer complex.....	20
3. Objectives.....	23
4. References.....	26
5. Main text: Interrogating the Potential of Helical Aromatic Foldamers for Protein Recognition.....	29
6. Supporting Information: Interrogating the Potential of Helical Aromatic Foldamers for Protein Recognition.....	36
6.1 Supplementary Figures.....	37
6.2 Experimental section.....	46
6.3 X-ray crystallographic analysis of compound 2.....	50
6.4 Supplementary Data.....	54
6.5 Reference.....	59
7. Main text: Display selection of peptide ligands for helical aromatic foldamers.....	60
8. Supplementary Information: Display Selection of Macrocyclic Peptides as Diastereoselective Ligands of Helical Aromatic Foldamers.....	68
8.1 Supplementary figures.....	69
8.2 Experimental section.....	77
8.3 Experimental procedures for chemical synthesis.....	84

8.4 Synthesis of peptide macrocycles	102
8.5 HPLC profiles, HRMS (ESI ⁺) and ¹ H NMR spectra.....	105
8.6 CD spectra of foldamers 1d-1e, 2c-2d	136
8.7 References.....	137
9. Main text: Structure-based Design of an Aromatic Helical Foldamer-Protein Interface	138
10. Supplementary Information: Structure-based Design of an Aromatic Helical Foldamer-Protein Interface	150
10.1 Supplementary Figures and Tables	151
10.2 Computational studies.....	175
10.3 Crystallography.....	176
10.4 Biophysical measurements.....	178
10.5 Chemical Synthesis.....	184
10.6 References.....	226
11. Summary and Perspective	228
11.1 Summary for published/submitted work.....	228
11.2 Challenges for protein surface recognition and future designs	229
12. Acknowledgement.....	232

List of Publication

Accepted:

Sunbum Kwon, ‡ Vasily Morozov, ‡ Lingfei Wang, ‡ Pradeep K. Mandala, Stéphane Chaignepain, Céline Douat, and Ivan Huc*, Interrogating the potential of helical aromatic foldamers for protein recognition, *Org. Biomol. Chem.*, 2024, **22**, 9342

‡ These authors contributed equally

Lingfei Wang, Joseph M. Rogers, Simon J. Dawson, Line Langhorn, Ryan T. Howard, Sunbum Kwon, Céline Douat, Hiroaki Suga and Ivan Huc*, Display selection of macrocyclic peptides as ligands of helical aromatic foldamers, *Org. Biomol. Chem.*, 2025

Submitted:

Lingfei Wang, Céline Douat, Johannes Sigl, Post Sai Reddy, Lucile Fischer, Béatrice Langlois d'Estaintot, Zhiwei Liu, Vojislava Pophristic, Yuwei Yang, Yingkai Zhang and Ivan Huc*, Structure-based design of an aromatic helical foldamer-protein interface, *Chem. Sci.*, 2025

1. Abstract

Proteins and peptides are essential biomolecules that regulate human homeostasis and maintain normal cellular functions. As key targets in disease research, understanding their interactions with various molecules is crucial for guiding drug design and development. However, most established principles/rules in drug design focus primarily on small molecules, there is a blank area in guiding the development of middle-size molecules like peptides. With the rapid advancements in peptide-based and antibody-drug therapeutics, expanding drug research into mid-sized molecular systems is increasingly important. Many types of aromatic oligoamides as structurally constrained oligomers, serve as valuable model systems due to their stable conformations in both aqueous and organic solutions. For a group of molecules with confined, folded structures, the term "foldamer" is used to describe them. In recent years, growing attention has been directed toward their interactions with biological targets. Moreover, the focus has gradually shifted from fundamental model systems to practical applications, highlighting their potential in drug discovery and development.

In this work, helical aromatic oligoamide foldamers, based on quinoline (Q), benzene (B), and pyridine (P) rings as building blocks, are synthesized and studied in the field of recognition of peptide macrocycles as well as proteins. We first explore the potential of our foldamers with different side-chain presentations on the surface to randomly recognize a library of proteins (cell lysate). Submicromolar binding was observed, but *P*- and *M*-conformers did not differentiate from each other concerning binding affinities. Thanks to the stable helical structure of the foldamer, the interface, which is constituted by side chains of building blocks, can be precisely designed and has the potential to interact with other molecules. We then designed two series of foldamers with the same chemical formula but with different interfaces consisting of five biogenic side chains. The peptide macrocycle library built by the RaPID system (Random non-standard Peptide Integrated Discovery) was then selected against target foldamers. The selected peptide macrocycles showed high selectivity for a specific arrangement of foldamer side chains. This specificity was further proven by the fact that the selected peptides could exclusively bind one handedness of the helical foldamer, strongly suggesting that recognition might take place on the foldamer surface. Furthermore, we targeted a protein surface (HCA II) by tethering the foldamer with a nanomolar-binding protein ligand. Several new monomers were designed under the guidance of computational tools and synthesized to interact with the protein surface via their biogenic side chains. With four crystal structures of the complex between the foldamer and protein, we could prove that main chains are interchangeable in the context of the foldamer–protein complex. Side chains could be inserted into the foldamer structure without affecting the overall complex structures, which is different from peptide structure design.

In conclusion, the development of side chains on Q and B monomers enables the design of various foldamer surfaces. The robustness of the foldamer helical structure could be further applied to structure-

based design for covering large protein surface areas. The automatic synthesis of foldamers (on solid state) also brings the possibility of providing sequence libraries within a relatively short time compared with manual synthesis. These results present the high potential of aromatic oligoamide foldamers in recognizing peptide macrocycles and proteins.

2. Introduction

2.1 Target protein surface – approach for drug discovery

Proteins are essential macromolecules in our bodies, known for their versatility and vital functions. They are referred to different names depending on their roles: when catalyzing biological processes, they are called enzymes;^[1] when dealing with transmitting signals between cells, tissues, and systems, they are called neuron-transmitters^[2]; when providing cellular structural support to cells, they are cytoskeleton.^[3] Consequently, proteins are a major focus in drug development. The chemical and physical properties of a protein largely depend on its side chains, which vary in size, shape, charge and hydrophobicity.^[4] The backbone of the peptide chain folds through hydrogen bonds and other non-covalent bonds, adopting a specific folding pattern that leads to the secondary, tertiary, and quaternary structures. Surfaces of proteins can be recognized by a variety of molecules, leading to the inhibition or activation of their function, which in turn modulates downstream biochemical reactions. A deep understanding of how to target protein surface by *ab initio* design, as well as obtaining specific structural information of interactions between specific anionic and cationic residues to protein surface, are crucial for unraveling disease mechanisms and developing new drug candidates. Protein surfaces can vary in shape, from flat and grooved to irregular, which poses challenges in designing complementary molecules that can effectively interact with. The design of many lead compounds, known as the first identified binding candidates, starts from analyzing the so-called hot spots on a protein surface. Hot spots are defined as the residues that cause an increase of more than 2 kcal/mol in binding energy when subjected to an alanine scan,^[5] and are considered as starting point of targeting in many cases of protein surface recognition.^[6] Designing lead compounds to precisely target hot spots on proteins surface remains a daunting task for researchers. There are many aspects to consider when developing protein surface targeting compounds: 1. Shape complementary; 2. entropy and entropy change; 3. non-covalent bond interactions. Many lead compounds are selected from the natural products, which showed preliminary effect, and then undergo multiple rounds of structural refinements before they become a marketable drug. Besides the libraries of compounds from nature, researchers have developed structure-based design like fragment-based lead discovery (FBLD) also named fragment-based drug discovery (FBDD),^[7] which are widely used to find lead compounds in early drug discovery process. FBDD focuses more on dealing with small compound fragments less than 300 Da in molecular weight. High throughput screening (HTS) is preferred when key structural information is missing and combined with FBDD.^[8] The main difference between these two methods is summarized in **Figure 1**.^[9]

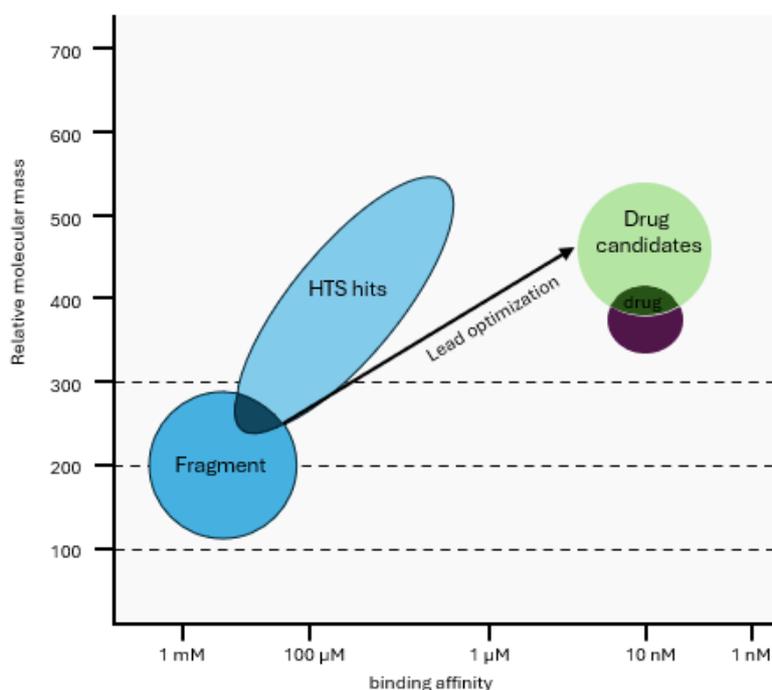


Figure 1:^[9] schematic comparison of FBDD and HTS in drug research. FBDD covers the area of low binding affinity (mM range to high μM) area with low molecular weight (120-250 Da), as a starting point for the rounds of improvement (toxicity, affinity, bioavailability).

2.2 Protein-protein interaction (PPI)

2.2.1 Introduction

PPIs play a fundamental role in nearly all biological processes. From a structural perspective, based on the complexity of the interaction interface, PPIs can be broadly categorized into three types: (1) interactions involving small interfaces and short peptide sequences, which focus on parts of secondary structure; (2) interactions where the secondary structures of the partner proteins bind to a hydrophobic groove and (3) complex interactions that involve multiple points of contact on both sides. While most of the traditional drug targets are related to enzymes,^[10] ion channels^[11] or receptors on membranes.^[12] PPIs have become key targets for drug development in the last few decades.^[13] PPIs can be broadly divided by two different modes of action: activation and inhibition. In the following discussion, inhibition will be the primary focus (**Figure 2**). Understanding these protein assemblies and their structural information can pave the way for discovering new active pharmaceutical ingredients (APIs). However, PPIs have often been considered as ‘undruggable’ in traditional medicinal chemistry, as the interface of the target protein is typically too large and flat for a small ligand to fit in, creating significant challenges for rational design of therapeutics.^[14] The PPI interface is often highly hydrophobic and involves large contact area (reaching 1500 – 3000 Å)^[15] or flat grooves, which are infeasible for small molecules to occupy. Structural studies of known PPIs reveal that, in some cases, only a small fraction of the complementary protein region participates in the binding process. However, this fraction of

peptide sequences cannot be directly applied. The reason is that when only the complementary region is used to reproduce the binding with the target protein, the affinity is, in most cases, significantly reduced, potentially due to the lack of well-defined peptide conformation of the synthesized protein fragment. This synthesized protein loses structural constraints compared with the original residue within protein and thus shows a lower binding affinity. However, these protein-protein interfaces, once structurally identified, can be used as a starting point for the development of potential therapeutic agents.

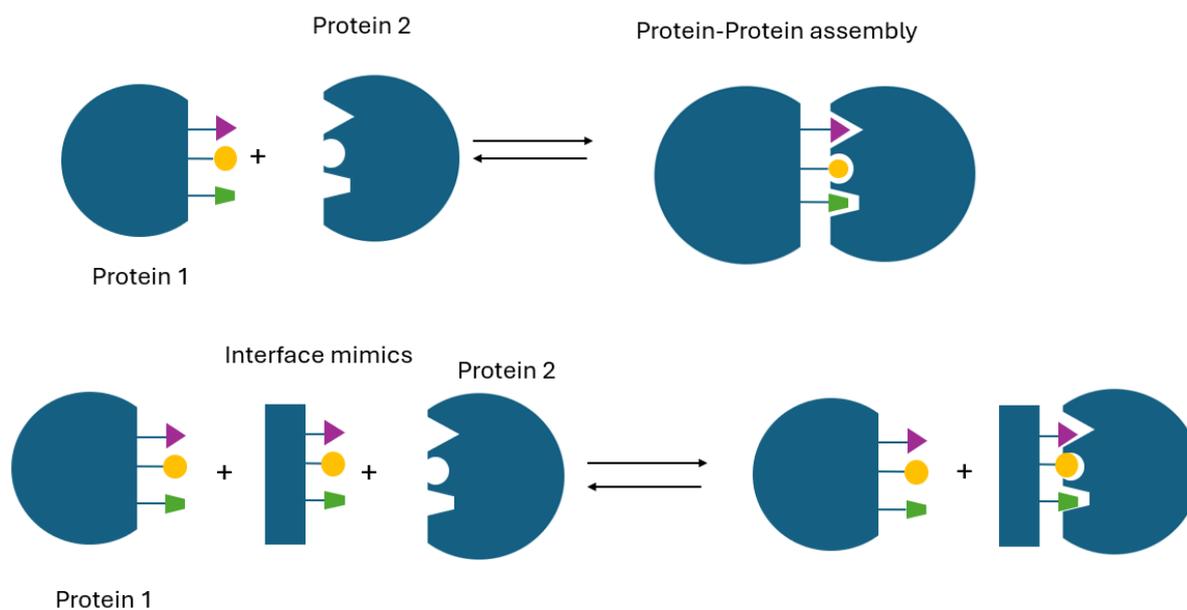


Figure 2:^[16] schematic illustration of PPI and interface mimic inhibition of PPI

Small molecules have the advantage of being relatively easy to synthesize and meet the criteria of ‘Rule of Five’ (Ro5), introduced by Christopher A. Lipinski in 1997.^[17] Large molecules have no significant bioavailability advantage over small molecules and are challenging to redesign and synthesize. However, when it comes to targeting protein surface, large molecules like peptides,^[18] antibodies^[19] and nucleic acids^[20] are good candidates. In the past 20 years, significant progress has been made in targeting different types of PPIs and some drug candidates have entered clinical research phases. Lu *et al.* have reported some recent PPI modulators in clinical trials.^[13] Most of them focus on cancer treatment, for example, the MDM2/p53 complex has been widely studied in acute myeloid leukemia,^[21] metastatic melanoma^[22] and solid tumor.^[23] P53 is a tumor suppressor protein, which could prevent the formation of cancer and activate DNA repair.^[24] In normal cells, the level of p53 is relatively low, when cell homeostasis is changed, the expression of protein 53 will be activated to a high level, leading to various posttranslational modifications, such as phosphorylation, acetylation, ubiquitination, neddylation, sumoylation, and methylation.^[24-25] The interface of p53-MDM2 involves 3-4 helical turns and the three hot-spot residues of P53 responsible to the binding are hydrophobic: Phe(19), Trp(23) and Leu(26) (see **Figure 3a**).^[26] Another PPI as target in cancer treatment is Bcl-2/BIM interaction. Bcl-2 (B-cell lymphoma 2) and BIM (Bcl-2-interacting mediator) are both cell apoptosis-regulated proteins.

Bcl-2 is responsible for preventing the cell apoptosis, while BIM functions as a pro-apoptotic protein (BH3-only protein).^[27] The binding domain of Bcl-2/BIM contains four residue side chains of the BIM peptide positioned themselves over six helical turns length and all of them point deep into the surface of Bcl-2 (**Figure 3b**).

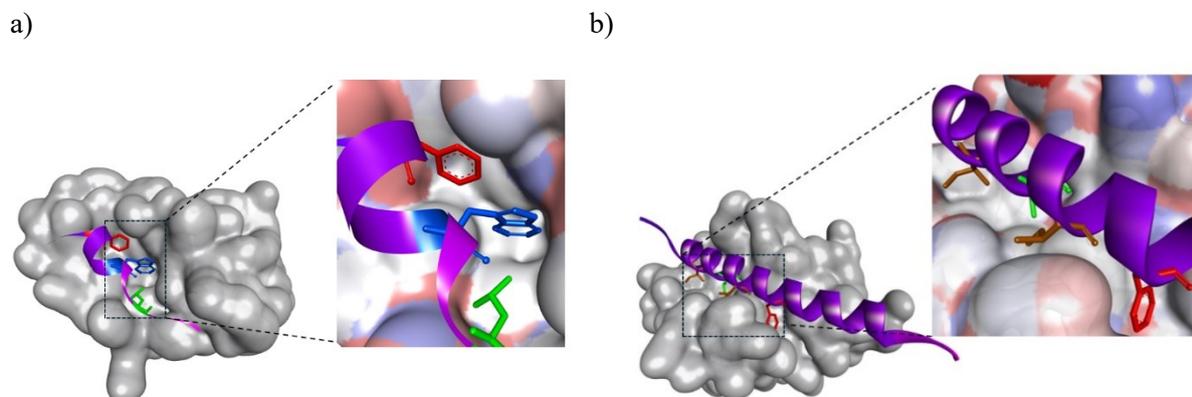


Figure 3: Examples of α -helix protein secondary structure involved in PPIs. Isolated α -helical segment are labelled in violet. Different key residues involved in interaction were labelled in different colors. Phe: red; Trp: blue; Leu: green; Ile: brown; Glu: light blue; a) p53/MDM2 interaction (PDB: 1YCR) b) Bcl-xL/BIM interaction (PDB: 1PQ1).

2.2.2 Technologies to identify PPI

Researchers in both chemistry and biology are actively engaged in ‘protein-targeted’ drug research from two different perspectives. Chemists aim to synthesize compounds based on the structure-activity relationship (SAR), while biologists focus on uncovering the specific functions of proteins and the mechanisms of their activation and inhibition. Despite their different approaches, both fields share common interests: understanding the structures of drug candidates and their binding sites. There are many techniques involved in the structural characterization of ligand-protein interactions and PPIs. Current methods such as NMR, X-ray crystallography, and cryogenic electron microscopy are commonly used to elucidate these structures. X-ray crystallography^[28] and cryo-electron microscopy^[29] provide detailed structural information, while NMR spectroscopy can identify the protein binding site in solution based on the perturbation of the chemical shift values of the unbound species. Comparing the attempts to obtain X-ray crystal structures, NMR spectroscopy is suitable for quantifying the binding affinity of one of the two partners by titration experiments.^[30] The chemical shift perturbation is a common technique for illustration of ligand binding to proteins,^[31] while X-ray structure potentially shows one snapshot of the binding event in the solid state. ITC (isothermal titration calorimetry) and SPR (surface plasmon resonance) are two biophysical methods commonly used to determine the binding affinity in solution between the two partners. ITC measures the heat released during the interaction process and is commonly used for natural proteins that are well-soluble. However, ITC is less suitable for substances with poor water solubility, such as unnatural peptides or hydrophobic ligands. Another limitation of ITC is that the measurements are only reliable for binding processes that exhibit

a significant change of enthalpy; when this change is intrinsic low, the result is sometimes not reliable. SPR and BioLayer Interferometry (BLI), which operate on similar principles, are widely used to measure binding interactions over a large detectable range of concentrations. In these methods, proteins are typically immobilized on sensor chips (SPR) or tips (BLI), and the surface regeneration process is reproducible. Understanding ligand-protein mode of recognition and assembly of ligand-protein interactions is crucial for finding suitable starting points for structure-based design. Taking good advantage of the structural information of protein-ligand or protein-protein complex can benefit the research of drug development since proteins are predominantly chosen as targets.

2.3 Biological display technologies and RaPID system

In vitro selection methods are useful tools and approaches to screen a substance library against target compound or protein. Many technologies have been developed in the past decades like pull-down assay, phage display, mRNA display, and ribosome display. In the following chapter, pull-down assay as a screening method and mRNA display will be introduced.

2.3.1 Pull-down Assay

The pull-down assay is a technique commonly used to identify binding partners of a target compound, specifically those that physically interact with it. This technique shares similarities with co-immunoprecipitation, as both rely on a ligand exhibiting high binding affinity to capture interacting proteins.^[32] The general procedure involves immobilizing a bait compound onto a solid support, which is then exposed to a cell lysate or compound library. After incubation and thorough washings, prey proteins that show high affinity for the immobilized bait molecule bind to the solid support, while those with lower affinity are removed through centrifugation or filtration. Repeating this process can increase the accumulation of bait-protein complexes, facilitating subsequent analysis.

The resulting binary complexes are next released from the solid support using an appropriate buffer system and analyzed by techniques such as LC-MS/MS or SDS-PAGE to identify the proteins selected from the cell lysate. To avoid false positives, control experiments are crucial. For example, the stability of the bait throughout incubation and washings should be confirmed. Additionally, it is important to account for nonspecific binding, ensuring that proteins binding only to the solid support and other nonspecific sites are excluded from the analysis.

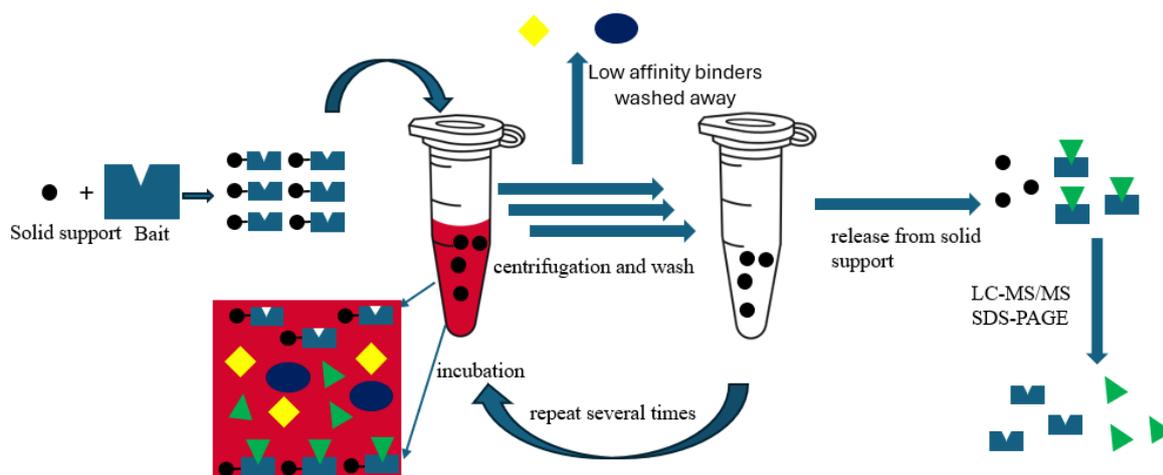


Figure 4: general procedure of pull-down assay.

2.3.2 mRNA display technique

The mRNA display, first introduced by Roberts and Szostak,^[33] is an *in vitro* selection technique that has been extensively developed over the past two decades. The principle of this technique is that the translated peptides or proteins are covalently bound to their mRNA progenitor via a puromycin linkage. Additionally, the power of this technique lies in the stability of this chemical linkage, which allows for the selection of trillions of variants (10^{12} - 10^{14}) under versatile conditions. The puromycin is an antibiotic whose chemical structure mimics the aminoacyl end of tRNA (**Figure 5**), and is covalently attached to the 3' terminal of an mRNA. Briefly, as the ribosome moves along mRNA sequences during the translation process, puromycin residue occupies the A site of ribosome and covalently binds to the C-terminus of the translated polypeptide. Following by the reverse transcription to generate the mRNA/cDNA-fusion peptide library, tighter peptide binders are selected against the immobilized target, and the DNA library is subsequently amplified by PCR (**Figure 5**). After several rounds, this DNA library contains the genetic information for the selected peptides showing the highest affinity for the immobilized target compared to the original DNA library. In pull down assay, the immobilized complexes will be released after several round of incubation and separation while in mRNA display selection, the peptide-cDNA-mRNA complex is released from solid support in each selection round and cDNA library will be used for next round. What makes mRNA display distinct from other *in vitro* selection techniques is that it enables the expansion of the 20 natural amino acids to non-proteinogenic amino acid residues or short abiotic segments by the possibility to introduce genetic code reprogramming methods, which will be discussed below.

2.3.3 FIT and RaPID technologies

In lead-compound screening, compound pools are essential for achieving promising results. Most of the big pharmaceutical companies possess a huge number of substance libraries. Therefore, enriching the substance libraries by developing biological techniques is drawing increasing attention among

researchers. As mentioned in Chapter 2.2.1, molecules satisfied the “Rule of 5” are typically ineffective for targeting PPIs. Consequently, peptide template becomes a new drug modality to target protein surfaces. However, traditional peptides have the disadvantage of showing rapid proteolytic degradation and often lack of cell-permeability.^[34] The linear peptides are less conformational constrained compared to peptide macrocycles, leading to higher entropic loss during the binding process.^[35] Additionally, peptide macrocycles have also been defined as having higher capabilities to spontaneously penetrate eukaryotic cells (*i.e.* without resorting to a conjugation with cell-penetrating peptides (CPP)).^[34a] With the aim to rapidly and efficiently screen for peptide macrocycles against a defined biological target, large libraries of cyclic peptides are requested. The flexizymes developed in the group of Prof. Suga enable the building of cyclic and non-natural peptide libraries by relying on a flexible *in vitro* translation (FIT) system.^[36] Flexizymes are non aminoacyl-tRNA synthetases, which catalyse the tRNA acylation by identifying the 3’ end of tRNA.^[36] With the assistance of a given flexizyme, “preactivated” non-standard amino acids can be loaded on the tRNA, which are then inserted into peptide sequences using the ribosome translation system. Remarkably, peptide macrocyclization can be achieved by starting the translation with a N-terminus chloroacetamide residue, forming a spontaneous thioether bond with a cysteine residue introduced downstream in the peptide sequence.^[37] Flexizymes enable the building of a pool of modified codon-anticodon library (genetic code reprogramming) and following this modified coding, non-canonical amino acids can be delivered to the ribosome, thus non-natural peptide sequences can be achieved under the genetic code reprogramming.

FIT system as a new tool is compatible with mRNA display technology, which is referred as RaPID system (Random non-standard Peptide Integrated Discovery). The first advantage of this technology is to enable high diversity of peptide libraries. RaPID system enables the *in vitro* building of peptide libraries encompassing unnatural building blocks and serve the purpose of peptide selection against a target compound. **Figure 5** illustrates the general procedure of RaPID system selected against one foldamer (See Chapter 2.4) target immobilized on a streptavidin beads. The foldamer as target could be changed by other biological molecules of interest like proteins. Dengler *et al.* have successfully incorporated aromatic foldamers segments to tRNA by using flexizyme and built two series of foldamer-peptide macrocycle libraries serving as candidates to be selected against the C-lobe domain of E6AP HECT domain.^[38] The foldamer-peptide macrocycle with highest binding affinity (K_D in the nanomolar range) was crystallized with the target protein domain and revealed that both the peptide and foldamer segments were helically folded in an intriguing reciprocal stapling fashion.

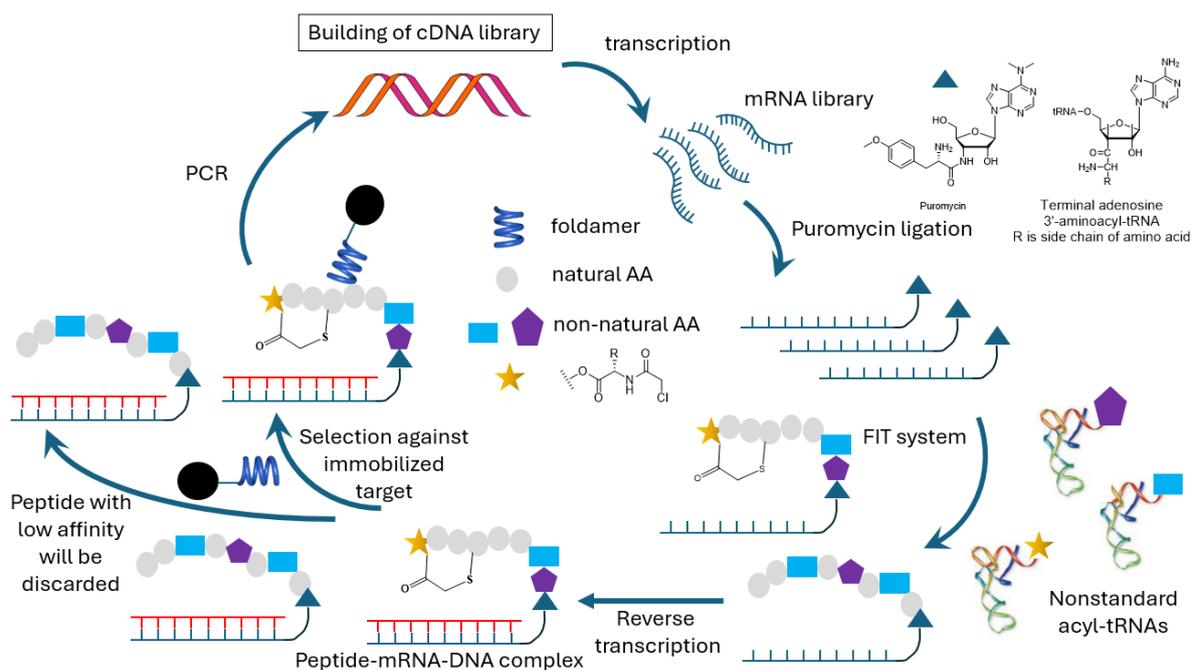


Figure 5: schematic representation of RaPID system applied to selection against foldamers.

2.4 Foldamers as templates to target protein surface

To have a better understanding of the structural information in PPIs and identify potential binding sites of protein surface, molecules with highly predicted structure are required to collect such data and facilitate further iterative structural modifications. Folding is a natural phenomenon shared by many biological macromolecules like proteins, DNA and RNA.^[39] From secondary to quaternary protein structure, folding controls the molecular shape at different levels of assemblies. Protein folding enables the processes of enzyme activation or inhibition through conformational changes. Nucleic acid chain folds to store the genetic information.^[39b, 40] Inspired by these examples how natural compounds take advantage of folding to achieve complex biological function, chemists endeavor to create and discover novel structures beyond to what Nature provides.^[41] The concept of foldamer was brought up by Prof. Samuel Gellman, who defined it as ‘any polymer with a strong tendency to adopt a specific compact conformation’.^[42] Among the design of foldamer architectures, the conception of backbones is fundamental since backbone largely determines the spatial configuration of foldamer. Thus, foldamer can be divided into two main families, biotic and abiotic foldamers, according to the folding principle. Biotic foldamers shared a folding principle similar to biopolymers while abiotic foldamers adopt different and remote modes of folding from what Nature offers.^[43]

In protein secondary structure, α -helix consists of more than 30% because of its compact, tightly wound structures when compared to β -sheets and turns.^[44] This helical structure is stabilized by hydrogen bonds and folds as a right-handed helix (in α -peptides). A full α -helix turn contains 3.6 residues on average which might be a slightly shifted upon chain elongation.^[45] Many biotic scaffolds have been

developed to mimic α -helix conformation of peptides. As a well-studied and historical example is β -amino acid oligomers (β -peptides) which provide stable and well-defined secondary structures. β -peptides showed a slightly stretched helical structure comparing with α -helix due to the additional CH_2 group, bringing some flexibility as well.^[46] Gellman *et al.* have thoroughly studied α/β peptides, heterogeneous-backbone oligomers.^[47] Studies showed that this new type of artificial peptides (α/β^3 , **Figure 6**) adopted well-folded structures and was applied to many disease area like HIV.^[48] In α/β peptides, replacing β^3 with β^2 amino acid showed changes in the helical structure as well as in the binding affinity improvement with Bcl-x_L.^[49] Extending the β -peptide backbone to one additional unit leads to γ -peptides. These biotic foldamers are less studied because increasing the number of flexible carbon atoms of sp^3 hybridization induced an overall instability with respect to α - and β -peptides.^[50]

Concurrently, abiotic foldamers take advantage of aromatic systems as backbones, also showing a stable conformation in organic and aqueous solvent.^[51] Aromatic foldamers are constituted upon oligomerization from monomers, forming a stable three-dimensional structure guided by hydrogen bonds, hydrophobic interactions, and other non-covalent interactions. They represent a powerful tool for creating synthetic molecules that can show the structural and functional properties of proteins.^[52] Aromatic foldamers consisting of different building blocks combine not only the advantages of natural macromolecules, but also numerous possibilities brought by design. Their ability to form larger, more complex structures, with controlled side chain orientation, opens up new possibilities in targeting protein surface,^[6a, 53] for example, mimicking the large interface of side chains projection in α -helix,^[54] which can serve the aim of modulating PPIs. The surfaces of these oligomers can be precisely decorated with proteinogenic side chains, enabling them to effectively mimic the side chains projection of protein region, which participates in the binding process. This unique design capability makes aromatic foldamers as valuable tools for gaining deeper insights into these interactions. Examples are foldamers based on terphenyl scaffold, oligophenyl scaffold, terpyridine backbone, reproducing the i , $i+4$, $i+7$ projecting residues of α -helix^[55]. Hamilton and coworkers discovered the oligoanthranilamides composed of anthranilic acid and pyridine-2,6-dicarboxylic acid monomers (**Figure 6**) forming stable five- or six-membered ring hydrogen bonds between adjacent amides function.^[56] Foldamer research, as a significant subset of supramolecular chemistry, is now undergoing continuous expansion.

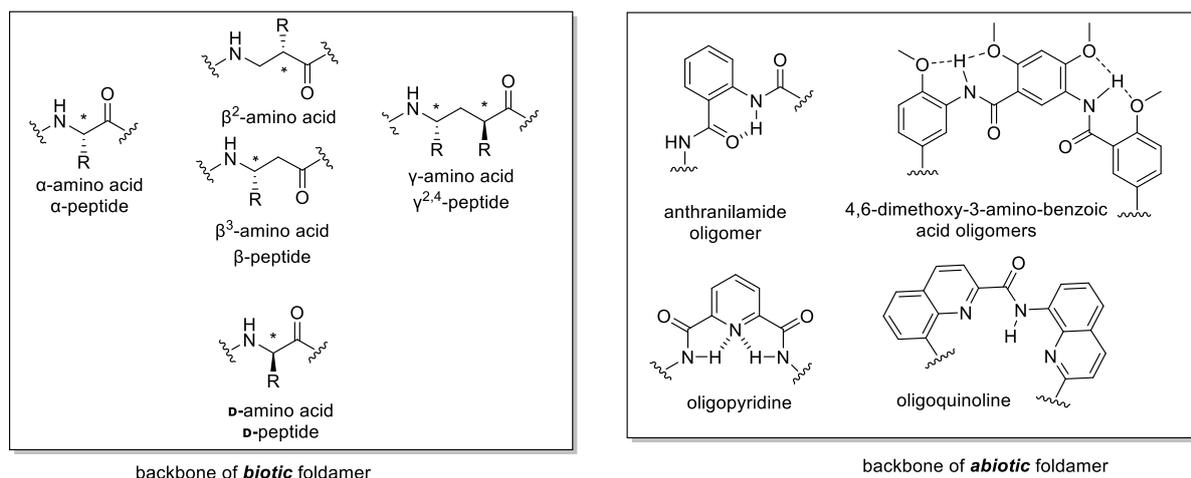


Figure 6: Chemical structures of backbones of biotic/abiotic foldamers.

Among those successful designed foldamers, following different or similar folding principle from what Nature does, aromatic oligoamide foldamers (AOFs) have been one of the most successful examples. One of most commonly used building block in AOFs is the 8-aminoquinoline-2-carboxylic acid, which folds into a stable, highly predictable 2.5-helix structure in organic/inorganic solvents upon oligomerization, achieving many advanced functions like protein surface recognition,^[57] or mimicking double-stranded DNA (with ^mQ monomer).^[58] AOFs can be assembled starting from a variety of monomers, offering many possibilities in terms of shapes and folding. **Figure 7b** depicts the intramolecular interactions of a dimer composed of two quinoline-type (**Q**) monomers. The electrostatic repulsion between the endocyclic nitrogen and the carbonyl of the amide moiety restricts the rotation of the amide bond which results in a *trans* conformation. Consequently, a pentamer (**Q**₅) forms two helical turns with a pitch of the thickness of one aromatic ring (3.4 Å) further stabilized by the aromatic packing. In hetero-oligomers (foldamers built from different aromatic building blocks), the helical curvature depends mostly on the angle between C- and N-termini of the monomers. The 7-amino-2-quinolinecarboxylic acid monomer (**Q**^H) recently applied in aqueous media by Teng *et al.* forms helix with a larger diameter since the angle between the C- and N-termini is equivalent to 120° (60° for regular 8-amino quinoline unit).^[59] Additionally, when the amino group is introduced in 6 position, the angle reaches 180 degree, and X-ray crystal structure revealed that oligomer consisted of 6-aminoquinoline-2-carboxylic acid forms a linear rod.^[60]

Due to the folding propensity of aromatic foldamers, the hydrophobic backbones are decorated by the side chains, which are mainly exposed to the solvent. The choice of side chains to decorate the helix surface easily facilitate the switching of the solubility from organic to aqueous medium.^[61] Recently, a variety of proteinogenic side chains have been successfully introduced either in position 4, 5 or 6 of the quinoline ring with the aim to reproduce the side chains pattern found in α -helix.^[54]

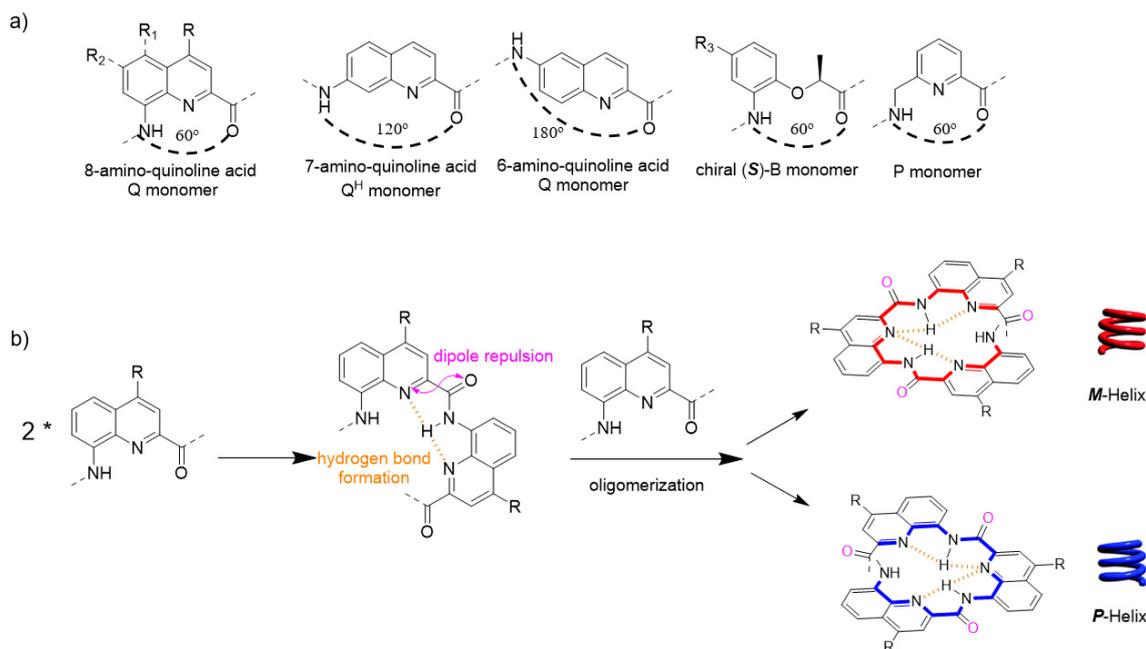


Figure 7: a) Common building blocks for AOFs. R, R₁, R₂, R₃ represent different side chains. The angle between N- and C-termini determines the curvature of the helix. **B** and **P** monomers allow flexibility to the aromatic backbone since the sp³ carbon is free to rotate in the overall structure. b) helix folding elucidation of homo-quinoline oligomers. When there is no chiral residue in the sequence, the oligoquinoline is achiral and the *P* and *M* helices do exist in a 1:1 ratio in solution.

Helical AOFs are achiral. They do not possess any stereogenic center. The right-handed (*P*) helix and the left-handed (*M*) helix exist under equilibrium in solution.^[62] Many parameters influence the dynamic of interconversion between *P* and *M* helices. Shorter sequences have a shorter half-life of helix handedness inversion in the same solvent (heptamer of 30 minutes, octamer of 2 hours at 30 °C, in *n*-hexane/chloroform (75:25 v/v)).^[63] Moreover, the polarity of solvents also has a strong effect on the half-time of helix handedness inversion: longer sequences (up to eight quinoline units) undergo no equilibrium in protic solvent.^[64] The so-called **P** monomer (*i.e.* aminomethyl pyridine unit structure in **Figure 7a**) has been designed to bring flexibility to the helical backbone of the oligoquinolines. Compared with **Q** monomer, the CH₂ group in 5 position of pyridine allows the rotation of the chemical bond, leading to a much faster *P/M* helix interconversion of sequences rich in **P** monomer with respect to those depleted. These dynamic AOF sequences enable the use as sensors when interacting with other chiral substances, resulting in a handedness bias.^[65] When a chiral group is incorporated within the aromatic oligoamide backbone, it can also bias the *P/M* helix equilibrium to one handedness with the respect to the other. X-ray crystal structure of camphanic acid (Camph.) and oxazolylaniline unit (Oxaz.) coupled trimer were obtained.^[62, 66] Recently, a new chiral aromatic monomer, referred as **B** monomer, deriving from 2-(2-aminophenoxy)acetic acid showed quantitative handedness bias in aqueous solvent, with *S*-chirality promoting (*P*)-handedness.^[67] The **B** monomer can indeed be incorporated in the middle of the AOF sequence, which offers the advantage in sequence design to be free from the restriction of

incorporating the chiral moiety at the very end of the sequence (*i.e.* principally the N-terminus). The **B** monomer, which is also a δ -amino acid, brings additional flexibility to the backbone but shares the same angle as a **Q** monomer, globally not perturbing the 2.5-helical fold of **Q**-oligomers.

One important chemical feature of AOFs is their straightforward synthesis. The main chain can be elongated by stepwise or iterative approaches in solution phase.^[51] Previously, acid-chloride activation was commonly applied in solution for the amide bond formation.^[68] Oxalyl chloride is one of the common reagent employed to form the acid chloride in a chemical laboratory. But the strong activation and release of hydrogen chloride are not compatible with acid-labile protecting group introduced for side chain protections, like the tert-butyloxycarbonyl protecting (Boc) and tert-Butyl (*t*Bu) groups. Ghosez's reagent has been chosen as an alternative to form acid chloride with mild conditions but still, its use requires strict conditions, with substantial activation and drying times, to remove any trace of Ghosez's reagent. Otherwise, the excess of Ghosez's reagent will react with the amine group, capping the sequences and preventing further coupling. A new efficient way for AOFs synthesis was highly required.

Peptide synthesis is nowadays performed in routine thanks to the development of solid peptide phase synthesis (Nobel prize in 1984 to Bruce Merrifield) and has been automated with different peptide synthesizers available on the market. The introduction of orthogonal protecting groups for the side chains of amino acids and the development of efficient coupling reagents makes it possible to obtain chemically synthesized peptides in high purity and yield. AOFs composed of different monomer units can be assembled following the same principle on solid support.^[69] Upon chain elongation, the aromatic amine on the resin-bound quinoline unit is poorly nucleophilic and suffers from steric hindrance. Consequently, a stronger activation of the carboxylic acid of the following **Q** unit is required to allow high coupling yields ($\approx 99\%$) and therefore the possibility to synthesize oligoquinolines on solid support. Corvaglia *et al.* have recently achieved helical aromatic oligoamide foldamers synthesis with a commercial peptide synthesizer, by adapting the Appel's reaction with an *in-situ* acid chloride activation and iterative process for efficient automation of the solid phase synthesis.

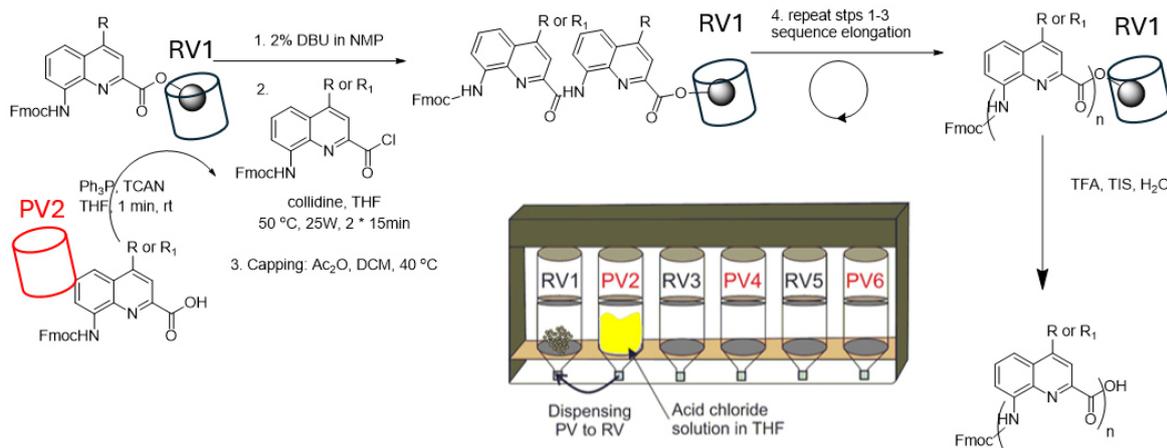


Figure 8: general SPFS (solid phase foldamer synthesis) protocols developed for aromatic oligoamides (PV: pre-activation vessel, RV: reaction vessel)

2.5 Approaches to obtaining structural information of a model protein-foldamer complex

The highly predictable spatial orientation of the quinoline side chains enables the design of foldamers capable of interacting with large areas on protein surface. Human carbonic anhydrases (HCA) are zinc-coated metalloenzymes, catalyzing the interconversion between carbon dioxide and bicarbonate.^[70] There are in total 14 types of isomers which are distributed in cytosol, membrane-bound and mitochondria.^[71] HCAII, especially, plays a crucial role in modifying disease conditions such as glaucoma, epilepsy, edema, high altitude sickness, and renal disorders.^[72] HCAII is easy to produce and co-crystallize with ligand containing sulphamate and sulphonamide moieties, which have shown nanomolar binding affinity to the zinc-pocket.^[73] Gaining structural data between aromatic oligoamide foldamers and HCAII protein could provide us valuable information for the precise protein-ligand design. Proteinogenic side chains can be designed and decorated on the surface of short oligomers and attempted to explore structural information between foldamer and protein. (see the structure in **Figure 9**). However, without a reasonable binding affinity at first place, it would be difficult to have a starting point for foldamer-protein interaction. In this respect, Buratto *et al.* have firstly coupled sulfonylbenzamide moiety on the N-terminal of short AOFs and obtained nanomolar binding between AOF sequences and HCAII. The X-ray crystal structure of protein-foldamer complex showed that aromatic foldamer was in close proximity to the protein surface with the help of sulfonylbenzamide ligand.

As introduced in Chapter 2.4, short aromatic foldamers helices undergo a fast *P* and *M* conformers interconversion in solution. Upon binding with the protein, when one handedness is favored, the equilibrium will be biased to one direction. Appearance of circular dichroism (CD) signal above 300

nm can be regarded as a sign of protein-foldamer interaction. The reason is that the protein does not absorb above 300 nm, thus the occurring CD band indicates only the handedness bias of quinoline backbone of foldamers. Therefore, CD has facilitated a preliminary screening method to monitor the binding and the short length of the screened foldamers allowed handedness inversion to be complete within days. (*P*)-helix was found to have higher affinity for the protein surface, with positive band in CD above 300 nm.^[74]

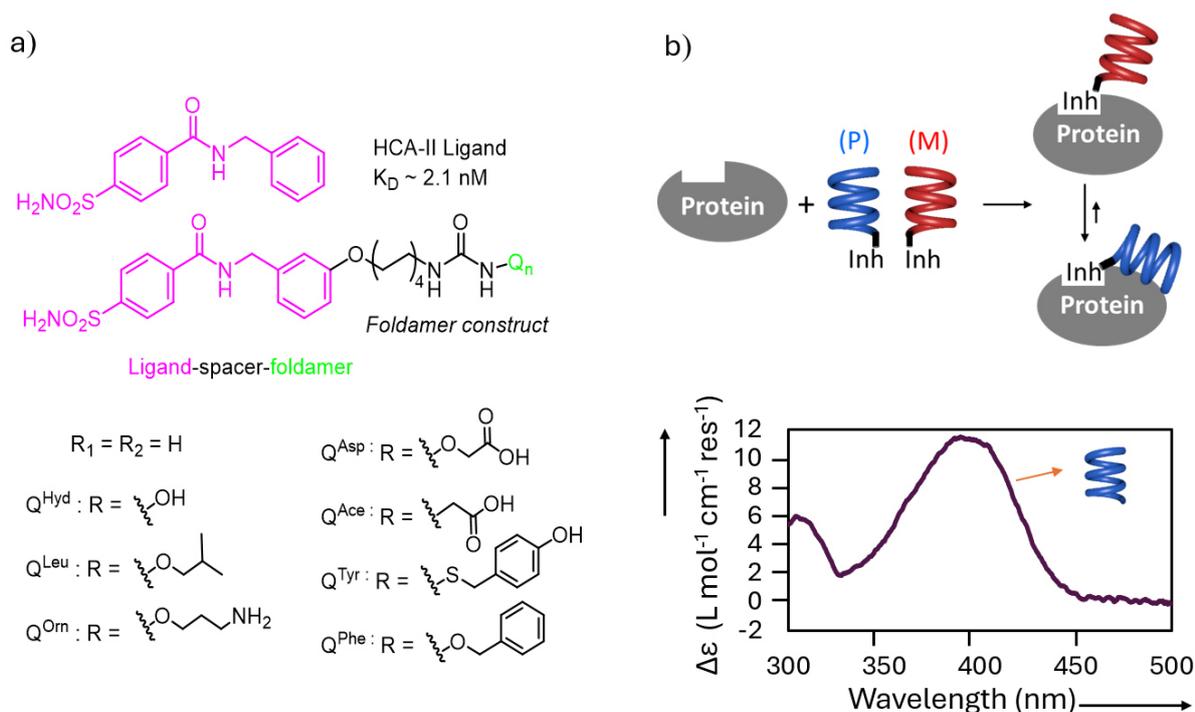


Figure 9: a) chemical structure of HCAII ligand, general complex of ligand-foldamer construct and side chains on the 4th position of quinoline. b) schematic picture to illustrate how CD signal emerge upon binding.

Since short sequences can only reach limited surface area of protein, longer foldamer sequences were needed for exploring larger protein surface. However, longer quinoline oligomers showed almost no *P* and *M* conformers interconversion in aqueous solvent.^[65] To facilitate the CD screening, the foldamer needs to possess certain structural flexibility. **P** monomers were therefore incorporated into the AOF sequence at given positions and the resulting **P/Q** hybrids sequences (up to fourteen units) allowed for fast helix handedness inversion in aqueous environment.^[65] Based on this information, Saireddy *et al.* have successfully crystallized the complex of HCAII and **P/Q** hybrids foldamers of 9 units and 14 units (see the X-ray crystal of a 14mer-protein complex in **Figure 10**).^[53] The structure showed extensive protein-foldamer hydrophobic contacts and foldamer-foldamer interactions. Unexpectedly, the proteinogenic side chains were not exposed to HCAII surface, instead, they pointed toward the solvent and the **P** units were positioned in a smooth surface groove driven by shape complementarity and hydrophobic effect. Right-handed *P* helix still prevails upon the binding. The shape complementarity between foldamer and protein as well as the hydrophobic contacts are dominant factors to build the

binary complex. This structural elucidation, however, could provide a good starting point for further protein–foldamer interaction since the foldamer fits itself well in the flat groove and the binding between foldamer and protein is tight. The ultimate objective now is to fulfil the binding of foldamer to protein surface without the help of sulfonylbenzamide moiety by iterative modification of proteinogenic side chains of quinoline units. The related work will be introduced in Chapter 9.

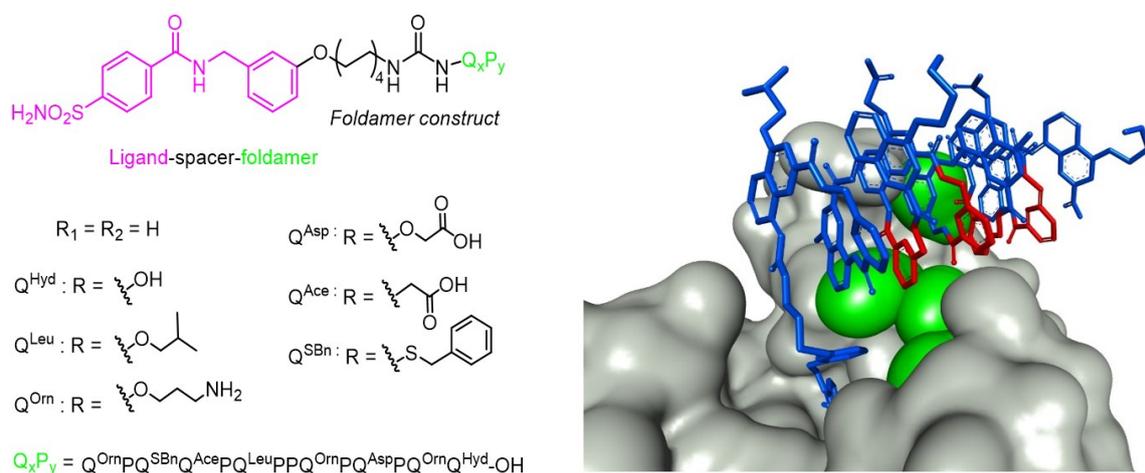


Figure 10: Building blocks of 14mer and crystal structure of 14mer and HCAII (PDB# 6Q9T). The surface of protein is presented in grey and green area are the surface of hydrophobic grooves consisted of Phe20, Pro21, Ile22, Val 34, Pro201 and Leu203. The foldamer colored in blue and P units pointing to the surface of proteins colored in red.

3. Objectives

This thesis aimed to extend our knowledge of protein-AOFs, peptide macrocycle-AOFs interaction and to gain structural information of the resulting binary complexes. Acquiring a good knowledge of the mode of interactions between foldamers and a given protein could be eventually useful to target sizeable surface areas involved in PPI. Taking into consideration precedents in the group, various foldamers composed of **Q**, **P** and **B** monomers were synthesized via solid phase synthesis to finely tune foldamer side-chain composition and elaborate long AOFs showing good to high affinity for model protein surface.

Classical ways of selection methods to seek protein-binding partners consist in first defining a protein target and then use a compound library to screen for potential binders. As mentioned before, long AOFs can be now easily accessible in the laboratory thanks to the automation of SPFS and a given sequence can be decorated with proteinogenic or abiotic side chains, to design potential good candidates to recognize and bind large protein surface. Hence, in the context of this PhD thesis, we first sought to conceive long helical AOFs carrying diverse side chains to test their ability to randomly recognize proteins without prior side chains design. The length of foldamer is a critical factor since short sequence could only form limited area of interface while long sequences might encounter the difficulty of synthesis. After the determination of foldamer length, it could be applied as a bait to fish out prey proteins in the cell lysate, a natural protein library. This so-called pull-down assay (see chapter 2.3.1) enabled us to screen over thousands of protein candidates simultaneously. From the pull-down assay, 74 proteins were consistently observed. Next, considering the difficulty of protein expression and stability, in the further binding assay experiments, DNA repair and homologous recombination protein (Rad52),^[75] RNA binding protein (SGN1)^[76] and the coenzyme Q9 homolog protein (COQ9)^[77] were selected as model proteins to assess their binding to the AOF. As we have defined in chapter 2.4, the aromatic foldamer helix is achiral, so in the performed pull-down assay, what we actually screened were the two handedness (*P* and *M* enantiomers) against a pool of proteins. If a tight recognition took place on the surface of a prey protein, we should observe a different binding affinity between *P*- and *M*- helix since the surface of *P* and *M* helices is different as mirror image to each other. For binding assay, we turned to BLI kinetic experiments due to the possibility to immobilize each helix handedness (ligand) on the sensor tip and have the protein in solution (analyte). After successful, chiral HPLC purification in reversed phase mode, biotinylated *P* and *M* enantiomers were individually loaded on the SA sensor tips, and binding was investigated in the presence of three proteins introduced above. Rad52 showed nanomolar binding with both *P* and *M*-helix conformers. Since foldamer possesses high binding affinity to the protein in cell lysate, assays of how foldamer affects the cellular life activities could be considered. Research on peptide drugs has evolved rapidly in the past decade and is the focus of interest by pharmaceutical companies. Their medium size is perfectly suitable to target protein-protein interfaces

which is unpractical with small molecules. Peptides possess the advantages of higher potential for cell penetration, stable metabolic properties and lower costs of production.^[78] Besides natural peptides, non-natural peptides consisting of standard amino acids exhibit novel properties.^[79] The use of flexizymes enables tRNA charging with the non-standard amino acids, which can be applied in *in vitro* translation system and mRNA display. With the help of flexizymes, the peptide library was expanded for RaPID system (FIT and RaPID system, see details in chapter 2.3.3). As earlier introduced, AOFs can exhibit a surface comparable to a small protein, so the second objective of this thesis was therefore to challenge the RaPID technology by using a helical AOF as a target, instead of a protein, to select peptide macrocycles as good AOF binders. The arrangement of side chains on a oligoquinoline foldamers can be represented as a five-pointed star and allowed for the design of two faces: one face projected to the solvent with water-solubilizing side-chains and the other for specific recognition (**Figure 11** shows an example of 8mer crystal structure).^[80] Several peptide macrocycle candidates were selected by RaPID system against two foldamers, which shared the same types and number of side chains but differed in the position along the sequence, forming two different surfaces. Thus, each sequence could be regarded as a negative control to the other. Because of the instability of one monomer, new sequences were synthesized and binding affinity to peptide macrocycle was measured by SPR experiment. We showed that the replacing the instable side chain with other residues had no dramatic effect on the binding affinity to the selected peptide macrocycle, which indicated that the degradation of foldamer had no dramatic effect on the selection process or the side-chain remained stable during the course of the selection and that the selected peptide macrocycles were true binders. Binding constants were determined by SPR and BLI tests which validate the specific binding, namely one peptide macrocycle showed significant binding difference between sequences with different surfaces. We also wanted to investigate the selective interaction mode between enantiomers, since peptide macrocycles are intrinsic chiral substance, which means that the recognition of foldamer to these peptides should also be diastereoselective.

Based on previous work in our group in terms of foldamer-protein interaction,^[65, 81] we wanted to pave the way for structure-based design of AOFs tightly binding to a specific protein. HCAII was chosen because it is easy to crystallize and commercially available. The helical foldamer sequences tethering with a nanomolar binding ligand of HCA could cover large surface of HCAII.^[53] The conjugation with the arylsulfonamide ligand could bring the foldamer in direct proximity to the protein surface and X-ray crystal structure elucidation of protein-foldamer complex unveiled a high number of hydrophobic

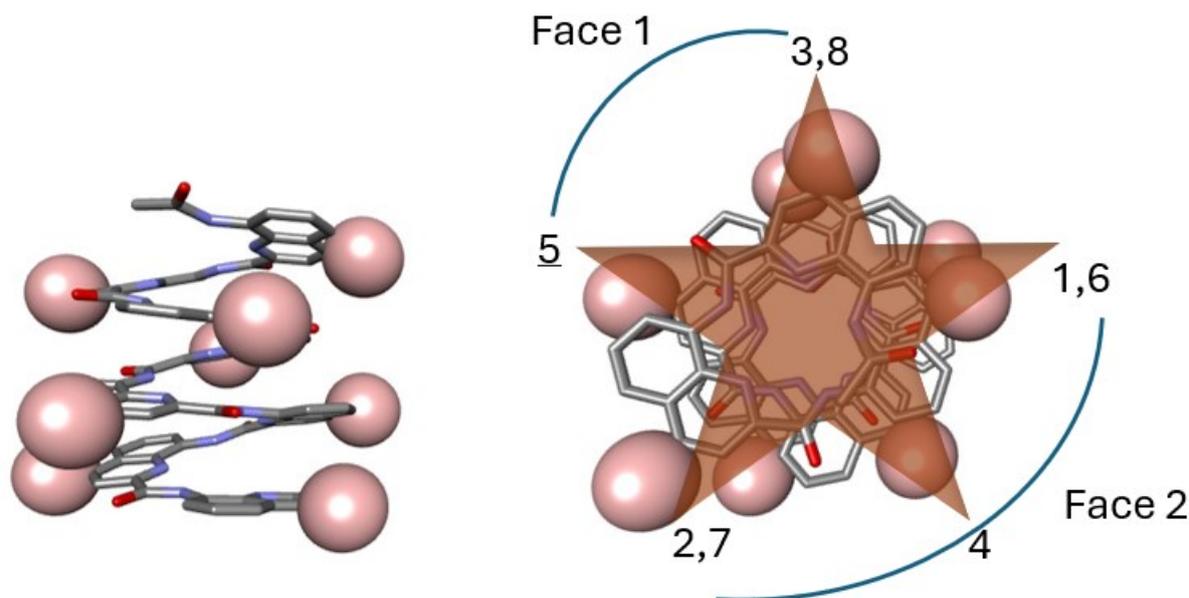


Figure 11: Front view (left) and top view (right) of crystal structure of a homo-quinoline octamer;^[80] the hydrogen bonds were omitted for clarification; side chains were simplified and represented as pink balls. The arrangement of side chains can be divided into two faces and to facilitate the design.

contacts. What we wanted to achieve, during the course of this thesis, was to conceive foldamers long enough to satisfy a good binding affinity to HCAII mainly driven by the side chains of the monomers. The design and choice of side chains residues was determined by Alphaspace 2.0, a recently developed computational analysis tool for topographical mapping of biomolecular concavities,^[82] and by the feasibility of synthesis. Based on the previous crystal structure of protein-foldamer complex, many areas on the protein surface are observed as reachable by the side chain proposals from Alphaspace 2.0. Considering the flexibility of helix and preservation of the backbone position, the replacement of **Q** with **P** or **B** unit needs to be planned accordingly. We basically used X-ray crystal structure elucidation to analyse the newly targeted protein surface interaction. Consequently, the AOF sequence design followed certain principles since we would like to preserve the helical handedness of the foldamer. Indeed, the *P* helical conformation of foldamer was known to favour the interactions on HCAII surface. Therefore, the introduction of a chiral **B** unit had to satisfy the same handedness bias toward the *P* helix. We have introduced the newly discovered side chains in an iterative manner on a lead foldamer sequence^[53] and sought to observe a cumulative effect of this side chain implementation on the binding affinity. To gather binding affinity data, we first endeavoured to use SPR then later BLI, and we finally set-up a fluorescence spectroscopy experiment to precisely quantify the binding affinity since the k_{off} of the foldamer sequences bound to HCAII revealed to be extremely low preventing the determination of the K_D by direct fitting of the kinetic curves recorded by SPR and BLI ($K_D = k_{\text{off}}/k_{\text{on}}$).

4. References

- [1] a) M. Bilal, J. Cui, H. M. N. Iqbal, *International Journal of Biological Macromolecules* **2019**, *130*, 186-196; b) M. Bilal, T. Rasheed, Y. Zhao, H. M. N. Iqbal, J. Cui, *International Journal of Biological Macromolecules* **2018**, *119*, 278-290.
- [2] S. Bartoszewska, J. F. Collawn, *Cellular & Molecular Biology Letters* **2020**, *25*, 18.
- [3] a) P. Gagneux, *Journal of Heredity* **2004**, *95*, 274-274; b) B. J. Pieters, M. B. van Eldijk, R. J. Nolte, J. Mecinovic, *Chem Soc Rev* **2016**, *45*, 24-39.
- [4] E. J. Stollar, D. P. Smith, *Essays in Biochemistry* **2020**, *64*, 649-680.
- [5] a) T. C. a. J. A. Wells, *Science* **1995**, *267*, 383-386; b) I. S. Moreira, P. A. Fernandes, M. J. Ramos, *J Comput Chem* **2007**, *28*, 644-654; c) O. Keskin, B. Ma, R. Nussinov, *J Mol Biol* **2005**, *345*, 1281-1294.
- [6] a) H. Yin, A. D. Hamilton, *Angew Chem Int Ed Engl* **2005**, *44*, 4130-4163; b) W. L. DeLano, *Curr Opin Struct Biol* **2002**, *12*, 14-20.
- [7] M. Moinul, S. Khatun, S. A. Amin, T. Jha, S. Gayen, *Biochem Pharmacol* **2022**, *206*, 115301.
- [8] V. Blay, B. Tolani, S. P. Ho, M. R. Arkin, *Drug Discovery Today* **2020**, *25*, 1807-1821.
- [9] D. C. Rees, M. Congreve, C. W. Murray, R. Carr, *Nat Rev Drug Discov* **2004**, *3*, 660-672.
- [10] S.-K. Kim, *International Journal of Molecular Sciences* **2023**, *24*, 3258.
- [11] G. J. Kaczorowski, O. B. McManus, B. T. Priest, M. L. Garcia, *J Gen Physiol* **2008**, *131*, 399-405.
- [12] H. Yin, A. D. Flynn, *Annu Rev Biomed Eng* **2016**, *18*, 51-76.
- [13] H. Lu, Q. Zhou, J. He, Z. Jiang, C. Peng, R. Tong, J. Shi, *Signal Transduct Target Ther* **2020**, *5*, 213.
- [14] X. Xie, T. Yu, X. Li, N. Zhang, L. J. Foster, C. Peng, W. Huang, G. He, *Signal Transduct Target Ther* **2023**, *8*, 335.
- [15] M. C. Smith, J. E. Gestwicki, *Expert Rev Mol Med* **2012**, *14*, e16.
- [16] V. Azzarito, K. Long, N. S. Murphy, A. J. Wilson, *Nat Chem* **2013**, *5*, 161-173.
- [17] C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, *Adv Drug Deliv Rev* **2001**, *46*, 3-26.
- [18] A. Monti, L. Vitagliano, A. Caporale, M. Ruvo, N. Doti, *Int J Mol Sci* **2023**, *24*.
- [19] M. L. Jones, M. A. Alfaleh, S. Kumble, S. Zhang, G. W. Osborne, M. Yeh, N. Arora, J. J. C. Hou, C. B. Howard, D. Y. Chin, S. M. Mahler, *Scientific Reports* **2016**, *6*, 26240.
- [20] E. Jankowsky, M. E. Harris, *Nat Rev Mol Cell Biol* **2015**, *16*, 533-544.
- [21] C. Lehmann, T. Friess, F. Birzele, A. Kiialainen, M. Dangel, *J Hematol Oncol* **2016**, *9*, 50.
- [22] D. Sun, Z. Li, Y. Rew, M. Gribble, M. D. Bartberger, H. P. Beck, J. Canon, A. Chen, X. Chen, D. Chow, J. Deignan, J. Duquette, J. Eksterowicz, B. Fisher, B. M. Fox, J. Fu, A. Z. Gonzalez, F. Gonzalez-Lopez De Turiso, J. B. Houze, X. Huang, M. Jiang, L. Jin, F. Kayser, J. J. Liu, M. C. Lo, A. M. Long, B. Lucas, L. R. McGee, J. McIntosh, J. Mihalic, J. D. Oliner, T. Osgood, M. L. Peterson, P. Roveto, A. Y. Saiki, P. Shaffer, M. Toteva, Y. Wang, Y. C. Wang, S. Wortman, P. Yakowec, X. Yan, Q. Ye, D. Yu, M. Yu, X. Zhao, J. Zhou, J. Zhu, S. H. Olson, J. C. Medina, *J Med Chem* **2014**, *57*, 1454-1472.
- [23] P. Holzer, K. Masuya, P. Furet, J. Kallen, T. Valat-Stachyra, S. Ferretti, J. Berghausen, M. Bouisset-Leonard, N. Buschmann, C. Pissot-Soldermann, C. Rynn, S. Ruetz, S. Stutz, P. Chene, S. Jeay, F. Gessier, *J Med Chem* **2015**, *58*, 6348-6358.
- [24] S. Surget, M. P. Khoury, J. C. Bourdon, *Onco Targets Ther* **2013**, *7*, 57-68.
- [25] D. W. Meek, C. W. Anderson, *Cold Spring Harb Perspect Biol* **2009**, *1*, a000950.
- [26] P. H. Kussie, S. Gorina, V. Marechal, B. Elenbaas, J. Moreau, A. J. Levine, N. P. Pavletich, *Science* **1996**, *274*, 948-953.
- [27] K. Wang, Yin, X.M., Chao, D.T., Milliman, C.L., Korsmeyer, S.J., *Genes & Development* **1996**, *10*, 2859-2869.
- [28] Y. Urakubo, T. Ikura, N. Ito, *Protein Sci* **2008**, *17*, 1055-1065.
- [29] S. Malhotra, A. P. Joseph, J. Thiyagalingam, M. Topf, *Nat Commun* **2021**, *12*, 3399.
- [30] J. Mbianda, M. Bakail, C. André, G. Moal, M. E. Perrin, G. Pinna, R. Guerois, F. Becher, P. Legrand, S. Traoré, C. Douat, G. Guichard, F. Ochsenbein, *Science Advances* **2021**, *7*, eabd9153.

- [31] M. P. Williamson, in *Modern Magnetic Resonance* (Ed.: G. A. Webb), Springer International Publishing, Cham, **2018**, pp. 995-1012.
- [32] J. S. Lin, E. M. Lai, *Methods Mol Biol* **2017**, *1615*, 211-219.
- [33] D. S. Wilson, A. D. Keefe, J. W. Szostak, *Proc Natl Acad Sci U S A* **2001**, *98*, 3750-3755.
- [34] a) G. Appiah Kubi, P. G. Dougherty, D. Pei, *Methods Mol Biol* **2019**, *2001*, 41-59; b) L. Gentilucci, R. De Marco, L. Cerisoli, *Curr Pharm Des* **2010**, *16*, 3185-3203.
- [35] A. A. Vinogradov, Y. Yin, H. Suga, *J Am Chem Soc* **2019**, *141*, 4167-4181.
- [36] H. Murakami, A. Ohta, H. Ashigai, H. Suga, *Nat Methods* **2006**, *3*, 357-359.
- [37] Y. Goto, A. Ohta, Y. Sako, Y. Yamagishi, H. Murakami, H. Suga, *ACS Chemical Biology* **2008**, *3*, 120-129.
- [38] S. Dengler, R. T. Howard, V. Morozov, C. Tsiamantas, W. E. Huang, Z. Liu, C. Dobrzanski, V. Pophristic, S. Brameyer, C. Douat, H. Suga, I. Huc, *Angew Chem Int Ed Engl* **2023**, *62*, e202308408.
- [39] a) G. Zemora, C. Waldsich, *RNA Biol* **2010**, *7*, 634-641; b) J. D. C. Watson, F. H. Nature **1953**, *171*, 737-738.
- [40] C. B. Anfinsen, *Science* **1973**, *181*, 223-230.
- [41] a) C. M. Goodman, S. Choi, S. Shandler, W. F. DeGrado, *Nat Chem Biol* **2007**, *3*, 252-262; b) C. J. Knill, J. F. Kennedy, *International Journal of Biological Macromolecules* **2008**, *42*, 82.
- [42] S. H. Gellman, *Acc. Chem. Res* **1997**, *31*, 173-180.
- [43] G. Guichard, I. Huc, *Chem Commun (Camb)* **2011**, *47*, 5933-5941.
- [44] C. N. Pace, J. M. Scholtz, *Biophys J* **1998**, *75*, 422-427.
- [45] P. Y. Chou, G. D. Fasman, *Biochemistry* **1974**, *13*, 211-222.
- [46] I. V. Korendovych, A. Senes, Y. H. Kim, J. D. Lear, H. C. Fry, M. J. Therien, J. K. Blasie, F. A. Walker, W. F. DeGrado, *Journal of the American Chemical Society* **2010**, *132*, 15516-15518.
- [47] H. M. Werner, W. S. Horne, *Curr Opin Chem Biol* **2015**, *28*, 75-82.
- [48] W. S. Horne, L. M. Johnson, T. J. Ketas, P. J. Klasse, M. Lu, J. P. Moore, S. H. Gellman, *Proc Natl Acad Sci U S A* **2009**, *106*, 14751-14756.
- [49] G. A. Eddinger, S. H. Gellman, *Angew Chem Int Ed Engl* **2018**, *57*, 13829-13832.
- [50] W. S. Horne, S. H. Gellman, *Accounts of Chemical Research* **2008**, *41*, 1399-1408.
- [51] D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes, J. S. Moore, *Chemical Reviews* **2001**, *101*, 3893-4012.
- [52] a) S. De, B. Chi, T. Granier, T. Qi, V. Maurizot, I. Huc, *Nat Chem* **2018**, *10*, 51-57; b) D. Mazzier, S. De, B. Wicher, V. Maurizot, I. Huc, *Angew Chem Int Ed Engl* **2020**, *59*, 1606-1610.
- [53] P. S. Reddy, B. Langlois d'Estaintot, T. Granier, C. D. Mackereth, L. Fischer, I. Huc, *Chemistry* **2019**, *25*, 11042-11047.
- [54] M. Zwillinger, P. S. Reddy, B. Wicher, P. K. Mandal, M. Csekei, L. Fischer, A. Kotschy, I. Huc, *Chemistry* **2020**, *26*, 17366-17370.
- [55] J. M. Davis, L. K. Tsou, A. D. Hamilton, *Chem Soc Rev* **2007**, *36*, 326-334.
- [56] a) Y. Hamuro, S. J. Geib, A. D. Hamilton, *Journal of the American Chemical Society* **1996**, *118*, 7529-7541; b) Y. Hamuro, S. J. Geib, A. D. Hamilton, *Angewandte Chemie International Edition in English* **2003**, *33*, 446-448.
- [57] E. R. G. P. S. Shirude, S. Ladame, F. Godde, K. Shin-ya, I. Huc, *J. Am. Chem. Soc.* **2007**, 11890-11891.
- [58] 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.
- [59] B. Teng, P. K. Mandal, L. Allmendinger, C. Douat, Y. Ferrand, I. Huc, *Chem Sci* **2023**, *14*, 11251-11260.
- [60] K. Aftahy, P. Arrasate, P. V. Bashkirov, P. I. Kuzmin, V. Maurizot, I. Huc, V. A. Frolov, *Journal of the American Chemical Society* **2023**, *145*, 25150-25159.
- [61] M. Pasco, C. Dolain, G. Guichard, in *Comprehensive Supramolecular Chemistry II*, **2017**, pp. 89-125.
- [62] A. M. Kendhale, L. Poniman, Z. Dong, K. Laxmi-Reddy, B. Kauffmann, Y. Ferrand, I. Huc, *J Org Chem* **2011**, *76*, 195-200.

- [63] N. Delsuc, T. Kawanami, J. Lefeuvre, A. Shundo, H. Ihara, M. Takafuji, I. Huc, *Chemphyschem* **2008**, *9*, 1882-1890.
- [64] T. Qi, V. Maurizot, H. Noguchi, T. Charoenraks, B. Kauffmann, M. Takafuji, H. Ihara, I. Huc, *Chem Commun (Camb)* **2012**.
- [65] M. Vallade, P. Sai Reddy, L. Fischer, I. Huc, *European Journal of Organic Chemistry* **2018**, *2018*, 5489-5498.
- [66] L. Yang, C. Ma, B. Kauffmann, D. Li, Q. Gan, *Org Biomol Chem* **2020**, *18*, 6643-6650.
- [67] D. Bindl, E. Heinemann, P. K. Mandal, I. Huc, *Chem Commun (Camb)* **2021**, *57*, 5662-5665.
- [68] V. Maurizot, C. Dolain, Y. Leydet, J.-M. Léger, P. Guionneau, I. Huc, *Journal of the American Chemical Society* **2004**, *126*, 10049-10052.
- [69] B. Baptiste, C. Douat-Casassus, K. Laxmi-Reddy, F. Godde, I. Huc, *J Org Chem* **2010**, *75*, 7175-7185.
- [70] M. R. B. a. G. D. Price, *Annu. Rev. Plam Physioi. Plam Mol. Biol* **1994**, *45*, 369-392.
- [71] C. B. Mishra, M. Tiwari, C. T. Supuran, *Med Res Rev* **2020**, *40*, 2485-2565.
- [72] C. T. Supuran, *Nat Rev Drug Discov* **2008**, *7*, 168-181.
- [73] J. S. C. Chu-Young Kim, Jeffrey B. Doyon, Teaster T. Baird, Carol A. Fierke, Ahamindra Jain and David W. Christianson, *J. Am. Chem. Soc.* **2000**, *122*, 12125-12134.
- [74] 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 Engl* **2014**, *53*, 883-887.
- [75] a) J. Deveryshetty, R. Chadda, J. R. Mattice, S. Karunakaran, M. J. Rau, K. Basore, N. Pokhrel, N. Englander, J. A. J. Fitzpatrick, B. Bothner, E. Antony, *Nat Commun* **2023**, *14*, 6215; b) N. V. Khade, T. Sugiyama, *PLoS One* **2016**, *11*, e0158436.
- [76] A. Gaspary, R. Laureau, A. Dyatel, G. Dursuk, Y. Simon, L. E. Berchowitz, *J Cell Biol* **2023**, *222*.
- [77] C. H. He, D. S. Black, T. P. Nguyen, C. Wang, C. Srinivasan, C. F. Clarke, *Biochim Biophys Acta* **2015**, *1851*, 1227-1239.
- [78] a) P. G. Dougherty, A. Sahni, D. Pei, *Chem Rev* **2019**, *119*, 10241-10287; b) E. A. Villar, D. Beglov, S. Chennamadhavuni, J. A. Porco, Jr., D. Kozakov, S. Vajda, A. Whitty, *Nat Chem Biol* **2014**, *10*, 723-731; c) L. Wang, N. Wang, W. Zhang, X. Cheng, Z. Yan, G. Shao, X. Wang, R. Wang, C. Fu, *Signal Transduct Target Ther* **2022**, *7*, 48.
- [79] Y. Li, P. A. Dalby, *Biosci Rep* **2022**, *42*.
- [80] X. Hu, S. J. Dawson, P. K. Mandal, X. de Hatten, B. Baptiste, I. Huc, *Chem Sci* **2017**, *8*, 3741-3749.
- [81] a) J. M. Alex, V. Corvaglia, X. Hu, S. Engilberge, I. Huc, P. B. Crowley, *Chem Commun (Camb)* **2019**, *55*, 11087-11090; b) M. Jewginski, T. Granier, B. Langlois d'Estaintot, L. Fischer, C. D. Mackereth, I. Huc, *J Am Chem Soc* **2017**, *139*, 2928-2931.
- [82] a) D. Rooklin, C. Wang, J. Katigbak, P. S. Arora, Y. Zhang, *J Chem Inf Model* **2015**, *55*, 1585-1599; b) J. Katigbak, H. Li, D. Rooklin, Y. Zhang, *J Chem Inf Model* **2020**, *60*, 1494-1508.

5. Main text: Interrogating the Potential of Helical Aromatic Foldamers for Protein Recognition

Authors: Sunbum Kwon,[‡] Vasily Morozov,[‡] Lingfei Wang,[‡] Pradeep K. Mandala, Stéphane Chaignepain, Céline Douat, and Ivan Huc*

Author contribution: The project was conceptualized by I. Huc. S. Kwon performed the synthesis of foldamer firstly. V. Morozov performed the protein expression. I improved the synthetic route of foldamers and re-synthesized the sequences and performed the chiral HPLC separation together with C. Douat. P. K. Mandala crystallized and solved the structure of foldamer. S. Chaignepain performed the pull-down assay. C. Douat co-supervised the project and assisted the chemical synthesis. The manuscript was written by C. Douat and I. Huc.

Cite this: *Org. Biomol. Chem.*, 2024, **22**, 9342Received 2nd September 2024,
Accepted 29th October 2024

DOI: 10.1039/d4ob01436g

rsc.li/obc

Interrogating the potential of helical aromatic foldamers for protein recognition†

Sunbum Kwon,^{†a,b} Vasily Morozov,^{†a} Lingfei Wang,^{†a} Pradeep K. Mandal,^{†a} Stéphane Chaignepain,^c Céline Douat^{†a} and Ivan Huc^{†a}*

A biotinylated helical aromatic oligoamide foldamer equivalent in size to a 24mer peptide was designed without any prejudice other than to display various polar and hydrophobic side chains at its surface. It was synthesized on solid phase, its *P*- and *M*-helical conformers were separated by HPLC on a chiral stationary phase, and the solid state structure of a non-biotinylated analogue was elucidated by X-ray crystallography. Pull-down experiments from a yeast cell lysate using the foldamer as a bait followed by proteomic analysis revealed potential protein binding partners. Three of these proteins were recombinantly expressed. Biolayer interferometry showed submicromolar binding demonstrating the potential of a given foldamer to have affinity for certain proteins in the absence of design considerations. Yet, binding selectivity was low in all three cases since both *P*- and *M*-conformers bound to the proteins with similar affinities.

With their main chain amide functions, their variety of hydrophobic, polar neutral, cationic and anionic side chains arranged at defined positions in space, and their complex shapes, proteins are ideally suited to selectively interact with one another. An illustration of this fact resides in the very propensity of the α -peptide backbone to adopt folded conformations – folding can be viewed as a kind of self-recognition. It follows that peptides, particularly peptide macrocycles,¹ and proteins themselves are prime candidates for the selective recognition of protein surfaces. Purposely developed for that task are naturally occurring antibodies as well as recombinant

proteins amenable to *in vitro* display selection, *e.g.* phage-, ribosome-, or mRNA-display, when their structure is stable enough to withstand the randomization of multiple surface residues.² To bind a protein surface, being protein-like is not a requirement, as nucleic acid aptamers illustrate,³ but it is certainly an advantage.⁴ There is thus currently great interest in the design of proteomimetics, synthetic molecules that would reproduce some features of proteins, in particular in the context of molecular recognition.⁵

We and others have been developing aromatic amide foldamers that adopt stable, predictable, helically folded conformations.⁶ Recent designs show that such synthetic objects may reach the size of small proteins.⁷ Although their aromatic backbone and folding propensity are remote from those of proteins, their surface can be decorated with proteinogenic side chains at precise locations.⁸ Their ability to recognize protein surfaces is currently being explored,⁹ with emphasis on targeting amyloid proteins¹⁰ as well as DNA-binding proteins.¹¹ It is hoped that such large foldamers may become efficient tools to bind sizeable protein surface areas and serve as competitive inhibitors of protein–protein and protein–nucleic acid interactions for pharmacological or even therapeutic applications, complementing antibodies and aptamers in such tasks. Indeed, abiotic foldamers may even overcome some of the disadvantages of proteins that can hamper their practical use. For example, proteins may be susceptible to proteolytic degradation and to denaturation which can in turn cause precipitation under certain conditions. In contrast, helical aromatic oligoamides are not degraded by proteases and, provided they are sufficiently hydrophilic, their clear solutions are unchanged after months of storage at 4 °C. However, designing a foldamer that specifically recognizes a given protein surface remains a challenging task and only multiple such endeavours will reveal the general suitability of the approach. Here, we sought to estimate the chances of a given proteomorphous aromatic foldamer to have affinity for any of the proteins present in a cell lysate. Using pull-down experiments and proteomic analysis, we observed significant enrichments and

^aDepartment of Pharmacy, Ludwig-Maximilians-Universität, Butenandtstraße 5–13, D-81377 München, Germany. E-mail: ivan.huc@cup.lmu.de

^bDepartment of Chemistry, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul 06974, Republic of Korea

^cCBMN (UMR5248), Univ. Bordeaux—CNRS—IPB, 2 rue Robert Escarpit, 33600 Pessac, France

† Electronic supplementary information (ESI) available: Supplementary figures and data, detailed experimental protocols and characterisation of new compounds. CCDC 2280177. For ESI and crystallographic data in CIF or other electronic format see DOI: <https://doi.org/10.1039/d4ob01436g>

* These authors contributed equally.



subsequently determined submicromolar dissociation constants (K_D), in the absence of any specific design. The capacity of a single foldamer to interact with proteins hints at a potential to interfere with their functions, although this was not investigated here. These results thus bode well for future developments and encourage further screening of various foldamers from different backbones and with different side chain arrangements, not only for binding, but also for their ability to modify interesting cellular phenotypes.

Foldamer 2 and its biotinylated analogue 1 consist of a dodecaamide of 8-amino-2-quinoline carboxylic acid (Fig. 1) bearing side chains in position 4. In water, such oligomers fold into stable aromatic helices having 2.5 units per turn and a pitch equal to the thickness of one aromatic ring.⁸ Since each δ -amino acid monomer is equivalent in size to a dipeptide, the helices of 1 and 2 reach the size of a 24mer peptide. These sequences, including final biotinylation, were synthesized on solid support using recently optimized protocols (Fig. S1 in the ESI†) and purified by reversed-phase HPLC.¹² The anionic, cationic, polar neutral, or hydrophobic nature of the side chains and their location on the helix surface were chosen among the building blocks available at the time to promote various types of intermolecular interactions but without any prejudice of which protein could be recognized.

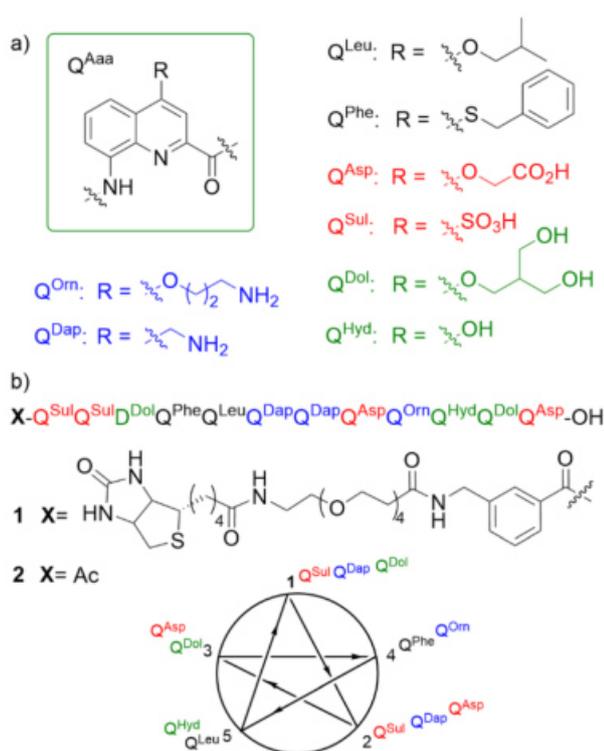


Fig. 1 (a) Foldamer aromatic building blocks colour coded according to their side chain nature: hydrophobic (black), polar neutral (green), cationic (blue) and anionic (red). (b) Sequence of dodecamer 2 used for X-ray crystallographic studies and its biotinylated version 1 used in the pull-down assay. The five-pointed star helix wheel representation depicts the side-chain positioning around the oligoquinoline backbone.

The arrangement of the side chains in space avoids that different parts of the helix surface resemble and thus enhance the chances that some proteins bind one area of the helix or another *via* shape recognition and complementary hydrophobic and electrostatic contacts. Sequences 1 and 2 contain several instances of contiguous residues that are identical or have similar polar or hydrophobic features. Yet, due to the high helix curvature – a monomer spans 0.4 helix turn – contiguous residues project their side chains towards different sides of the helix (Fig. 1b). Putting two identical residues contiguous in the sequence in fact contributes to the diversity of environments at the surface of the foldamer. The presentation of diverse arrays of side chains on the different sides of the helix is highlighted by the helix wheel in Fig. 1b.

High-quality crystals of compound 2 were obtained and diffracted at atomic resolution, allowing for the elucidation of its solid state structure (Fig. 2a). The top view of the structure illustrates the distribution of side chains. It also reveals a slight deviation of helix curvature from the usual 2.5 units per turn, possibly due to crystal packing interactions (Fig. S2†). The electrostatic charge potential shown on the solvent accessible surface of 2 (Fig. 2b) provides a graphical illustration of its proteomorphous nature. Despite the quinoline monomers being coarse, *i.e.* twice as large as an α -amino acid, the display of the side chains at the surface of the helix makes it look

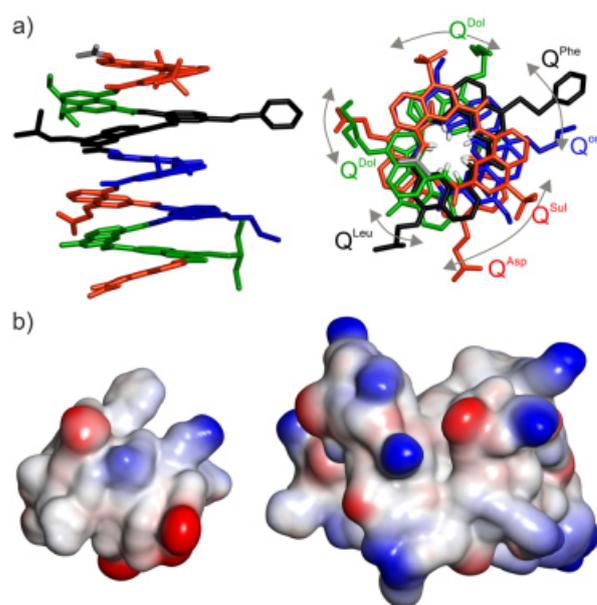


Fig. 2 (a) Solid state structure of foldamer 2 determined by X-ray crystallography. Each monomer in stick representation is colour-coded as in Fig. 1a. Included solvent molecules and hydrogen atoms are omitted for clarity, except in the top view at right where amide NH protons are shown in white at the inner rim of the helix. Bent arrows indicate side chains that should be aligned in the top view if curvature was exactly 2.5 units per turn (see Fig. 1b). (b) Comparison of the size, shape and electrostatic potential (blue: cationic, red: anionic, white: neutral) of foldamer 2 (left) and a small protein shown at the same scale (Sac7D, PDB #1AZQ, 66 residues)¹⁴ illustrating the mini-protein-like aspect of the foldamer.



protein-like. Its size remains small, comparable to that of a 24mer peptide. If required, substantially larger (as well as smaller) foldamers would be accessible through solid phase synthesis.

For the pull-down experiments, biotinylated foldamer **1** was immobilized on streptavidin-coated magnetic beads and used as a bait.¹³ Since the main chain helix of **1** does not contain stereogenic centers, it exists as a 1 : 1 mixture of right-handed (*P*) and left-handed (*M*) conformers. It is thus two, mirror-imaged helices that were used as baits. The beads were then incubated in a solution of cell lysate of *S. cerevisiae* (Fig. 3) to allow the foldamer to interact with putative prey proteins. After a thorough wash with phosphate buffered saline, bound prey proteins were eluted off from the beads (Fig. S3†) and subsequently digested with trypsin. Extracted peptides were then subjected to LC-MS/MS proteomic analysis. To ensure high fidelity of LC-MS/MS and assess statistical significance, the solution of extracted peptides was divided into three portions and analysed separately. In addition, the whole triplicate pull-down assay was repeated in three independent experiments. Proteomic analysis by LC-MS/MS spectrometry identified approximately 2000 proteins (Fig. S4 and ESI†). The abundance of proteins was compared to that obtained from control experiments with non-modified magnetic beads (*i.e.* streptavidin without foldamer as a bait). Identified prey proteins were then sorted based on criteria such as fold change and confidence. Across the three experiments, 153, 166 and 214 proteins with a fold change >2.0 were identified. Among those, 74 proteins were consistently observed with a fold change >2.0 in all three independent experiments, and another 75 proteins in two out of three experiments, highlighting the reproducibility of both protein pull-down and relative abundance (Fig. S5†). It should be pointed here that the conditions of the pull-down assay were not stringent, as reflected in the total number of proteins identified with high confidence and still large number of proteins with a fold change >2.0. Our priority was to establish the reproducibility of the assay and we started

without knowing what the outcome would be. In future experiments, one may consider using wash buffers more efficient at disrupting weak interactions than the phosphate buffered saline used here.

Among several candidates of prey proteins that were *a priori* amenable to simple recombinant expression in *E. coli* and were devoid of posttranslational modification, we chose to express four different proteins and assess their affinity for **1**. The DNA repair and homologous recombination protein Rad52¹⁵ showed reproducible, moderate enrichment and high confidence. Through the three pull-down assays, the fold changes of Rad52 were ranked in 20th, 22nd, and 22nd place among all identified proteins. The same stands for the RNA binding protein SGN1¹⁶ (ranked at the 31st, 25th and 43rd positions in the three pull-downs), the coenzyme Q9 homolog (COQ9) protein¹⁷ (ranked at the 50th, 5th and 32nd positions) and the splicing factor Mud2 (ranked at the 14th, 27th and 6th positions).¹⁸ All four proteins showed calculated *p*-values from Student's *T*-test that were below the threshold of 0.05, indicating a nominal statistical significance (Table 1).

All four proteins were recombinantly expressed in *E. coli*. However, Mud2 proved to be somewhat problematic (poor overexpression, propensity to precipitate) and was not considered further. Binding studies using biolayer interferometry (BLI) were performed with the three remaining proteins to measure both binding kinetics and dissociation constants (K_D). This technique was preferred because it allowed us to immobilize the biotinylated foldamer on the sensor and overcome problems associated with its aggregation or even poor solubility at concentrations that would be relevant for K_D determination in solution. Isothermal titration calorimetry, for example, was attempted but gave poor results, on top of requiring a lot of material. To assess binding selectivity, we set to measure binding to *P*-1 and *M*-1 separately. These two com-

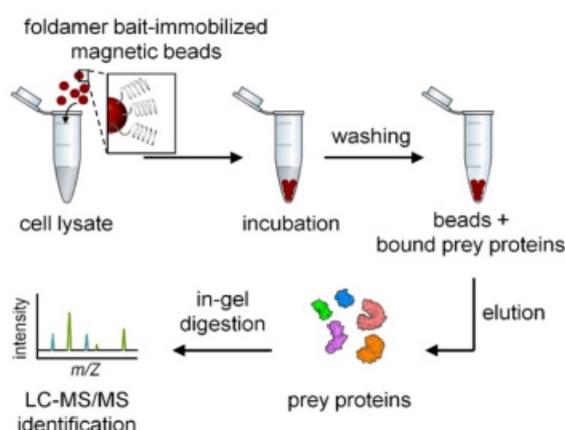


Fig. 3 Schematic representation of the pull-down assay and proteomic analysis.

Table 1 Fold change and statistical evaluation of Rad 52, SGN1, COQ9 and Mud2 from three independent pull-down experiments from a *S. cerevisiae* cell-lysate using **1** as a bait

Exp. number	Fold change (rank) ^a	<i>T</i> -test ^b
Rad52		
1 st	7.57 (20 th)	4.45×10^{-3}
2 nd	6.44 (22 nd)	8.63×10^{-3}
3 rd	11.39 (22 nd)	1.98×10^{-6}
SGN1		
1 st	5.63 (31 st)	2.31×10^{-3}
2 nd	5.89 (25 th)	9.65×10^{-3}
3 rd	6.51 (43 rd)	2.92×10^{-6}
COQ9		
1 st	4.01 (50 th)	4.71×10^{-3}
2 nd	17.85 (5 th)	8.43×10^{-3}
3 rd	8.51 (32 nd)	1.36×10^{-3}
Mud2		
1 st	10.19 (14 th)	9.69×10^{-4}
2 nd	5.38 (27 th)	3.99×10^{-5}
3 rd	27.89 (6 th)	1.64×10^{-4}

^a Enrichment in abundance when the bait was replaced from streptavidin to foldamer **1**. ^b *T*-test uses the two-tailed distribution for samples with unequal variances.



pounds may in principle serve as ideal controls for one another in that their physical properties are essentially identical – thus alleviating non-specific effects – but the projection of their side chains are mirror images and may thus not lead to identical recognition properties. We thus endeavoured in the separation of *P*-1 and *M*-1 by chiral HPLC. Note that the handedness of helices of such length is kinetically inert in water at 25 °C: despite being conformational isomers, *P*-1 and *M*-1 do not interconvert to a detectable extent.¹⁹

In previous studies, we have demonstrated the possibility to separate the *P*- and *M*-helical conformers of organo-soluble aromatic oligoamides of moderate sizes (from hexamer to hexadecamer) by chiral HPLC using a Chiralpak™ IA stationary phase under normal phase conditions (*i.e.* an *n*-hexane/chloroform mixture).²⁰ Here, we sought for a stationary phase and eluting conditions allowing for a separation in reversed-phase mode. We selected the Chiralpak™ QN-AX, a quinine-based stationary phase initially developed for the separation of chiral α -amino acids.²¹ Finding elution conditions took some screening and optimization but we eventually found that a solvent mixture composed of 30% acetonitrile in triethylammonium acetate buffer (150 mM, pH 7.21) resulted in an impressive separation of the two helical conformers with a separation factor of 3.7 between *M*-1 and *P*-1. The assignment of the handedness was made using a circular dichroism (CD) detector connected after the UV/Vis detector (Fig. S6†) and based on previously published absolute handedness assignment of such aromatic helices.²² Thanks to the remarkable peak separation, we could directly use the analytical column to successfully isolate both helix conformers from the *M*-1/*P*-1 mixture. We finally further confirmed the chiral purity by reinjecting the collected pure fractions on the same column (Fig. S7†). Even an octamer shows no detectable handedness inversion after long incubation in water.^{19,20,23}

The biotinylated *P*- and *M*-helices of 1 were next independently immobilized on BLI streptavidin sensors tips. Loading was performed at a concentration of 2 $\mu\text{g mL}^{-1}$. The sensors were then exposed to a range of protein concentrations for time course monitoring of the association before being dipped in a buffer solution to record the dissociation. The interaction between Rad52 and between *P*-1 and *M*-1 was assessed in multiple runs with a protein concentration ranging from 125 nM to 7.75 nM. Real-time BLI sensorgrams fit very well to a 1 : 1 kinetic binding model. Rapid association (large k_a) and slow dissociation (low k_d) were calculated by curve fitting (Fig. 4, Table 2) yielding a remarkably low nanomolar K_D . The values obtained for *M*-1 and *P*-1 differed marginally (1.3 and 1.5 nM, respectively), suggesting that the chiral features of 1 are not critical for binding. The lack of selectivity came as a surprise given the low K_D values. The exact binding mode was not further investigated. Nevertheless, one may recall that yeast Rad52 is a 490 amino acid protein. A 90-residue domain near the N-terminus is known to form a stable decamer^{15a} with an inner cavity of 32 nm. Most of the remaining sequence is not seen experimentally by cryo-electron microscopy^{15a} and is predicted by AlphaFold not to belong to a stable tertiary structure

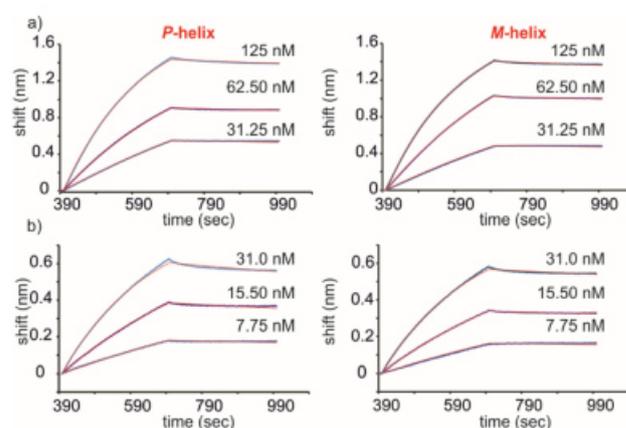


Fig. 4 Distinct BLI experiments report on the 1 : 1 binding affinity of protein Rad52 to *P*-1 (left graphs) and *M*-1 (right graphs) yielding $K_D = 1.5$ and 1.3 nM, respectively. Both biotinylated chiral helices were immobilized onto the BLI SA sensors tips ($3 \times$ *P*-helix and $3 \times$ *M*-helix tips at $2 \mu\text{g mL}^{-1}$ for the loading and two reference tips were used). The BLI tips were then dipped into solutions containing varying concentrations of Rad52 in buffer (20 mM Hepes (pH 7.4), 150 mM NaCl, 0.02% Tween-20®) to record the concentration-dependent association over 240 s: (a) 125–31.25 nM and (b) 31–7.75 nM. Next the sensors were dipped in pure buffer solution to record dissociation events over 300 s. Experimental curves are shown in blue and calculated curves according to a 1 : 1 binding isotherm are shown in red.

Table 2 Kinetic binding constant (k_a), kinetic dissociation constant (k_d), and equilibrium dissociation constant (K_D) derived from BLI measurements performed for three proteins as analytes and either *P*-1 or *M*-1 as a ligand immobilized on streptavidin tips

Protein/ <i>P</i> -1 or <i>M</i> -1	k_a ($10^3 \text{ M}^{-1} \text{ s}^{-1}$)	k_d (10^{-6} s^{-1})	K_D (nM)
Rad52			
<i>P</i> -1	125 ± 0.36	187 ± 1.72	1.49 ± 0.015
<i>M</i> -1	124 ± 0.31	163 ± 1.51	1.31 ± 0.013
SGN1			
<i>P</i> -1	—	—	104.9 ± 10.7^a
<i>M</i> -1	—	—	115.5 ± 9.2^a
COQ9			
<i>P</i> -1	4.10 ± 0.033	2034 ± 15.22	$\approx 500^b$
<i>M</i> -1	3.38 ± 0.026	1695 ± 12.88	$\approx 500^b$

^aThe K_D values have been calculated assuming an equilibrium level of the SGN1 binding to the sensor tips (Langmuir's equation). ^bCurve fitting to a 1 : 1 binding model was not ideal. We thus choose not to report precise K_D values but rather an order of magnitude (see ESI†).

(Fig. S8†). In solution, a 10 : 1 foldamer binding ratio of the foldamer to the 10mer may be considered unless the foldamer binds in the central cavity of the 10mer. This is precluded in the BLI experiments where the foldamer is immobilized on the sensor. One may still envisage that a certain degree of multivalency (or "avidity"), *e.g.* two or more Rad52 subunits of the 10mer binding to two or more foldamers on the sensor, contribute to the low K_D values. We performed additional BLI experiments decreasing the loading of foldamer on the sensor to prevent multivalency and observed no change in the shape



of the response curve, which suggests the binding observed is indeed 1 : 1.

The binding of SGN1 was estimated in a similar manner. Rapid association and dissociation were recorded but, after the initial binding event when a steady state should be reached, a gradual linear climbing of the association step was noticed as if SGN1 was further associating to itself on top of its binding to **1** on the sensor. Adding 2% Bovine Serum Albumin (BSA) to the buffer reduced this phenomenon but did not completely remove it, making the curve fitting to a 1 : 1 kinetic binding model inaccurate. Consequently, the K_D values of SGN1 binding to *M-1* and *P-1* were calculated with the Langmuir's equation and assuming that equilibrium was reached after the initial climb of the signal (steady state). Binding in the three-digit nanomolar range was observed and again revealed similar association to *M-1* and *P-1* (K_D of 105 and 115 nM, respectively, Fig. S9†). Unlike for Rad52, not much is known about the structure of yeast SGN1 other than that it contains an RNA binding domain.¹⁶ AlphaFold predicts that only a small part of its 250 residues belong to a stable tertiary fold (Fig. S8†) but such a prediction may be taken with caution. In contrast, the structure of the 228-residue yeast COQ9 is not known but AlphaFold predicts a fully folded protein similar to the structure of human COQ9 which contains ten α -helices (Fig. S8†). For BLI measurement with COQ9, we screened a range of buffers with variable amounts of BSA (1–2%) and/or Tween-20® as a detergent (0.05–0.1%) to reduce protein aggregation and eventually obtained a good, albeit not ideal, fit to a 1 : 1 kinetic binding model. K_D is therefore provided as an order of magnitude (Table 2). The best binding behaviour was again obtained with HEPES buffer. Binding of COQ9 to **1** was weaker than for the two other proteins, although still in the submicromolar range, with again no preference for the *P*- or *M*-helix (Table 2 and Fig. S10†).

Conclusions

A single helical aromatic foldamer, not resulting from a specific design but simply displaying different side chains at its surface, was shown to possess nanomolar affinity for certain proteins identified from a cell lysate. These results bode well for the use of such medium-sized molecules to interfere with cellular functions. While efforts towards the structure-based design of helical foldamers to recognize protein surfaces are already under way,^{9–11} these new results suggest that simple screening approaches may also yield promising results. The current study focused on the detection of binding and led to the identification of protein binders that do not necessarily play an important role in diagnostics or therapeutic intervention. For that reason, further investigations on the selectivity of binding, on the thermodynamics involved – are associations entropy or enthalpy driven – on the structural elucidation of the interactions involved, or on whether these interactions result in interference with protein function were not undertaken. Future efforts will instead focus on screening

the effects of various foldamers for relevant cell phenotypes ensuring that what is detected is not solely binding but also function.

Author contributions

SK, VM, LW, PKM, SC, and CD performed the experiments. IH supervised the research. All authors contributed to the writing, reviewing and editing of the manuscript and approved its final version.

Data availability

The data supporting this article have been included as part of the ESI. Crystallographic data for **2** has been deposited at the CCDC under # 2280177 and can be obtained from <https://www.ccdc.cam.ac.uk>.†

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) through project CRC1309-C7 (project ID 325871075) to I. H. We would like to thank Prof. T. Sugiyama for having kindly offered us the Rad52 construct.

References

- (a) A. A. Vinogradov, Y. Yin and H. Suga, *J. Am. Chem. Soc.*, 2019, **141**, 4167; (b) T. A. F. Cardote and A. Ciulli, *ChemMedChem*, 2016, **11**, 787; (c) P. Timmerman, J. Beld, W. C. Puijk and R. H. Melen, *ChemBioChem*, 2005, **6**, 821; (d) C. Heinis, T. Rutherford, S. Freund and G. Winter, *Nat. Chem. Biol.*, 2009, **5**, 502; (e) K. Maola, J. Wilbs, J. Touati, M. Sabisz, X.-D. Kong, A. Baumann, K. Deyle and C. Heinis, *Angew. Chem.*, 2019, **131**, 11927; (f) Z. C. Adams, A. P. Silvestri, S. Chiorean, D. T. Flood, B. P. Balo, Y. Shi, M. Holcomb, S. I. Walsh, C. A. Maillie, G. K. Pierens, S. Forli, K. J. Rosengren and P. E. Dawson, *ACS Cent. Sci.*, 2023, **9**, 648; (g) A. Krzyzanowski, L. M. Esser, A. Willaume, R. Prudent, C. Peter, P. Hart and H. Waldmann, *J. Med. Chem.*, 2022, **65**, 15300.
- (a) J. Helma, M. C. Cardoso, S. Muyldermans and H. Leonhardt, *J. Cell Biol.*, 2015, **209**, 633; (b) D. Schumacher, J. Helma, A. F. L. Schneider, H. Leonhardt and C. P. R. Hackenberger, *Angew. Chem., Int. Ed.*, 2018, **57**, 2314; (c) F. Y. Frejd and K.-T. Kim, *Exp. Mol.*



- Med.*, 2017, **49**, e306; (d) S. Schlehuber and A. Skerra, *Expert Opin. Biol. Ther.*, 2005, **5**, 1453.
- 3 (a) P. S. Pendergrast, H. N. Marsh, D. Grate, J. M. Healy and M. Stanton, *J. Biomol. Tech.*, 2005, **16**, 224; (b) M. Takahashi, *Biochimie*, 2018, **145**, 63; (c) Z. Zhang, J. Li, R. Amini, A. Mansfield, J. Gu, J. Xia, J. D. Brennan and Y. Li, *Analysis Sensing*, 2023, **3**, e202300001.
- 4 L.-G. Milroy, T. N. Grossmann, S. Hennig, L. Brunsveld and C. Ottmann, *Chem. Rev.*, 2014, **114**, 4695.
- 5 W. S. Horne and T. N. Grossmann, *Nat. Chem.*, 2020, **12**, 331.
- 6 (a) D.-W. Zhang, X. Zhao, J.-L. Hou and Z.-T. Li, *Chem. Rev.*, 2012, **112**, 5271; (b) I. Huc, *Eur. J. Org. Chem.*, 2004, 17.
- 7 (a) D. Mazzier, S. De, B. Wicher, V. Maurizot and I. Huc, *Angew. Chem., Int. Ed.*, 2020, **59**, 1606; (b) S. De, B. Chi, T. Granier, T. Qi, V. Maurizot and I. Huc, *Nat. Chem.*, 2018, **10**, 51.
- 8 X. Hu, S. J. Dawson, P. K. Mandal, X. de Hatten, B. Baptiste and I. Huc, *Chem. Sci.*, 2017, **8**, 3741.
- 9 (a) J. M. Alex, V. Corvaglia, X. Hu, S. Engilberge, I. Huc and P. B. Crowley, *Chem. Commun.*, 2019, **55**, 11087; (b) P. S. Reddy, B. Langlois d'Estaintot, T. Granier, C. D. Mackereth, L. Fischer and I. Huc, *Chem. – Eur. J.*, 2019, **25**, 11042; (c) M. Vallade, M. Jewginski, L. Fischer, J. Buratto, K. Bathany, J.-M. Schmitter, M. Stupfel, F. Godde, C. D. Mackereth and I. Huc, *Bioconjugate Chem.*, 2019, **30**, 54; (d) J. Buratto, C. Colombo, M. Stupfel, S. J. Dawson, C. Dolain, B. Langlois d'Estaintot, L. Fischer, T. Granier, M. Laguerre, B. Gallois and I. Huc, *Angew. Chem., Int. Ed.*, 2014, **53**, 883.
- 10 (a) S. Kumar, M. Birol, D. E. Schlamadinger, S. P. Wojcik, E. Rhoades and A. D. Miranker, *Nat. Commun.*, 2016, **7**, 11412; (b) S. Kumar, A. Henning-Knechtel, I. Chehade, M. Magzoub and A. D. Hamilton, *J. Am. Chem. Soc.*, 2017, **139**, 17098; (c) J. Ahmed, T. C. Fitch, C. M. Donnelly, J. A. Joseph, T. D. Ball, M. M. Bassil, A. Son, C. Zhang, A. Ledreux, S. Horowitz, Y. Qin, D. Paredes and S. Kumar, *Nat. Commun.*, 2022, **13**, 2273.
- 11 (a) K. Ziach, C. Chollet, V. Parissi, P. Prabhakaran, M. Marchivie, V. Corvaglia, P. P. Bose, K. Laxmi-Reddy, F. Godde, J. M. Schmitter, S. Chaignepain, P. Pourquier and I. Huc, *Nat. Chem.*, 2018, **10**, 511; (b) V. Corvaglia, D. Carbajo, P. Prabhakaran, K. Ziach, P. K. Mandal, V. D. Santos, C. Legeay, R. Vogel, V. Parissi, P. Pourquier and I. Huc, *Nucleic Acids Res.*, 2019, **47**, 5511; (c) V. Kleene, V. Corvaglia, E. Chacin, I. Forne, D. B. Konrad, P. Khosravani, C. Douat, C. F. Kurat, I. Huc and A. Imhof, *Nucleic Acids Res.*, 2023, **51**, 9629.
- 12 (a) V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat and I. Huc, *Chem. – Eur. J.*, 2023, e202300898; (b) B. Baptiste, C. Douat-Casassus, K. Laxmi-Reddy, F. Godde and I. Huc, *J. Org. Chem.*, 2010, **75**, 7175.
- 13 J. Maryas, J. Faktor, L. Čapkova, P. Müller, P. Skladal and P. Bouchal, *Cancer Genomics Proteomics*, 2018, **15**, 395.
- 14 S. Su, Y.-G. Gao, H. Robinson, Y.-C. Liaw, S. P. Edmondson, J. W. Shriver and A. H.-J. Wang, *J. Mol. Biol.*, 2000, **303**, 395.
- 15 (a) J. Deveryshetty, R. Chadda, J. R. Mattice, S. Karunakaran, M. J. Rau, K. Basore, N. Pokhrel, N. Englander, J. A. J. Fitzpatrick, B. Bothner and E. Antony, *Nat. Commun.*, 2023, **14**, 6215; (b) N. V. Khade and T. Sugiyama, *PLoS One*, 2016, **11**, e0158436.
- 16 A. Gasparly, R. Laureau, A. Dyatel, G. Dursuk, Y. Simon and L. E. Berchowitz, *J. Cell Biol.*, 2023, **222**, 1.
- 17 C. H. He, D. S. Black, T. P. T. Nguyen, C. Wang, C. Srinivasan and C. F. Clarke, *Biochim. Biophys. Acta*, 2015, **1851**, 1227.
- 18 (a) J.-C. Rain and P. Legrain, *EMBO J.*, 1997, **16**, 1759; (b) R. Minocha, V. Popova, D. Kopytova, D. Misiak, S. Hüttelmaier, S. Georgieva and K. Sträßer, *Nucleic Acids Res.*, 2018, **46**, 9749.
- 19 T. Qi, V. Maurizot, H. Noguchi, T. Charoenraks, B. Kauffmann, M. Takafuji, H. Ihara and I. Huc, *Chem. Commun.*, 2012, **48**, 6337.
- 20 N. Delsuc, T. Kawanami, J. Lefeuvre, A. Shundo, H. Ihara, M. Takafuji and I. Huc, *ChemPhysChem*, 2008, **9**, 1929.
- 21 M. Lämmerhofer and W. Lindner, *J. Chromatogr. A*, 1996, **741**, 33.
- 22 A. M. Kendhale, L. Poniman, Z. Dong, K. Laxmi-Reddy, B. Kauffmann, Y. Ferrand and I. Huc, *J. Org. Chem.*, 2011, **76**, 195.
- 23 S. J. Dawson, Á. Mészáros, L. Petho, C. Colombo, M. Csékei, A. Kotschy and I. Huc, *Eur. J. Org. Chem.*, 2014, 4265.



6. Supporting Information: Interrogating the Potential of Helical Aromatic Foldamers for Protein Recognition

Supporting Information

Interrogating the potential of helical aromatic foldamers for protein recognition

Sunbum Kwon,^{‡a,b} Vasily Morozov,^{‡a} Lingfei Wang,^{‡a} Pradeep K. Mandal,^a Stéphane Chaignepain,^c Céline Douat,^a and Ivan Huc^{*a}

^a Department of Pharmacy, Ludwig-Maximilians-Universität, Butenandtstraße 5-13, D-81377 Munich, Germany

^b Department of Chemistry, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul 06974, Republic of Korea

^c CBMN (UMR5248), Univ. Bordeaux—CNRS—IPB, 2 rue Robert Escarpit, 33600 Pessac, France

[‡] These authors contributed equally

6.1 Supplementary Figures

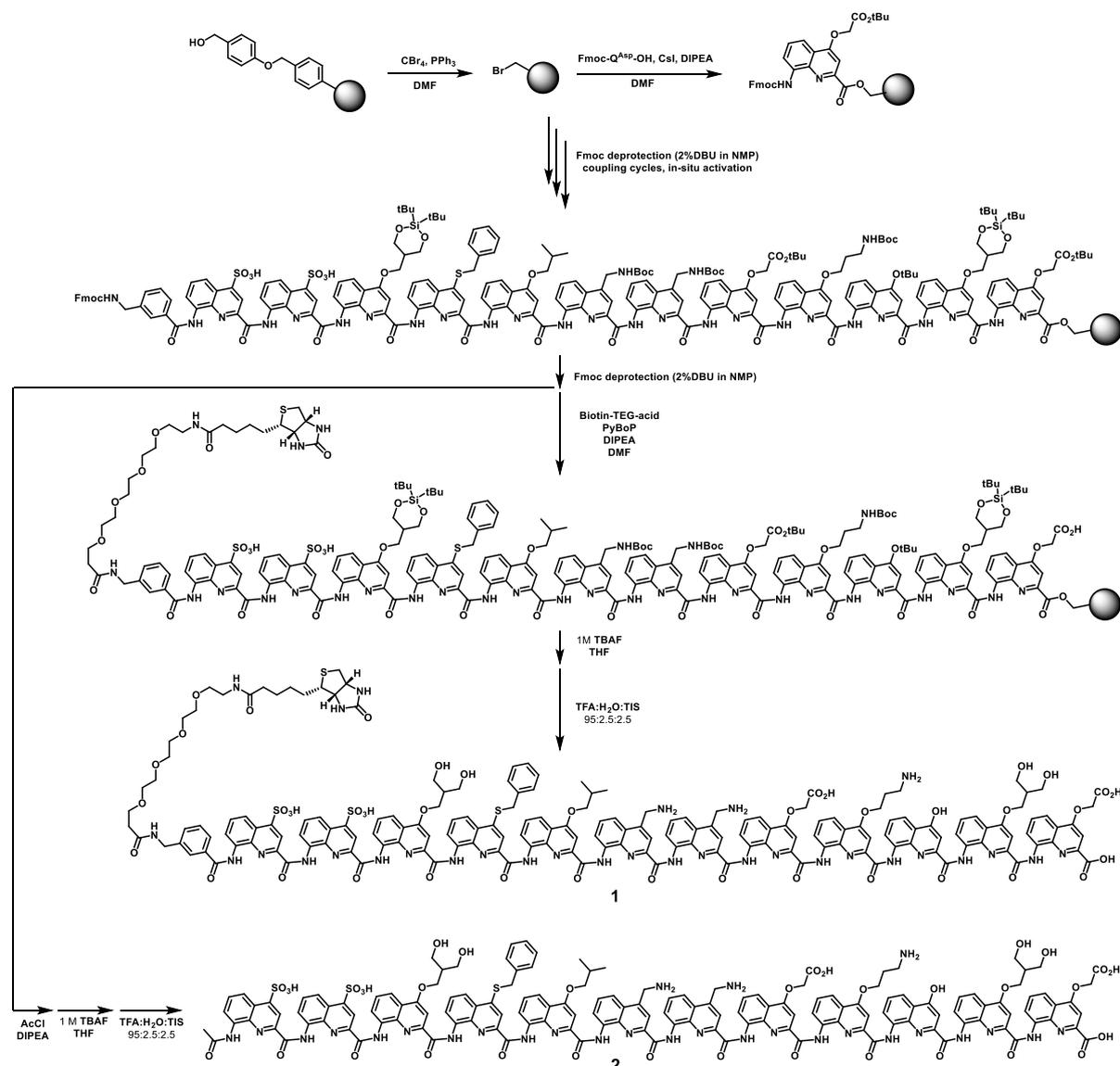


Figure S1. Solid phase synthesis of foldamers 1 and 2.

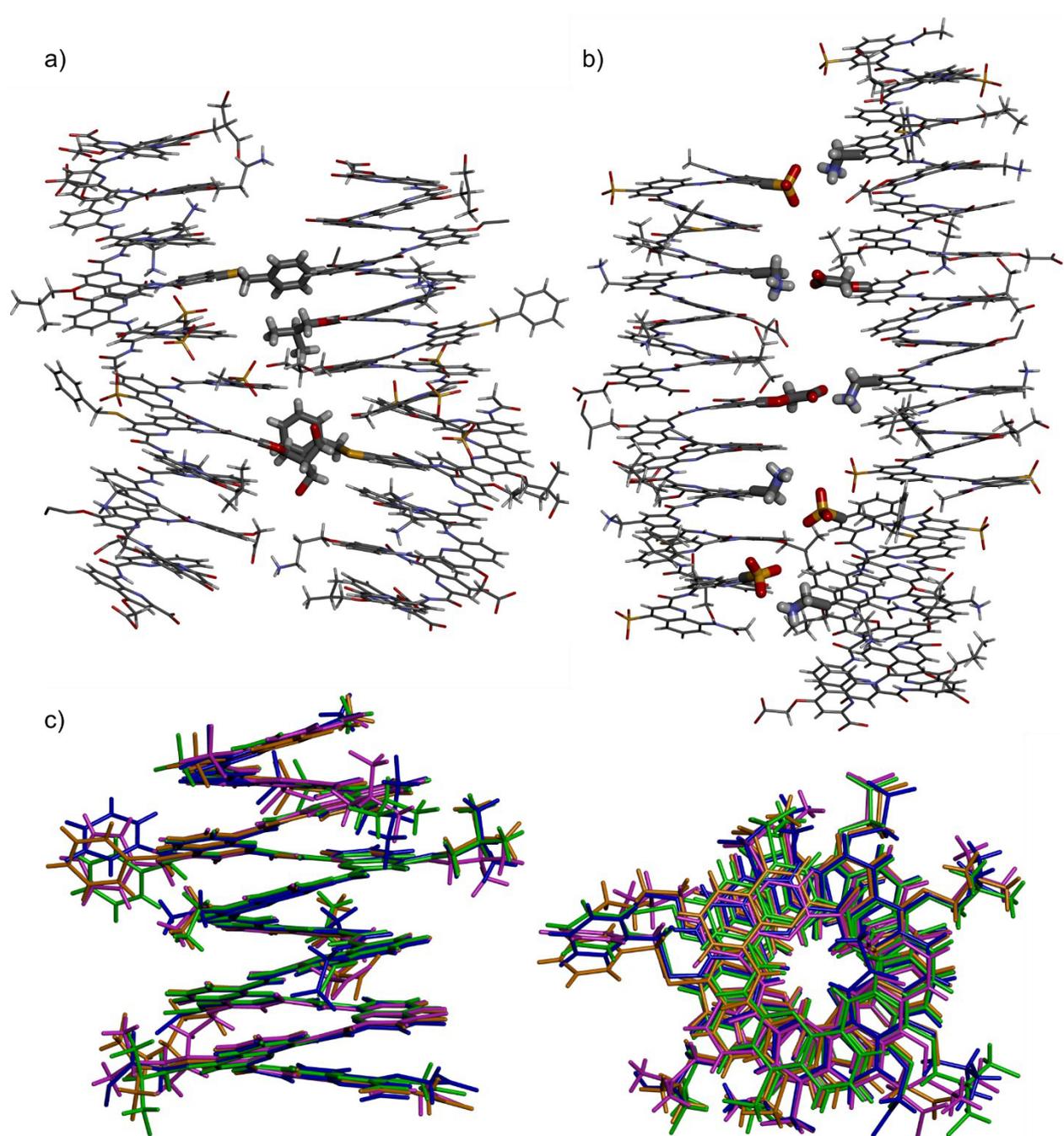


Figure S2. Crystal structure of **2**. a) Helix-helix contacts mediated by hydrophobic side chains (Q^{Phe} and Q^{Leu} or Q^{Dol}). b) Helix-helix interactions mediated by salt bridges between Q^{Dap} ammonium groups and Q^{Sul} sulfonate groups or Q^{Asp} carboxylate groups. In a) and b), the relevant side chains are shown in tube representation. c) Overlay of the main chain helix of the four crystallographically independent molecules found in the asymmetric unit.

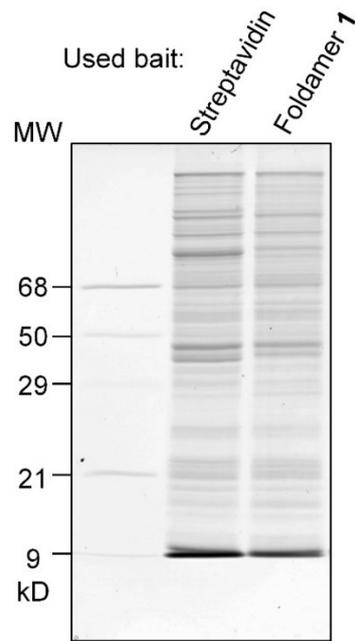


Figure S3. 10% SDS-PAGE of captured protein samples eluted from magnetic beads.

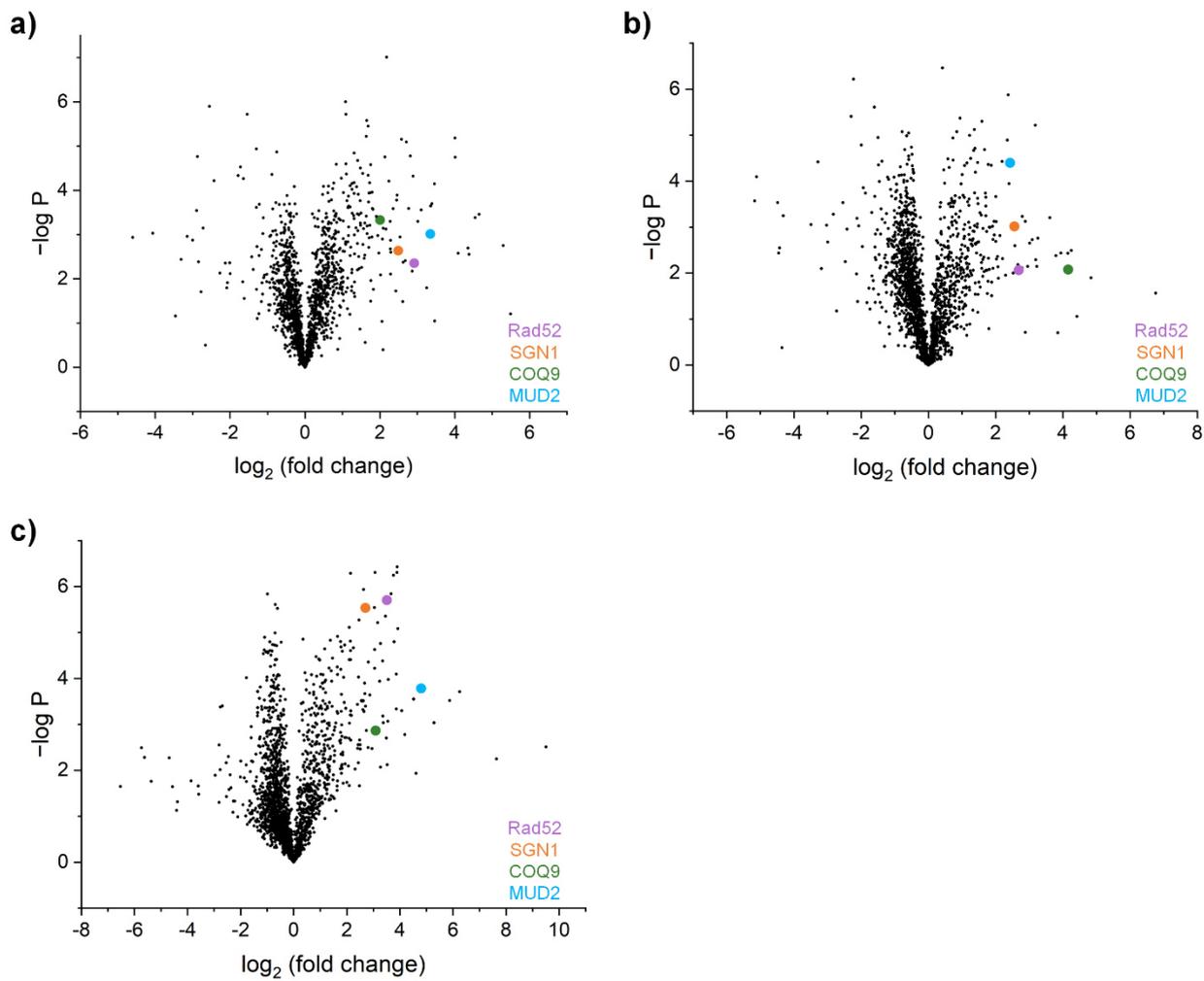


Figure S4. Volcano plots showing the significance (y -axis, $-\log P$ value) versus the enrichment (x -axis, \log_2 fold change) for each identified protein. Purple, orange, green and cyan dots indicate Rad52, SGN1, COQ9 and MUD2 proteins, respectively. Plots were obtained from (a) 1st, (b) 2nd and (c) 3rd pull-down experiments with foldamer 1, respectively.

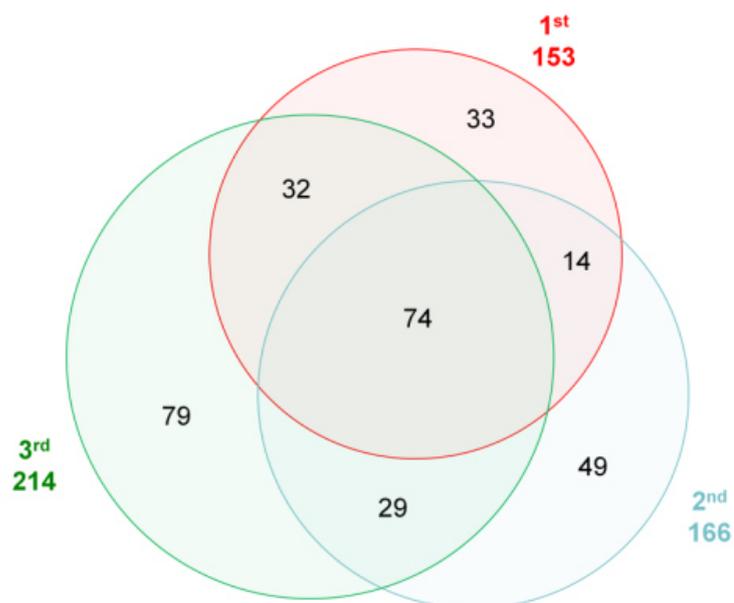
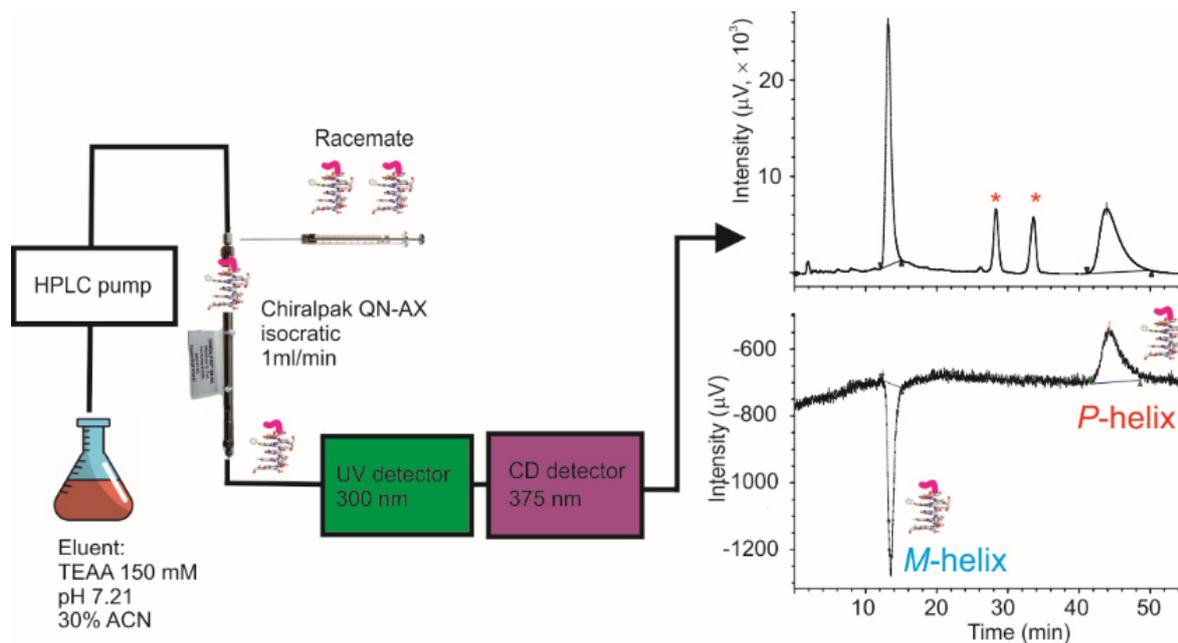


Figure S5. A Venn diagram showing the number of proteins that were enriched (fold change >2.00) when the foldamer bait was present. A total of 74 proteins exhibited enrichment across all three experiments.



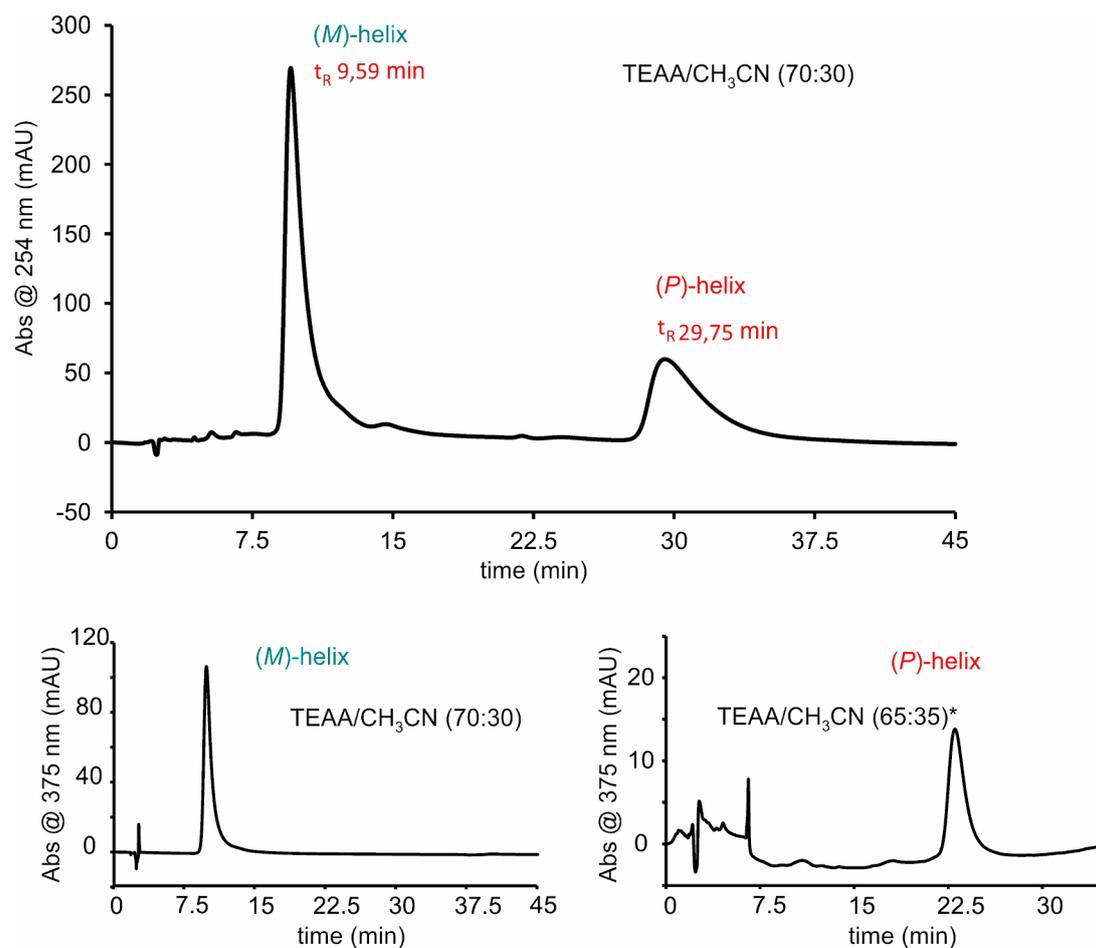


Figure S7. Chiral semi-preparative purification of racemic **1** performed on an Ultimate 3000 Thermo HPLC line at a flow rate of 1 mL min^{-1} with detection at 254 nm at 20°C in the column oven. The HPLC profiles shown below correspond to the purified *M*- and *P*-helices. To sharpen the peak of the *P*-conformer, we increased the ratio of CH₃CN (30% \rightarrow 35%) and consequently, the *P*-helix of **1** was eluted earlier.

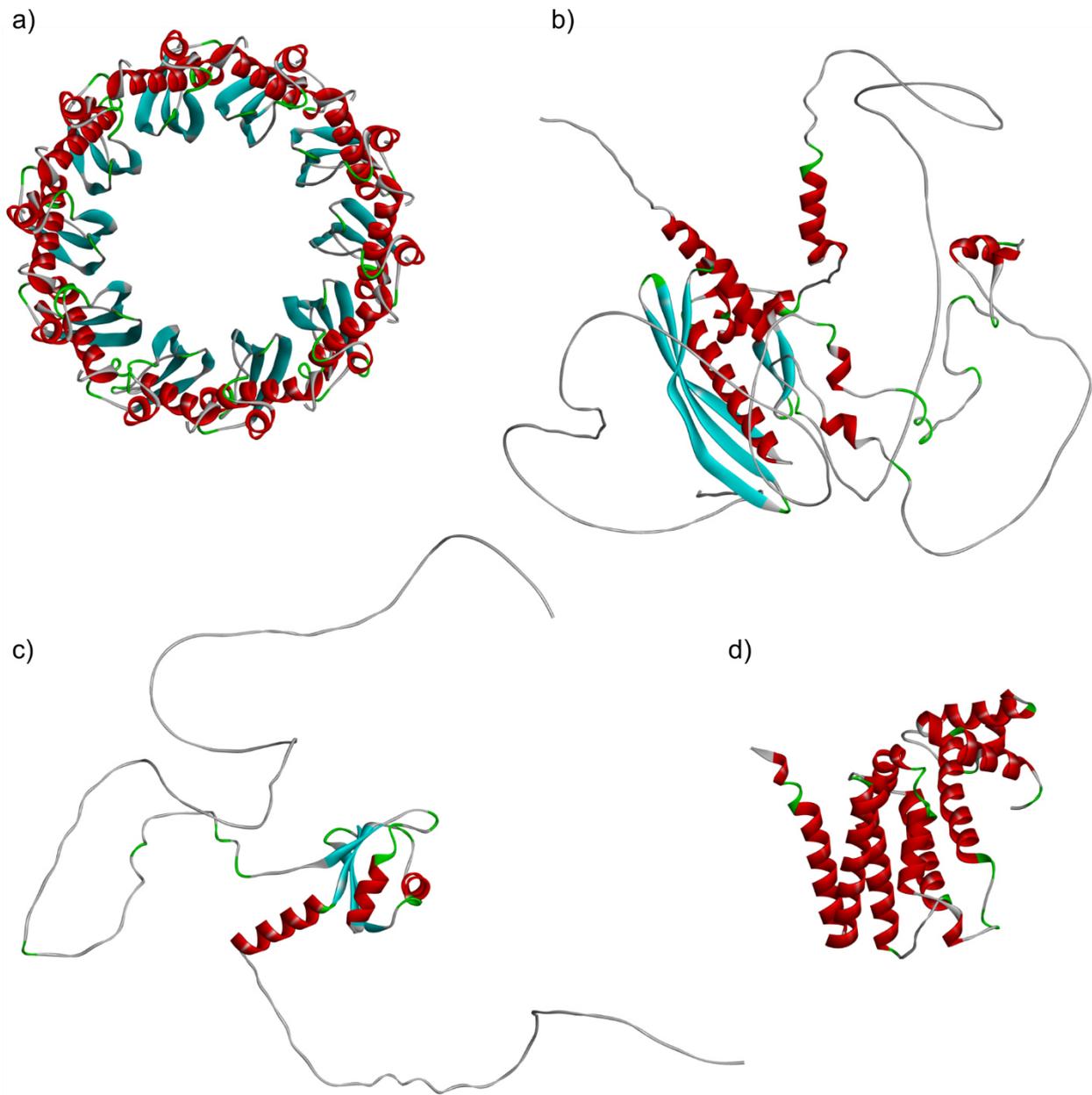


Figure S8. a) Cryo-EM structure of the decamer of yeast RAD52 (PDB #8G3G). b) AlphaFold model prediction of a full length monomer of yeast RAD52. c) AlphaFold model prediction of full length yeast SGN1. d) AlphaFold model prediction of full length yeast COQ9.

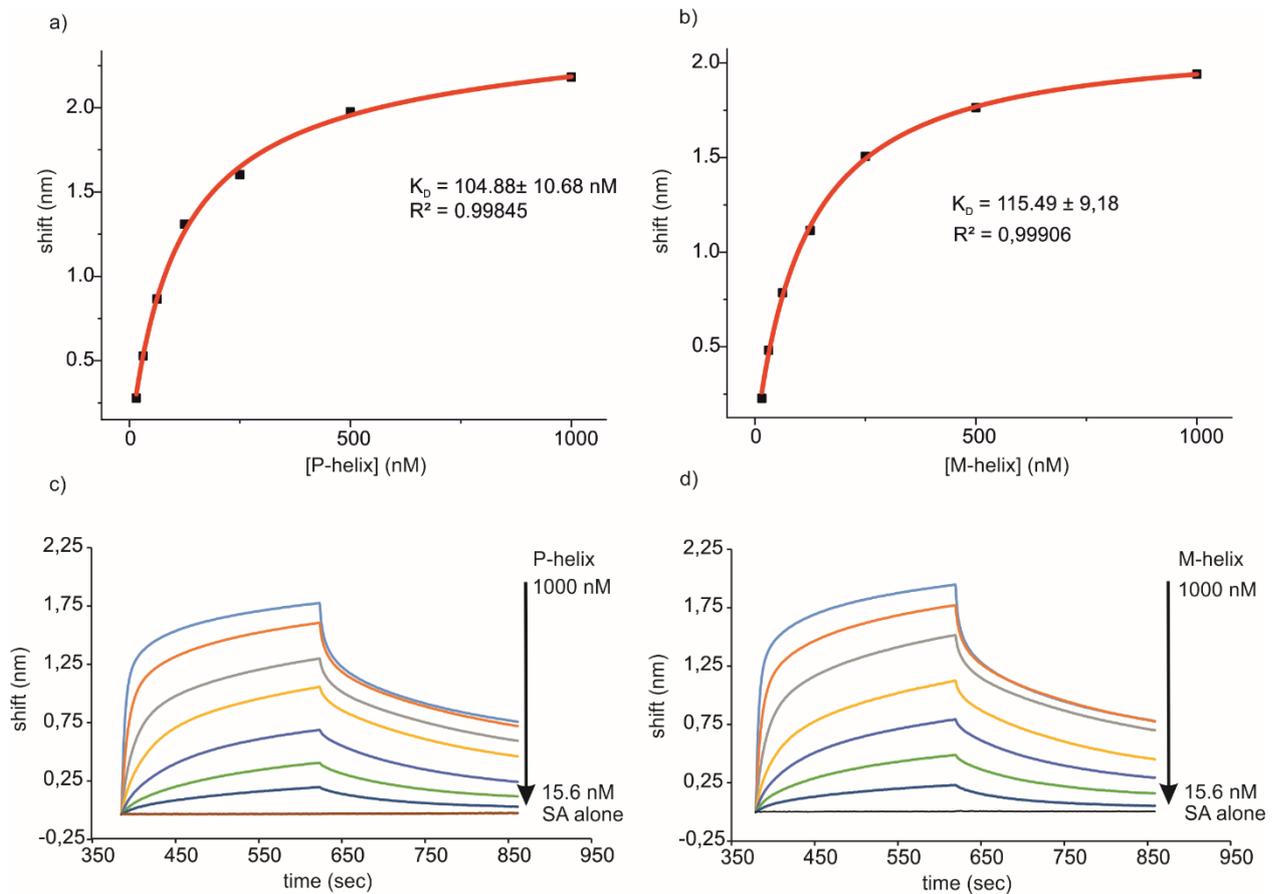


Figure S9. BLI assessment of the binding of SGN1 to *P-1* (a) and *M-1* (b) immobilized on streptavidin sensor tips. The graphs at the top show the curve fitting according to a 1:1 binding isotherm of the experimentally measured equilibrium response plotted against [SGN1]. A series of twofold dilutions of SGN1 was performed starting from 1000 nM. The real-time binding kinetics of both *M-1*/SGN1 and *P-1*/SGN1 interactions were characterized by rapid association (k_a) and dissociation (k_d) (graphs at the bottom). However, the fact that the curves keep climbing after the initial steep climb indicates some additional binding events. The dissociation constants (K_D) were estimated using the steady state model by fitting the signal response (y) as a function of protein concentration (x) to Langmuir's equation:

$$y = (R_{max} * x)/(K_D + X) + N_s * x$$

where N_s represents the slope of the linear component corresponding to second phase binding event and R_{max} , the maximum response.

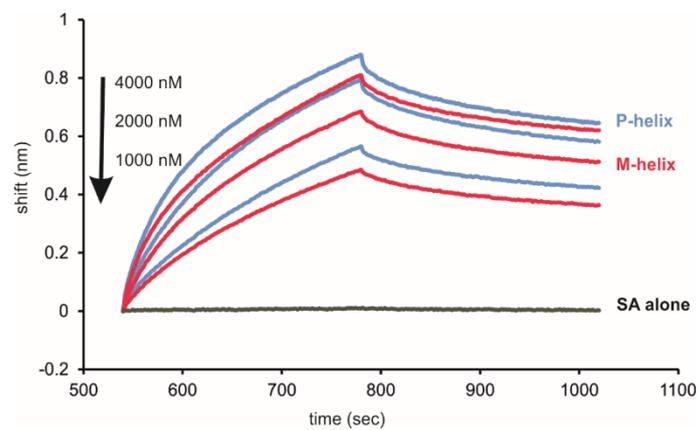


Figure S10. BLI assessment of the binding of COQ9 to *P-1* (blue) and *M-1* (red) immobilized on streptavidin sensor tips. Three different concentrations of COQ9 were assayed (4000 to 2000 nM) for each compound. A reference sensor tip was used to subtract the baseline and another to verify the absence of unspecific binding of COQ9 with the streptavidin (SA alone trace).

6.2 Experimental section

6.2.1 Materials and methods for chemical synthesis and characterizations

General: Chemical reagents were purchased from commercial suppliers (Sigma-Aldrich, Alfa-Aesar, and TCI) and used without further purification. Low loading Wang resin (100-200 mesh, manufacturer's loading: 0.41 mmol g⁻¹) was purchased from Sigma-Aldrich. 15-[D-(+)-Biotinylamino]-4,7,10,13-tetraoxapentadecanoic acid was purchased from Iris Biotech. Tetrahydrofuran (THF) and dichloromethane (DCM) were dried over alumina columns. *N,N*-diisopropylethylamine (DIPEA) was distilled over CaH₂ prior to use. Solid phase syntheses of foldamers were performed manually in open vessel mode using a CEM Discover microwave oven. HPLC grade acetonitrile and Milli-Q water were used for RP-HPLC analyses and purification. RP-HPLC analyses and purification were performed with JASCO HPLC systems (PU-2089 Plus, UV-2077 Plus, HV-2080-01, and AS-2055 Plus for analytical HPLC; DG-2080-53, PU-2086 Plus, and UV-2075 Plus for semi-preparative HPLC). 12.5 mM aqueous NH₄OAc-NH₄OH adjusted to pH 8.5 (solvent A) and pure acetonitrile (solvent B) were used as the mobile phase. RP-HPLC analyses were carried out on a Macherey-Nagel Nucleodur C18 HTec column (4×100 mm, 5 μm) at a flow rate of 1 mL min⁻¹. Semi-preparative RP-HPLC purifications were carried out on a Macherey-Nagel Nucleodur C18 HTec column (10×250 mm, 5 μm) at a flow rate of 3 mL min⁻¹. Eluate from column was monitored by UV detection at 254 nm and 300 nm using a diode array detector. High-resolution electrospray mass spectra were recorded on a Thermo Exactive orbitrap instrument.

Chiral HPLC analyses were performed in reverse mode with JASCO HPLC systems (PU-2080-53 Plus, UV-2075 Plus, HV-2080-01, and AS-2055 Plus). A chiral QN-AX from Daicel was used for chiral separation of foldamer **1** at a flow rate of 1ml/min using a solvent mixture of 70 % of 120 mM TEAA buffer at pH 7.22 (solvent A) and 30% of pure acetonitrile. UV detection was recorded at 300 nm. A CD 1595 detector was mounted after the UV/Vis detector and CD detection was recorded at 400 nm.

Preparation of quinoline amino acid monomers for solid phase synthesis: Fmoc-protected 8-aminoquinoline-2-carboxylic acid monomers were prepared by using the methods previously reported.¹⁻²

Solid phase foldamers synthesis: Conversion of low-loading Wang resin to bromomethyl Wang resin, loading of the first Fmoc-QAsp monomer, acid chloride activation of the monomers using Appel's reaction, automation of the SPFS and TFA cleavage were performed by using recently reported methodology.³ Of note, the use of 2% DBU in NMP for Fmoc deprotection was recently optimized to two times 3 min (previously 2 × 10 min) and these conditions allowed us to skip the resin washing with 20% DIPEA in NMP after the Fmoc deprotection of the Q_{Sul} monomer that was prescribed in reference 1.

6.2.2 Methods for pull-down assay

Cell culture: Yeast cells BY4742 were grown aerobically at 28°C in minimal medium (0.175% yeast nitrogen base without amino acids and ammonium sulfate, 0.5% ammonium sulfate, 0.1% potassium phosphate, 0.2% Drop-Mix, 0.01% of auxotrophic amino acids and nucleotide, pH 5.5), supplemented with 2% glucose as a carbon source. Cell growth was followed by optical density at 600 nm. For preparation of cell lysates, 5×10^7 cells were broken with glass beads in a buffer containing 0.6 M mannitol, 10 mM Tris maleate, 2 mM EGTA, pH 6.8 plus protease inhibitor cocktail (Roche); lysates were centrifuged 10 min at 800 g. Lysates (supernatant) were kept frozen at -80°C as aliquots until further uses.

Immobilization of biotinylated foldamers: 20 µL of resuspended Dynabeads™ MyOne™ Streptavidin T1 (Invitrogen) was transferred to a microcentrifuge tube and washed with 200 µL of PBS buffer (pH 7.4) four times. The beads were then incubated with 40 µL of 20 µM biotinylated foldamer for 60 min while shaking. The supernatant was subsequently discarded, and the beads were washed with 40 µL of PBS.

Enrichment of protein binders for foldamers: 15.39 µL of the yeast cell lysate (7.8 g/L) was diluted with 34.61 µL of PBS and then added to the washed beads. The resulting mixture was incubated for 20 min while shaking. The beads were then washed with 200 µL of PBS five times. To maximize enrichment, incubation of cell lysate and washing were repeated five more times.

Protein elution and digestion for mass spectrometry: Captured proteins were eluted off from the beads by incubation in 20 µL of 1× Laemmli sample buffer for 3 min at 100°C. The eluate was then loaded onto a 10% SDS-PAGE gel, and SDS-PAGE was run for 5 min at 150 V. The resulting gel was stained with Coomassie blue for 1 h and subsequently destained with water. For reduction and alkylation, lanes for each replicate were first cut into small cubes (1×1×1 mm), and then incubated with destaining solution (25 mM ammonium bicarbonate in 50% CH₃CN) until bands were no longer visible. Destained gel pieces were collected and were subsequently incubated with 30 µL of 10 mM DTT solution for 30 min at 56°C. The DTT solution was discarded, and the gels were then incubated with 30 µL of 100 mM iodoacetamide in 100 mM ammonium bicarbonate for 30 min in the dark. The iodoacetamide solution was removed, and the gels were subsequently dehydrated with CH₃CN. For in-gel digestion, the dehydrated gel pieces were submerged in a trypsin solution (5 µg of Trypsin (Promega) in 50 µL of 1 mM HCl and 450 µL of 50 mM ammonium bicarbonate) and incubated overnight at 37°C. After the overnight digestion, 500 µL of 50 mM ammonium bicarbonate was added to the digestion mixture. The resulting supernatant was collected after 10 min incubation. 500 µL of an extraction solution (formic acid/CH₃CN/water, 5.0/47.5/47.5, v/v/v) was then added to gels, and the resulting supernatant was collected after 10 min incubation and combined with previously obtained supernatant. This process was repeated once again with 250 µL of the extraction solution. Extracted peptides were dried under reduced

pressure and redissolved in 30 μ L of 6% formic acid. Solution of extracted peptides was divided into three portions and analyzed separately.

nLC-MS/MS analysis and Label-Free Quantitative Data Analysis: Peptide mixture was analyzed on an Ultimate 3000 nanoLC system (Dionex, Amsterdam, The Netherlands) coupled to an Electrospray Orbitrap Fusion™ Lumos™ Tribrid™ Mass Spectrometer (Thermo Fisher Scientific, San Jose, CA). 10 μ L of peptide digests were loaded onto the system. Peptides were separated on an analytical 75-mm id x 50-cm C18 Pep-Map column (LC Packings) with a 5–27.5% linear gradient of solvent B in 105 min (solvent A was 0.1% formic acid in water and solvent B was 0.1% formic acid in CH₃CN:water (8:2, v/v)) followed by a 10 min gradient from 27.5% to 40% solvent B. The mass spectrometer operated in positive ion mode and data were acquired using Xcalibur 4.1 software in a data-dependent mode. MS scans (m/z 375-1500) were recorded at a resolution of R = 120 000 (at m/z 200) with a dynamic exclusion set to 60 s. Fragmentation was limited to +2 to +7 charged ions and performed in HCD mode.

Database search and results processing: Data were searched by SEQUEST through Proteome Discoverer 1.4 (Thermo Fisher Scientific Inc.) against the *Saccharomyces cerevisiae* Reference Proteome Set (from Uniprot 2017-10; 5991 entries). Spectra from peptides higher than 5000 Da or lower than 350 Da were rejected. The search parameters were as follows: mass accuracy of the monoisotopic peptide precursor and peptide fragments was set to 10 ppm and 0.6 Da respectively. Only b- and y-ions were considered for mass calculation. Oxidation of methionines (+16 Da) was considered as variable modification and carbamidomethylation of cysteines (+57 Da) as fixed modification. Two missed trypsin cleavages were allowed. Peptide validation was performed using Percolator algorithm⁴ and only “high confidence” peptides were retained corresponding to a 1% False Positive Rate at peptide level.

Label-Free Quantitative Data Analysis: Label-free quantitation was performed thanks to Progenesis QI for Proteomics 2.0 (Nonlinear Dynamics Ltd, Newcastle, U.K). Calculation of protein abundance was the sum of the volume of corresponding peptides. A statistical test (ANOVA) was calculated for each group comparison and proteins were filtered based on p-value < 0.05. Noticeably, only non-conflicting features and unique peptides were considered for calculation at protein level. Quantitative data were considered for proteins quantified by a minimum of 2 peptides. Relative quantitation was achieved by calculating the ratio of captured proteins (capture performed with foldamer bait) over control samples (same capture system but without foldamer bait).

Da Dalton

MS Mass Spectrometry

ppm Part per million

HCD Higher-energy Collisional Dissociation

SDS-PAGE Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis

6.2.3 Recombinant protein expression and purification

Yeast DNA repair protein RAD52 homolog: The pET21a-yRad52 plasmid for yRad52 (UniProt accession number: P06778, aa 1 - 471) with a C-terminal hexa-histidine tag overexpression was obtained from Tomohiko Sugiyama (Ohio University). The full-length protein was expressed in E. coli BL21 Rosetta2 cells. Overnight pre-culture in Luria broth (LB) supplemented with 100 µg/ml ampicillin was diluted 1000-fold with fresh 4L LB media and grown at 37°C until OD600 reached 1. The expression was induced by addition of isopropyl 1-thio-β-D-galactopyranoside to make the final concentration 0.75 mM, and the culture was incubated for 4 hours at 27°C. Next, the cells were harvested at 6000 rpm (J-LITE® JLA-9.1000 Rotor, Beckman Coulter) and stored at -80°C. The following purification steps were carried out at 4°C. Briefly, the cells were resuspended in 60 mL of 50 mM PBS buffer pH 7.8 containing 2 mM 2-mercaptoethanol, 10% glycerol, 10 mM imidazole, and 1 mM PMSF. The cell were lysed by sonication and the lysate was clarified by centrifugation at 19000 rpm (JA-25.5 Rotor, Beckman Coulter) for 40 min. The supernatant was equilibrated with nickel-nitrilotriacetic acid (Ni-NTA) agarose beads by gentle mixing for 1 hour. The mixture was then applied to a gravity flow column, drained out, and washed with 50 mM Tris-HCl, pH 7.8 containing 100 mM imidazole. The protein was eluted with 50 mM Tris-HCl, pH 7.8, 250 mM imidazole, and immediately diluted with 50 mM Tris-HCl buffer to reduce the imidazole concentration to 100 mM. Finally, the protein was concentrated using a Vivaspin Turbo 4 column (100K MWCO), filtered through a 0.2-µm filter, and loaded onto the gel filtration HiLoad 16/600 Superdex 200 column. Peak fractions were pooled together, concentrated, and frozen in liquid nitrogen prior to BLI measurements.

Yeast ubiquinone biosynthesis protein COQ9: The recombinant yeast COQ9 protein (UniProt accession number: [insert accession number], aa 36-256) with a C-terminal His6 tag was molecularly cloned into the expression vector pET24a and subsequently expressed in E. coli BL21 RIP Codon Plus cells. Briefly, cells previously cultured overnight in LB medium were diluted 1000-fold with fresh LB medium and cultured at 37°C until an OD600 of 0.6-0.8 was reached. Induction of protein expression was initiated by the addition of 0.3 mM IPTG for 16 hours at 22°C. After induction, cells were harvested by centrifugation at 8000g and stored at -20°C until purification. Purification of the target protein was similar to that described for yRad52, except that TBS supplemented with 20mM imidazole was used for the washing step during IMAC purification.

Yeast RNA-binding protein SGN1: The gene encoding the SGN1 protein (UniProt accession number: P40561, amino acids 1 - 250) was cloned into the pMAL-c5e vector downstream of a maltose-binding protein (MBP), a His₁₀ tag, and an HRV 3C cleavage site. Protein expression proceeded as follows: BL21 cells were transformed with the expression vector. Overnight cultures of BL21 cells were then inoculated into fresh LB media (4 L) supplemented with 100 µg/ml ampicillin. Expression induction occurred upon reaching an optical density at OD600 of 0.6, with the addition of 0.3 mM IPTG. The protein was expressed at 22°C for 16 hours. Following expression, cells were harvested by

centrifugation at 8000g and resuspended in a buffer containing 50 mM TRIS pH 7.4, 500 mM NaCl, 20 mM imidazole, 1 mM PMSE, and 1x Halt™ protease inhibitor cocktail. Cells were lysed by sonication, and subsequent centrifugation was conducted to eliminate cellular debris. Protein was purified by IMAC column. The target SGN1 protein was cleaved from the MBP-His10 tag-SGN1 fusion protein by digestion with HRV 3C protease at 4°C overnight. Subsequently, a reverse IMAC column was employed to separate the cleaved MBP-His tag from the SGN1 protein. Finally, fractions containing the SGN1 protein from the reverse IMAC step were pooled and subjected to gel filtration chromatography using a HiLoad 16/600 Superdex 200 column.

6.2.4 Binding characterization by biolayer interferometry (BLI)

BLI experiments were performed on an Octet R8 instrument from Sartorius, following Sartorius recommendations. Prior to an assay, streptavidin (SA) sensors were soaked for at least 10 min in phosphate buffer saline (1 × PBS). The kinetic experiment always starts with a baseline step over 60 sec in 20 mM HEPES (pH 7.4, 150 mM NaCl, 0.02% Tween-20) buffer, followed by the loading of *P*- or *M*-foldamer isolated by chiral RP-HPLC isolated at 2 µg/mL over 120 sec in HEPES. After foldamer ligand immobilization, the sensors were washed for 60 sec in the same buffer, before to record a second baseline for 120 sec, again in HEPES. Serial column dilutions (× 2) of the different proteins in HEPES were analysed, keeping the last well of the association column free of protein for referencing. Association lasted 240 sec, followed by dissociation for another 240 sec. The curves were fitted to binding models using the Octet analysis studio 13.0 software and replotted in Excel. Of note, the absence of unspecific binding of the proteins to streptavidin was confirmed by running a kinetic assay in a single well on a SA sensor with no immobilized foldamer at the highest screened protein concentration. For SGN1, the K_D value was calculated with the Langmuir's equation assuming a 1:1 binding model (see caption of Figure S8)

6.3 X-ray crystallographic analysis of compound 2

Lyophilized powder of **2** was dissolved using water and ammonium bicarbonate to a final concentration of 2mM. Crystallization trials were made using standard sitting drop vapor diffusion method at 293 K. X-ray quality crystals were obtained after three weeks by the addition of 0.75 µl of **2** and 1.25 µl of crystallization reagent composed of 50% w/v (+/-)-2-methyl-2,4-pentanediol, 50 mM HEPES buffer at pH 7.0, 80 mM potassium chloride, 10 mM magnesium sulphate in the reservoir. For low temperature diffraction measurement, a crystal was fished using a micro loop and plunged into liquid nitrogen. The mother liquor served as cryo-protectant for the crystal.

The X-ray diffraction data was collected at the micro-focus, fixed energy beamline ID30b⁵ in European Synchrotron Radiation Facility (ESRF), Grenoble with a Dectris PILATUS3 X 2M detector. Diffraction data was measured at $T = 100$ K, $\lambda = 0.8000$ Å. The crystals were exposed for 0.02 s and 0.1° oscillation

per frame. Diffraction data was processed using the program *XDS*⁶. The crystal belonged to the space group *P2₁/n* with unit cell parameters: $a = 35.903 (7) \text{ \AA}$, $b = 65.428 (13) \text{ \AA}$, $c = 36.158 (7) \text{ \AA}$, $\alpha = 90^\circ$, $\beta = 93.93 (3)^\circ$, $\gamma = 90^\circ$; $V = 84738 (30) \text{ \AA}^3$ and 4 molecules per asymmetric unit ($Z = 16$, $Z' = 4$). The structure was solved with the program *SHELXT*⁷ and refined by full-matrix least-squares method on F^2 with *SHELXL-2014*⁸ within *Olex2*⁹. After each refinement step, visual inspection of the model and the electron-density maps were carried out using *Olex2*⁹ and *Coot*¹⁰ using $2F_o - F_c$ and difference Fourier ($F_o - F_c$) maps. The initial structure revealed all main-chain atoms of **2**. Few of the side chain atoms were refined with full or partial occupancy. AFIX, DFIX, SADI and FLAT instructions were used to improve the geometry of molecules. Restraints on anisotropic displacement parameters were implemented with RIGU and EADP instructions. All non-H atoms of the backbones were refined with anisotropic displacement parameters. From the difference Fourier map a molecule of (+/-)-2-Methyl-2,4-pentanediol (MPD) was identified (from crystallization reagent). After several attempts to model the disordered side chains, the SQUEEZE¹¹ procedure was used to flatten the electron density map. Very disordered side chains and solvent molecules were removed. Calculated total potential solvent accessible void volume and electron count per cell are 34,860.6 \AA^3 and 11,283 respectively. Hydrogen atoms were placed at idealized positions.

Statistics of data collection and refinement of **2** are described in Table S1. The final cif file of **2** was examined in IUCr's *checkCIF* algorithm. Due to the large volume fractions of disordered solvent molecules, weak diffraction intensity and poor resolution, a number of A- and B- level alerts remain in the *checkCIF* file. These alerts are inherent to the data set and refinement procedures. They are listed below and were divided into two groups. The first group demonstrates weak quality of the data and refinement statistics when compared to those expected for small molecule structures from highly diffracting crystals. The second group is concerned to decisions made during refinement and explained below. Atomic coordinates and structure factors of **2** was deposited in the Cambridge Crystallographic Data Centre (CCDC) with accession code 2280177. The data is available free of charge upon request (www.ccdc.cam.ac.uk/).

CheckCIF validation of **2**:

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

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

Calculated sin(theta_max)/wavelength = 0.5208

PLAT082_ALERT_2_A High R1 Value

0.22 Report

PLAT084_ALERT_3_A High wR2 Value (i.e. > 0.25) 0.56 Report

PLAT097_ALERT_2_B Large Reported Max. (Positive) Residual Density 1.63 eA-3

PLAT201_ALERT_2_A Isotropic non-H Atoms in Main Residue(s) 16 Report

PLAT241_ALERT_2_B High 'MainMol' Ueq as Compared to Neighbors of

PLAT242_ALERT_2_B Low 'MainMol' Ueq as Compared to Neighbors of

PLAT315_ALERT_2_B Singly Bonded Carbon Detected (H-atoms Missing)

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

PLAT410_ALERT_2_A Short Intra H...H Contact

PLAT412_ALERT_2_B Short Intra XH3 .. XHn

PLAT430_ALERT_2_A Short Inter D...A Contact

PLAT733_ALERT_1_A Torsion Calc

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

SHFSU01_ALERT_2_A The absolute value of parameter shift to su ratio > 0.20

Additional cycles of refinement did not remove this alert.

PLAT016_ALERT_5_A No _shelx_fab_file Please Supply

Due to large file size, the file was separately supplied.

PLAT080_ALERT_2_A Maximum Shift/Error 5.09 Why?

Additional cycles of refinement did not remove this alert.

PLAT202_ALERT_3_A Isotropic non-H Atoms in Anion/Solvent 84 Check

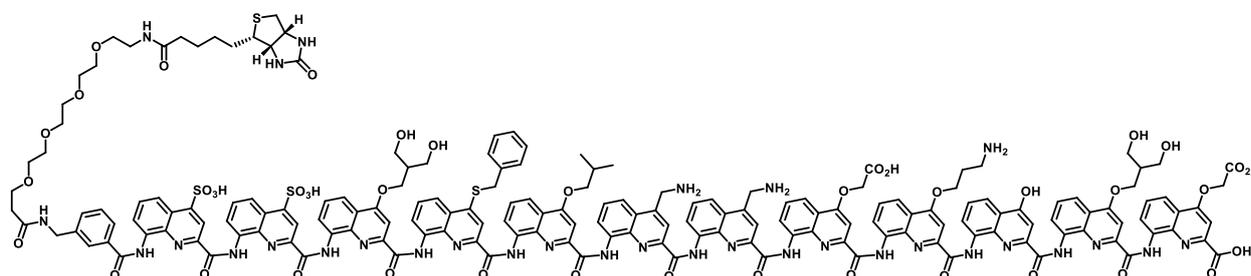
Dummy O atom was introduced into refinement.

Table S1. Crystallographic data and refinement details for **2**.

Compound	2
Empirical formula	C _{150.5} H _{111.1} N _{26.4} O _{33.8} S ₃
Formula weight	2925.2
Temperature	100 K
Wavelength	0.8000 Å
Crystal system	Monoclinic
Space group	<i>P2₁/n</i>
Unit cell dimensions	<i>a</i> = 35.903 (7) Å <i>b</i> = 65.428 (13) Å <i>c</i> = 36.158 (7) Å <i>α</i> = 90° <i>β</i> = 93.93 (3)° <i>γ</i> = 90°
Volume	84738 (30) Å ³
<i>Z</i> , <i>Z'</i>	16, 4
Density (calculated)	0.917 g/cm ³
Absorption coefficient	0.128 μ/mm ⁻¹
Colour and shape	Yellow, blocks
Crystal size	0.120 x 0.100 x 0.005 mm
Index ranges	-36 ≤ <i>h</i> ≤ 37 -67 ≤ <i>k</i> ≤ 66 -37 ≤ <i>l</i> ≤ 37
Reflections collected	311527
<i>R</i> _{int}	0.1195
Data/restraints/parameters	97123/621/5266
Goodness-of-fit on <i>F</i> ²	2.426
Final <i>R</i> indexes [<i>I</i> > 2σ (<i>I</i>)]	<i>R</i> ₁ = 0.2180 <i>wR</i> ₂ = 0.5360
Final <i>R</i> indexes [all data]	<i>R</i> ₁ = 0.2462 <i>wR</i> ₂ = 0.5628
Largest diff. peak and hole	1.64/-1.33 e Å ⁻³
Total potential solvent accessible void volume from SQUEEZE	34860.6 Å ³
Electron count/cell	11283
CCDC #	2280177

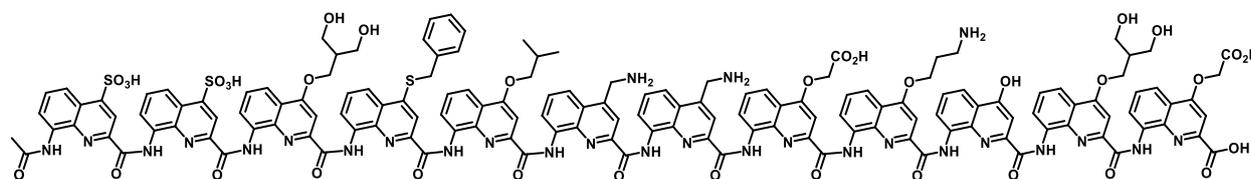
6.4 Supplementary Data

Foldamer 1. Biotin-3-Amb-Q^{Sul}-Q^{Sul}-Q^{Diol}-Q^{Phe}-Q^{Leu}-Q^{Dap}-Q^{Dap}-Q^{Asp}-Q^{Orn}-Q^{Hyd}-Q^{Diol}-Q^{Asp}-OH



Foldamer **1** was synthesized on a low loading Wang resin (7.19 μmol). The crude product obtained from cleavage was purified by RP-HPLC (22-28% solvent B, over 19 min) to afford the title compound as a yellow solid (6.5 mg, 25.7%, purity by RP-HPLC: >99%). HRMS (ESI⁻): m/z calcd for $\text{C}_{177}\text{H}_{161}\text{N}_{31}\text{O}_{42}\text{S}_4$ [M-2H]²⁻ 1760.5171 found 1760.5206; m/z calcd for $\text{C}_{177}\text{H}_{160}\text{N}_{31}\text{O}_{42}\text{S}_4$ [M-3H]³⁻ 1173.3423 found 1173.3473.

Foldamer 2. Ac-Q^{Sul}-Q^{Sul}-Q^{Diol}-Q^{Phe}-Q^{Leu}-Q^{Dap}-Q^{Dap}-Q^{Asp}-Q^{Orn}-Q^{Hyd}-Q^{Diol}-Q^{Asp}-OH



Foldamer **2** was synthesized on a low loading Wang resin (0.78 μmol). The crude product obtained from cleavage was purified by RP-HPLC (20-27% solvent B, over 19 min) to afford the title compound as a yellow solid (0.7 mg, 30.3%, purity by RP-HPLC: 98.34%). HRMS (ESI⁻): m/z calcd for $\text{C}_{150}\text{H}_{121}\text{N}_{27}\text{O}_{35}\text{S}_3$ [M-2H]²⁻ 1478.3863 found 1478.3890; m/z calcd for $\text{C}_{150}\text{H}_{120}\text{N}_{27}\text{O}_{35}\text{S}_3$ [M-3H]³⁻ 985.2551 found 985.2584.

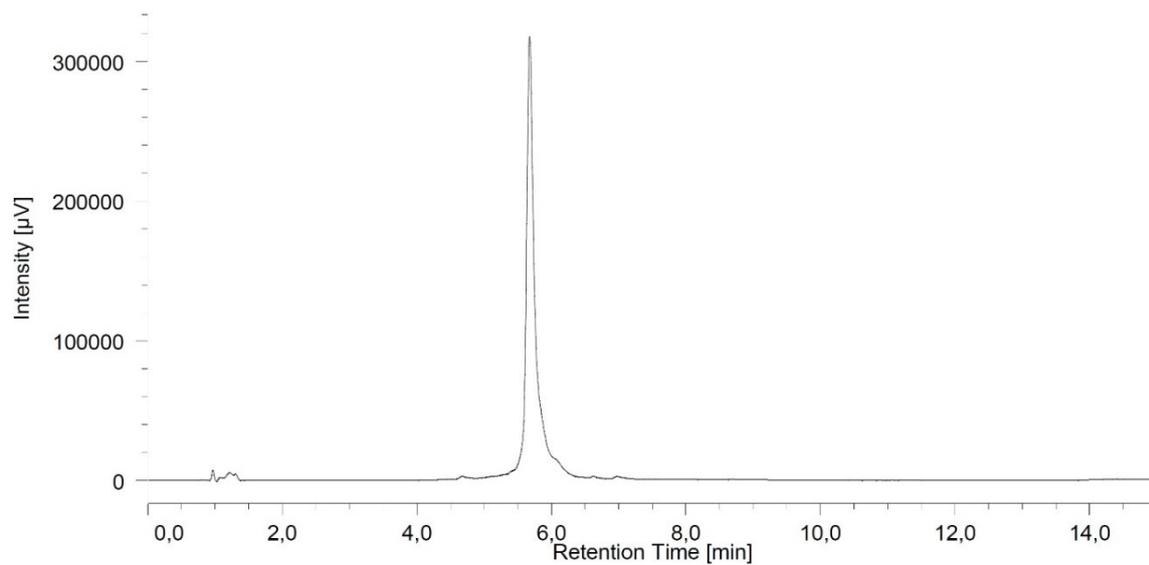


Figure S11. Analytical RP-HPLC (10-60% solvent B, over 12 min) trace of foldamer 1.

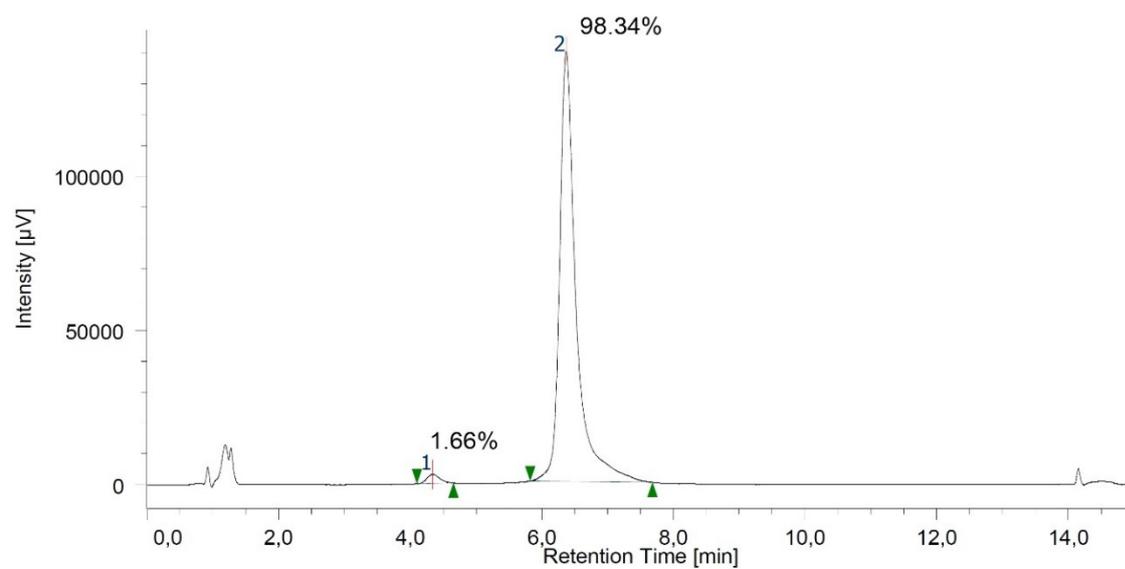
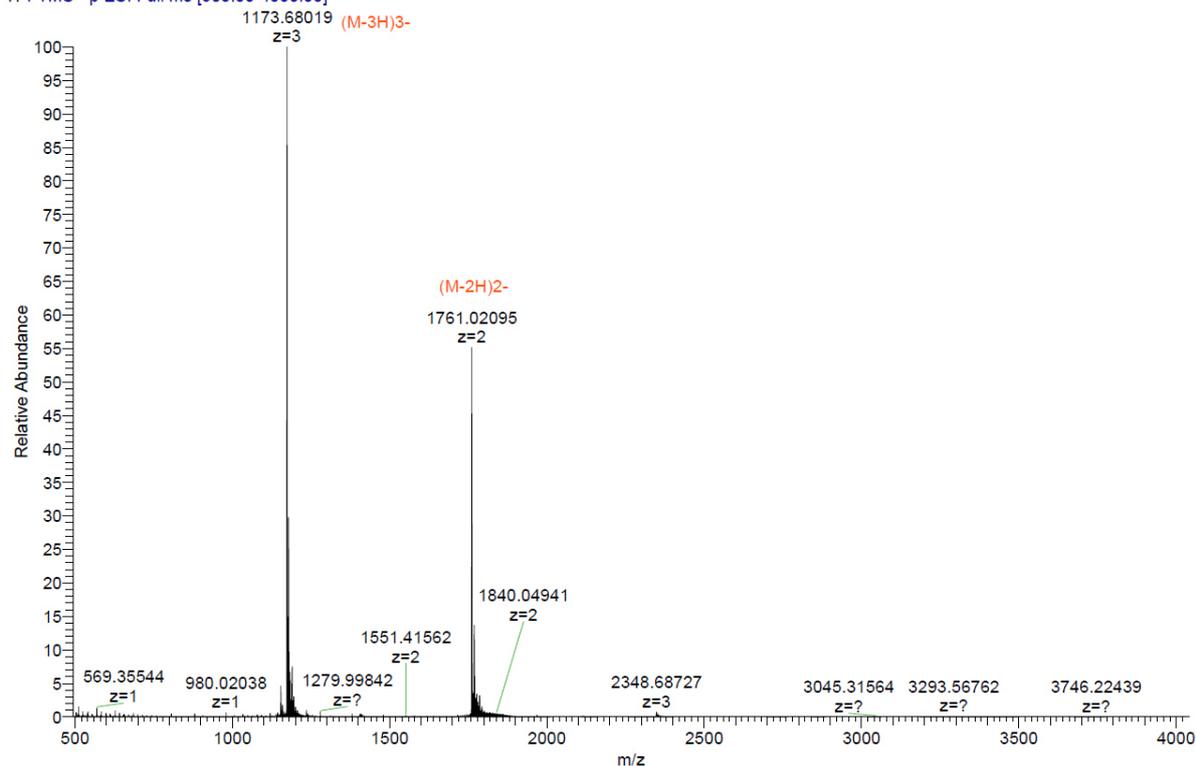
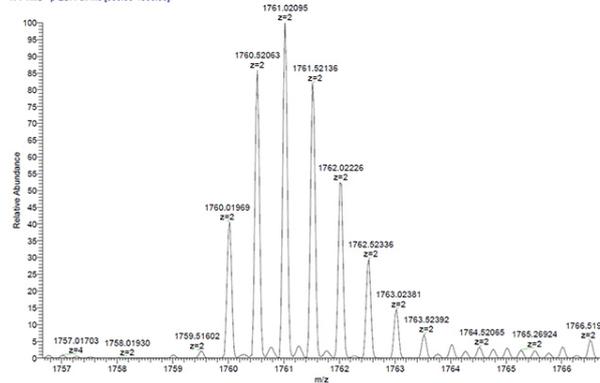


Figure S12. Analytical RP-HPLC (20-30% solvent B, over 12 min) trace of foldamer 2.

180130-Exac-5578-YF-SK-SK107 #79-125 RT: 1.91-2.89 AV: 44 NL: 3.08E5
T: FTMS - p ESI Full ms [500.00-4000.00]



180130-Exac-5578-YF-SK-SK107 #79-125 RT: 1.91-2.89 AV: 44 NL: 1.70E5
T: FTMS - p ESI Full ms [500.00-4000.00]



180130-Exac-5578-YF-SK-SK107 #79-125 RT: 1.91-2.89 AV: 44 NL: 3.08E5
T: FTMS - p ESI Full ms [500.00-4000.00]

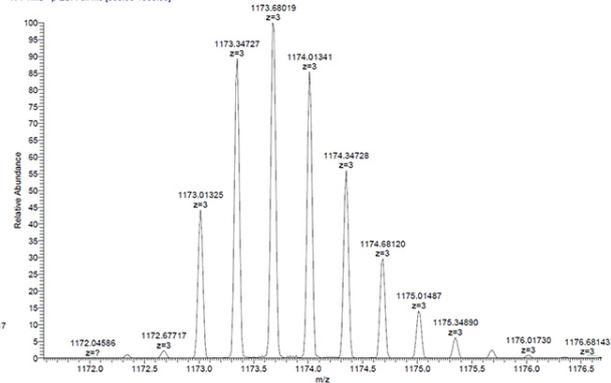
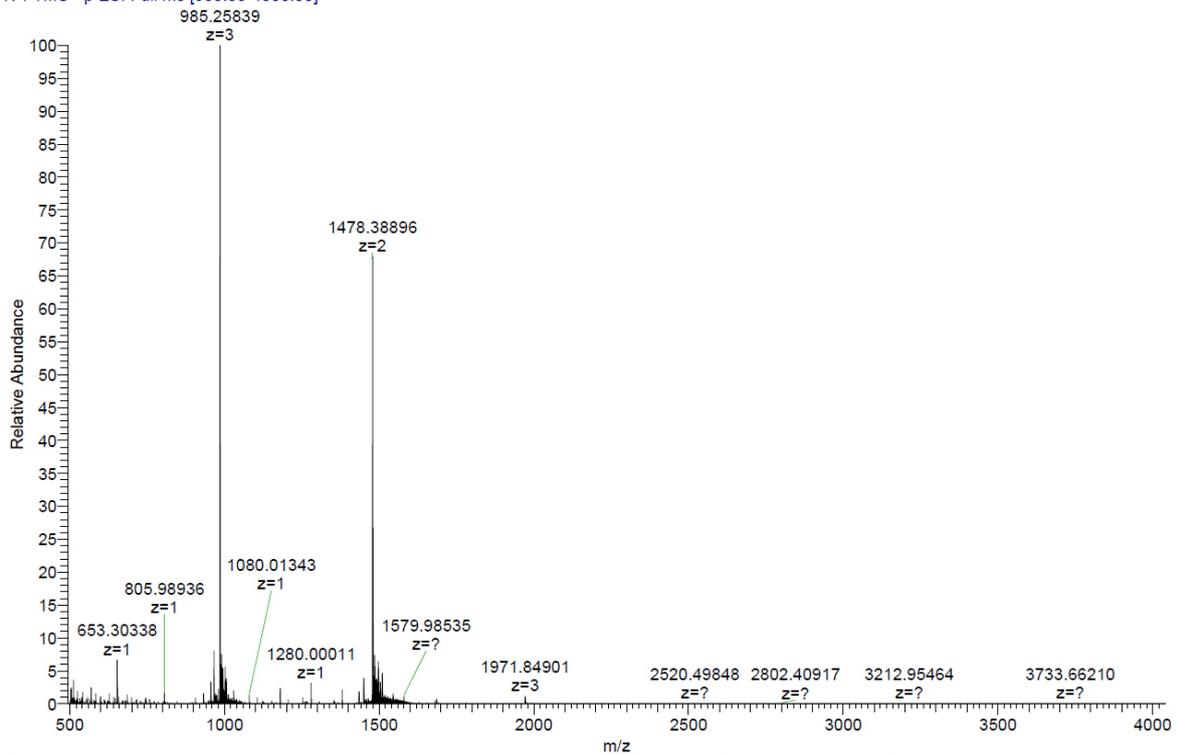
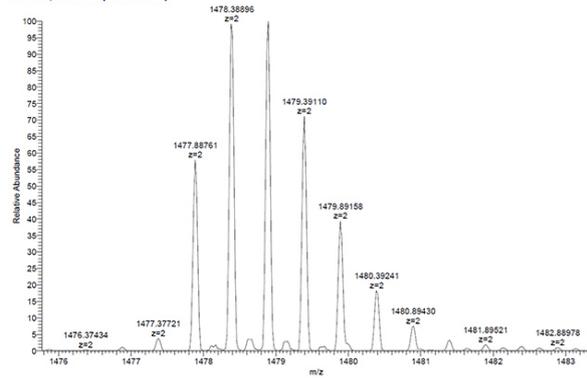


Figure S13. HR-MS spectrum of foldamer 1.

180130-Exac-5577-YF-SK-SK102 #45-80 RT: 1.13-1.86 AV: 33 NL: 7.61E4
T: FTMS - p ESI Full ms [500.00-4000.00]



180130-Exac-5577-YF-SK-SK102 #45-80 RT: 1.13-1.86 AV: 33 NL: 5.17E4
T: FTMS - p ESI Full ms [500.00-4000.00]



180130-Exac-5577-YF-SK-SK102 #45-80 RT: 1.13-1.86 AV: 33 NL: 7.61E4
T: FTMS - p ESI Full ms [500.00-4000.00]

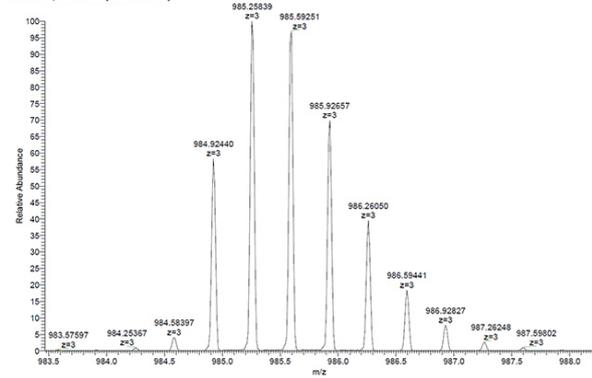


Figure S14. HR-MS spectrum of foldamer 2.

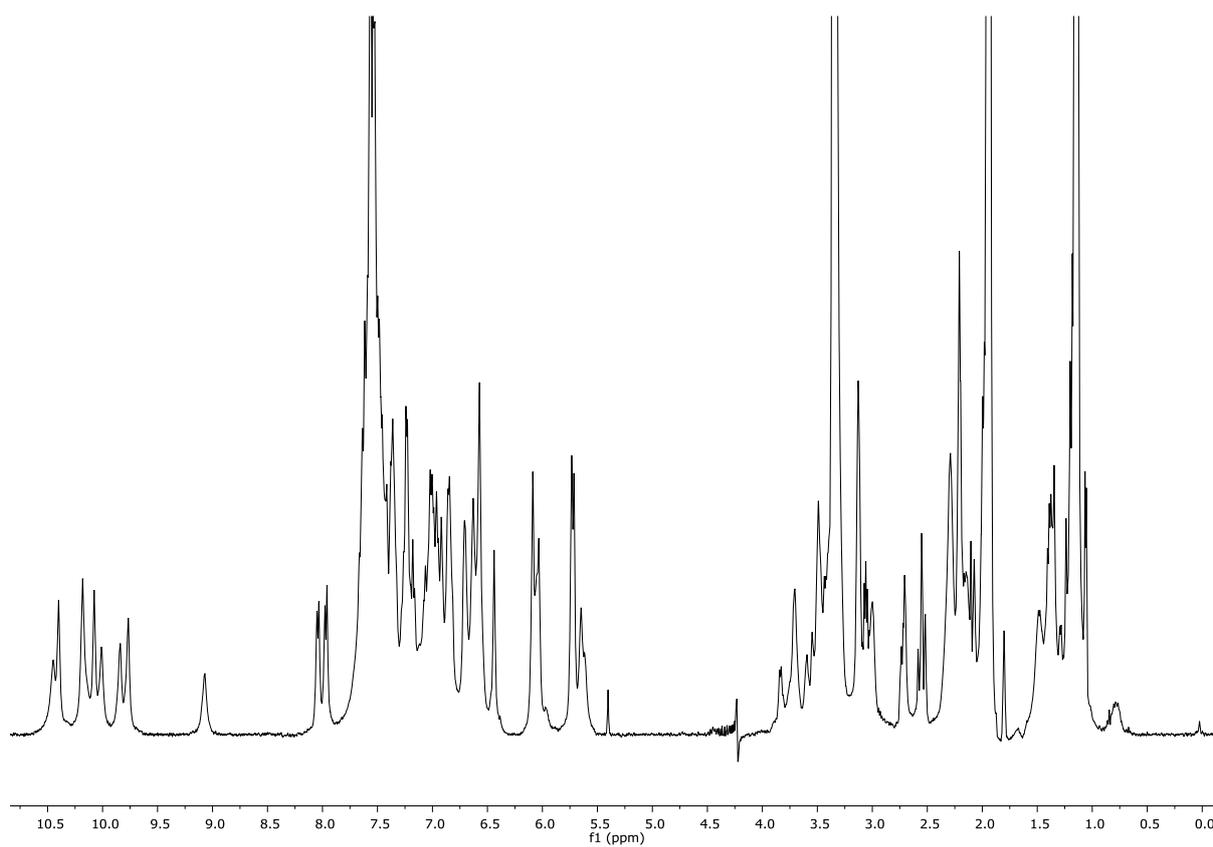


Figure S15 ¹H NMR spectrum of foldamer 1 (500 MHz, CD₃CN + 50% H₂O, water suppression), 25°C.

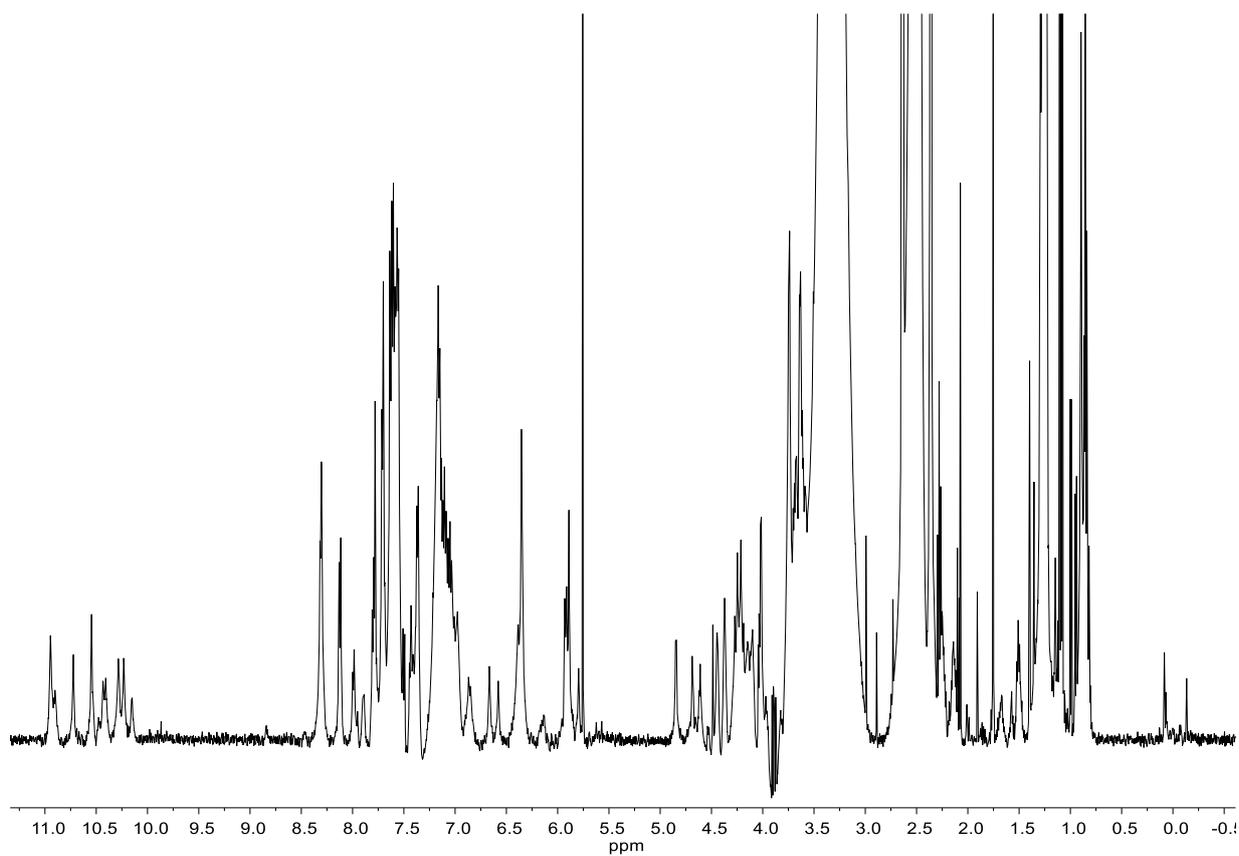


Figure S16 ¹H NMR spectrum of foldamer 2 (500 MHz, DMSO-*d*₆), 25°C.

6.5 Reference

1. Hu, S. J. Dawson, P. K. Mandal, X. de Hatten, B. Baptiste and I. Huc, *Chem. Sci.*, 2017, **8**, 3741-3749.
2. S. De, B. Chi, T. Granier, T. Qi, V. Maurizot and I. Huc, *Nat. Chem.*, 2018, **10**, 51-57.
3. a) V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat and I. Huc, *Chem. Eur. J.*, 2023, e202300898 ; b) B. Baptiste, C. Douat-Casassus, K. Laxmi-Reddy, F. Godde and I. Huc, *J. Org. Chem.*, 2010, **75**, 7175-7185.
4. L. Käll, J. D. Canterbury, J. Weston, W. S. Noble and M. J. MacCoss, *Nat. Methods*, 2007, **4**, 923-925.
5. A.A. McCarthy, R. Barrett, A. Beteva, H. Caserotto, F. Dobias, F. Felisaz, T. Giraud, M. Guijarro, R. Janocha, A. Khadrouche, M. Lentini, G.A. Leonard, M. Lopez Marrero, S. Malbet-Monaco, S. McSweeney, D. Nurizzo, G. Papp, C. Rossi, J. Sinoir, C. Sorez, J. Surr, O. Svensson, U. Zander, F. Cipriani, P. Theveneau and C. Mueller-Dieckmann, *J. Synchrotron Radiat.*, 2018, **25**, 1249-1260.
6. W. Kabsch, *Acta Crystallogr. Sect. D-Biol. Crystallogr.*, 2010, **66**, 125-132.
7. G.M. Sheldrick, *Acta Crystallogr. Sect. A*, 2015, **71**, 3-8.
8. G.M. Sheldrick, *Acta Crystallogr. Sect. C-Struct. Chem.*, 2015, **71**, 3-8.
9. O.V. Dolomanov, L.J. Bourhis, R.J. Gildea, J.A.K. Howard and H. Puschmann, *J. Appl. Crystallogr.*, 2009, **42**, 339-341.
10. P. Emsley, B. Lohkamp, W.G. Scott and K. Cowtan, *Acta Crystallogr. Sect. D-Biol. Crystallogr.*, 2010, **66**, 486-501.
11. A.L. Spek, *Acta Crystallogr. Sect. D-Struct. Biol.*, 2009, **65**, 148-155.

7. Main text: Display selection of peptide ligands for helical aromatic foldamers

Authors: Lingfei Wang, Joseph M. Rogers, Simon Dawson, Line Langhorn, Ryan T. Howard, Sunbum Kwon, Céline Douat*, Hiroaki Suga* and Ivan Huc*

Author contribution: The project was planned by I. Huc in collaboration with H. Suga. L. Langhorn improved and synthesized the Q^{Tyr} monomer. S. Kwon performed the degradation test of para-hydroxybenzyl side-chain. S. Dawson performed one sequence synthesis for first round of selection. J. M. Rogers performed the *in vitro* translation and RaPID selection. I performed the most sequences and peptide synthesis. C. Douat co-supervised the project and assisted with chemical synthesis. The manuscript was written by me, C. Douat and I. Huc.



Cite this: DOI: 10.1039/d5ob00228a

Received 9th February 2025,
 Accepted 3rd April 2025

DOI: 10.1039/d5ob00228a

rsc.li/obc

Display selection of peptide ligands for helical aromatic foldamers†

Lingfei Wang,^a Joseph M. Rogers,[‡] Simon J. Dawson,^c Line M. Langhorn,^a Ryan T. Howard,^a Sunbum Kwon,[§] Céline Douat,^{‡*} Hiroaki Suga^{‡*} and Ivan Huc^{‡*}

Helical aromatic oligoamide foldamers with a cluster of five biogenic side chains at their surface were designed and synthesized. Display selection of thioether macrocyclic peptides against these targets generated low micromolar binders that are highly selective for the side-chain arrangement.

Aromatic oligoamides – oligoamides having aryl groups in their main chain – are privileged scaffolds that interact with biopolymers.¹ This large class of compounds includes natural products,² old drugs like suramin that can bind to a range of proteins,³ and rod-like⁴ and helically folded^{5,6} synthetic oligomers equipped with biogenic side chains. Our own efforts have been focused on the latter grouping that includes sizeable (1–15 kDa) and structurally well-defined molecules potentially able to bind to proteins through large contact surface areas.^{5c} Methods to identify helical aromatic oligoamides that interact with a given protein target have included screening of small libraries^{5c–e} and mimicry of α -helix side chain presentation^{5b} or of B-DNA shape and negative charge distribution.⁶ Aromatic amino acids[¶] may also be combined with α -amino acids in hybrid sequences,⁷ some of which are compatible with *in vitro* ribosomal peptide expression⁸ and amenable to display selection against protein targets.⁹ To further explore the scope of aromatic oligoamide recognition of proteins or peptides, it is relevant not only to screen or design synthetic aromatic oligoa-

mides with affinity for a given target, but also to screen potential binders to a given aromatic oligoamide target. If a large library of proteins can be used to find a specific foldamer binder, then a large library of foldamers may as well be used to find a specific protein binder. For example, we recently reported that pull-down experiments allowed for the identification of nanomolar protein binders to a given helical aromatic foldamer out of the pool of proteins contained in a cell lysate.¹⁰ Similarly, phage display selection allowed for the identification of affimers – protein binders – for α -helix mimicking foldamers.¹¹ However, it is not yet known if these aromatic oligoamides are capable of interacting with short peptides.

Here, we report on the successful display selection of thioether-macrocyclic peptides (teMPs) using the random non-standard peptides integrated discovery (RaPID)¹² system for the selective recognition of aromatic foldamer helices having a deliberately reduced target area. Reducing the binding area on the target molecule created a challenge for peptide macrocycle display selection. Indeed, unlike for nucleic acid aptamers¹³ and in contrast to the success of the display selection of peptides against protein targets^{12,14} or other large substrates like polymers and solids,¹⁵ examples of successful display selections of peptides against small molecules are rare.¹⁶ Our results thus highlight both the potential of aromatic helical foldamers to recognize peptides and the potential of cyclic peptides to recognize a smaller target.

The folding of oligoamides of 8-amino-2-quinoline carboxylic acid Q^X (Fig. 1b) into stable aromatic helices has been thoroughly described.¹⁷ Their conformational stability is such that the equilibrium between the right-handed (*P*) and the left-handed (*M*) enantiomeric helical conformers of an octamer is kinetically inert in water at 25 °C.¹⁸ Methods to introduce a variety of biogenic side chains on Q^X monomers without impacting helix stability are also available.^{5b,19} We synthesized an aromatic oligoamide target potentially suitable for interaction with a peptide ligand: dodecameric sequence **1a** which contains five monomers carrying biogenic side chains (hydro-

^aDepartment Pharmazie, Ludwig-Maximilians-Universität München, Butenandtstr. 5-13, 81377 München, Germany. E-mail: celine.douat@cup.lmu.de, ivan.huc@cup.lmu.de

^bDepartment of Chemistry, School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-0033, Japan. E-mail: hsuga@chem.s.u-tokyo.ac.jp

^cCBMN (UMR5248), Univ. Bordeaux-CNRS-INP, Institut Européen de Chimie et Biologie, 2 rue Escarpit, 33600 Pessac, France

† Electronic supplementary information (ESI) available: Supplementary figures, schemes and data, detailed experimental protocols and characterisation of new compounds. See DOI: <https://doi.org/10.1039/d5ob00228a>

‡ Current address: Department of Drug Design and Pharmacology, University of Copenhagen, Copenhagen 2100, Denmark.

§ Current address: Department of Chemistry, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul 06974, Republic of Korea.



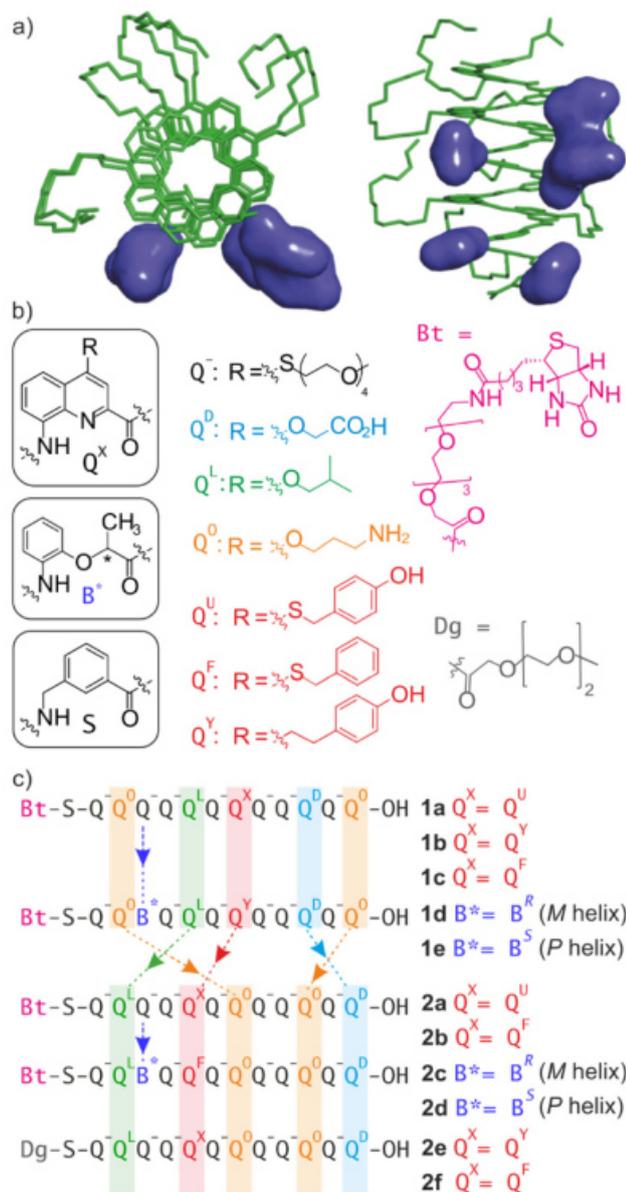


Fig. 1 (a) Top and side views of an energy-minimized model (MMFFs in Maestro)²² of the aromatic helix of **1a**. The five clustered biogenic side chains are shown as solvent accessible surfaces. The main chain and tetraethylene glycol side chains are shown in green stick representation. (b) Formula of aromatic δ -amino acid residues (Q) colour coded according to their side chain: hydrophilic (black), hydrophobic (green), cationic (orange), blue (anionic), and red (aromatic). The chiral B* residue with R or S configuration, the spacer (S), and the chemical formula of the N-terminal functionalization of the foldamers (Bt or Dg) are also defined. (c) Synthesized foldamer sequences **1a–e** and **2a–f**. A one letter code has been used to define the side-chain or the chirality of the aromatic α -amino acids.

phobic leucine-like for Q^L, anionic aspartate-like for Q^D, cationic ornithine-like for Q^O, and aromatic tyrosine-like for Q^U), arranged in such a way that they form a cluster, *i.e.* a potential binding spot on one side of the helix (Fig. 1a and c). The seven other Q⁻ residues bear tetraethylene glycol side chains to

provide water solubility and also to prevent peptide interactions outside of the binding spot.²⁰ We also synthesized sequence **2a**, an analogue of **1a** where the same biogenic residues are presented differently on the helix surface, to use as an alternative target and test the selectivity of peptide-foldamer interactions. Both sequences end with a 3-aminomethyl benzoic acid residue used as a spacer (S) separating the aromatic backbone from a biotin tail (Bt) intended for immobilization on streptavidin (SA). These foldamers are achiral and thus exist as a racemic mixture of *P*- and *M*-helical conformers. Both sequences were built on a solid support starting from a low-loading bromo-Wang resin and using *in situ* activation, as recently reported.²¹ Crude purity was good and final purification was achieved using RP-HPLC (see the ESI† for details).

Sequences **1a** and **2a** were individually loaded on magnetic SA-beads and two independent RaPID selections were carried out in parallel to identify teMP binders. In the first selection, **1a** was the target with a counter-selection against the streptavidin on the solid support. In the second selection, **2a** was the target and **1a** on SA-beads was used for counter-selection. Thus, the second selection should in principle exclude teMPs that bind to areas present in both **1a** and **2a** like the helix cross-sections (the area decorated with Q⁻ is also shared by the sequences but it is not considered to be a plausible binding area). Since both the *P* and *M* helices are present, selections were simultaneously performed against two targets. The peptide precursor DNA library contained a coding region of ATG-(NNK)_{*n*}-TGC sequences, N is all 4 bases, K = T or G, and *n* = 8–15. This was transcribed to mRNA for *in vitro* translation, where the AUG start codon was reprogrammed with *N*-chloroacetylated-(L)-Trp (ClAc-W) for selection against **1a** and with chloroacetylated-(D)-Trp (ClAc-w) for selection against **2a**. The repeats of NNK mixed codons encoded random α -amino acids and the TGC codon encoded a mandatory Cys to form a thioether macrocycle with the N-terminal chloroacetyl group. Of note, TGT was also present in the random NNK codons so additional Cys residues may appear earlier in the sequence. After the mandatory Cys, that is, nearer to the C-terminus, all peptides had a GSGSGS segment to act as a spacer between the cyclic peptide and the site where the encoding mRNA is attached.

Five iterative rounds of selection were performed before sequence analysis of the cDNA library (Fig. 2, S1 and S2†). The first selection led to strong convergence with twelve out of the top twenty ranked sequences showing more than 85% similarity, and the second to less extensive convergence. In both cases, the most amplified sequence (P1 and P2 in Fig. 2) contained two Cys residues. Based on previous studies,^{7b,23} we considered that the smallest possible thioether macrocycle, that is, the macrocycle involving the Cys nearest to the N terminus, formed first during the spontaneous post-translational cyclization, giving rise to a lariat structure as found in other RaPID selections.²⁴ P1 thus consisted of an -Ac-WSTC- macrocycle with a linear peptide extension, and P2 of an -Ac-wSYGWC- macrocycle with a linear peptide extension. In both cases, the linear extensions contained several other hydro-



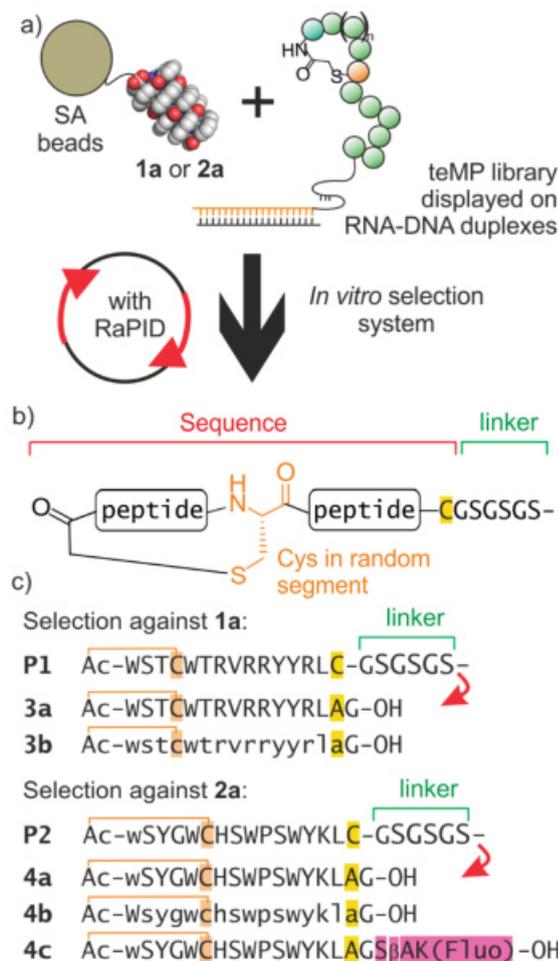


Fig. 2 (a) Overview of the *in vitro* selection of teMP binders for aromatic helical foldamers using the RaPID system. The biotinylated foldamers **1a** and **2a** were immobilized on SA-beads. (b) General structure of the selected teMPs. Peptide macrocyclization occurs via thioether formation from an N-terminal chloroacetyl group and a cysteine thiol which may be the mandatory Cys residue (highlighted in gold) or an earlier Cys in the random peptide segment. (c) Sequences of **P1** and **P2**, the most enriched peptides in the selections against **1a** and **2a**, respectively. These teMPs have a small macrocycle involving the first Cys in the sequence and a linear extension. Formulas of the synthesized lariat peptides **3a**, **3b**, **4a**, **4b**, and **4c**. Lowercase letters correspond to D amino acids. The orange line depicts the thioether linkage between the Cys thiol and N-terminal acetyl group. In these peptides, the C-terminal Cys was replaced with an alanine.

phobic residues (W, Y, P or L), suggesting that binding is driven by hydrophobic effects. For subsequent K_D determination, and to avoid possible issues with thiol oxidation, the C-terminal Cys was replaced with an Ala residue. Thus peptides **3a** and **4a** were prepared using standard solid phase peptide synthesis of a chloroacetylated precursor followed by resin cleavage and side-chain deprotection and by cyclization in a basic medium. The two teMPs were purified by semi-preparative RP-HPLC and characterized by LC-MS analysis.

At this stage, an unanticipated complication required our attention as we found that foldamers **1a** and **2a** slowly

degraded. Upon standing in an aqueous medium, the *para*-hydroxybenzyl moiety of Q^U underwent 1,6-elimination to give 4-mercapto-quinoline. This degradation had been overlooked in sequences made previously that contained Q^U .^{5d} We investigated side-chain instability using model pentamer **5** synthesized for that purpose (Fig. 3, S3 and S4[†]). The rate of elimination was found to vary with pH. It was slow enough for pH values between 5 and 7 to consider that elimination had been limited during the teMP RaPID selection. Nonetheless, the original target sequences **1a** and **2a** were judged unsuitable for physical investigations as degradation would reduce accuracy and reproducibility. We therefore prepared new foldamer sequences **1b–1e** and **2b–2f** in which Q^U was replaced either with Q^F , which bears a phenylalanine-like side chain similar to that of Q^U but lacking its hydroxy group, or with Q^Y which bears a tyrosine-like side chain with an ethylene connector instead of a thioether (Fig. 1b). Q^F and Q^Y do not differ from Q^U in the same way and it was not known whether these differences would matter so both were included in this study. The synthesis of Fmoc- $Q^{Y(Bu)}$ -OH and Fmoc- Q^F -OH are reported in the ESI (Schemes S1 and S2[†]).

Peptide–foldamer interactions were first assessed using bio-layer interferometry (BLI, Table 1, Fig. 4a, b and S5[†]). Biotinylated foldamers were independently loaded on SA-sensors and peptides were used as analytes in Tris buffer saline with Tween 20 (0.05%) and DMSO (0.1%) (TBST-D). The procedure is presented in detail in the ESI.[†] The observed kinetic curves showed a steady state regime, *i.e.* they reached a plateau value, and K_D values were determined using the Langmuir equation. Peptide **3a** from the first selection showed similar micromolar affinity for **1b** and **1c**, two analogues of the first selection target **1a**, suggesting that Q^F and Q^Y reproduce essential features of Q^U in this context. In contrast, in the

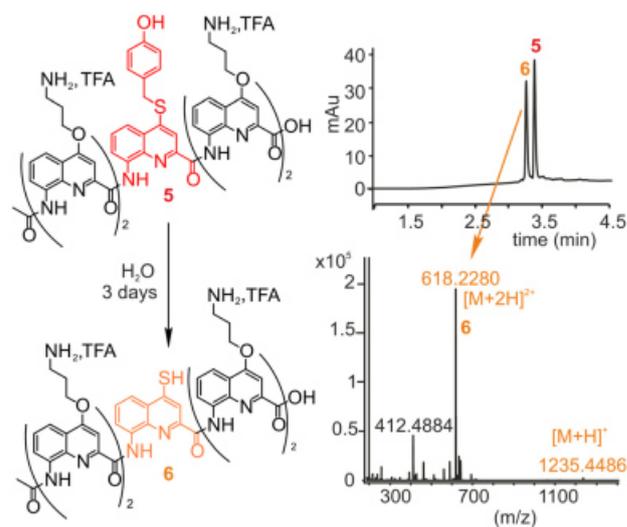


Fig. 3 Monitoring of the *para*-hydroxybenzyl side-chain elimination of Q^U in pure water on model pentamer **5**. The formation of the 4-mercapto-quinoline-containing pentamer **6** was confirmed by LC-MS analysis.



Table 1 K_D values (μM) determined by BLI

Immobilized foldamer	Lariat peptide analyte			
	3a	3b	4a	4b
1b	12.8 ± 0.3	— ^b	NBD ^a	— ^b
1c	19 ± 0.6	— ^b	— ^b	— ^b
1d (M)	9.5 ± 1.8	NBD	— ^b	— ^b
1e (P)	NBD ^a	10.3 ± 1.9	— ^b	— ^b
2b	NBD ^a	— ^b	16.1 ± 3.0	— ^b
2c (M)	— ^b	— ^b	NBD ^a	14.7 ± 2.2
2d (P)	— ^b	— ^b	9.7 ± 2.9	NBD ^a

^a NBD indicates no binding detected. ^b A dash indicates that this measurement was not performed.

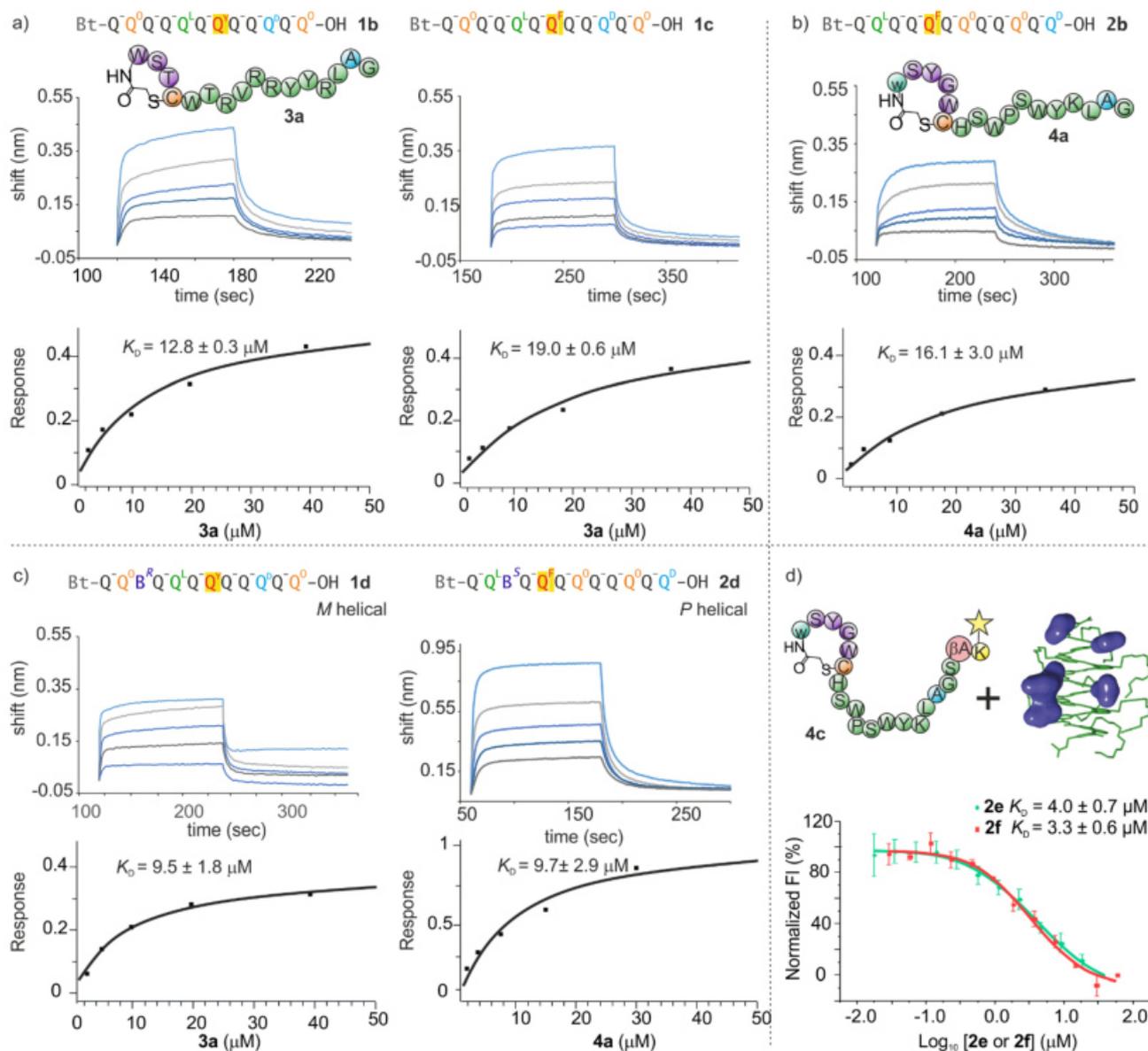


Fig. 4 BLI sensorgrams and steady state curve fittings to determine K_D values of the binding of (a) 3a with 1b and 1c loaded on SA-sensor tips, (b) 4a with 2b loaded on SA-sensor tips, (c) 3a with 1d loaded on SA-sensor tips, and 4a with 2d loaded on SA-sensor tips. (d) Fluorescence intensity monitoring of the binding of fluorescein-labelled teMP 4c and foldamers 2e and 2f. The binding assay results are from three independent experiments. Error bars represent the standard error of the mean of $N = 3$ experimental replicates.



same concentration range,** peptide **3a** showed no binding to **2b**, an analogue of the second selection target **2a**. This suggests that **3a** recognizes surface features unique to **1b** and **1c**, which excludes the aromatic helix cross-section also present in **2b**, and this is despite the absence of a counter-selection step against **2b** in the first selection. Similarly, peptide **4a** from the second selection showed micromolar affinity for **2b** and no binding to **1b** and **1c**. The observed selectivity thus points to the peptides targeting the small areas in which the biogenic side chains of their targets are clustered.

To consolidate these results, peptide–foldamer interactions were also assessed using a fluorescence assay as a second physical technique requiring no immobilization on a surface (Fig. 4d). Peptide **4c** was synthesized as an analogue of **4a** extended at its C-terminus using a Ser- β -Ala spacer and a fluorescein-labelled Lys. Concurrently, foldamers **2e** and **2f**, two analogues of **2a** having no biotin and either Q^F or Q^Y instead of Q^U, were prepared. Fluorescence was found to drop upon binding of the foldamers to the peptides. Peptide **4c** bound to **2e** and **2f** with K_D values of 4.0 μ M and 3.4 μ M, respectively. These values are slightly lower than those measured by BLI, and again show no significant differences between Q^F and Q^Y.

Next, we assessed the diastereoselectivity of the peptide foldamer interactions. The five Q^O ($\times 2$), Q^L, Q^D and Q^{Y/F} biogenic side chains are not only projected differently in space at the surface of **1b** and **2b** but, for both compounds, the side chains are also projected differently in their *M* and *P* helical conformations, both of which were present during selection (Fig. S6†). We thus synthesized four new biotinylated foldamers, **1d**, **1e**, **2c** and **2d** (Fig. 1c). In each of them, the third residue (Q⁻) was replaced with a chiral 2-(2-aminophenoxy)-propionic acid monomer (B* in Fig. 1b) with either (*R*) or (*S*) configuration. B^R and B^S have been shown to quantitatively bias the handedness of the aromatic helix to *M* (in **1d** and **2c**) and *P* (in **1e** and **2d**), respectively.²⁵ The extent of handedness bias was confirmed by ¹H NMR spectra where a single set of signals corresponding to only one (*R-M* or *S-P*) diastereomeric conformer is seen. The sign of helix handedness was controlled by circular dichroism spectroscopy. Each new foldamer was then individually immobilized on SA-sensor tips and their binding to peptides was assessed by BLI. These unveiled that the interactions are diastereoselective in the range of peptide concentrations used. teMP **3a** binds to *M*-helical **1d** and not to *P*-helical **1e**, while teMP **4a** binds to *P*-helical **2d** and not to *M*-helical **2c** (Table 1, Fig. 4c and S7†). It is noteworthy that the first selection has generated a binder selective for the *M* helix whereas the other selection yielded a binder selective for the *P* helix. The K_D values calculated from steady-state curve fitting are comparable to those obtained during the first round of BLI measurements on *P/M* helix mixtures, *i.e.* in the low micromolar range. Finally, we further validated the diastereoselective interactions in experiments where both the foldamer handedness and the peptide chirality were inverted. For that purpose, we prepared peptides **3b** and **4b**, the enantiomers of **3a** and **4a**, respectively. We found that the binding of **4b** to *M*-helical **2c** was similar to that of **4a** to *P*-helical **2d** (Table 1 and Fig. S8†). In the case of **3b**,

measurements were first hampered by nonspecific binding of this *D*-peptide to the reference SA-sensor tips. This effect decreased sufficiently when adding 0.05% BSA to the TBST-D buffer and it could be again verified that the binding of **3b** to *P*-helical **1e** was similar to the binding of its enantiomer **3a** to *M*-helical **1d** (Table 1 and Fig. S8†).

Conclusions

In conclusion, despite the relatively small potential binding spot on the surface of the target helical aromatic oligoamide foldamers (only five biogenic side chains), the low molecular weight of the target (1.5 kDa, significantly smaller than typical RaPID protein targets),¹² and the highly non-proteinogenic chemistry, RaPID selections successfully delivered low micromolar binders that showed high selectivity for a specific arrangement of foldamer side chains. The specificity for the sequence of the foldamer biogenic side chains suggests that the cluster of biogenic side chains constitutes the teMP binding site. This specificity was further emphasized by the peptides' ability to selectively bind one handedness of the helical foldamer. Furthermore, the synthetic tractability of the peptides allowed for the synthesis of enantiomers to target the other handedness. Because of the moderate binding affinities in the low μ M range, we did not invest further efforts in structural studies or systematic investigations of the sequence dependence of the interactions through, *e.g.*, Ala-scans. Nevertheless, our results highlight the high potential of aromatic foldamers equipped with biogenic side chains to interact with α -peptidic objects.

Author contributions

LW, JMR, SJD, LL, SK, and CD performed the experiments. CD, HS and IH supervised the research. All authors contributed to the writing, reviewing and editing of the manuscript and approved its final version.

Data availability

The data supporting this article have been included as part of the ESI.†

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) through project CRC1309-C7 (project ID 325871075) to I. H. and the JSPS Grant-in-Aid for



Specially Promoted Research (JP20H05618) to H. S. We thank Ivan Alonso for assistance with peptide synthesis and purification.

References

†Aromatic groups bearing amine and carboxylic acid substituents. Amino acids with aryl groups in their side chains such as Phe, Tyr, Trp are also called "aromatic amino acids" but are not relevant here.

‡Changing the configuration of the N-terminal α -amino acid often results in the selection of different sequences (see ref. 12d). To explore new sequence space when targeting 2a, we used "w" instead of "W".

**BLI experiments were typically performed up to teMP concentrations of 50 μ M. At higher concentrations, teMP aggregation and nonspecific binding to SA-biosensors made it difficult to acquire reliable data.

- 1 T. Seedorf, A. Kirschning and D. Solga, *Chem. – Eur. J.*, 2021, **27**, 7321.
- 2 S. Baumann, J. Herrmann, R. Raju, H. Steinmetz, K. I. Mohr, S. Huttel, K. Harmrolfs, M. Stadler and R. Muller, *Angew. Chem., Int. Ed.*, 2014, **53**, 14605; E. Michalczyk, K. Hommernick, I. Behroz, M. Kulike, Z. Pakosz-Stepien, L. Mazurek, M. Seidel, M. Kunert, K. Santos, H. von Moeller, B. Loll, J. B. Weston, A. Mainz, J. G. Heddle, R. D. Sussmuth and D. Ghilarov, *Nat. Catal.*, 2023, **6**, 52; S. M. Hashimi, *J. Antibiot.*, 2019, **72**, 785; P. G. Baraldi, C. Nunez, A. Espinosa and R. Romagnoli, *Curr. Top. Med. Chem.*, 2004, **4**, 231; Y. Hiraku, S. Oikawa and S. Kawanishi, *Nucleic Acids Symp. Ser.*, 2002, **2**, 95.
- 3 J. Zeelen, M. van Straaten, J. Verdi, A. Hempelmann, H. Hashemi, K. Perez, P. D. Jeffrey, S. Halg, N. Wiedemar, P. Maser, F. N. Papavasiliou and C. E. Stebbins, *Nat. Microbiol.*, 2021, **6**, 392; G. H. M. Salvador, T. R. Dreyer, A. A. S. Gomes, W. L. G. Cavalcante, J. I. Dos Santos, C. A. Gandin, M. de Oliveira Neto, M. Gallacci and M. R. M. Fontes, *Sci. Rep.*, 2018, **8**, 10317; L. Jiao, S. Ouyang, M. Liang, F. Niu, N. Shaw, W. Wu, W. Ding, C. Jin, Y. Peng, Y. Zhu, F. Zhang, T. Wang, C. Li, X. Zuo, C. H. Luan, D. Li and Z. J. Liu, *J. Virol.*, 2013, **87**, 6829; M. T. Murakami, E. Z. Arruda, P. A. Melo, A. B. Martinez, S. Calil-Elias, M. A. Tomaz, B. Lomonte, J. M. Gutierrez and R. K. Arni, *J. Mol. Biol.*, 2005, **350**, 416; G. H. M. Salvador, T. R. Dreyer, W. L. Cavalcante, F. F. Matioli, J. I. Dos Santos, A. Velazquez-Campoy, M. Gallacci and M. R. Fontes, *Acta Crystallogr.*, 2015, **71**, 2066; E. Mastrangelo, M. Pezzullo, D. Tarantino, R. Petazzi, F. Germani, D. Kramer, I. Robel, J. Rohayem, M. Bolognesi and M. Milani, *J. Mol. Biol.*, 2012, **419**, 198.
- 4 I. Saraogi, J. A. Hebda, J. Becerril, L. A. Estroff, A. D. Miranker and A. D. Hamilton, *Angew. Chem., Int. Ed.*, 2010, **49**, 736; T. Flack, C. Romain, A. J. P. White, P. R. Haycock and A. Barnard, *Org. Lett.*, 2019, **21**, 4433; R. A. Dohoney, J. A. Joseph, C. Baysah, A. G. Thomas, A. Siwakoti, T. D. Ball and S. Kumar, *ACS Chem. Biol.*, 2023, **18**, 1510; I. Arrata, C. M. Grison, H. M. Coubrough, P. Prabhakaran, M. A. Little, D. C. Tomlinson, M. E. Webb and A. J. Wilson, *Org. Biomol. Chem.*, 2019, **17**, 3861; G. M. Burslem, H. F. Kyle, A. L. Breeze, T. A. Edwards, A. Nelson, S. L. Warriner and A. J. Wilson, *ChemBioChem*, 2014, **15**, 1083; S. Kumar and A. D. Hamilton, *J. Am. Chem. Soc.*, 2017, **139**, 5744; J. P. Plante, T. Burnley, B. Malkova, M. E. Webb, S. L. Warriner, T. A. Edwards and A. J. Wilson, *Chem. Commun.*, 2009, **34**, 5091.
- 5 (a) J. M. Alex, V. Corvaglia, X. Hu, S. Engilberge, I. Huc and P. B. Crowley, *Chem. Commun.*, 2019, **55**, 11087; (b) M. Zwillinger, P. S. Reddy, B. Wicher, P. K. Mandal, M. Csékei, L. Fischer, A. Kotschy and I. Huc, *Chem. – Eur. J.*, 2020, **26**, 17366; (c) P. S. Reddy, B. Langlois d'Estaintot, T. Granier, C. D. Mackereth, L. Fischer and I. Huc, *Chem. – Eur. J.*, 2019, **25**, 11042; (d) J. Buratto, C. Colombo, M. Stupfel, S. J. Dawson, C. Dolain, B. Langlois d'Estaintot, L. Fischer, T. Granier, M. Laguerre, B. Gallois and I. Huc, *Angew. Chem., Int. Ed.*, 2014, **53**, 883; (e) S. Kumar, M. Birol, D. E. Schlamadinger, S. P. Wojcik, E. Rhoades and A. D. Miranker, *Nat. Commun.*, 2016, **7**, 11412; (f) J. Ahmed, T. C. Fitch, C. M. Donnelly, J. A. Joseph, T. D. Ball, M. M. Bassil, A. Son, C. Zhang, A. Ledreux, S. Horowitz, Y. Qin, D. Paredes and S. Kumar, *Nat. Commun.*, 2022, **13**, 2273.
- 6 V. Corvaglia, D. Carbajo, P. Prabhakaran, K. Ziach, P. K. Mandal, V. D. Santos, C. Legeay, R. Vogel, V. Parissi, P. Pourquier and I. Huc, *Nucleic Acids Res.*, 2019, **47**, 5511; V. Kleene, V. Corvaglia, E. Chacin, I. Forne, D. B. Konrad, P. Khosravani, C. Douat, C. F. Kurat, I. Huc and A. Imhof, *Nucleic Acids Res.*, 2023, **51**, 9629; D. Deepak, J. Wu, V. Corvaglia, L. Allmendinger, M. Scheckenbach, P. Tinnefeld and I. Huc, *Angew. Chem., Int. Ed.*, 2025, **64**, e202422958.
- 7 (a) M. Kudo, V. Maurizot, B. Kauffmann, A. Tanatani and I. Huc, *J. Am. Chem. Soc.*, 2013, **135**, 9628; (b) S. Dengler, C. Douat and I. Huc, *Angew. Chem., Int. Ed.*, 2022, **61**, e202211138.
- 8 C. Tsiamantas, S. Kwon, J. M. Rogers, C. Douat, I. Huc and H. Suga, *Angew. Chem., Int. Ed.*, 2020, **59**, 4860; C. Tsiamantas, S. Kwon, C. Douat, I. Huc and H. Suga, *Chem. Commun.*, 2019, **55**, 7366; J. M. Rogers, S. Kwon, S. J. Dawson, P. K. Mandal, H. Suga and I. Huc, *Nat. Chem.*, 2018, **10**, 405; O. Ad, K. S. Hoffman, A. G. Cairns, A. L. Featherston, S. J. Miller, D. Söll and A. Schepartz, *ACS Cent. Sci.*, 2019, **5**, 1289; T. Katoh and H. Suga, *J. Am. Chem. Soc.*, 2020, **142**, 16518.
- 9 T. Katoh and H. Suga, *J. Am. Chem. Soc.*, 2022, **144**, 2069; S. Dengler, R. T. Howard, V. Morozov, C. Tsiamantas, W.-E. Huang, Z. Liu, C. Dobrzanski, V. Pophristic, S. Brameyer, C. Douat, H. Suga and I. Huc, *Angew. Chem., Int. Ed.*, 2023, **62**, e202308408.
- 10 S. Kwon, V. Morozov, L. Wang, P. K. Mandal, S. Chaignepain, C. Douat and I. Huc, *Org. Biomol. Chem.*, 2024, **22**, 9342.
- 11 I. Arrata, A. Barnard, D. C. Tomlinson and A. J. Wilson, *Chem. Commun.*, 2017, **53**, 2834.
- 12 (a) Y. Goto and H. Suga, *Acc. Chem. Res.*, 2021, **54**, 3604; (b) Y. Yamagishi, I. Shoji, S. Miyagawa, T. Kawakami, T. Katoh, Y. Goto and H. Suga, *Chem. Biol.*, 2011, **18**, 1562;



- (c) M. Nawatha, J. M. Rogers, S. M. Bonn, I. Livneh, B. Lemma, S. M. Mali, G. B. Vamisetti, H. Sun, B. Bercovich, Y. Huang, A. Ciechanover, D. Fushman, H. Suga and A. Brik, *Nat. Chem.*, 2019, **11**, 644;
- (d) T. Passioura, K. Watashi, K. Fukano, S. Shimura, W. Saso, R. Morishita, Y. Ogasawara, Y. Tanaka, M. Mizokami, C. Sureau, H. Suga and T. Wakita, *Cell Chem. Biol.*, 2018, **25**, 906.
- 13 A. Ruscito and M. C. DeRosa, *Front. Chem.*, 2016, **4**; Y. Mu, Z. Chen, J. Zhan and J. Zhang, *Anal. Sens.*, 2024, **4**, e202400027.
- 14 S. You, G. McIntyre and T. Passioura, *Expert Opin. Drug Discovery*, 2024, **19**, 961.
- 15 T. Serizawa, T. Sawada, H. Matsuno, T. Matsubara and T. Sato, *J. Am. Chem. Soc.*, 2005, **127**, 13780; C. A. Stevens, F. Bachtiger, X.-D. Kong, L. A. Abriata, G. C. Sosso, M. I. Gibson and H.-A. Klok, *Nat. Commun.*, 2021, **12**, 2675; J. Bang, H. Park, J. Yoo, D. Lee, W. I. Choi, J. H. Lee, Y.-R. Lee, C. Kim, H. Koo and S. Kim, *Sci. Rep.*, 2020, **10**, 10576.
- 16 T. Sawada, Y. Okeya, M. Hashizume and T. Serizawa, *Chem. Commun.*, 2013, **49**, 5088.
- 17 H. Jiang, J. M. Leger and I. Huc, *J. Am. Chem. Soc.*, 2003, **125**, 3448; C. Dolain, A. Grelard, M. Laguerre, H. Jiang, V. Maurizot and I. Huc, *Chem. - Eur. J.*, 2005, **11**, 6135;
- T. Qi, V. Maurizot, H. Noguchi, T. Charoenraks, B. Kauffmann, M. Takafuji, H. Ihara and I. Huc, *Chem. Commun.*, 2012, **48**, 6337.
- 18 S. J. Dawson, Á. Mészáros, L. Pethő, C. Colombo, M. Csékei, A. Kotschy and I. Huc, *Eur. J. Org. Chem.*, 2014, 4265.
- 19 X. Hu, S. J. Dawson, P. K. Mandal, X. de Hatten, B. Baptiste and I. Huc, *Chem. Sci.*, 2017, **8**, 3741; M. Zwillinger, P. Sőregi, F. Sanchez, C. Douat, M. Csékei, I. Huc and A. Kotschy, *J. Org. Chem.*, 2025, **90**, 3043.
- 20 C. Bernhard, S. J. Roeters, J. Franz, T. Weidner, M. Bonn and G. Gonella, *Phys. Chem. Chem. Phys.*, 2017, **19**, 28182.
- 21 S. Dengler, P. K. Mandal, L. Allmendinger, C. Douat and I. Huc, *Chem. Sci.*, 2021, **12**, 11004; V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat and I. Huc, *Chem. - Eur. J.*, 2023, **29**, e202300898.
- 22 *Maestro*, Schrödinger, LLC, New York, NY, 2021.
- 23 K. Iwasaki, Y. Goto, T. Katoh and H. Suga, *Org. Biomol. Chem.*, 2012, **10**, 5783.
- 24 H. Yu, P. Dranchak, Z. Li, R. MacArthur, M. S. Munson, N. Mehzabeen, N. J. Baird, K. P. Battalie, D. Ross, S. Lovell, C. K. Carlow, H. Suga and J. Inglese, *Nat. Commun.*, 2017, **8**, 14932.
- 25 D. Bindl, E. Heinemann, P. K. Mandal and I. Huc, *Chem. Commun.*, 2021, **57**, 5662.



8. Supplementary Information: Display Selection of Macrocyclic Peptides as Diastereoselective Ligands of Helical Aromatic Foldamers

Supplementary Information

Display selection of macrocyclic peptides as diastereoselective ligands of helical aromatic foldamers

Lingfei Wang,^a Joseph M. Rogers,^{b†} Simon J. Dawson,^c Line Langhorn,^a Ryan T. Howard,^a Sunbum Kwon,^{a‡} Céline Douat,^{*a} Hiroaki Suga^{*b} and Ivan Huc^{*a,c}

^a Department Pharmazie, Ludwig-Maximilians-Universität, Butenandtstr. 5–13, 81377 München (Germany).

^b Department of Chemistry, School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo, 113-0033 (Japan).

^c CBMN (UMR5248), Univ. Bordeaux-CNRS-INP, Institut Européen de Chimie et Biologie, 2 rue Escarpit, 33600 Pessac (France).

† Current address: Department of Drug Design and Pharmacology, University of Copenhagen, Copenhagen 2100 (Denmark).

‡ Current address: Department of Chemistry, Chung-Ang University, 84 Heukseokro, Dongjakgu, Seoul 06974 (Republic of Korea).

8.1 Supplementary figures

1a / counterselection SA; 6 th ; L-Trp	2a / counterselection 1a; 5 th ; D-Trp
1. x STCWTRVRRYYRLC*	1. y SYCWCHSWPSWYKLC*
2. x HWFCTAYRPC*	2. y WKRTFWWHLHKC*
3. x TTPWNNFLCRRTYPC*	3. y WLCESMLFRCRLIYAAAAAAA
4. x WNEWVWITWRLLVRC*	4. y RQRQRQLGRCAEKICRAHV* T
5. x HWISWRPVTCLFAAAAAART	5. y SYCWCHSWPSWYKLRQRQRQL
6. x SACWTRVRRYYRLC*	6. y ICTGC-TLWNAKFWWC*
7. x WTECKIYNSWVTYWNC*	7. y SWCAQFIHNGCIYSC*
8. x STCWTRVRRYYRLC*	8. y LSLQYWRSRTFALC*
9. x STCWTRVRRHYRLC*	9. y WNCRIIFCCRLIIAAAAAAA
10. x AYAQKRYSRAYELC*	10. y QHCH-VvCANLGFYYC*
11. x STCWTRVRRYYRLC*	11. y NDREYQSVQYvWC*
12. x STCWTRVRRYYRLC*	12. y RVWVHVYSC*
13. x STCWTRVRRYYRSC*	13. y SYCWCHSWPSWYELC*
14. x STCWTRVRRYYRLC*	14. y LCTLMWKCQIFPLNTC*
15. x STCWTRVRRHYRLC*	15. y SYCWRHSWPSWYKLC*
16. x STCWTRVRRDYRLC*	16. y VNCWMLWHGKLC*
17. x WTKRRYCELWRLC*	17. y VTYCNYLI LAQFVSC*
18. x STCRTRVRRYYRLC*	18. y SYCWCHSWPSWYKPC*
19. x WLYRHSYRYFEVRVC*	19. y SYCWCHSWPSWYRLC*
20. x STRWTRVRRYYRLC*	20. y SYCWCHSWPSWYKLAAAAAA

Figure S1. Table results of the two selections performed with **1a** (left) with SA as a counterselection and **2a** (right) with **1a** as a counterselection. The 20 most abundant peptide sequences for both libraries, in which: (x) stands for the CH₂CO-Trp, (y) for the CH₂CO-(D)-Trp (*) for the GSGSGS linker followed by the amber stop codon (TAG) and (-) in the middle of the sequence (sequence 6 in the right library) indicates the appearance of the amber stop codon (TAG). The GSGSGS linker in some cases was mutated to AAAAAA and RQRQRQ due to indels that occurred during the initial oligo synthesis, or during PCR, transcription or reverse transcription. Codons for methionine (M) cannot be translated as methionine as this amino acid was not added to the translation system, likely this was translated as Isoleucine.

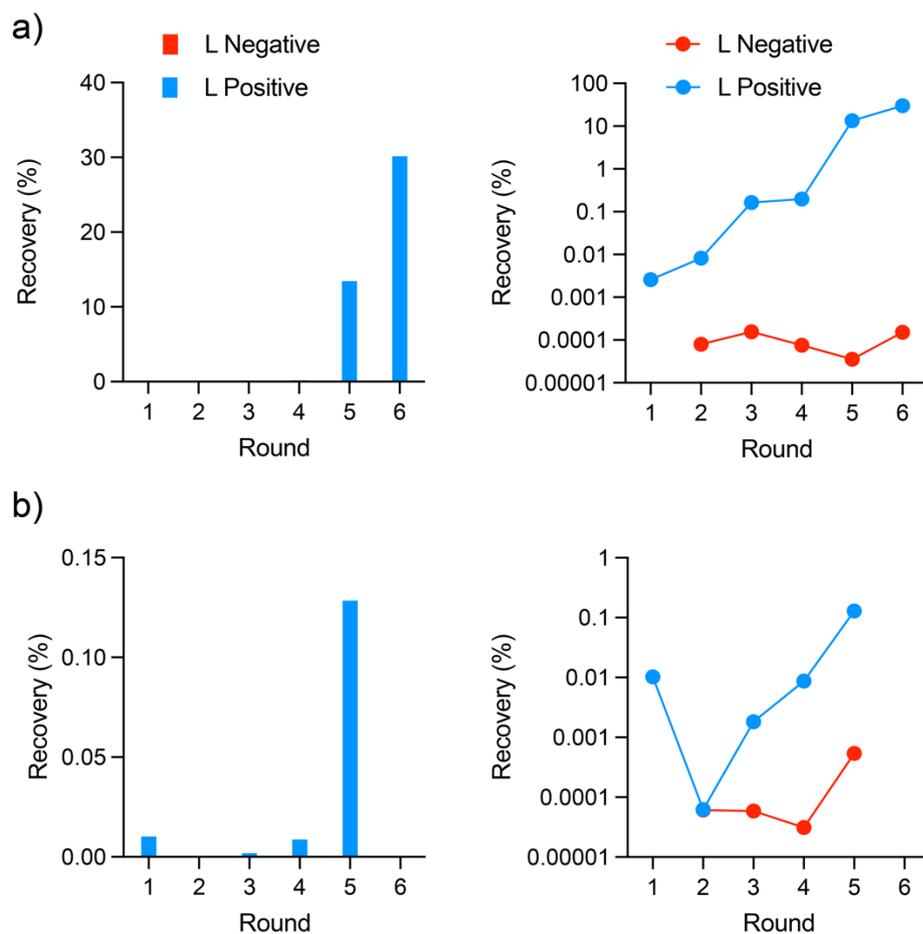


Figure S2. RaPID selection against aromatic oligoamides **1a** and **2a** resulted in an increase in DNA recovery suggesting an increase in the binding affinity of the library as particular peptides become enriched. a) DNA recovery, as measured by qPCR, for RaPID against **1a** (positive) with streptavidin as a counterselection (negative). Higher recovery is observed for positive, suggesting binders to the solid support are not being recovered. b) DNA recovery for RaPID against **2a** (positive) with **1a** as a counterselection (negative).

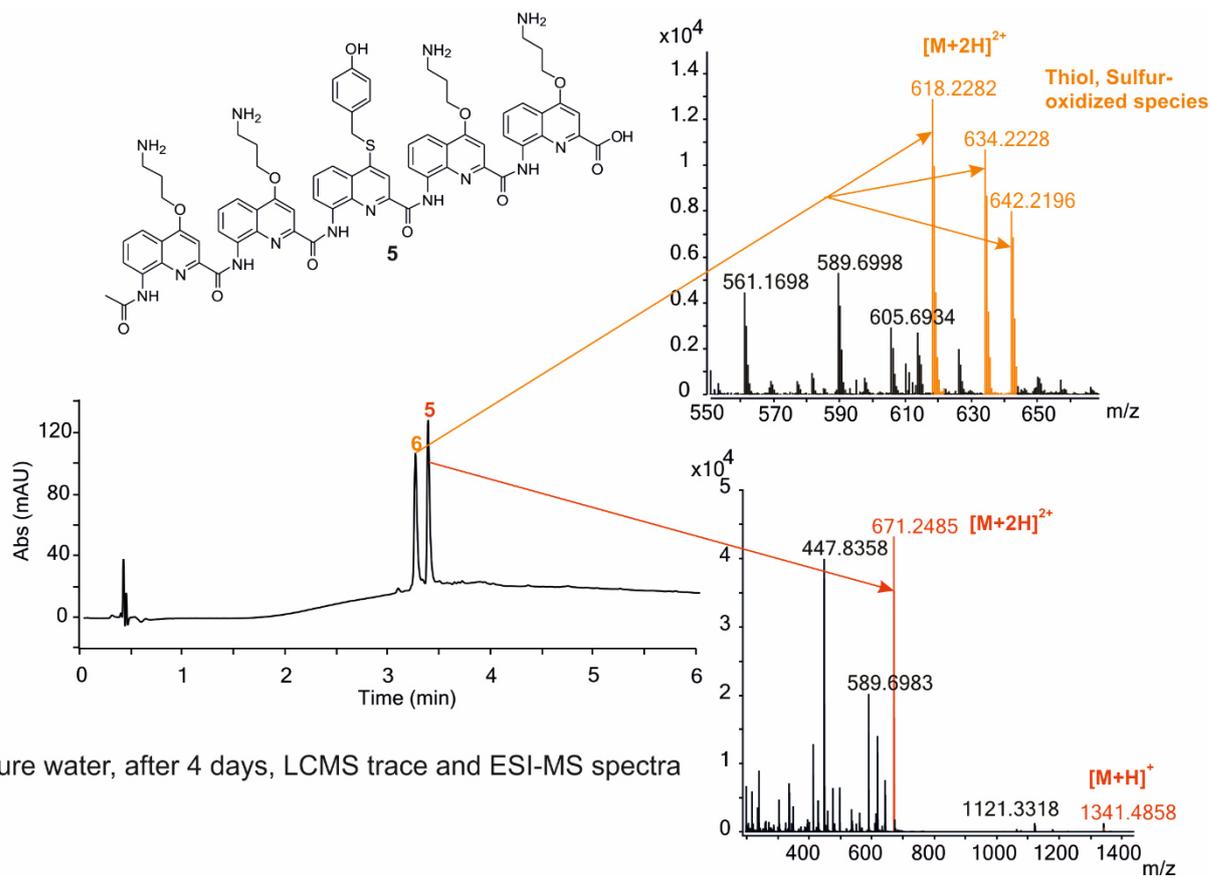


Figure S3. Monitoring of the degradation of model Ac-Q₅-OH pentamer **5** in pure water.

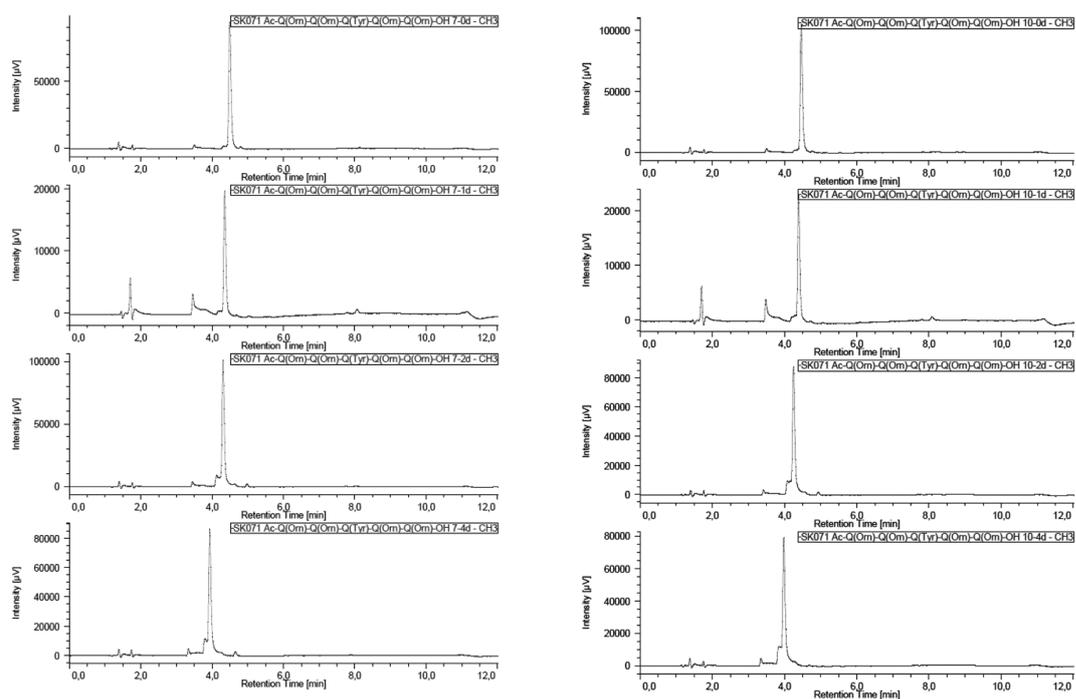


Figure S4. Examples of the monitoring of the stability of Ac-Q₅-OH (5) at two different pH values of phosphate buffer over 4 days: pH 6.4 (left) and pH 7 (right).

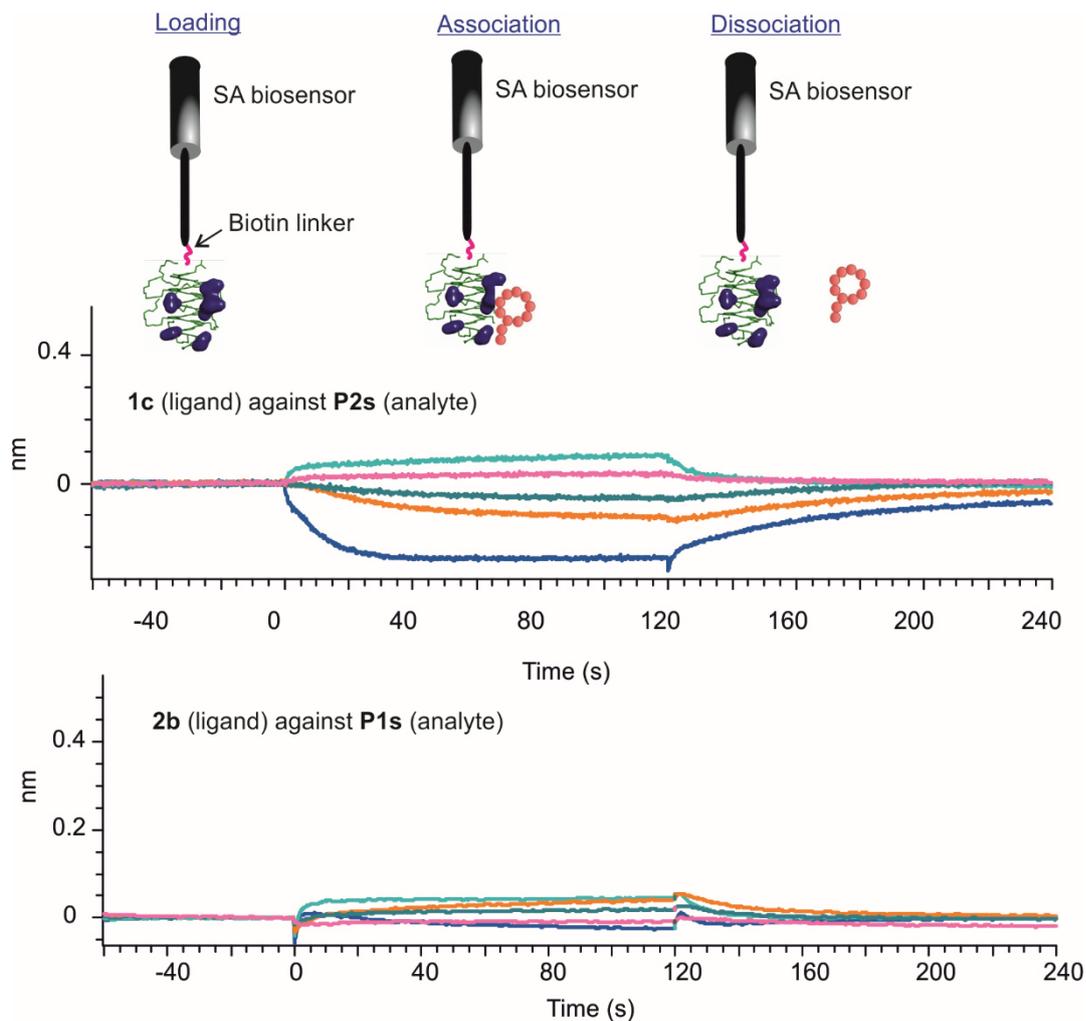


Figure S5. BLI control experiments to confirm the selective and specific binding of each teMP **3a** and **4a** to their corresponding foldamer sequence **1c** and **2b** respectively. For each experiment, five concentrations of peptide were used, ranging from 39-2.5 μM for **3a** and **4a**. In both assays, a double referencing was performed to remove any nonspecific binding of the peptide macrocycle to the SA-surface.

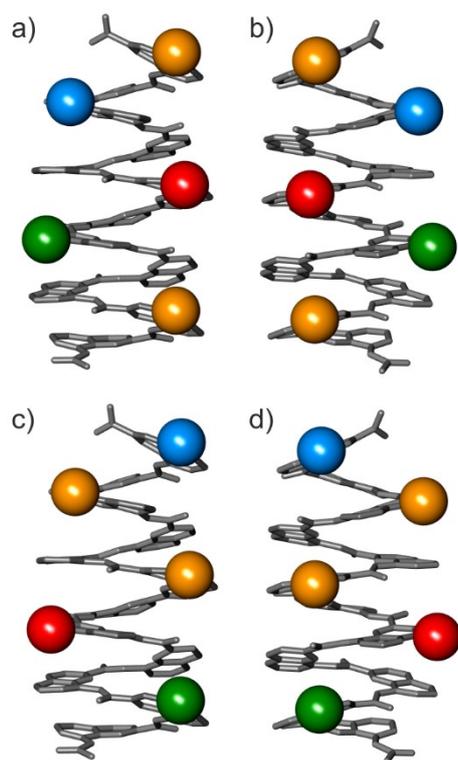


Figure S6. Biogenic side chain presentation at the surface of the *M* helix (a) and the *P* helix (b) of **1a**, and at the surface of the *M* helix (c) and the *P* helix (d) of **1b**. The biogenic side chains are represented by a colored sphere. The color code is the same as in Fig. 1 (orange = Q^O, red = Q^U, green = Q^L, blue = Q^D). The helix backbone is shown in gray tube representation. Other side chains and hydrogen atoms have been removed for clarity.

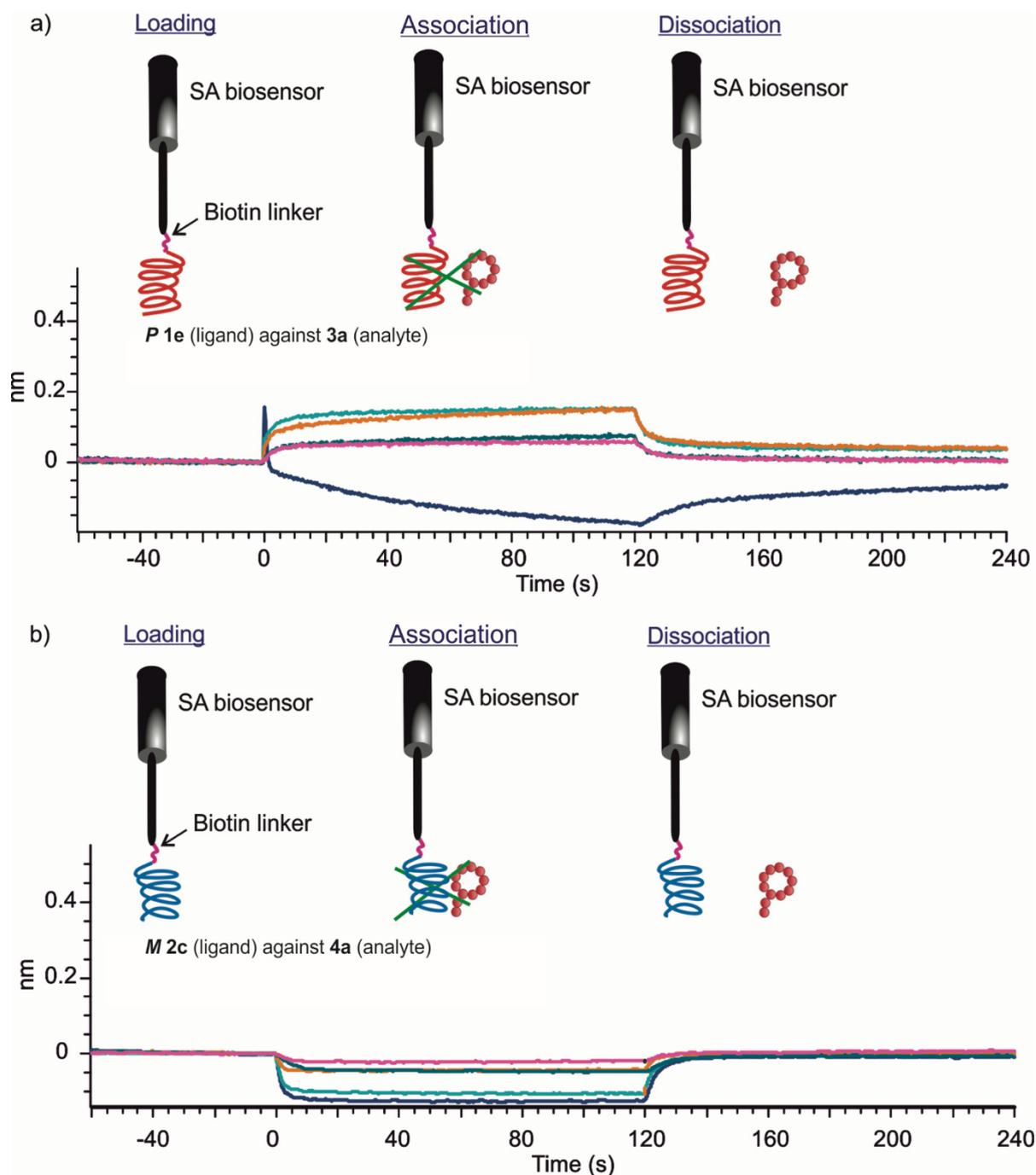


Figure S7. Negative control BLI experiments to confirm the selective binding of one aromatic helix handedness for the selected peptides: a) the *P* helix of foldamer **1** (**1e**) does not bind to **3a** and b) the *M* helix of foldamer **2** (**2c**) does not recognize the **4a**. For each experiment, five concentrations of peptide were used, ranging from 39-2.5 μM for **3a** and from 35-2.2 μM for **4a**. In both assays, a double referencing was performed to remove any nonspecific binding of the macrocyclic peptides to the SA-biosensors.

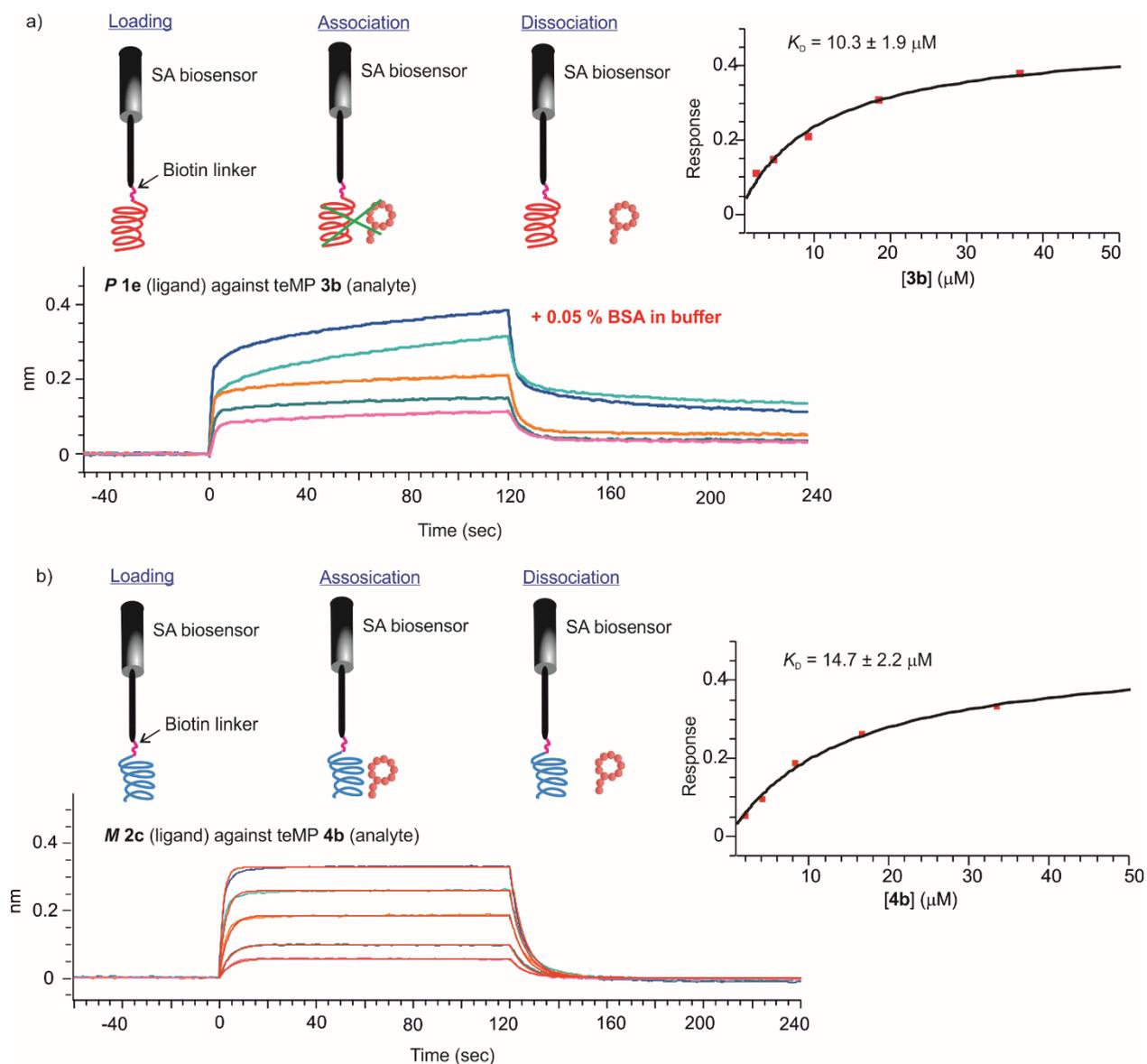


Figure S8. BLI experiments to confirm the selective binding of the opposite aromatic helix handedness for the other peptide enantiomer (**3b**) or (**4b**): a) the *P* helix of foldamer **1 (1e)** binds to **3b** and b) the *M* helix of foldamer **2 (2c)** does recognize **4b**. For each experiment, five concentrations of peptide were used ranging, from 35-2.2 μM for **3b** and from 33-2.1 μM for **4b**. In both assays, a double referencing was performed to remove any nonspecific binding of the peptide macrocycle to the SA-surface.

8.2 Experimental section

8.2.1 Experimental procedures for RaPID system

Two RNA libraries, consisting of 8–15 NNK codons, were prepared as previously described.¹ Briefly, RNA molecules were synthesized by T7 RNA polymerase reactions from DNA templates assembled by PCR and purified by PAGE. Furthermore, puromycin-linked mRNA was prepared by incubation with puromycin-linked oligonucleotide and T4 RNA ligase and was purified by phenol/chloroform extraction and ethanol precipitation. Oligonucleotides for both libraries and puromycin-linked oligonucleotide were previously reported.²

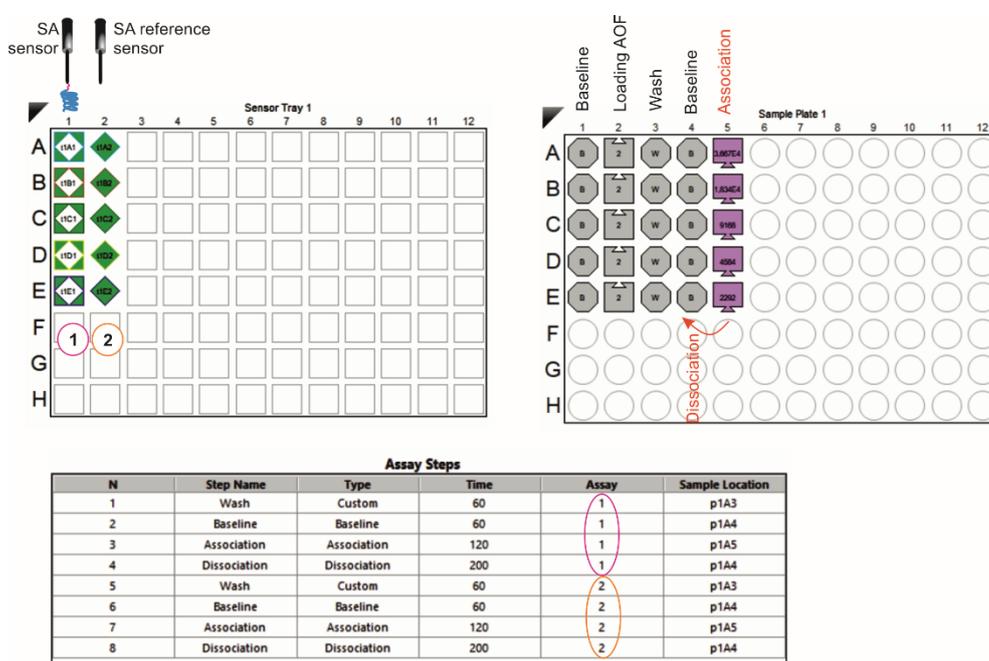
Ribosomal synthesis of the macrocyclic peptide libraries was performed as previously described.³ In brief, for the initial selection, 1.2 μM puromycin-linked mRNA library was translated in a Met-deficient translation system reaction containing 25 μM of ClAc^{L(D)}Trp-tRNA^{fMet} for 30 min at 37 °C (150 μL scale for first round, 5 μL scale for subsequent rounds). The reaction was incubated at 25 °C for 12 min before disruption of the ribosome–mRNA complex by incubation at 37 °C for 30 min in the presence of 20 mM EDTA. The resulting peptide-linked mRNAs were then reverse-transcribed using RNase H-reverse transcriptase (Promega) for 1 h at 42 °C. For the first selection, affinity screening was performed by three serial passages (counterselections, 30 min each at 4 °C) of the library over Streptavidin Dynabeads (Life Technologies) followed by affinity selection against 200 nM **1a** immobilized on the same beads for 30 min at 4 °C. For the second selection, affinity screening was performed by three serial passages (counterselections, 30 min each at 4 °C) of the library over Streptavidin Dynabeads loaded with **1a** (Life Technologies) followed by affinity selection against 200 nM **2a** immobilized on the same beads for 30 min at 4 °C. cDNA was eluted from the beads by heating to 95 °C for 5 min and fractional recovery from the final counterselection (negative control) and affinity selection step were assessed by quantitative PCR using Sybr Green I on a LightCycler thermal cycler (Roche) (Figure S8). Enriched DNA libraries were recovered by PCR and used as input for transcription reactions to generate the mRNA library for the subsequent round of screening. After five iterative rounds of library synthesis, affinity selection, and recovery, the final DNA library was sequenced to identify putative binders (Figure 2 and S2).

8.2.2 Binding characterization by BioLayer Interferometry (BLI)

BLI experiments were performed on an Octet R8 instrument from Sartorius, following Sartorius recommendations. Prior to an assay, streptavidin (SA) sensor tips were soaked for at least 10 min in phosphate buffer saline (1 \times PBS). The kinetic experiment always starts with a baseline step over 60 s in 1 \times TBST 0.1% DMSO (TBST-D) buffer, followed by the loading Bt-foldamer at 2 $\mu\text{g}/\text{mL}$ over 60 s in TBST-D. After foldamer ligand immobilization, the biosensors were washed for 60 s in the same buffer, before to record a second baseline for 120 s, again in TBST-D. Serial column dilutions (\times 2) of

the teMPs in TBST-D were analysed (five concentrations in total). Association lasted 120 s, followed by dissociation for 240 s. The curves were fitted to binding models using the Octet analysis studio 13.0 software. A double referencing was performed for each foldamer/teMP kinetic experiment. A second set of five sensors were hence used following the same series of kinetic steps for the kinetic curve at the exception of the loading step, which was replaced with pure TBST-D buffer. Each teMPs concentration was subtracted with its reference well/unloaded SA-sensor tip (see figure below). Of note, this double referencing was set-up to subtract any nonspecific binding of the teMPs to streptavidin. For all foldamer/teMPs binary complex, the dissociation constants (K_D) were calculated with the Langmuir's equation assuming, a 1:1 binding model (see equation below).

$$R = Rmax \times [teMP]/(K_D + [teMP])$$



8.2.3 Fluorescence polarization of foldamers 2e and 2f with 4c

In a F-bottom, black 96-well plate (Greiner, 738-0026), a serial dilution in triplicate of foldamer was prepared in RaPID selection TBST buffer (Tris-HCl (20 mM) pH 7.4, NaCl (150 mM), and Tween-20 (0.05% (v/v))). teMP 4c or 5(6)-carboxyfluorescein (Sigma-Aldrich, 21877) was added to each well to a final concentration of 20 nM and a final well volume of 200 μ L. Buffer-only and fluorophore-only (no foldamer) controls were also included on the plate. The plates were incubated at 4 $^{\circ}$ C for a minimum of 30 min before measurement. Polarization data were measured on a TECAN Infinite M1000 Pro at ambient temperature (24 $^{\circ}$ C) with the following parameters: λ_{ex} = 470 nm (bandwidth = 5 nm), λ_{em} = 525 nm (bandwidth = 5 nm), gain/z-position calculated from well with highest protein concentration, flashes = 12, settle time = 50 ms, G-factor (calibrated from 1 nM fluorescein standard in 10 mM NaOH) = 1.086. For the calculation of K_D values, dose-response data were fit in OriginPro 2019b software to the quadratic 1:1 binding model using the following equation.

$$Polarization = BOTTOM + (TOP - BOTTOM) \left(\frac{([4c] + K_D + [AOF]) - \sqrt{(A_0 + K_{dD} + [AOF])^2 - 4[4c][AOF]}}{2[4c]} \right)$$

Where $[AOF]$ is total foldamer concentration (μM), $BOTTOM$ is minimum polarization (mP, unbound), TOP is maximum polarization (mP, fully bound), K_D is the dissociation constant (μM), and $[4c]$ is total $4c$ concentration, constrained to $0.02 \mu\text{M}$.

To note, the concentration of foldamer was determined by UV on a NanoDrop instrument (Thermo Fisher) at $\lambda=375 \text{ nm}$ with an extinction coefficient (ϵ) of 32124 L/mol/cm .

8.2.4 Materials and methods for chemical synthesis and characterizations

Fmoc-Q-OH, Fmoc-Q^D-OH, Fmoc-Q^L-OH, Fmoc-Q^O-OH, Fmoc-Q^U-OH, Fmoc-Q^F-OH were synthesized by following the reported protocols.⁴ The synthesis of Fmoc-Q^{Y(Bu)}-OH and Fmoc-Q^F-OH are introduced in chapter 3.1. If not otherwise mentioned, chemical reagents were purchased from Sigma-Aldrich, and solvents from Fisher Scientific and used without further purification. Anhydrous tetrahydrofuran (THF) and anhydrous dichloromethane (DCM) were obtained from MBRAUN SPS-800 solvent purification system. Anhydrous chloroform (CHCl_3) and *N,N*-diisopropylethylamine (DIPEA) were distilled over CaH_2 prior to use. Exclusively ultrapure water was used. DMF and NMP (peptide grade) were purchased from Carlo Erba. Cl-MPA ProTide®, and low-loading Wang resins were purchased from CEM. Fmoc-N-protected amino acids, benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) and O-(1H-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HCTU) were purchased from IRIS. ¹H NMR spectra were recorded on Avance III HD 400 MHz Bruker BioSpin and Avance III HD 500 MHz Bruker BioSpin spectrometers. All chemical shifts (δ) are reported in ppm and calibrated against residual solvent signals of $\text{DMSO-}d_6$ (δ 2.50 ppm) and CDCl_3 (δ 7.26 ppm). High-resolution electrospray mass spectra for compounds **3-8** were recorded on a Thermo Finnigan LTQ FT Ultra Fourier Transform Ion Cyclotron Resonance Mass Spectrometer by direct infusion of the analyte dissolved in either DCM or aqueous media in positive or negative ionization mode. Mass spectra for foldamer **1a-1e**, **2a-2f** were recorded on a Bruker microTOF II from aqueous media in positive ionization mode. RP-HPLC analyses, as well as semi-preparative purification, were performed on an Ultimate 3000 HPLC System (ThermoFisher Scientific). Preparative RP-HPLC purification was performed on a Waters system with a 2707 Autosampler, a 2489 UV/Visible detector, a 2545 Quaternary Gradient Module and a Fraction Collector III. For analytical analysis, a Nucleodur C18 Gravity column ($4 \times 100 \text{ mm}$, $5 \mu\text{m}$, Macherey-Nagel) was used, and semi-preparative purifications were performed on a Nucleodur C18 EC column ($10 \times 250 \text{ mm}$, $5 \mu\text{m}$, Macherey-Nagel). When using acidic conditions 0.1% (v/v) TFA was added to the aqueous mobile phase (referred to as mobile phase A) and to acetonitrile (referred to as mobile phase B). For analytical RP-HPLC analysis, a flow rate of $1.0 \text{ mL}\cdot\text{min}^{-1}$ was applied, semipreparative

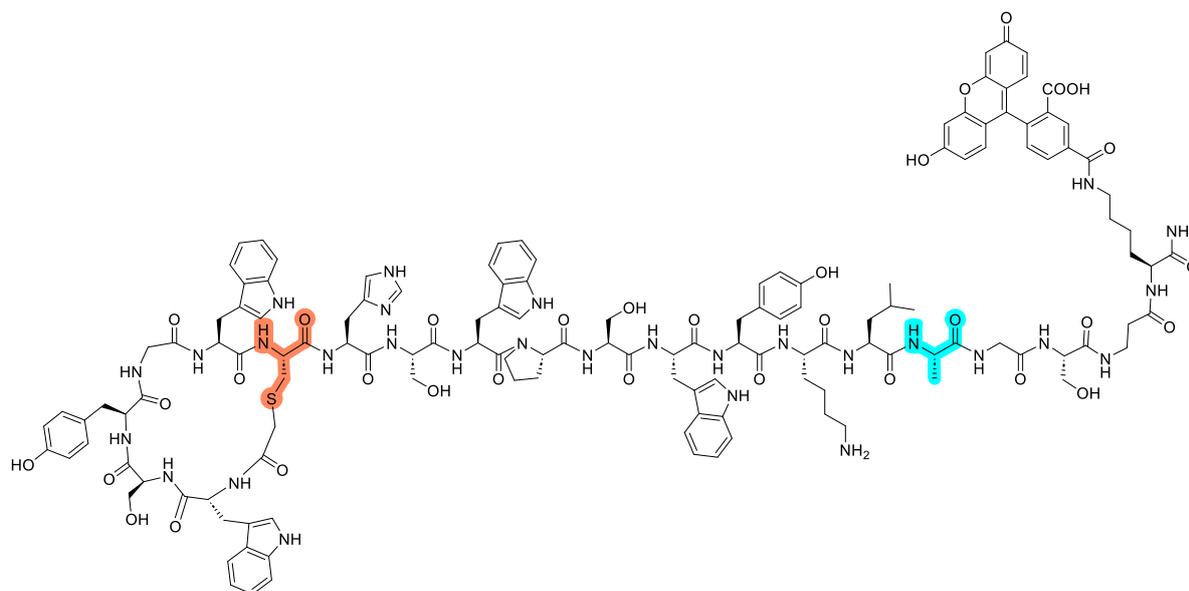
purification on RP-HPLC was performed at a flow rate of either 5.0 or 25 mL.min⁻¹ (on the Waters preparative instrument). The column eluent was monitored by UV detection at 214, 254, and/or 300 nm with a diode array detector.

The CD spectra of foldamers were recorded on a Jasco J-1500 spectrometer with 2 mm quartz cuvette. The following parameters were used: wavelength range from 450 to 250 nm. Scan speed: 200 nm/min; accumulation: 2; response time: 1.0 s; bandwidth: 2; temperature: 25°C; sensitivity: standard (100 mdeg); data pitch: 0.5 nm; nitrogen gas flow rate: 500 L/h. $\Delta\epsilon$ values (in M⁻¹.cm⁻¹) were obtained by using the formula: $\Delta\epsilon = m^\circ / (C \cdot l \cdot 32980)$ where m° = CD value in millidegrees; l = cuvette pathlength in cm; C = sample concentration in mol/L

8.2.4.1 Solid phase synthesis

The peptides were assembled by using a Liberty Blue CEM® synthesizer at a scale of 50 μ mol, using Fmoc-Gly-Wang PS resin. Microwave couplings were performed twice at 50 °C for 10 min with N-Fmoc-amino acid (6 equiv. relative to the resin loading), HCTU (6 equiv.), and NMM (12 equiv.) in DMF. Fmoc deprotection was performed twice with 20% piperidine in DMF at 75 °C (1 \times 30 sec. and 1 \times 180 sec.). The resin was washed with DMF (2 \times 2 mL) after each deprotection step and one time 3 mL after each coupling step. The procedure of N-terminal chloroacetylation, TFA cleavage and cyclization follow recently reported protocols.⁵ The crude peptide macrocycles were purified by semi-prep RP-HPLC to yield the pure compound.

SPPS of fluorescently-labelled 4c



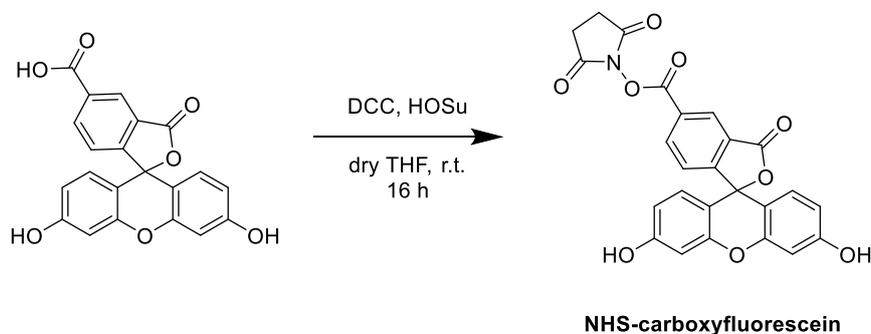
The resin-bound **Fmoc-4a-Ser-βAla-Lys(Alloc)** was synthesized on the Liberty Blue CEM® synthesizer at a scale of 50 μmol using a low-loading Rink amide MBHA resin. Fmoc-Lys(Alloc)-OH amino acid was first loaded using the same HCTU/NMM procedure as reported above. After the Lys(Alloc) coupling, a Fmoc-βAla-OH was coupled and served as a spacer. All the remaining amino acid residues, including the Fmoc-(*D*)-Trp(Boc)-OH were installed using the same coupling/deprotection cycle procedure.

Half of the resin (25 μmol) was then transferred in a syringe equipped with a filter, the Fmoc deprotection was performed manually at room temperature in the presence of 20% piperidine in DMF (1 × 3 min and 1 × 7 min). After several rounds of washings with DMF (3 × 3 mL) and then DCM (3 × 3 mL), chloroacetic anhydride (10 equiv.) was dissolved together with DIPEA (20 equiv.) in DCM (3 mL). The reaction mixture was directly added to the resin and shaken for 15 minutes at r.t. This coupling step was repeated once without any washing in between. The resin was then filtered off, washed with DCM (3 × 3 mL) and dried briefly under a nitrogen stream.

In a second time, the ε-Alloc protecting group was removed in the presence of PdP(Ph₃)₄ (0.1 equiv.) and Ph₃SiH (20 equiv.) in dry DCM (2mL) for 30 min under Ar atmosphere. This step was repeated once with washings with dry DCM (3 × 3 mL) in between. The deprotection of the Lys side-chain was qualitatively controlled with the TNBST test.

At last, the fluorescein was installed on the ε-NH₂ of the Lys residue. Freshly prepared NHS-carboxyfluorescein (3 equiv., see below) was added to the resin swollen in DMF (1.5 mL). The resin was next shaken for 16 hours. After washings with DMF (3 × 3 mL), the efficiency of the fluorescein coupling was monitored with the TNBS test, which proved to be negative.

After final TFA cleavage (TFA, TIS, H₂O, EDT / 92.5, 2.5, 2.5, 2.5, v/v/v), half of the crude **4c** (42.67 mg, 60%) was purified by using semi-preparative RP-HPLC to give **4c** as a yellow solid after lyophilization (1.34 mg, 3.7%). HRMS (ESI⁺): *m/z* calculated for C₁₃₉H₁₆₁N₂₉O₃₃S [M+2H]²⁺ 1399.0839 found 1399.1524.



NHS-carboxyfluorescein: 5-Carboxyfluorescein (250 mg, 0.66 mmol) was suspended in dry THF (2.5 mL) and a solution of *N*-hydroxysuccinimide (1.3 equiv., 99 mg, 0.86 mmol) dissolved in dry THF (0.5 mL) was added, followed by the addition of a solution of *N,N'*-dicyclohexylcarbodiimide (1.3 equiv., 177 mg, 0.86 mmol) in dry THF (0.5 mL). After 16 hours stirring at r.t. the mixture was filtered and the precipitate washed with Et₂O (2 × 10 mL) and EtOAc (1 × 10 mL). The filtrate was concentrated under reduced pressure and remaining solvents azeotroped with toluene (3 × 5 mL) providing **NHS-carboxyfluorescein** as an orange solid (270 mg, 90%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 10.18 (s, 2H), 8.54 (d, *J* = 1.7 Hz, 1H), 8.43 (dd, *J* = 8.1, 1.6 Hz, 1H), 7.56 (d, *J* = 8.1 Hz, 1H), 6.71 (s, 1H), 6.70–6.68 (m, 3H), 6.55 (d, *J* = 2.3 Hz, 1H), 6.53 (d, *J* = 2.3 Hz, 1H), 2.93 (s, 4H). HRMS (ESI⁺): *m/z* calculated for C₂₅H₁₅NO₉ [M+H]⁺ 474.0820 found 474.1079.

8.2.4.2 Solid Phase Foldamer Synthesis (SPFS)

The microwave-assisted solid phase synthesis of aromatic oligoamide foldamers **1a**, **2a** and **5** was carried out on a Discover Bio CEM® microwave oven in an opened vessel mode manually. The temperature of the reaction mixture within the reactor vessel was monitored with an optical fiber probe. The LL-Wang resin was first brominated, and the Fmoc-Q-OH unit was loaded using the CsI assisted reaction. The efficiency of the first quinoline monomer loading was determined by UV-dosing the dibenzofulvene-piperidine adduct at 301 nm with an ε = 7800 L/mol/cm: 80% (30.4 mmol/g, 60 μmol).⁶ The following Fmoc-Q/B units and Fmoc-3-Aminobenzoic acid were coupled with the *in-situ* activation protocol.⁷ N-terminal Biotinylation was performed on the resin-bound H-Amb-Q₁₂mer. The resin was suspended in 1 mL DMF (15 μmol scale) in a syringe equipped with a filter. Biotin-Peg-OH (2 equiv.), PyBOP (2.1 equiv.) were dissolved in another 1 mL DMF, followed by DIPEA (4 equiv.), then the solution was transferred to the resin and shaken overnight (reaction can be monitored by TNBS test to check whether all free amines have been consumed). The introduction of the PEG tail in **2e** and **2f** followed the same procedure.

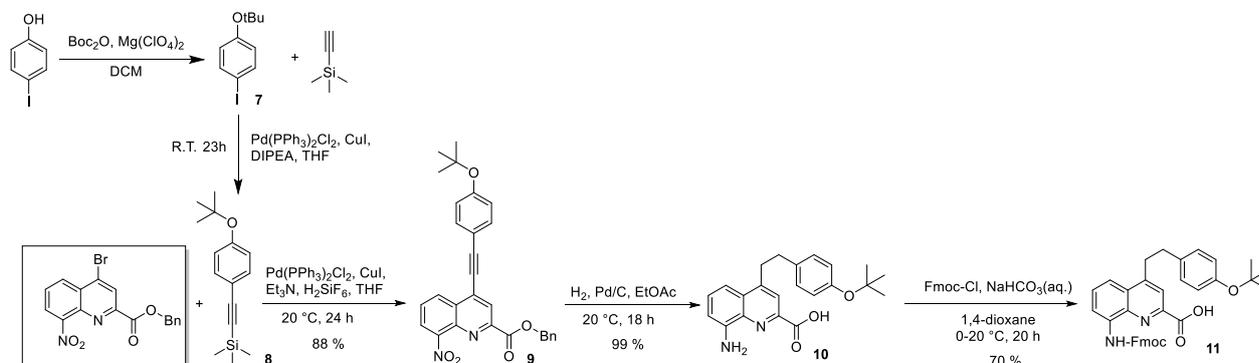
Foldamers **1b-1e** and **2b-2f** were prepared using the PurePep® Chorus synthesizer from Protein Technologies with LL-Wang resin or Cl-MPA ProTide® resin. The first loading on Cl-MPA ProTide® resin followed the recently published procedure⁸: Fmoc-Q^{D(O^tBu)}-OH or Fmoc-Q^{O(Boc)}-OH monomer (3.0 equiv.) was dissolved dry DMF and added to the resin together with a solution containing CsI (3.0 equiv.) and DIPEA (7.65 equiv.). The resin suspension was shaken overnight before to be washed first with DMF then with DCM, followed by loading determination.⁷ After automated SPFS⁶ and biotin moiety coupling, the foldamer was cleaved from the resin and deprotected with a solution of TFA/H₂O/TIS (95:2.5:2.5, v/v/v). The crude was then lyophilised before to be purified by semi-prep RP-HPLC to furnish the desired foldamer in high purity (> 95%).

Of particular note, some RP-HPLC chromatogram show an additional peak, upfront which corresponds to the oxidized-biotin foldamer conjugates. This biotin oxidation and its percentage varies with the foldamer sequence. In our hands, this biotin oxidation appeared to be inconsequential on the foldamer ligand loading to SA sensor tips. The oxidized-biotin foldamer conjugate is annotated with a red star on the RP-HPLC chromatograms.

8.3 Experimental procedures for chemical synthesis

8.3.1 Synthesis of Fmoc-Q^Y-OH

8.3.1.1 Scheme S1. Synthesis route of Fmoc-Q^{Y(tBu)}-OH (12)



Compound 7: This compound was prepared from the reported method.⁹ 4-Iodophenol (6.0 g, 27.3 mmol, 1.0 equiv.) was dissolved in anhydrous CH₂Cl₂ (60 mL) and Mg(ClO₄)₂ (1.22 g, 5.5 mmol, 0.2 equiv.) was added. Boc₂O (13.7 g, 62.7 mmol, 2.3 equiv.) was dissolved in CH₂Cl₂ (8 mL) and added dropwise to the first solution. The reaction was stirred under nitrogen atmosphere at room temperature for 16 hours. The reaction mixture was then washed with water (2 × 100 mL) followed by aqueous NaOH (2M, 2 × 100 mL). The organic phase was dried over Na₂SO₄ and concentrated *in vacuo* to yield compound 4 as oil (4.9 g, 60%). ¹H NMR (500 MHz, DMSO-*d*₆) δ 7.61-7.58 (m, 2H), 6.82-6.78 (m, 2H), 1.28 (s, 9H). ¹³C NMR (126 MHz, DMSO-*d*₆) δ 155.0, 137.7, 126.1, 86.9, 78.5, 28.4. Spectral data matched literature data.¹⁰

Compound 8: This compound was prepared from the reported method.¹¹ To a dry nitrogen-flushed Schlenk flask, 7 (2.0 g, 7.24 mmol, 1.0 equiv.), Pd(PPh₃)₂Cl₂ (51 mg, 72.40 μmol, 1.0 mol%), CuI (28 mg, 144.80 μmol, 2.0 mol%), anhydrous THF (40 mL), and anhydrous DIPEA (2.5 mL, 14.5 mmol, 2.0 equiv.) were added and degassed using the freeze-pump-thaw method of three cycles and finally back-flushed with nitrogen. Trimethylsilylacetylene (1.24 mL, 8.69 mmol, 1.2 equiv.) was added, and the reaction was stirred under nitrogen atmosphere for 23 hours. CH₂Cl₂ (50 mL) was added to the reaction mixture and the mixture was washed with water (2 × 100 mL) and brine (2 × 100 mL). The organic phase was dried over Na₂SO₄ and concentrated *in vacuo*. The product was purified by filtration over a plug of silica using CH₂Cl₂ as the eluent, which furnished compound 8 as an oil after concentration of the filtrate (1.6 g, 90%). ¹H NMR (500 MHz, DMSO-*d*₆) δ 7.39-7.32 (m, 2H), 6.99-6.92 (m, 2H), 1.31 (s, 9H), 0.21 (s, 9H). ¹³C NMR (126 MHz, DMSO-*d*₆) δ 155.9, 132.6, 123.3, 116.4, 105.2, 93.1, 78.8, 28.5, -0.03. Spectral data matched literature data.¹¹

Compound 9: To a dry nitrogen-flushed Schlenk flask, compound 8 (2.09 g, 5.41 mmol, 1.0 equiv.), Pd(PPh₃)₂Cl₂ (57 mg, 54.1 μmol, 1.0 mol%), CuI (31 mg, 0.11 mmol, 2.0 mol%), 4-bromo-8-nitro-

phenylmethyl ester (1.6 g, 6.49 mmol, 1.2 equiv.)^{4c} were dissolved in anhydrous THF (100 mL), and Et₃N (16.21 mL, 116 mmol, 20 equiv.) was added. The reaction mixture was immediately after degassed using the freeze-pump-thaw method of three cycles and finally back-flushed with nitrogen. Hexafluorosilicic acid (32% aq., 0.91 mL, 2.7 mmol, 0.5 equiv.) was added, and the reaction was stirred under nitrogen atmosphere for 24 hours. CH₂Cl₂ (80 mL) was added to the reaction mixture, and it was washed with citric acid (5% aq., 3 × 75 mL) and brine (2 × 100 mL) and the organic phase was dried over Na₂SO₄ and concentrated *in vacuo*. The product was purified by flash column chromatography using CH₂Cl₂ as the eluent. Compound **9** was obtained as a yellow solid (2.3 g, 88.5%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 8.70 (dd, *J* = 8.5, 1.3 Hz, 1H), 8.46 (dd, *J* = 7.5, 1.3 Hz, 1H), 8.36 (s, 1H), 8.01 (dd, *J* = 8.5, 7.5, Hz, 1H), 7.80-7.72 (m, 2H), 7.57-7.49 (m, 2H), 7.49-7.31 (m, 3H), 7.17-7.08 (m, 2H), 5.48 (s, 2H), 1.38 (s, 9H). ¹³C NMR (101 MHz, DMSO-*d*₆) δ 163.4, 157.4, 149.2, 148.7, 138.0, 135.6, 133.6, 131.3, 129.6, 129.0, 128.5, 128.3, 128.3, 128.1, 124.9, 124.4, 123.0, 114.4, 101.8, 83.5, 79.3, 67.2, 28.5. HRMS (ESI⁺): calcd. for C₂₉H₂₅N₂O₅ [M+H]⁺ 481.1758 found 481.1878.

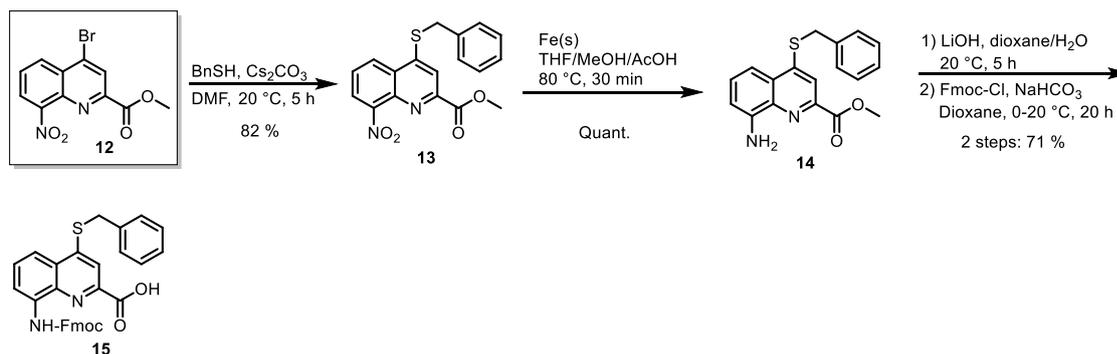
Compound 10: Compound **9** (1.46 g, 3.04 mmol, 1.0 equiv.) was dissolved in EtOAc (210 mL) containing DMF (10 mL). Pd/C (140 mg, 10% w/w) was added to the solution and the mixture was degassed for 15 min (with nitrogen balloon in an ultra sound bath), and finally the flask was backflushed with H₂. The reaction mixture was stirred under H₂-atmosphere at room temperature for 17 hours. The reaction mixture was then filtered over a pad of celite, which was washed several times with EtOAc. The filtrate was concentrated *in vacuo* (co-evaporation with toluene to remove DMF) to furnish **10** as a yellow solid (1.04 g, 95%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 12.72 (s, 1H), 7.85 (s, 1H), 7.43 (dd, *J* = 8.4, 7.7 Hz, 1H), 7.26 (dd, *J* = 8.4, 1.2 Hz, 1H), 7.15-7.12 (m, 2H), 6.88 (dd, *J* = 7.7, 1.2 Hz, 1H), 6.87-6.83 (m, 2H), 6.57 (s, 2H), 3.32 (dd, *J* = 9.4, 6.5 Hz, 2H), 2.95 (dd, *J* = 9.4, 6.5 Hz, 2H), 1.25 (s, 9H). ¹³C NMR (126 MHz, DMSO-*d*₆) δ 165.7, 153.1, 149.0, 147.2, 142.4, 135.7, 135.6, 130.3, 129.1, 128.9, 123.7, 119.5, 108.8, 108.7, 77.6, 34.7, 33.9, 28.5. HRMS (ESI⁺): calcd for C₂₂H₂₅N₂O₃ [M+H]⁺ 365.1859 found 365.1979.

Compound 11: Compound **10** (2.4 g, 6.59 mmol, 1.0 equiv.) was dissolved in 1,4-dioxane (80 mL) and aqueous NaHCO₃ (10%, 116 mL) was added. The solution was cooled down to 0 °C and a solution of Fmoc-Cl (2.21 g, 8.56 mmol, 1.3 equiv.) in 1,4-dioxane (90 mL) was added dropwise over 1 hour. The reaction mixture was stirred at 0 °C for another hour, and then at room temperature for 16 hours. The reaction was quenched with 1 M HCl to pH around 2. CH₂Cl₂ (150 mL) was added, and the phases were separated. The aqueous phase was extracted with CH₂Cl₂ (2 × 100 mL), and the organic phases were combined, dried over Na₂SO₄ and concentrated under reduced pressure. Final purification was done by silica gel column chromatography twice with an eluent of CH₂Cl₂/MeOH 100:0 to 95:5. Compound **11** was isolated as a green foam (2.7 g, 70%). ¹H NMR (400 MHz, DMSO-*d*₆) δ 13.58 (bs, 1H), 10.39 (s, 1H), 8.32 (bs, 1H), 8.01 (s, 1H), 7.93 (d, *J* = 7.5 Hz, 2H), 7.86 (dd, *J* = 8.5, 1.1 Hz, 1H), 7.77 (d, *J* = 7.5 Hz, 2H), 7.64 (t, *J* = 8.5, 1H), 7.49-7.40 (m, 2H), 7.38-7.34 (m, 2H), 7.19-7.08 (m, 2H), 6.88-6.79

(m, 2H), 4.61 (d, $J = 6.8$ Hz, 2H), 4.44 (t, $J = 6.8$ Hz, 1H), 3.43 (dd, $J = 9.2, 6.6$ Hz, 2H), 2.96 (dd, $J = 9.2, 6.6$ Hz, 2H), 1.25 (s, 9H). ^{13}C NMR (126 MHz, DMSO- d_6) δ 166.1, 153.3, 153.2, 149.9, 143.7, 140.8, 136.7, 136.0, 135.4, 128.9, 128.8, 128.1, 127.8, 127.2, 125.1, 123.7, 120.7, 120.2, 117.0, 115.4, 77.6, 66.4, 46.6, 35.0, 33.6, 28.5. HRMS (ESI $^+$): calcd for $\text{C}_{37}\text{H}_{35}\text{N}_2\text{O}_5$ $[\text{M}+\text{H}^+]^+$ 587.2540 found 587.2547.

8.3.2 Synthesis of Fmoc-Q^F-OH

8.3.2.1 Scheme S2. Synthesis route of Fmoc-Q^F-OH (15)



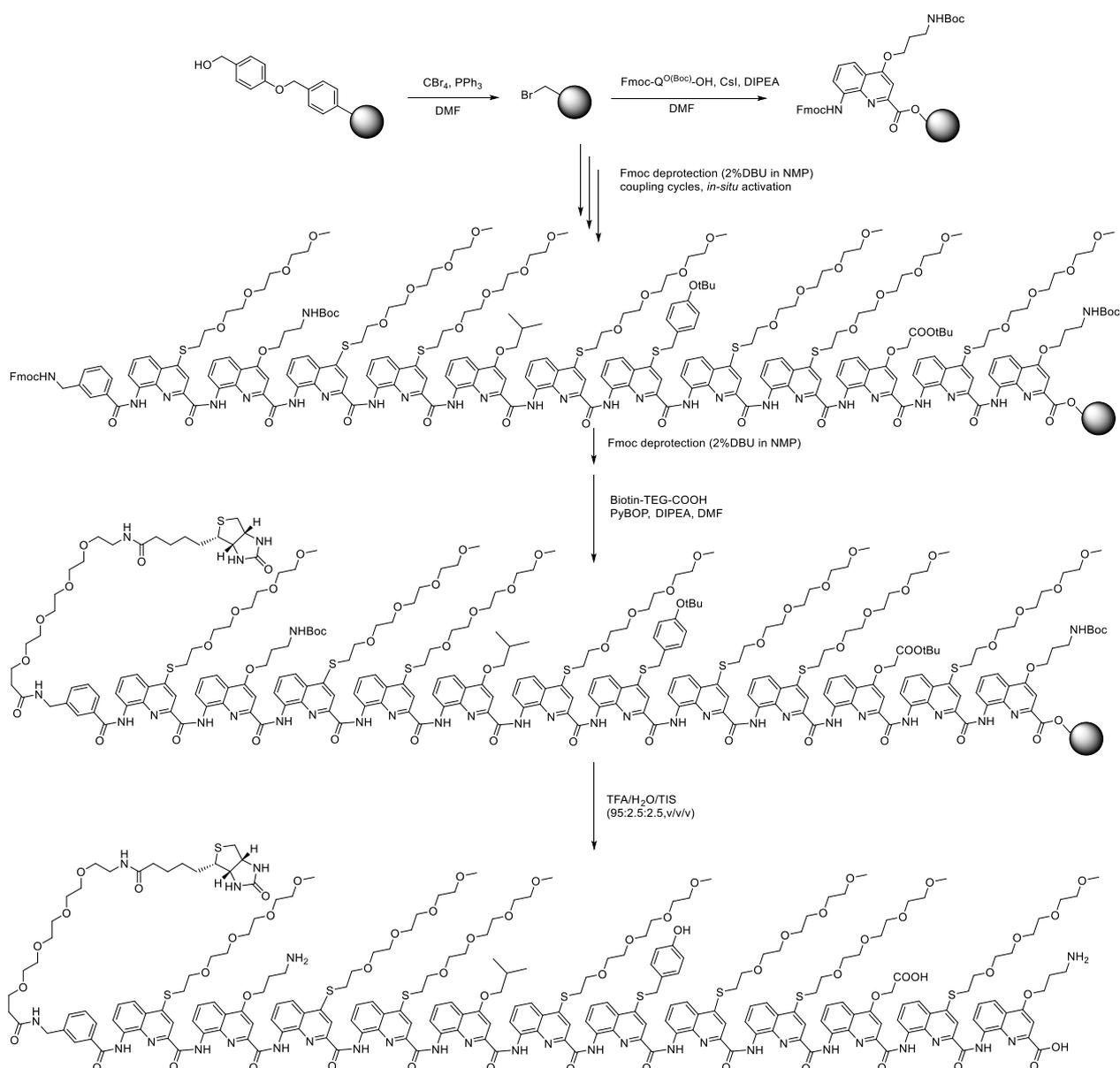
Compound 13: Compound **12** was synthesized according to the reported protocol.^{4b} **12** (6.0 g, 19.3 mmol, 1 equiv.) was then suspended in dry DMF (90 mL) and Cs₂CO₃ (9.43 g, 8.9 mmol, 0.5 equiv.) were added. While stirring under N₂, benzylthiol was added dropwise (2.15 mL, 18.3 mmol, 1 equiv.). The reaction mixture was heated to 55 °C for 5 h and cooled down to room temperature. EtOAc (150 mL) was added to the reaction mixture and the solution was washed with brine (3 × 100 mL). The organic layer was dried over MgSO₄, and concentrated under vacuum. The remaining solid was recrystallized from CH₂Cl₂/Et₂O and the **13** was isolated by filtration and washed with cold ether. Yield: 5.3 g (82%) ^1H NMR (500 MHz, CDCl₃) δ 8.35 (dd, $J = 8.6, 1.3$ Hz, 1H), 8.16 (s, 1H), 8.07 (dd, $J = 7.5, 1.3$ Hz, 1H), 7.68 (dd, $J = 8.5, 7.5$ Hz, 1H), 7.50 – 7.44 (m, 2H), 7.44 – 7.34 (m, 2H), 7.36 – 7.29 (m, 1H), 4.45 (s, 2H), 4.03 (s, 3H). ^{13}C NMR (126 MHz, CDCl₃) δ 165.42, 150.48, 149.15, 148.71, 138.54, 134.12, 129.11, 129.03, 128.26, 127.79, 127.40, 126.77, 124.62, 117.34, 53.42, 36.50. HRMS (ESI $^+$): calcd for $\text{C}_{18}\text{H}_{15}\text{N}_2\text{O}_4\text{S}$ $[\text{M}+\text{H}^+]^+$ 355.0747, found 355.0737.

Compound 14: Compound **13** (3.24 g, 9.14 mmol, 1 equiv.) was suspended in a solvent mixture composed of THF (100 mL), MeOH (95 mL), and AcOH (61 mL). The reaction mixture was heated up to 80 °C and Fe (2.55 g, 45.7 mmol, 5 equiv.) was added portionwise. After stirring at 80 °C for 30 min, the reaction mixture was let to cool down to r.t., and the yellow precipitate (Fe(CH₃CO₂)₂) was removed by filtration and washed with CH₂Cl₂. The washing and filtrate were combined and concentrated *in vacuo*. The crude product was purified by filtration over a plug of silica (eluent: CH₂Cl₂/MeOH increasing from 0 to 10 %) and **14** was recovered quantitatively (3 g). ^1H NMR (500 MHz, DMSO- d_6) δ 7.96 (s, 1H), 7.58 – 7.46 (m, 2H), 7.41 (dd, $J = 8.3, 7.7$ Hz, 1H), 7.40 – 7.31 (m, 2H), 7.34 – 7.26 (m, 1H), 7.14 (dd, $J = 8.3, 1.2$ Hz, 1H), 6.94 (dd, $J = 7.7, 1.2$ Hz, 1H), 6.16 (s, 2H), 4.54 (s, 2H), 3.94 (s,

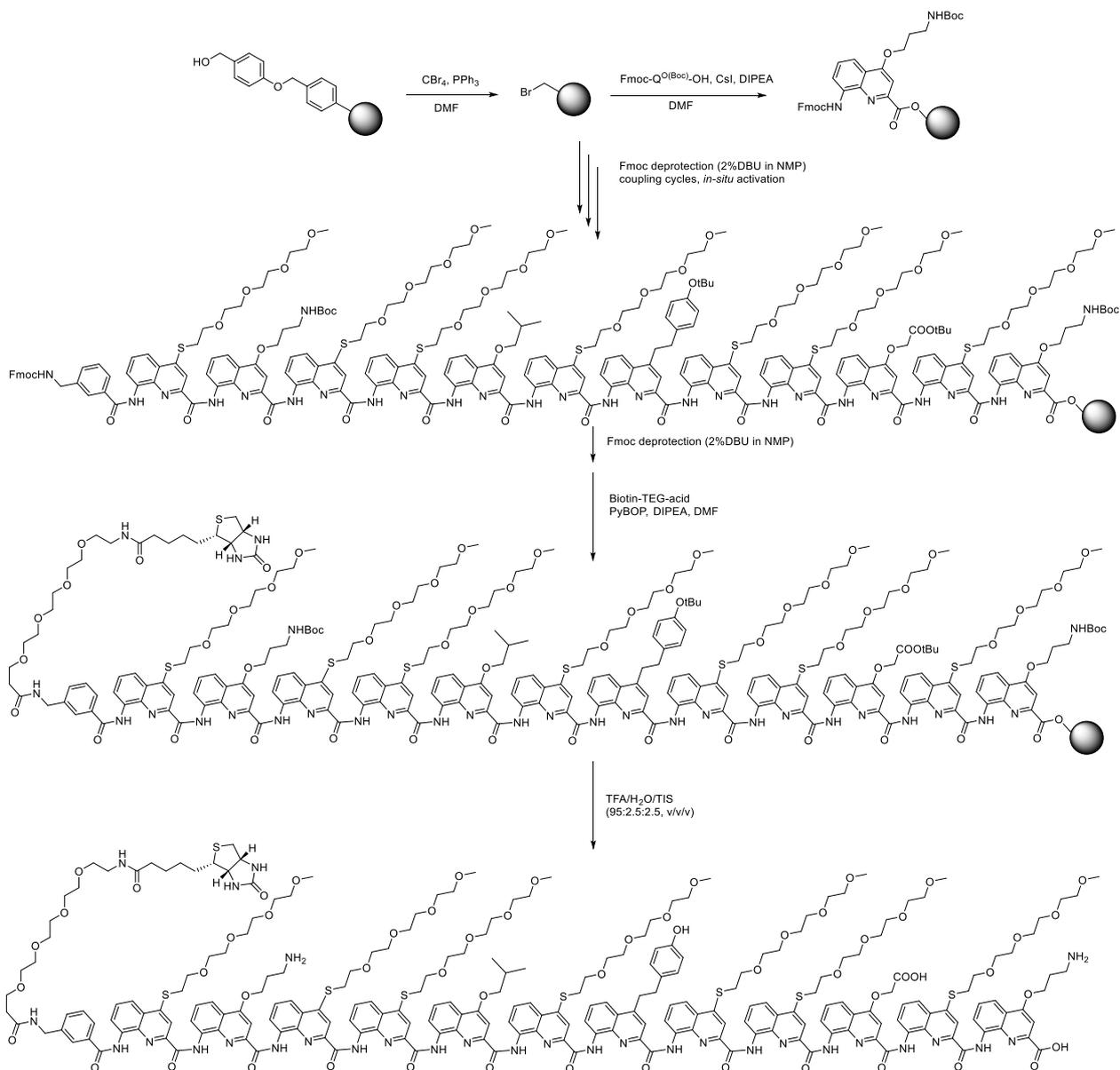
3H), 1.22 (s, 1H). ^{13}C NMR (126 MHz, DMSO- d_6) δ 165.75, 148.10, 147.38, 143.29, 136.19, 135.52, 130.68, 129.61, 129.11, 128.03, 127.40, 116.03, 110.21, 108.79, 53.07, 35.03. HRMS (ESI $^+$): calcd. for $\text{C}_{18}\text{H}_{17}\text{N}_2\text{O}_2\text{S}$ $[\text{M}+\text{H}^+]^+$ 325.1005, found 325.1005.

Compound 13: Compound **12** (3.4 g, 10.6 mmol, 1 equiv.) was dissolved in 1,4-dioxane (280 mL) and a solution of LiOH (0.40 g, 15.9 mmol, 1.5 equiv.) in water (70 mL) was added and the mixture was stirred for 5 hours. The reaction mixture was neutralized by dropwise addition of HCl (1M). Aqueous NaHCO_3 (10% v/v, 187 mL) was added to the reaction mixture, and the solution was cooled to 0 °C. A solution of Fmoc-Cl (3.6 g, 13.8 mmol, 1.3 equiv.) in 1,4-dioxane (75 mL) was prepared and added dropwise over 1 h. Afterwards the reaction was stirred at 0 °C for another 1 hour, and then at r.t. for 20 hours. The mixture was acidified by slow addition of aqueous HCl (1M, approx. 270 mL) and afterwards extracted with CH_2Cl_2 (2 x 300 mL). The organic phases were dried over MgSO_4 and concentrated *in vacuo*. The crude product was purified by flash column chromatography (gradient: CH_2Cl_2 : MeOH 100:0 to CH_2Cl_2 : MeOH 90:10). The fractions containing **13** were collected and concentrated and recrystallized from $\text{CH}_2\text{Cl}_2/\text{Et}_2\text{O}$, to furnish **13** in 71 % yield (4.0 g). ^1H NMR (400 MHz, DMSO- d_6) δ 13.51 (s, 1H), 10.43 (s, 1H), 8.34 (s, 1H), 8.12 (d, J = 1.2 Hz, 1H), 7.92 (d, J = 7.5 Hz, 2H), 7.77 (d, J = 7.5 Hz, 2H), 7.72 (dd, J = 8.5, 1.9 Hz, 1H), 7.65 (t, J = 8.1 Hz, 1H), 7.53 (dd, J = 7.9, 1.6 Hz, 2H), 7.47 – 7.26 (m, 8H), 4.64 – 4.58 (m, 4H), 4.44 (t, J = 6.8 Hz, 1H). ^{13}C NMR (101 MHz, DMSO- d_6) δ 165.76, 153.92, 150.28, 144.16, 141.29, 136.81, 136.09, 135.84, 129.89, 129.64, 129.16, 128.26, 128.13, 127.70, 126.86, 125.62, 120.72, 116.97, 116.50, 115.82, 66.93, 47.05, 35.12. HRMS (ESI $^+$): calcd. for $\text{C}_{32}\text{H}_{25}\text{N}_2\text{O}_4\text{S}$ $[\text{M}+\text{H}^+]^+$ 533.1530 found 533.1531.

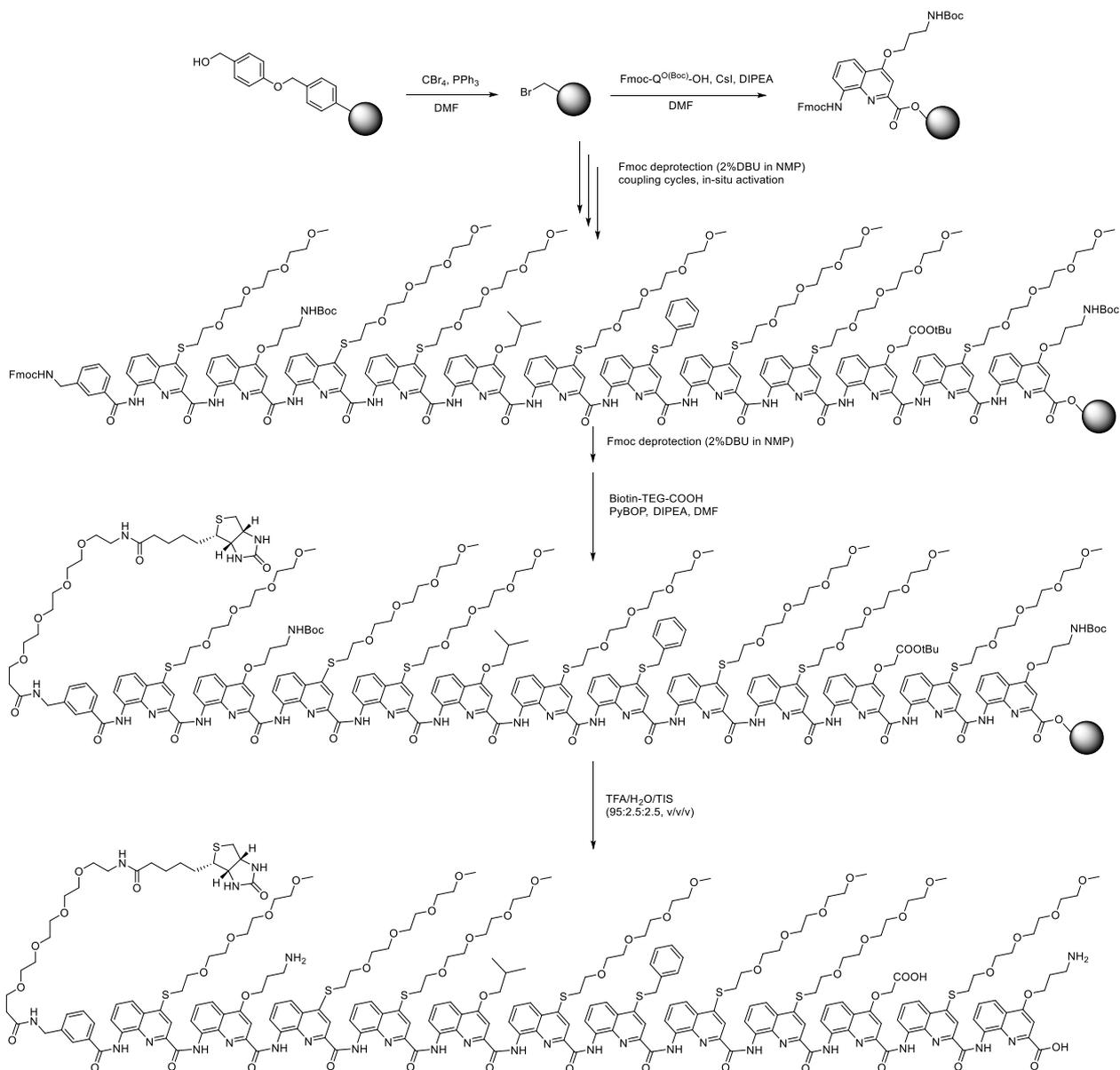
8.3.3 SPFS of compounds 1-5



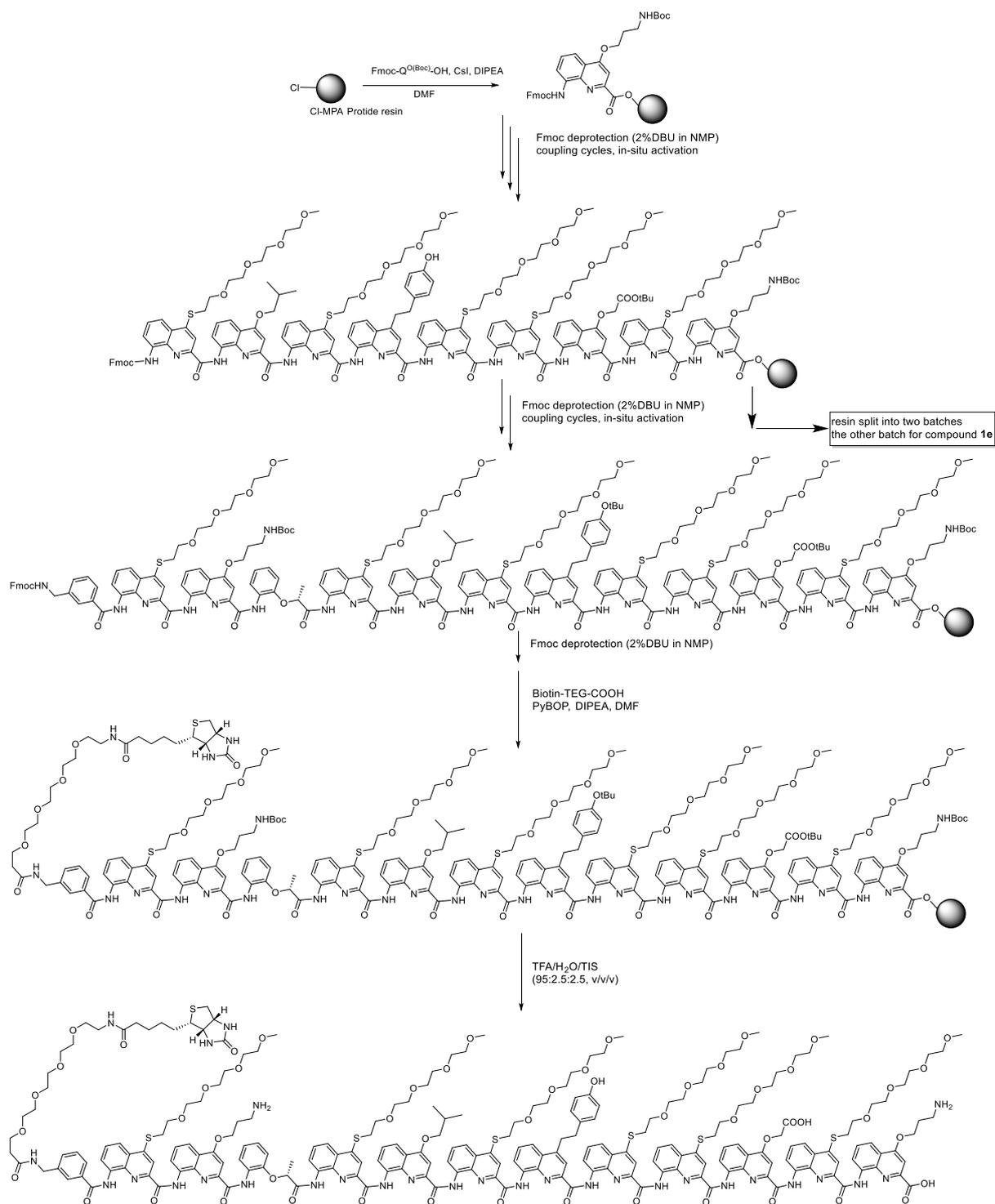
Compound 1a: Foldamer **1a** was synthesized on a low loading LL-Wang resin (19 μmol scale). After TFA cleavage and side chain deprotection, the crude foldamer was purified by semi prep RP-HPLC to furnish **1a** as yellow solid (1.7 mg, 1.9%). HRMS (ESI⁺): m/z calcd. for $\text{C}_{231}\text{H}_{272}\text{N}_{30}\text{O}_{56}\text{S}_9$ $[\text{M}+3\text{H}]^{3+}$ 1551.9055, found 1551.9388.



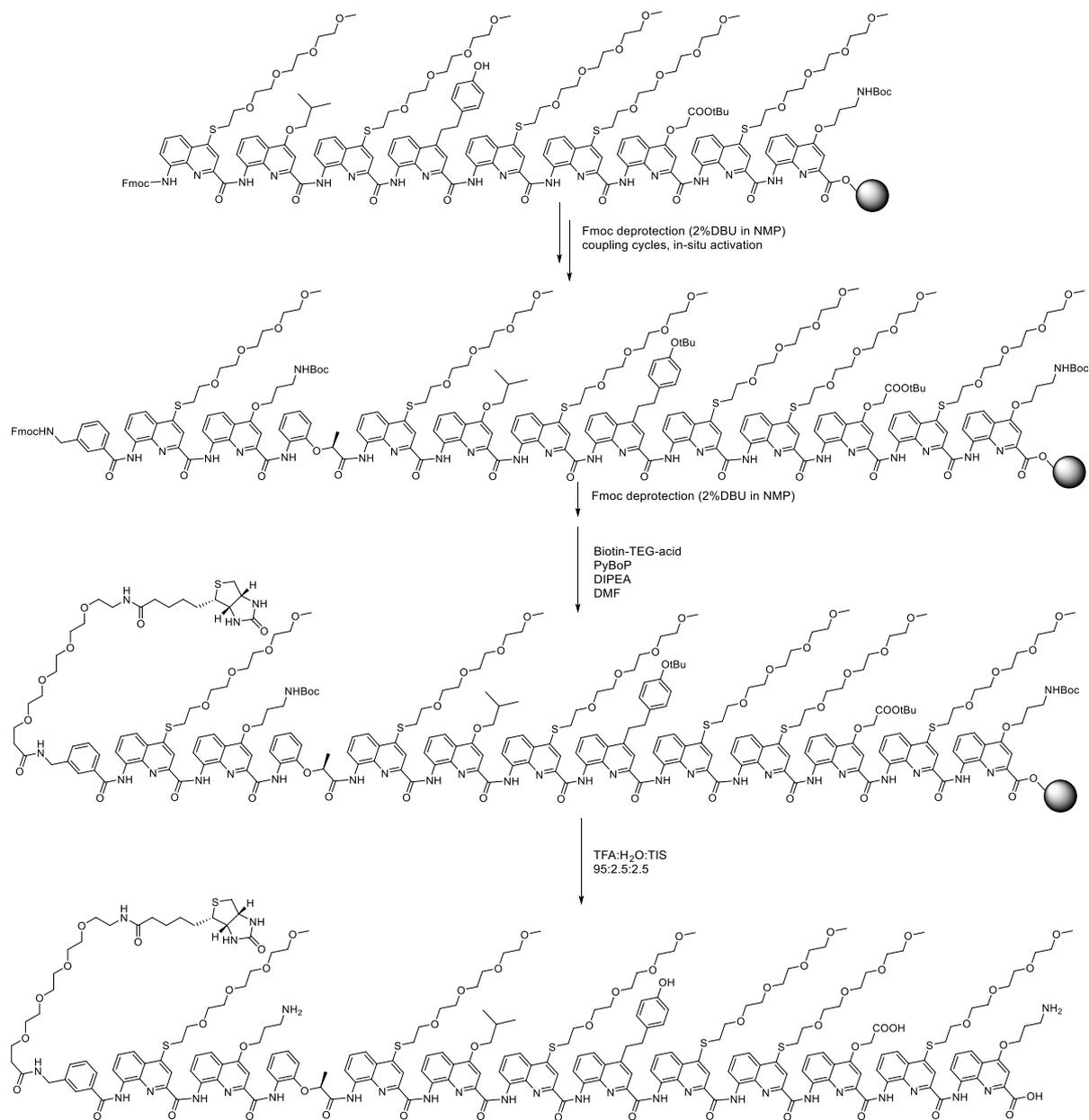
Compound 1b: Foldamer **1b** was synthesized on a low loading LL-Wang resin (15 μmol scale). After TFA cleavage and side chain deprotection, 14 mg of crude foldamer were recovered. The crude was then purified by semi prep RP-HPLC to furnish **1b** as an yellow solid (5 mg, 12%). HRMS (ESI⁺): m/z calcd. for $\text{C}_{232}\text{H}_{274}\text{N}_{30}\text{O}_{56}\text{S}_8$ $[\text{M}+2\text{H}]^{2+}$ 2318.3322, found 2318.3337.



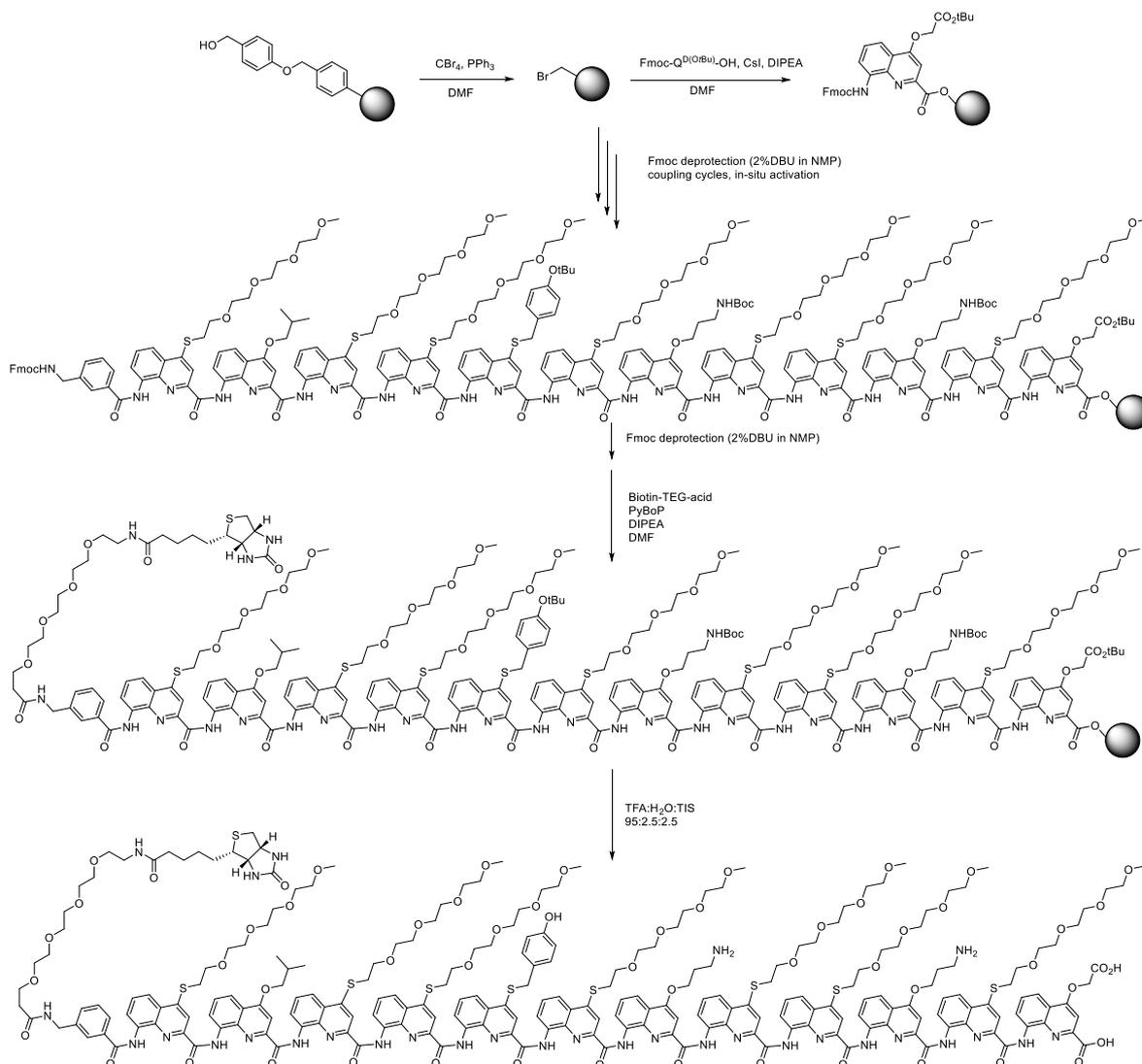
Compound 1c: Foldamer **1c** was synthesized on a LL-Wang resin (15 μmol scale). After TFA cleavage and side chain deprotection, 21 mg of crude foldamer was obtained. The crude was purified by semi prep RP-HPLC to furnish **1c** as an yellow solid (2 mg, 4.3%). HRMS (ESI⁺): m/z calcd. for $\text{C}_{231}\text{H}_{272}\text{N}_{30}\text{O}_{55}\text{S}_9$ $[\text{M}+2\text{H}]^{2+}$ 2319.3554, found 2319.3642.



Compound 1d: Foldamer **1d** was synthesized on a Cl-MPA ProTide® resin (20 μ mol scale). After the 9th quinoline coupling, the resin was divided into two batches. The other 7.5 μ mol was used for synthesizing **1e**. After TFA cleavage and side chain deprotection, 30 mg of crude foldamer was obtained. The crude was purified by semi prep RP-HPLC to furnish **1d** as yellow solid (7.4 mg, 22%). HRMS (ESI⁺): calcd. for C₂₂₂H₂₅₉N₂₉O₅₃S₇ [M+3H]³⁺ 1468.5575 found 1468.5728.

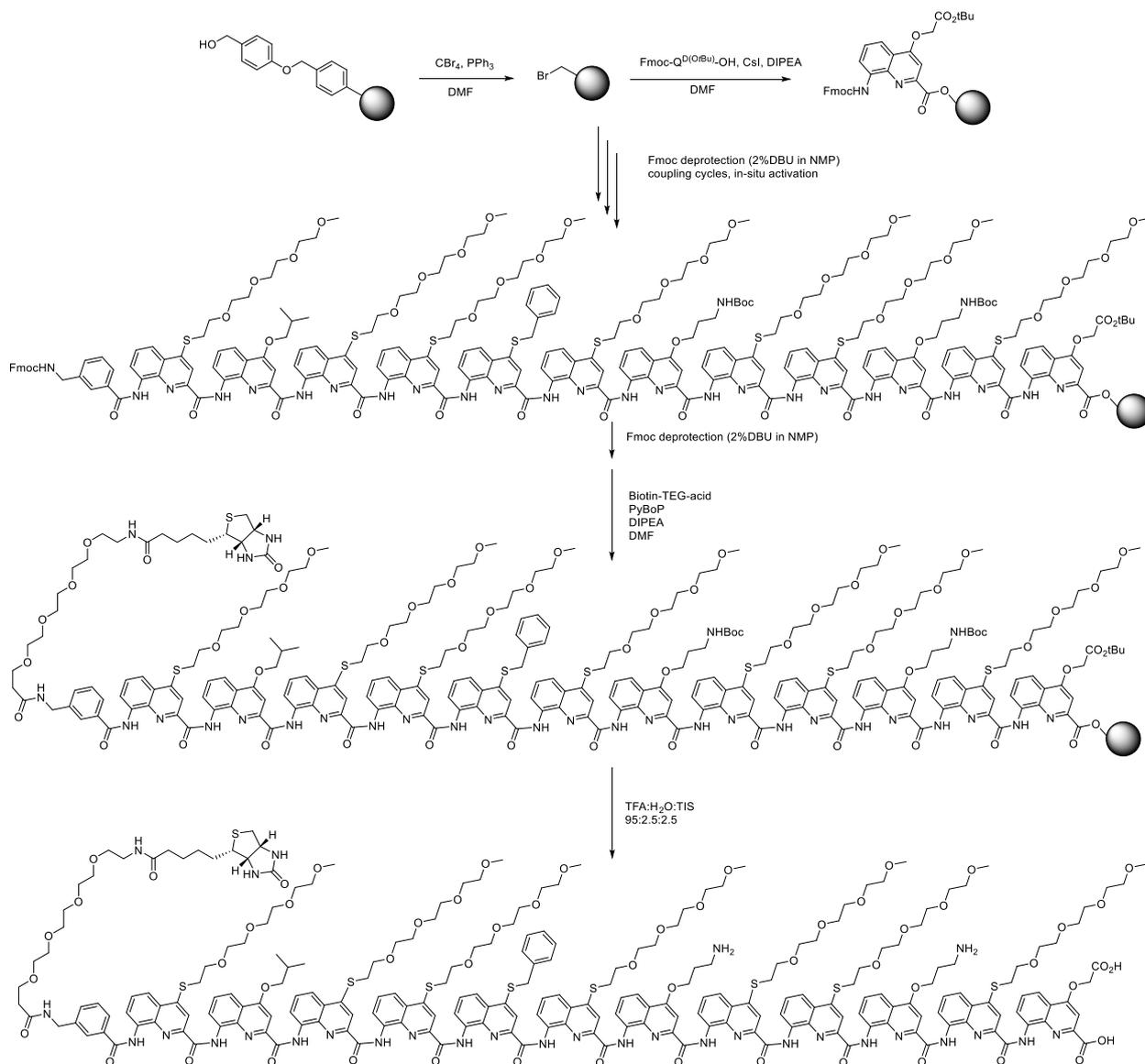


Compound 1e: The crude foldamer was purified by semi prep RP-HPLC to furnish **1e** as a yellow solid (1.6 mg, 5%). HRMS (ESI⁺): calcd. for C₂₂₂H₂₅₉N₂₉O₅₃S₇ [M+3H]³⁺: 1468.5575 found 1468.5850.



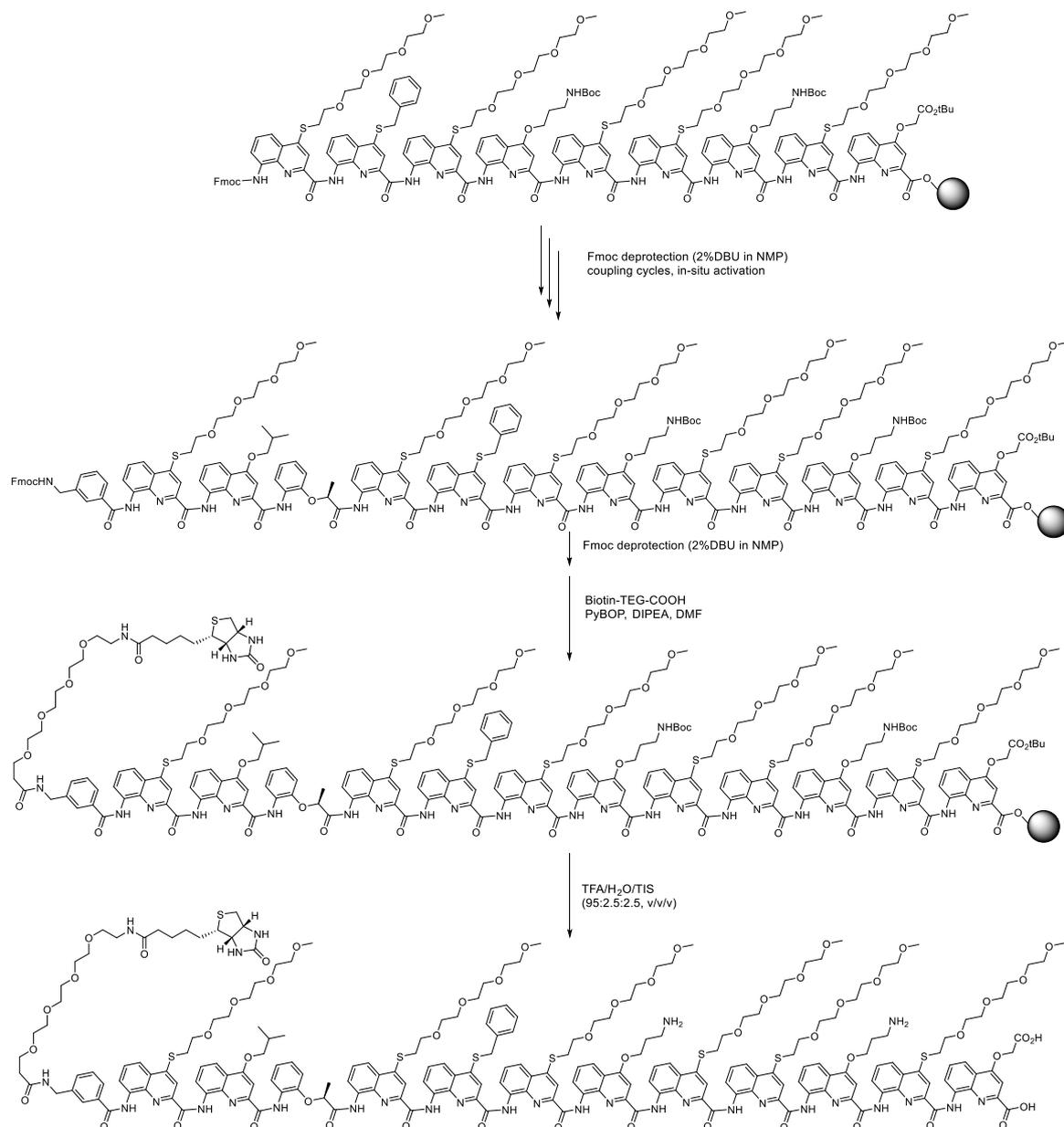
Compound 2a: Foldamer **2a** was synthesized on a LL-Wang resin (15 μ mol scale). After TFA cleavage and side chain deprotection, 15.2 mg of crude foldamer was obtained. The crude was purified by semi prep RP-HPLC to furnish **2a** as a yellow solid (5.7 mg, 12%). HRMS (ESI⁺): m/z calcd. for C₂₃₁H₂₇₂N₃₀O₅₅S₉ [M+3H]³⁺ 1551.9055, found 1551.9253.

¹H NMR (500 MHz, DMSO-*d*₆) δ 11.01 (s, 1H), 10.86 (s, 1H), 10.68 (s, 1H), 10.49 (s, 1H), 10.45 (s, 1H), 10.42 (s, 1H), 10.30 (s, 1H), 10.26 (s, 1H), 10.20 (s, 1H), 10.11 (s, 1H), 10.04 (s, 1H), 9.74 (s, 1H), 9.12 (s, 1H), 7.97 (d, $J = 7.2$ Hz, 1H), 7.92 (d, $J = 7.3$ Hz, 1H), 7.82 (s, 3H), 7.79 – 7.74 (m, 2H), 7.69 (d, $J = 7.2$ Hz, 1H), 7.63 (t, $J = 8.5$ Hz, 2H), 7.58 – 7.33 (m, 17H), 7.28 – 6.89 (m, 18H), 6.85 – 6.78 (m, 2H), 6.76 (s, 2H), 6.54 (s, 1H), 6.39 (s, 2H), 6.33 (d, $J = 9.6$ Hz, 2H), 6.28 (s, 1H), 6.18 – 6.08 (m, 5H), 4.14 – 4.01 (m, 2H), 3.98 – 3.90 (m, 2H), 3.86 – 3.43 (m, 62H), 3.40 – 3.35 (m, 9H), 3.25 (s, 3H), 3.19 (s, 3H), 3.13 – 3.03 (m, 2H), 2.78 (dd, $J = 12.4, 5.1$ Hz, 1H), 2.34 – 2.27 (m, 4H), 2.23 – 2.15 (m, 3H), 2.11 – 2.06 (m, 2H), 2.06 – 1.96 (m, 2H), 1.42 (m, 3H), 1.23 (s, 4H), 1.15 (t, $J = 6.2$ Hz, 7H).



Compound 2b: Foldamer **2b** was synthesized on a LL-Wang resin (17 μmol scale). After TFA cleavage and side chain deprotection, 50 mg crude was obtained. The crude was purified by semi prep RP-HPLC to furnish **2b** as yellow solid (5 mg, 8%). HRMS (ESI⁺): m/z calcd. for $\text{C}_{231}\text{H}_{272}\text{N}_{30}\text{O}_{55}\text{S}_9$ $[\text{M}+2\text{H}]^{2+}$ 2319.3554, found 2319.3724.

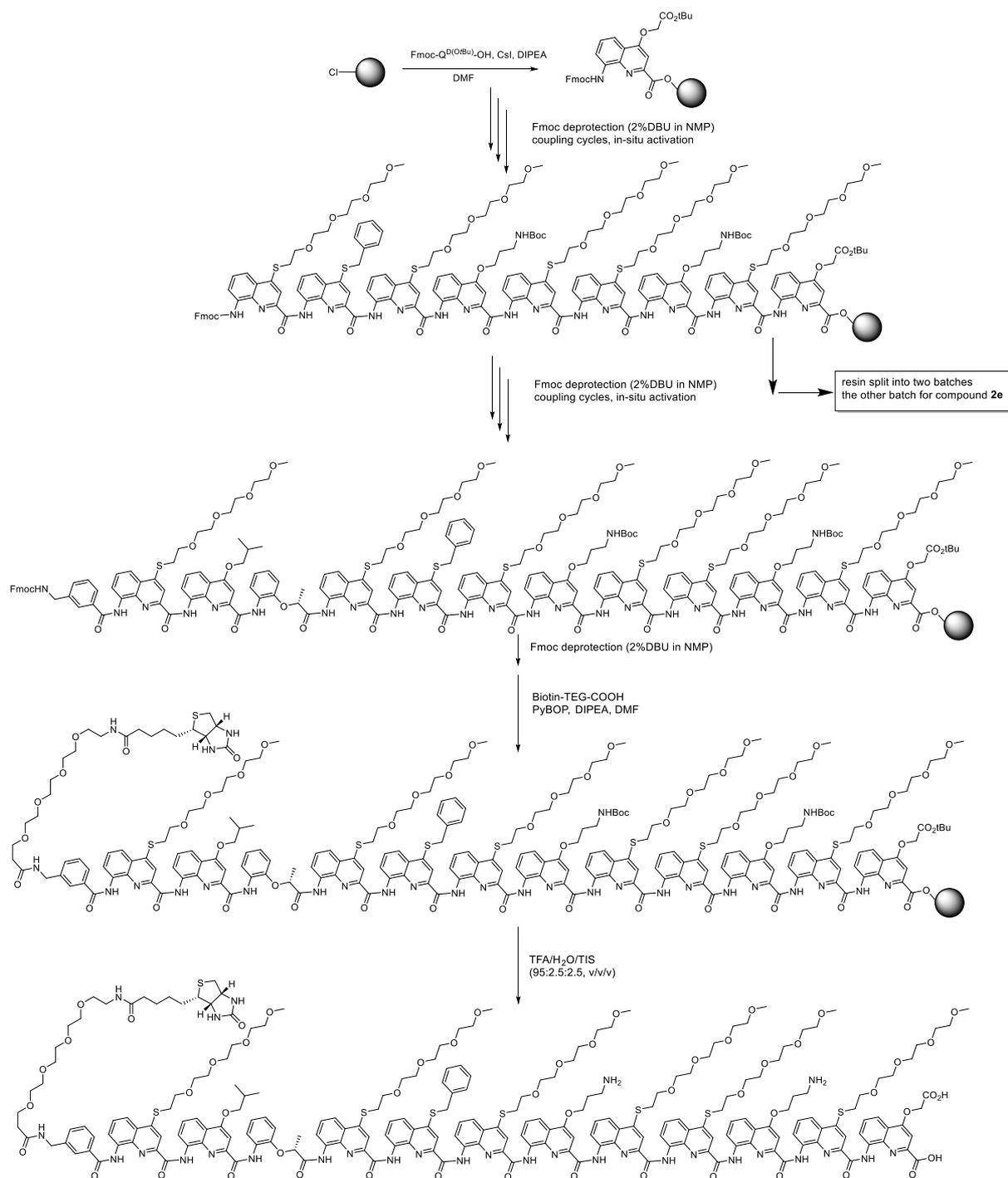
¹H NMR (500 MHz, DMSO-*d*₆) δ 10.89 (s, 1H), 10.68 (s, 1H), 10.62 (s, 1H), 10.49 (s, 1H), 10.45 (s, 1H), 10.35 (s, 1H), 10.29 (s, 1H), 10.24 (s, 1H), 10.22 (s, 1H), 10.12 (s, 2H), 9.11 (s, 1H), 7.98 (dd, $J = 22.5, 7.3$ Hz, 2H), 7.75 – 7.57 (m, 13H), 7.57 – 7.36 (m, 17H), 7.23 (q, $J = 7.2$ Hz, 6H), 7.17 (s, 4H), 7.10 (d, $J = 7.4$ Hz, 6H), 7.05 (s, 5H), 7.00 (s, 3H), 6.96 (d, $J = 7.3$ Hz, 2H), 6.92 (d, $J = 7.2$ Hz, 3H), 6.89 (s, 1H), 6.86 (d, $J = 7.4$ Hz, 2H), 6.80 (s, 1H), 6.76 (d, $J = 7.6$ Hz, 2H), 6.55 (s, 1H), 6.42 – 6.26 (m, 5H), 6.24 – 6.03 (m, 6H), 5.75 (s, 1H), 4.42 (s, 1H), 4.31 – 4.13 (m, 8H), 4.13 – 4.05 (m, 5H), 4.00 – 3.00 (peaks were overlapped with water solvent peak), 2.05 – 1.89 (m, 3H), 1.23 (s, 11H), 1.14 (t, $J = 6.3$ Hz, 9H), 0.85 (t, $J = 6.7$ Hz, 2H).



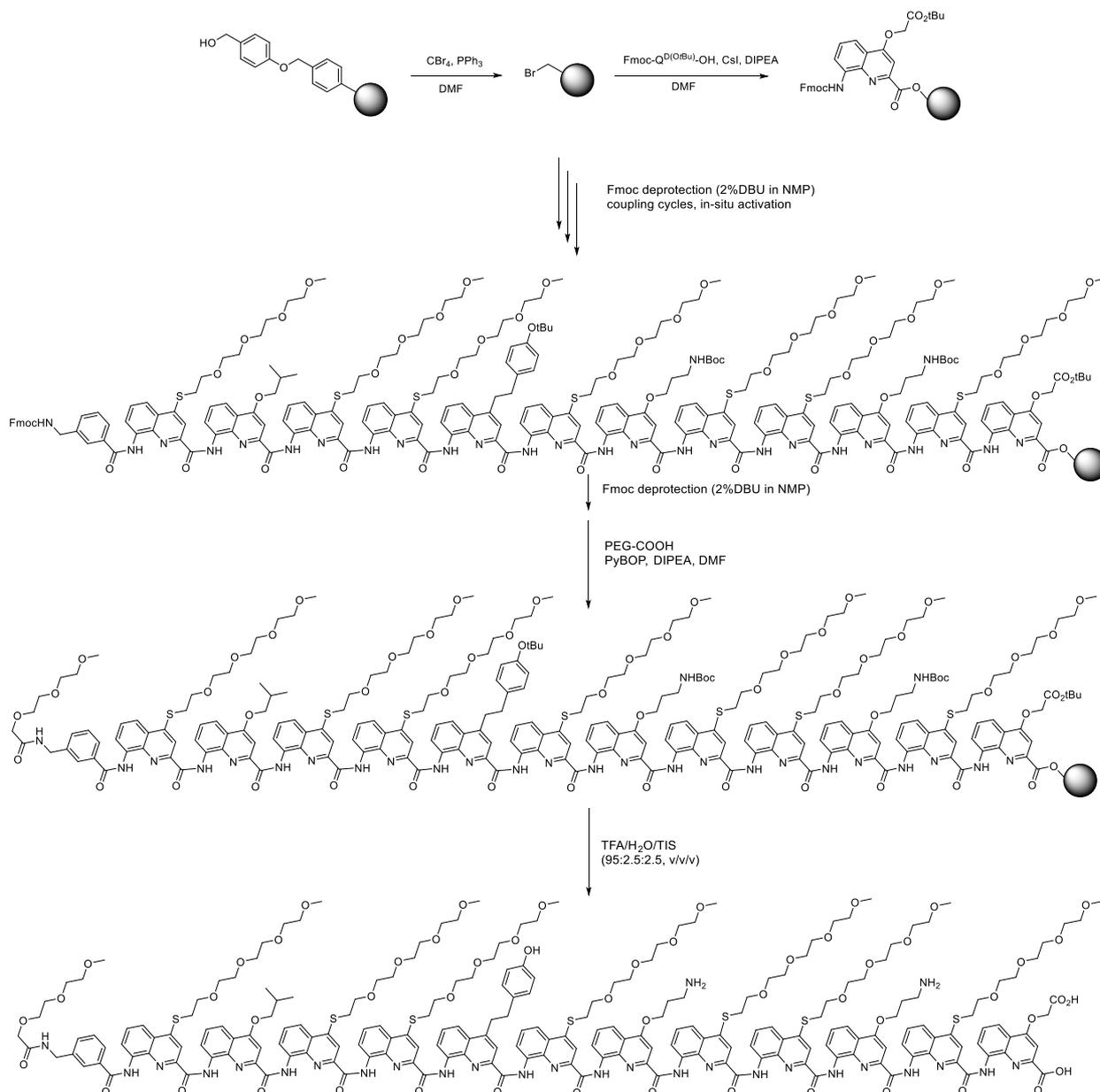
Compound 2c: Foldamer **2c** was synthesized on a Cl-MPA ProTide® resin (30 μ mol scale). After the 9th Q coupling, the resin was divided into two batches. The other 15 μ mol was used for **2d** synthesis (see below). After TFA cleavage and side chain deprotection, 30 mg of crude foldamer was obtained. The crude was purified by semi prep RP-HPLC to furnish **2c** as yellow solid (9.2 mg, 14%). HRMS (ESI⁺): calcd. for C₂₂₁H₂₅₉N₂₉O₅₂S₈ [M+2H]²⁺: 2204.8198, measured: 2204.9361.

¹H NMR (500 MHz, DMF-*d*₇) δ 11.18 (s, 1H), 10.91 (s, 1H), 10.71 (s, 1H), 10.67 (s, 1H), 10.61 (s, 1H), 10.53 (s, 1H), 10.42 (s, 1H), 10.24 (s, 1H), 10.22 (s, 1H), 9.90 (s, 1H), 9.14 (s, 1H), 8.87 (s, 1H), 8.47 (d, *J* = 7.4 Hz, 1H), 8.15 (dd, *J* = 19.4, 7.1 Hz, 2H), 7.86 (s, 1H), 7.80 (t, *J* = 6.7 Hz, 5H), 7.76 (s, 1H), 7.75 (d, *J* = 1.1 Hz, 1H), 7.74 – 7.67 (m, 5H), 7.67 – 7.60 (m, 6H), 7.60 – 7.51 (m, 6H), 7.51 – 7.43 (m, 5H), 7.41 (d, *J* = 7.4 Hz, 1H), 7.39 – 7.19 (m, 12H), 7.19 – 7.12 (m, 2H), 7.06 (d, *J* = 7.5 Hz, 1H), 7.01 (s, 1H), 6.96 (d, *J* = 7.5 Hz, 1H), 6.87 (d, *J* = 5.5 Hz, 2H), 6.77 (d, *J* = 7.9 Hz, 1H), 6.62 (d,

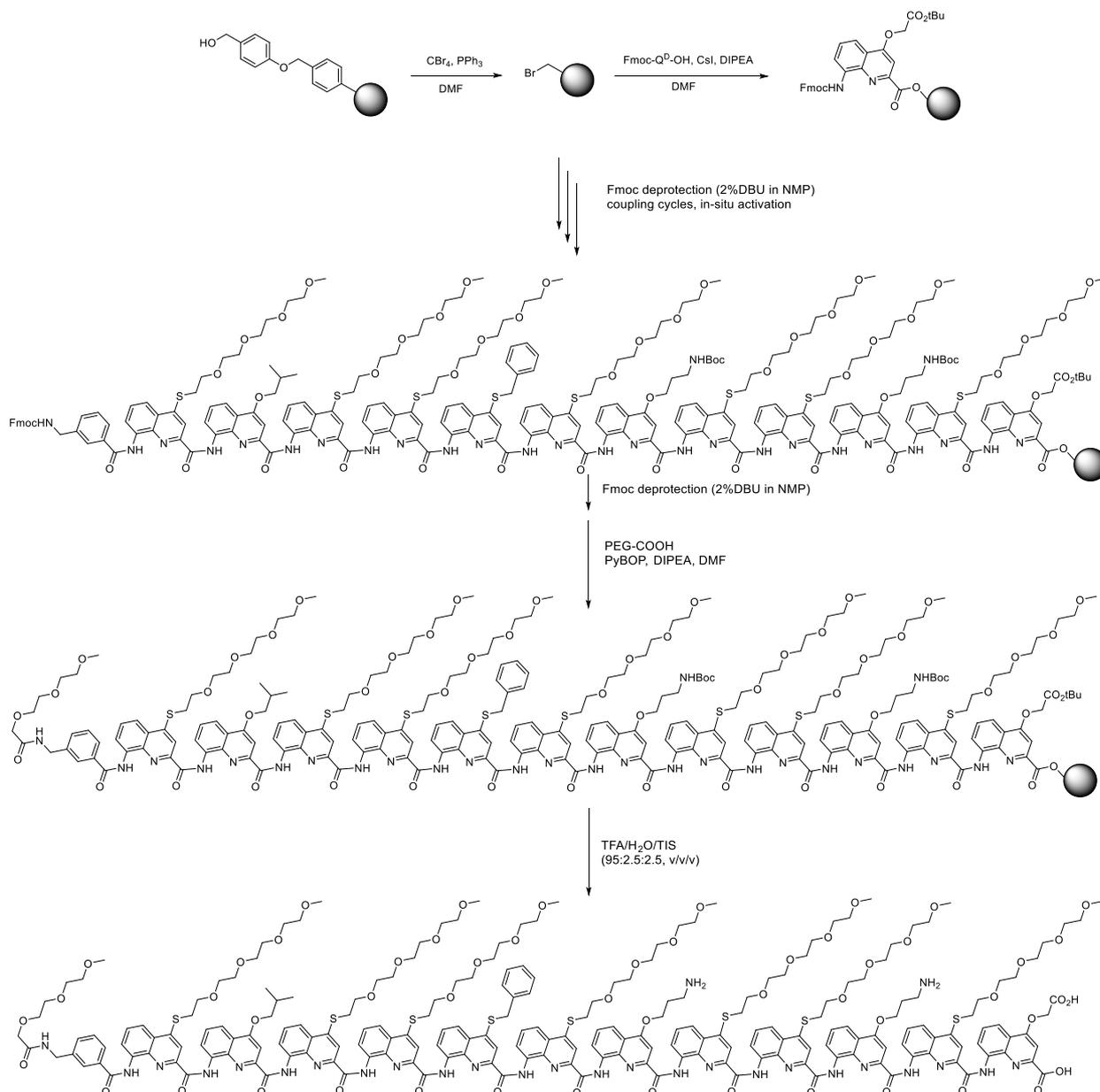
$J = 16.2$ Hz, 2H), 6.49 (s, 1H), 6.47 (s, 1H), 6.45 – 6.40 (m, 2H), 6.39 (s, 1H), 6.34 (s, 1H), 6.22 (s, 1H), 6.05 (s, 1H), 5.95 (s, 2H), 5.91 (t, $J = 7.3$ Hz, 1H), 4.87 – 4.69 (m, 2H), 4.62 (d, $J = 13.5$ Hz, 1H), 4.54 – 4.36 (m, 5H), 4.35 – 4.22 (m, 2H), 4.20 – 3.97 (m, 13H), 3.97 – 3.88 (m, 18H), 3.88 – 3.81 (m, 16H), 3.81 – 3.53 (m, 80H), around 3.5 (broad water solvent peak), 3.47 – 3.44 (m, 14H), 3.42 (d, $J = 3.5$ Hz, 12H), 3.39 (d, $J = 3.9$ Hz, 9H), 3.36 (s, 5H), 3.31 – 3.23 (m, 13H), 3.21 – 3.15 (m, 3H), 2.59 (d, $J = 9.9$ Hz, 5H), 2.31 (dt, $J = 14.1, 7.2$ Hz, 1H), 2.24 – 2.18 (m, 2H), 2.16 (d, $J = 1.6$ Hz, 4H), 1.77 – 1.69 (m, 1H), 1.62 – 1.51 (m, 4H), 1.39 (d, $J = 7.0$ Hz, 3H), 1.29 (d, $J = 9.2$ Hz, 4H), 1.23 – 1.17 (m, 7H), -0.33 (d, $J = 6.3$ Hz, 3H).



Compound 2d: The synthesis started from the resin described above (9mer, 15 μmol). After TFA/H₂O/TIS cleavage, 28 mg crude was obtained. The crude was purified by semi prep RP-HPLC to furnish **2d** as yellow solid (4.1 mg, 6%). HRMS (ESI⁺): calcd for C₂₂₁H₂₅₉N₂₉O₅₂S₈ [M+2H]²⁺ 2204.8198, found 2204.9232.



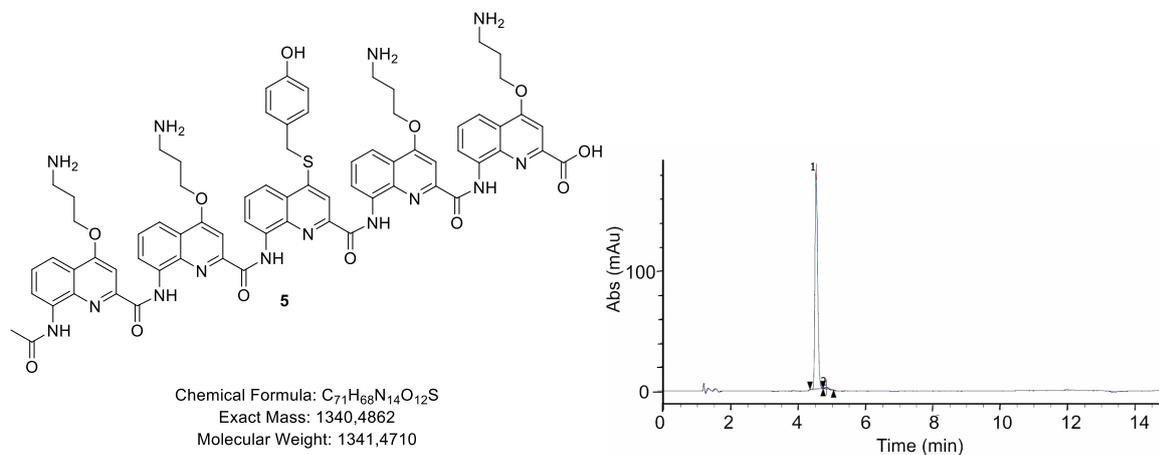
Compound 2e: Foldamer **2e** was synthesized on a LL-Wang resin (150 μmol scale). After TFA cleavage and side chain deprotection, the crude was purified by semi prep RP-HPLC to furnish **2e** as a yellow solid (85 mg, 62%). HRMS (ESI⁺) calcd. for C₂₁₈H₂₅₂N₂₇O₅S₇ [M+2H]²⁺ 2161.8030, found 2161.8397. **¹H NMR** (500 MHz, DMSO-*d*₆) δ 10.98 (s, 1H), 10.80 (s, 1H), 10.69 (s, 1H), 10.47 (s, 2H), 10.41 (s, 1H), 10.29 (s, 1H), 10.25 (s, 1H), 10.24 (s, 1H), 10.15 (s, 1H), 10.03 (s, 1H), 9.38 (s, 1H), 9.11 (s, 1H), 7.96 (d, *J* = 7.2 Hz, 1H), 7.90 (d, *J* = 7.2 Hz, 1H), 7.79 (s, 2H), 7.71 (d, *J* = 7.2 Hz, 1H), 7.65 (d, *J* = 7.8 Hz, 2H), 7.60 (dd, *J* = 13.1, 6.6 Hz, 4H), 7.56 (d, *J* = 9.3 Hz, 2H), 7.54 – 7.41 (m, 9H), 7.33 (d, *J* = 8.2 Hz, 3H), 7.29 – 7.21 (m, 5H), 7.20 – 6.96 (m, 16H), 6.90 (q, *J* = 7.5, 6.1 Hz, 5H), 6.80 (td, *J* = 16.2, 15.7, 7.4 Hz, 5H), 6.57 (s, 1H), 6.41 – 6.31 (m, 3H), 6.26 (s, 1H), 6.23 (d, *J* = 4.2 Hz, 2H), 6.20 (s, 1H), 6.15 (s, 1H), 6.10 (s, 1H), 5.74 (s, 1H), 4.21 – 4.09 (m, 1H), 4.09 – 3.88 (m, 9H), 3.89 – 3.43 (m, 116H), 3.40 – 3.36 (m, 8H), 3.25 (s, 7H), 3.19 – 3.04 (m, 20H), 3.01 (s, 5H), 2.35 – 2.27 (m, 5H), 2.16 (d, *J* = 7.7 Hz, 3H), 1.23 (s, 1H), 1.19 – 1.13 (m, 8H).



Compound 2f: Foldamer **2f** was synthesized on a LL-Wang resin (150 μmol scale). After TFA cleavage and side chain deprotection, the crude was purified by semi prep RP-HPLC to furnish **2f** as yellow solid (52 mg, 38%). HRMS (ESI⁺): m/z calcd for $\text{C}_{217}\text{H}_{250}\text{N}_{27}\text{O}_{52}\text{S}_8$ $[\text{M}+2\text{H}]^{2+}$ 2162.7838, found 2162.8219.

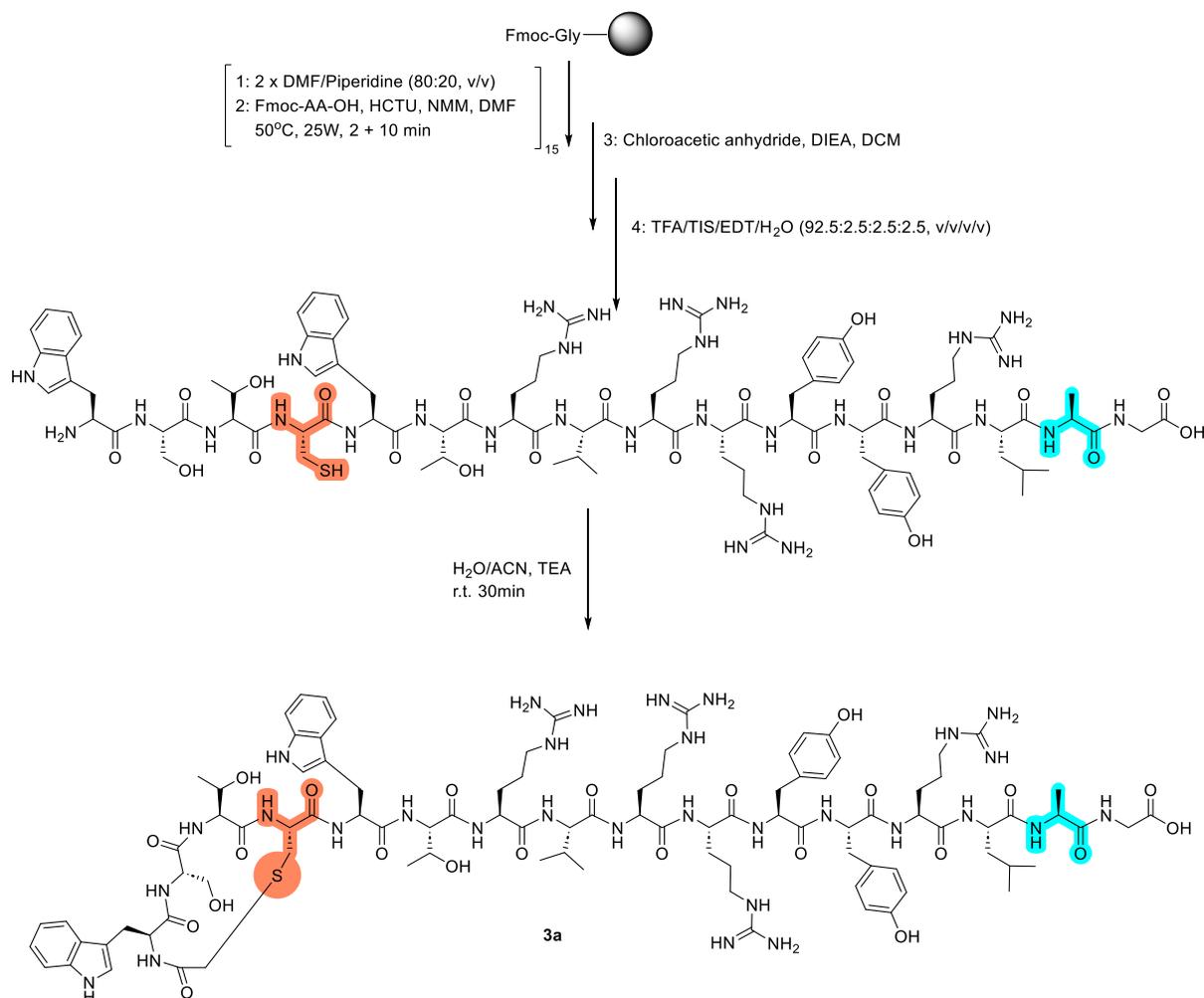
¹H NMR (500 MHz, DMSO-*d*₆) δ 13.34 (s, 1H), 11.81 (s, 1H), 11.02 (s, 1H), 10.89 (s, 1H), 10.66 (s, 1H), 10.49 (s, 1H), 10.44 (s, 2H), 10.33 (s, 1H), 10.24 (s, 1H), 10.17 (s, 1H), 10.11 (s, 1H), 10.08 (s, 1H), 9.10 (s, 1H), 8.04 (s, 4H), 7.96 (d, $J = 6.9$ Hz, 2H), 7.91 (d, $J = 7.3$ Hz, 2H), 7.83 (s, 4H), 7.76 (d, $J = 7.6$ Hz, 3H), 7.70 (d, $J = 7.2$ Hz, 2H), 7.64 (t, $J = 7.4$ Hz, 6H), 7.61 – 7.32 (m, 28H), 7.33 – 7.21 (m, 9H), 7.21 – 6.99 (m, 20H), 6.99 – 6.91 (m, 5H), 6.88 (d, $J = 4.5$ Hz, 3H), 6.79 (s, 5H), 6.54 (s, 1H), 6.43 (d, $J = 7.7$ Hz, 1H), 6.38 (s, 1H), 6.30 (d, $J = 4.6$ Hz, 2H), 6.21 – 6.09 (m, 6H), 5.73 (s, 1H), 4.71 (d, $J = 16.0$ Hz, 1H), 4.57 (d, $J = 16.0$ Hz, 1H), 4.19 (d, $J = 13.3$ Hz, 3H), 4.15 – 4.05 (m, 7H), 4.03 – 3.89 (m, 17H), 3.88 – 3.80 (m, 25H), 3.80 – 3.76 (m, 14H), 3.76 – 3.67 (m, 42H), 3.66 – 3.57 (m, 47H),

3.57 – 3.48 (m, 45H), 3.46 (m 18H), 3.37 (d, $J = 3.5$ Hz, 10H), 3.34 (s, 5H), 3.33 – 3.30 (m, 15H), 3.25 (s, 6H), 3.19 (s, 6H), 2.03 – 1.94 (m, 2H), 1.23 (s, 13H), 1.14 (t, $J = 6.4$ Hz, 8H), 1.06 – 1.01 (m, 2H), 0.98 (d, $J = 6.4$ Hz, 6H), 0.85 (t, $J = 6.7$ Hz, 2H).

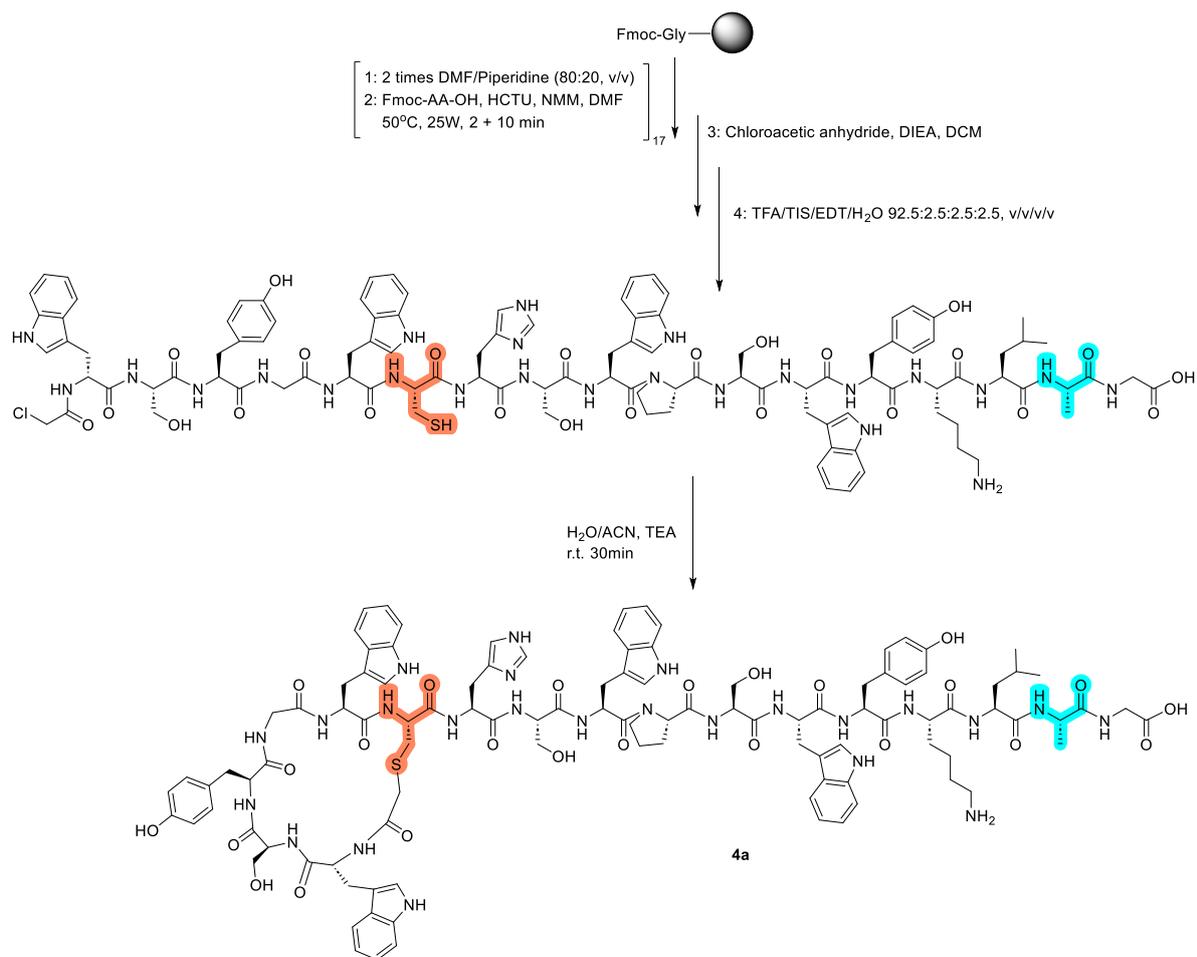


Compound 5: Pentamer **5** was synthesized on a LL-Wang resin (0.41 mmol/g, 10.25 μ mol scale). After TFA cleavage and RP-HPLC purification, **5** was recovered in 62 % yield (3 mg). HRMS (ESI⁺): calcd. for $C_{71}H_{68}N_{14}O_{12}S$ $[M+2H]^{2+}$ 671.3355, found 671.2494.

8.4 Synthesis of peptide macrocycles



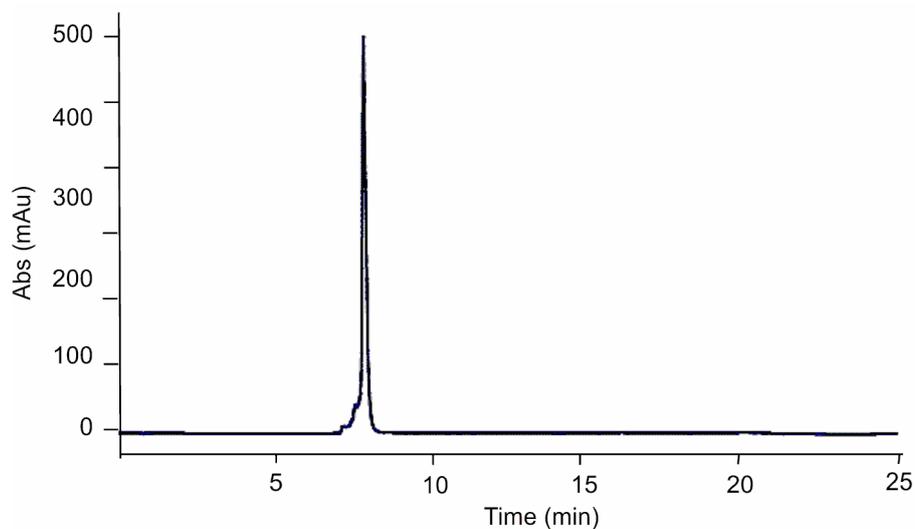
Peptide macrocycle 3a: The SPPS of linear peptide **3a** was performed Fmoc-Gly-Wang resin on a 50 μ mol scale. 73 mg of crude peptide was next dissolved in 5 mL of a CH₃CN/water mixture and TEA (375 μ L, 0.5 M) was added. The completion of cyclization was monitored by RP-HPLC, and after 30 min, the reaction was quenched by diluting the reaction mixture with water/ 0.1% TFA. After lyophilisation, the crude macrocyclic peptide was purified by semi-prep HPLC to furnish **3a** as a white powder (21 mg, 20%). HRMS (ESI⁺): m/z calcd. for C₉₆H₁₄₀N₃₀O₂₃S [M+H]⁺ 2115.0564, found 2115.0156.



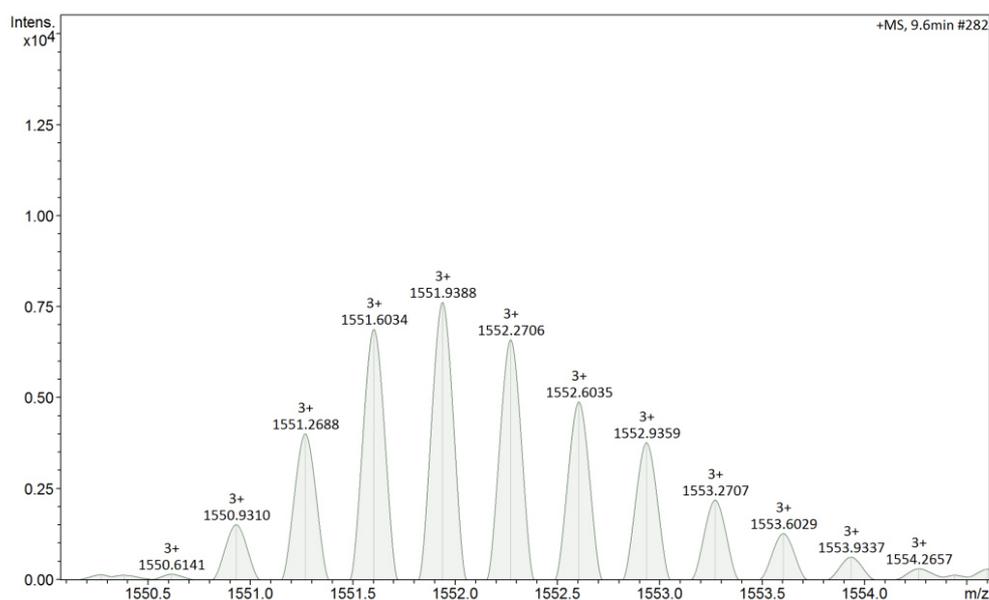
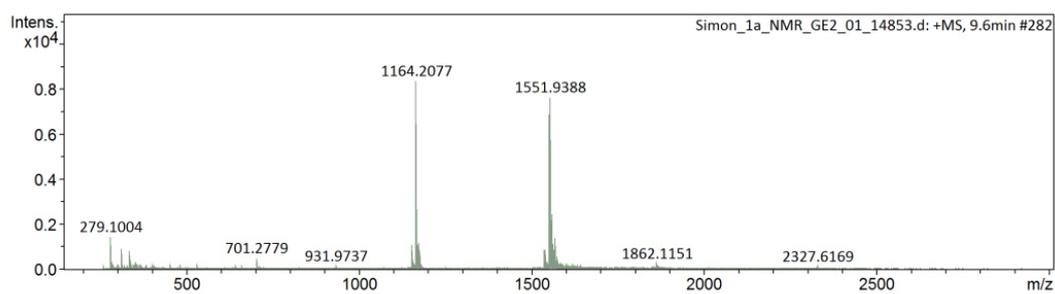
Peptide macrocycle 4a: The SPPS of linear peptide **4a** was performed on Fmoc-Gly-Wang resin on a 50 μmol scale. 81 mg of crude peptide was next dissolved in 5 mL of a CH_3CN /water mixture, and TEA (375 μL , 0.5 M) was added. The completion of cyclization was monitored by RP-HPLC, and after 30 min, the reaction mixture was quenched by pouring water/ 0.1% TFA. After RP-HPLC purification, **4a** was recovered as a white powder (29 mg, 27%). HRMS (ESI⁺): m/z calcd. for $\text{C}_{106}\text{H}_{128}\text{N}_{24}\text{O}_{24}\text{S}$ $[\text{M}+2\text{H}]^{2+}$ 1077.4699, found 1077.4771

3b and **4b** were synthesized following the same procedure using commercially available Fmoc-(*D*)-amino acid. The corresponding macrocyclic peptides were produced, and purified in a similar manner. To note in **4b**, the *N*-term (*D*)-Trp was replaced by a (*L*)-Trp.

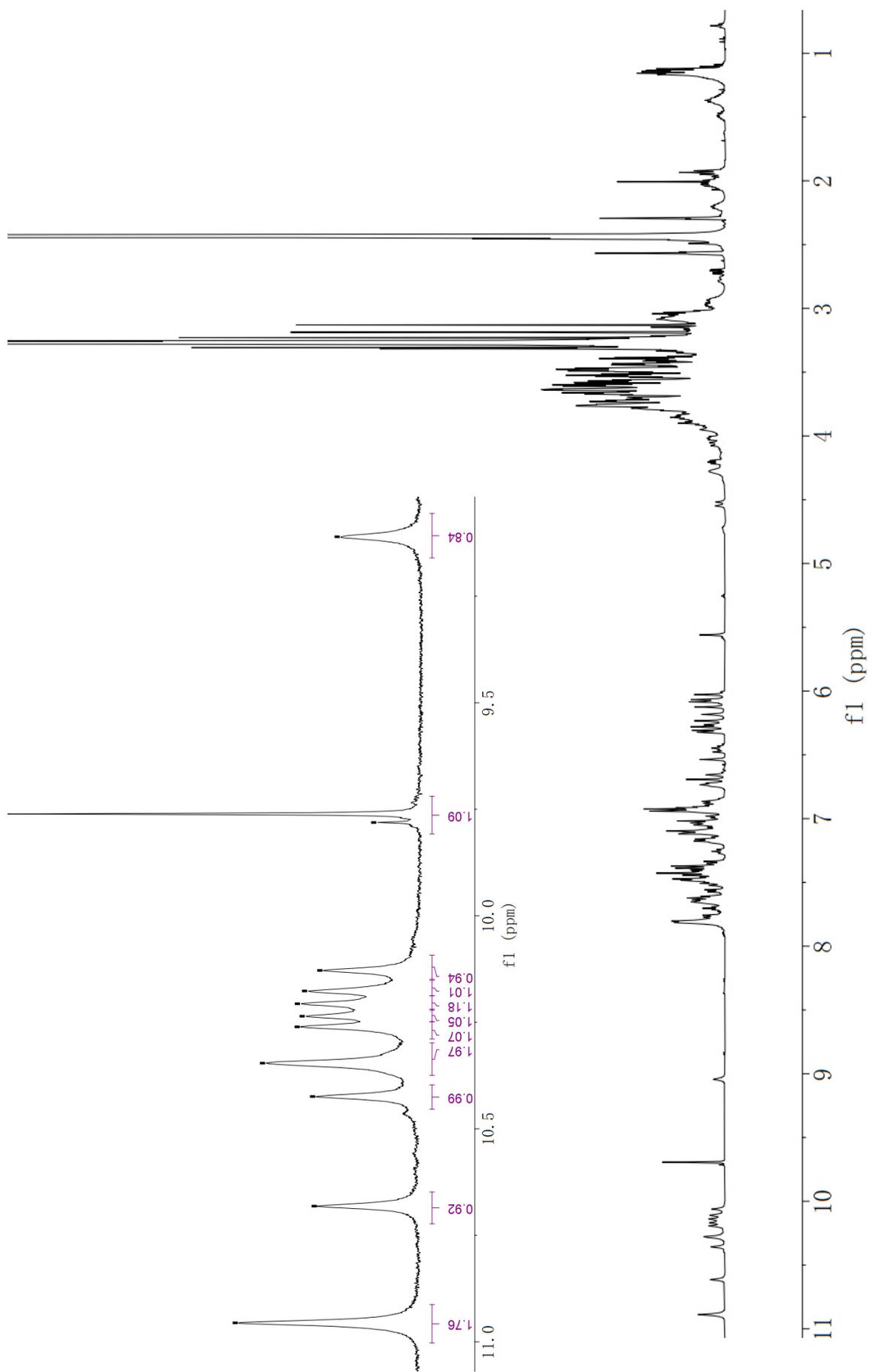
8.5 HPLC profiles, HRMS (ESI⁺) and ¹H NMR spectra



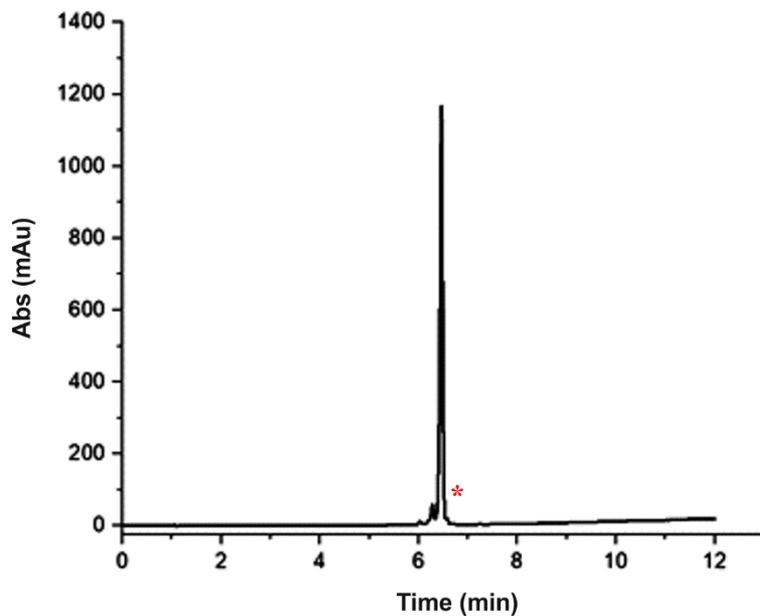
Analytical RP-HPLC profile of foldamer **1a** (Gradient: 5-100% B over 15 min, then 100% B for 10min, $\lambda = 254$ nm).



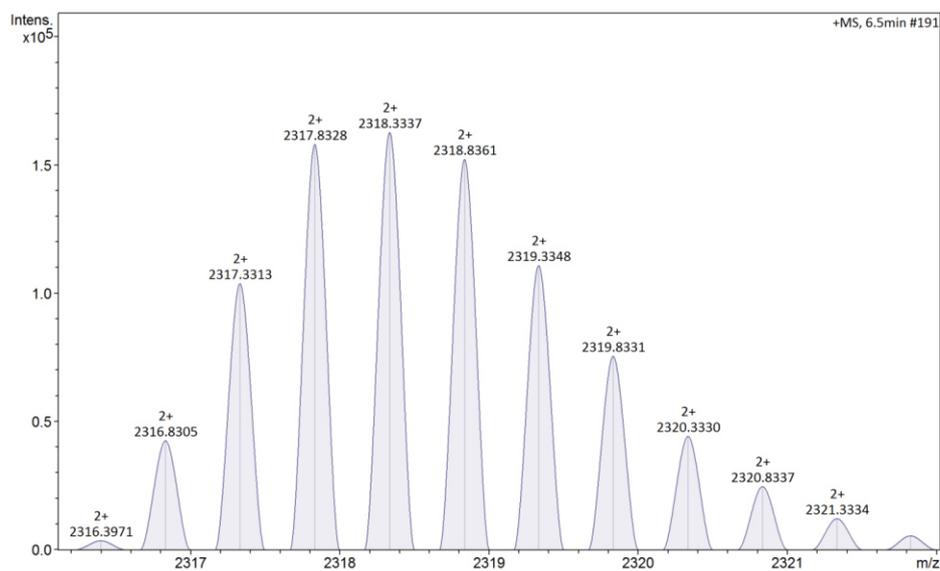
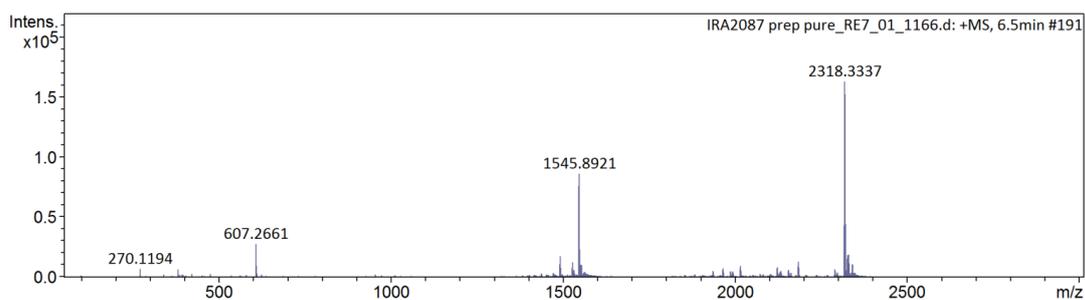
HRMS (ESI⁺) spectra of foldamer **1a**.



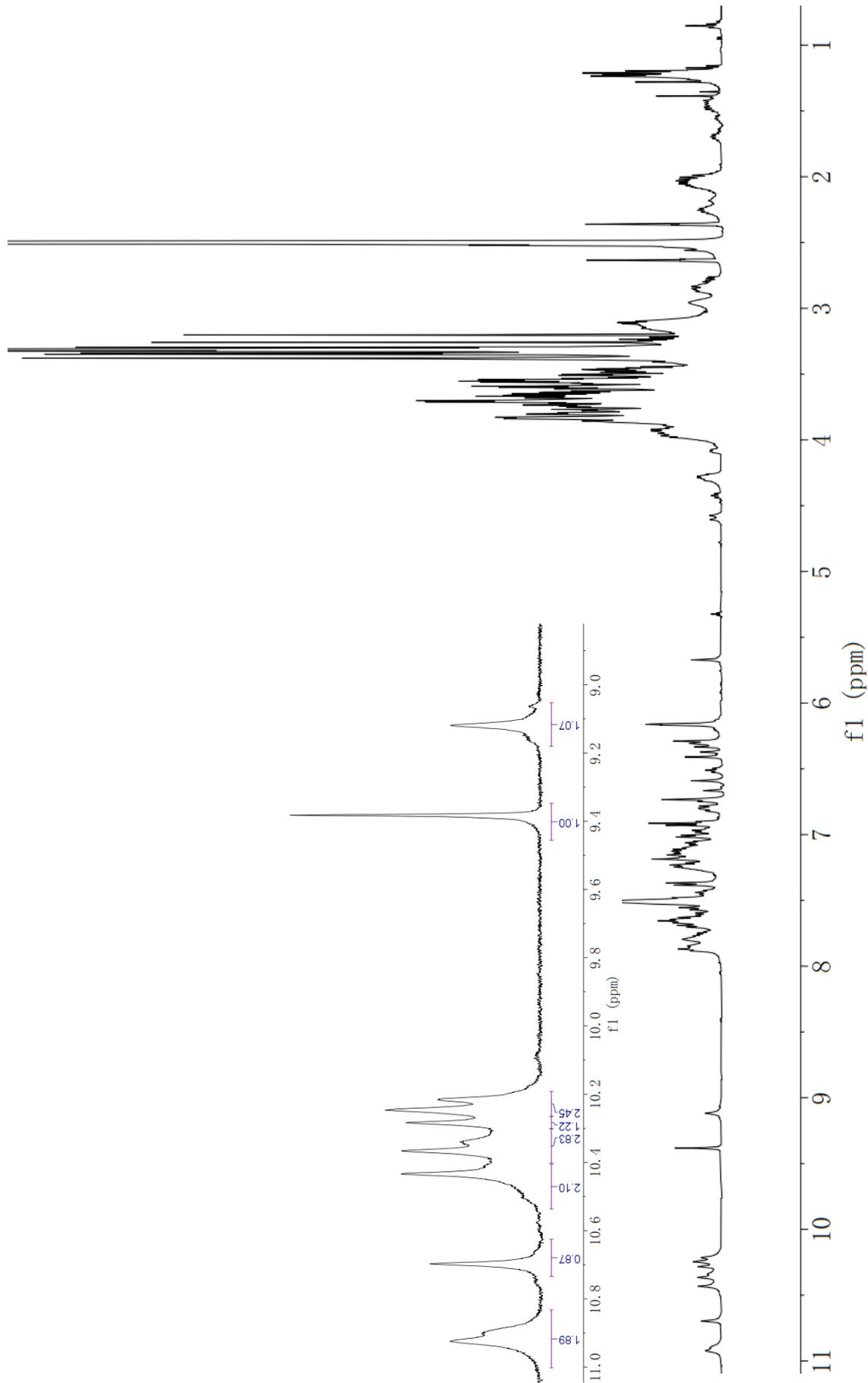
^1H NMR spectrum of foldamer **1a** and zoomed region (500 MHz, DMSO-d_6), 25°C .



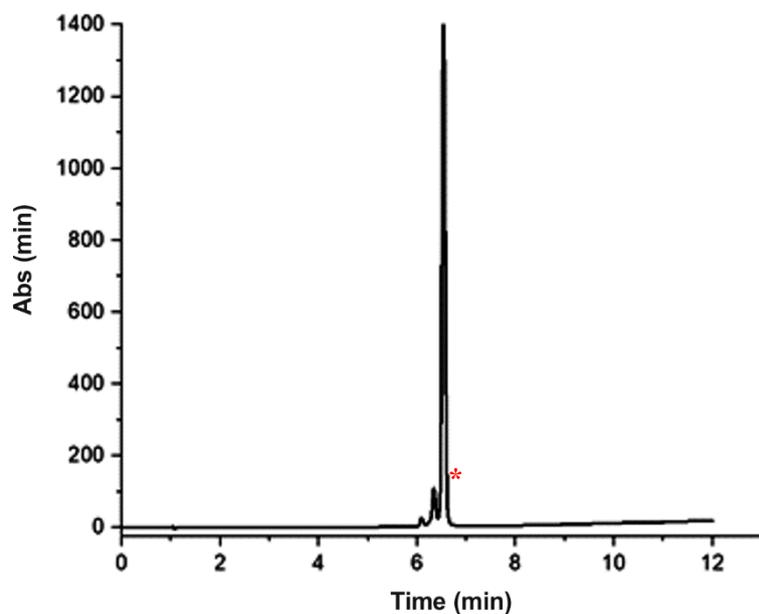
Analytical RP-HPLC profile of pure foldamer **1b** (Gradient: 10-100% B over 10 min, $\lambda = 254$ nm).



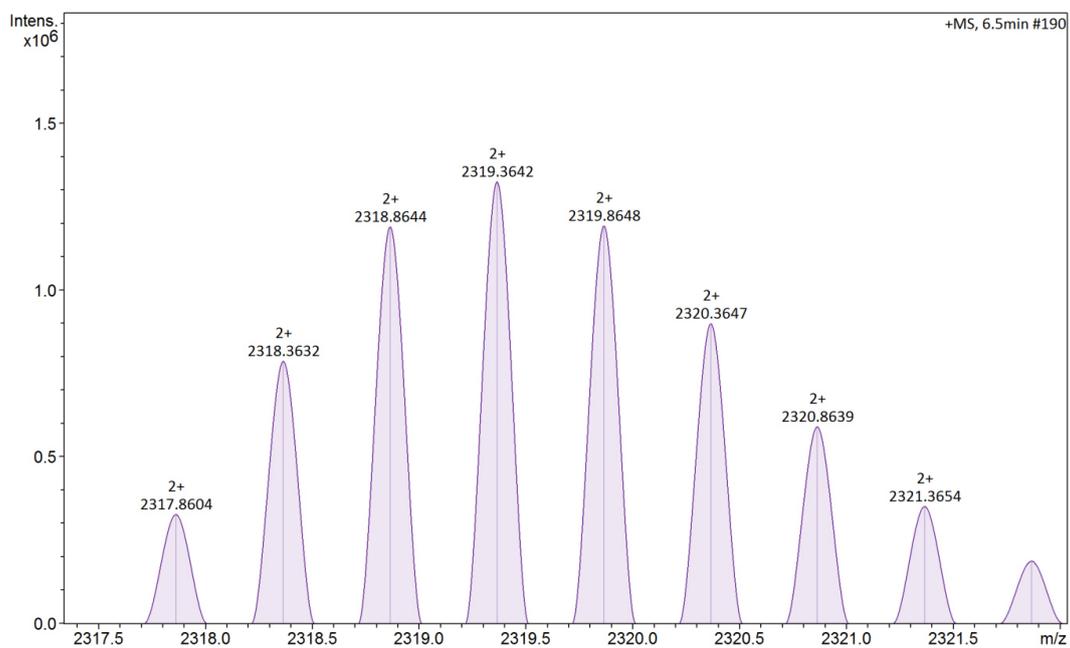
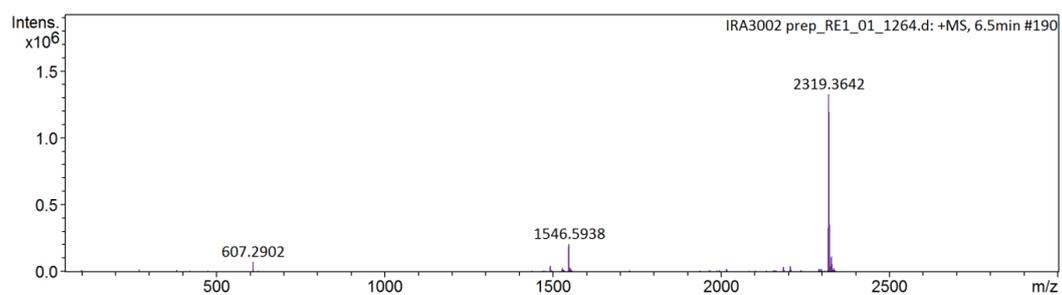
HRMS (ESI⁺) spectra of foldamer **1b**.



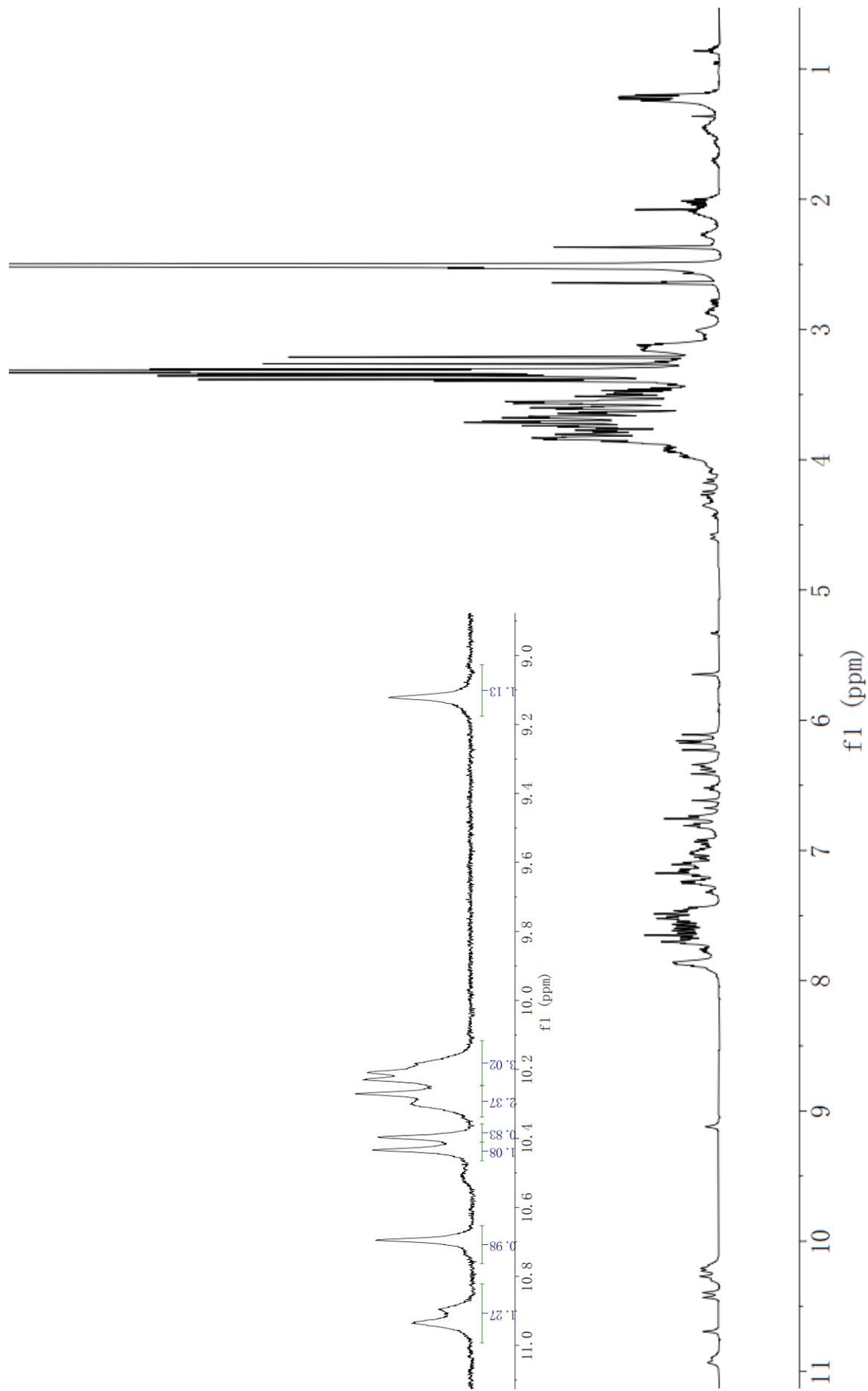
¹H NMR spectrum of foldamer **1b** and zoomed region (500 MHz, DMSO-*d*₆), 25 °C



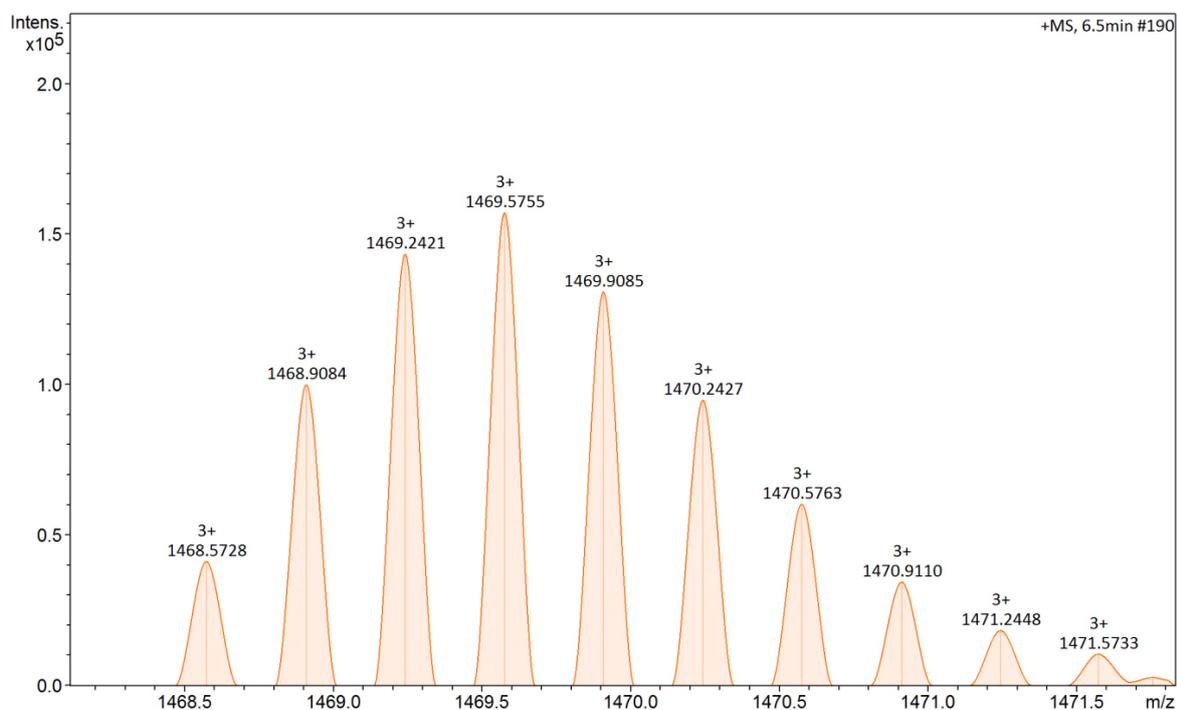
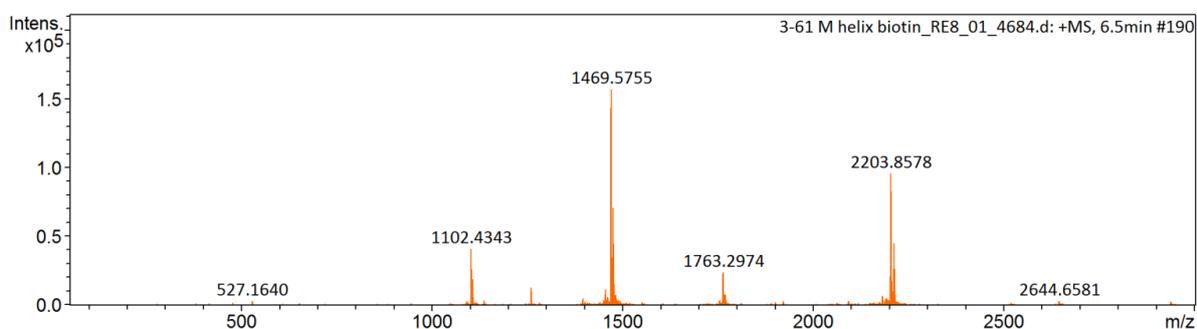
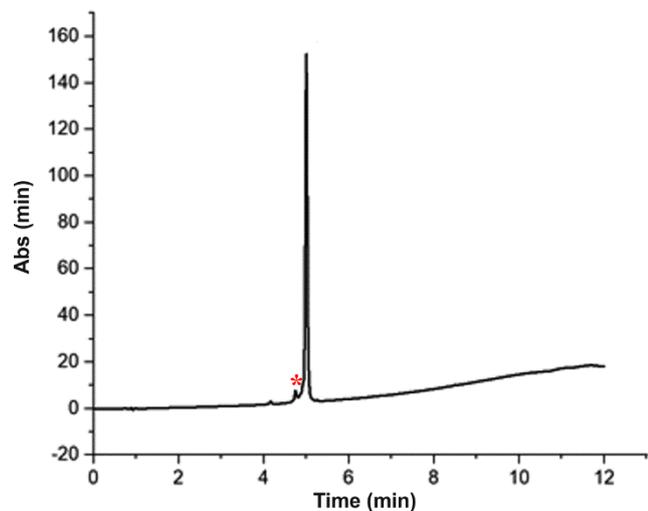
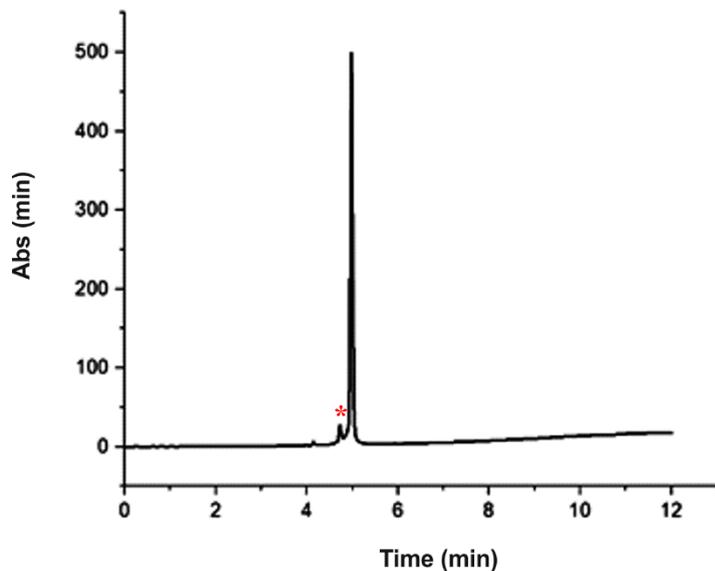
Analytical RP-HPLC profile of pure foldamer **1c** (Gradient: 10-100% B over 10 min, $\lambda = 254$ nm).



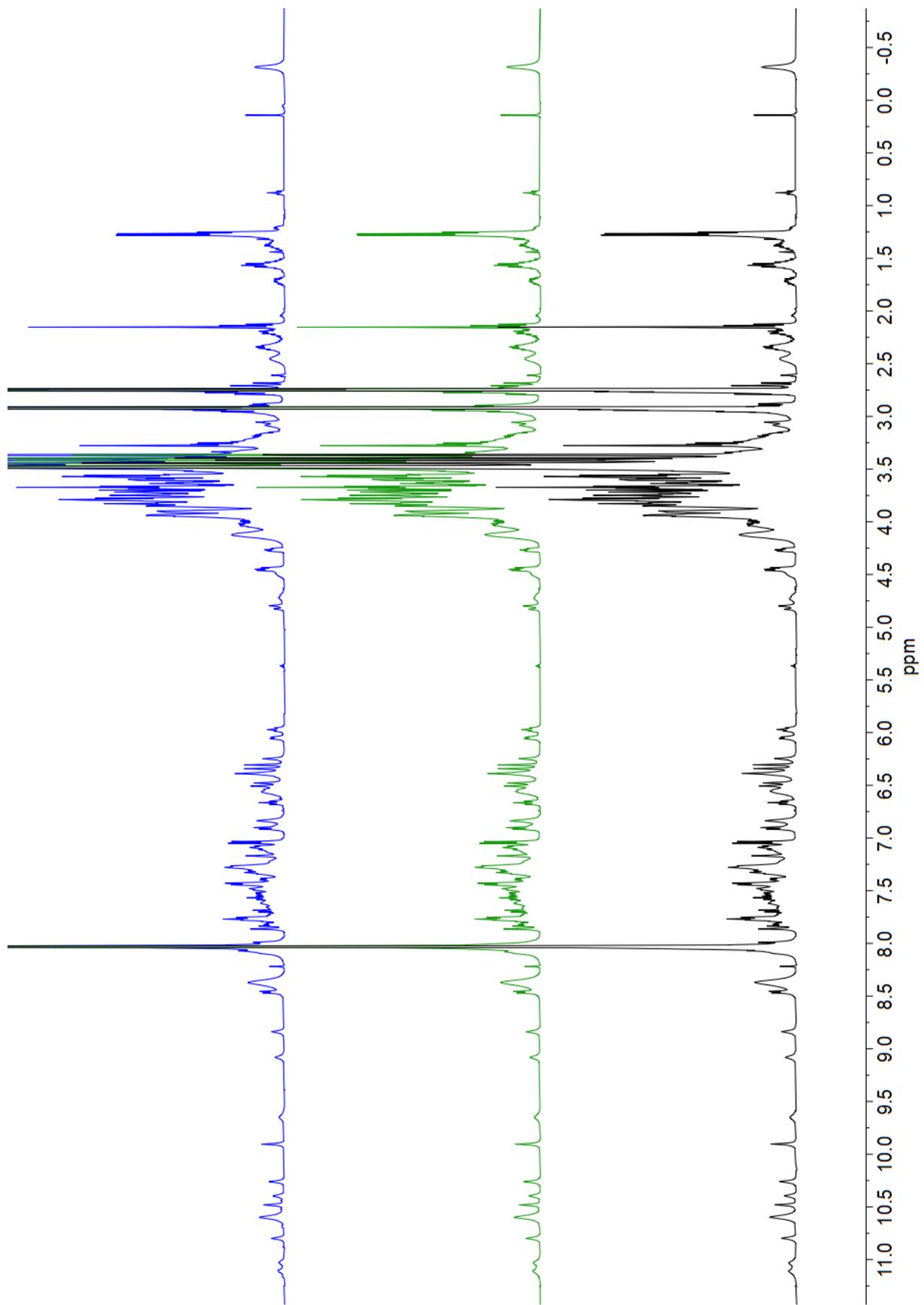
HRMS (ESI⁺) spectrum of foldamer **1c**.

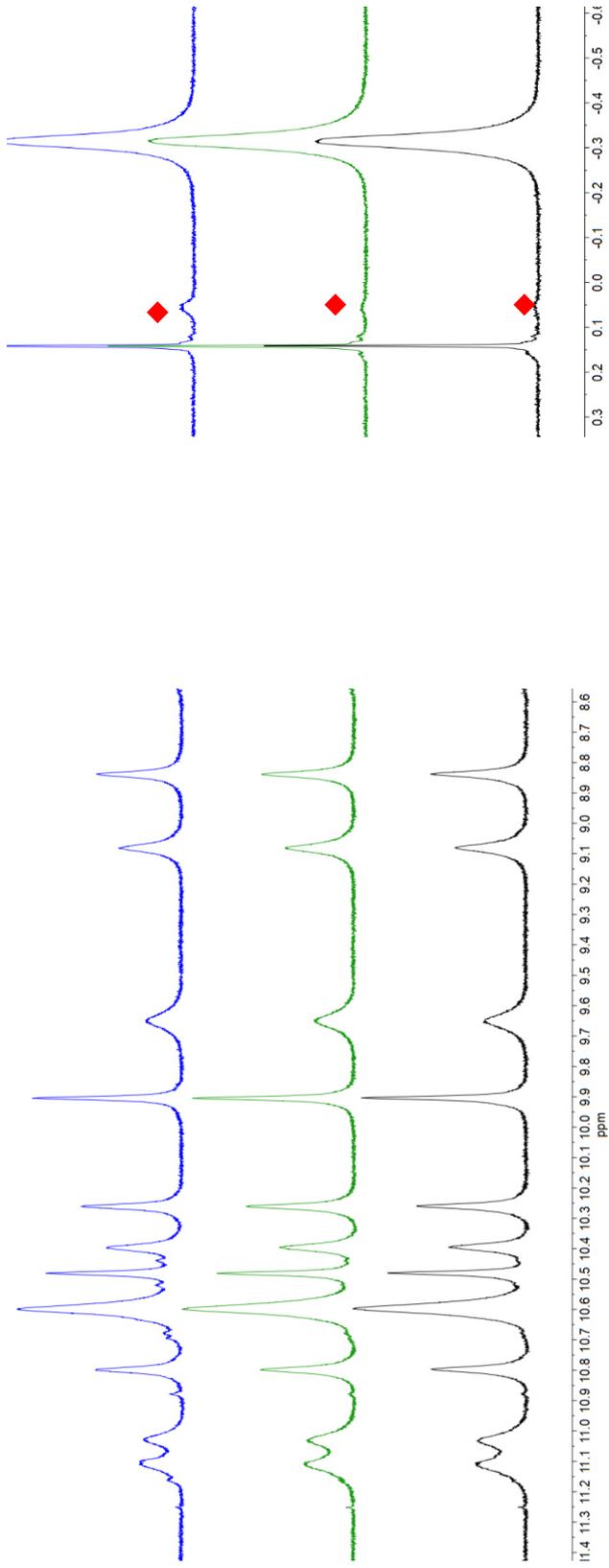


^1H NMR spectrum of foldamer **1c** and zoomed region (500 MHz, $\text{DMSO-}d_6$), $25\text{ }^\circ\text{C}$

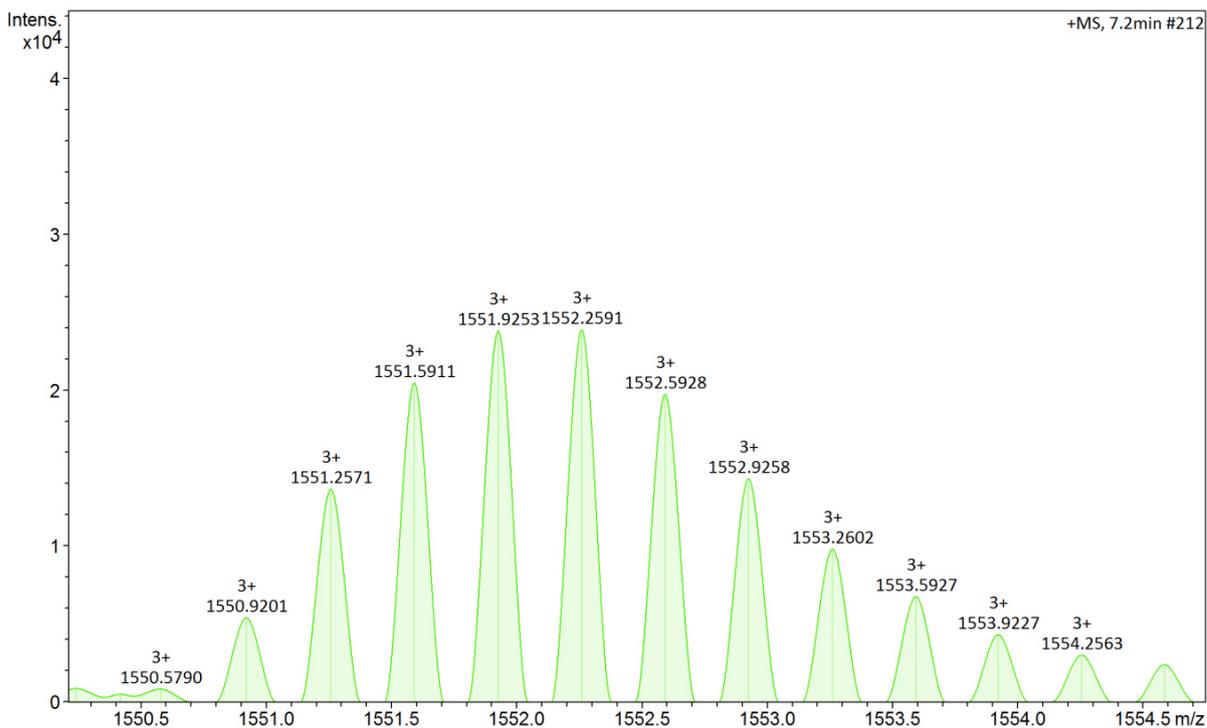
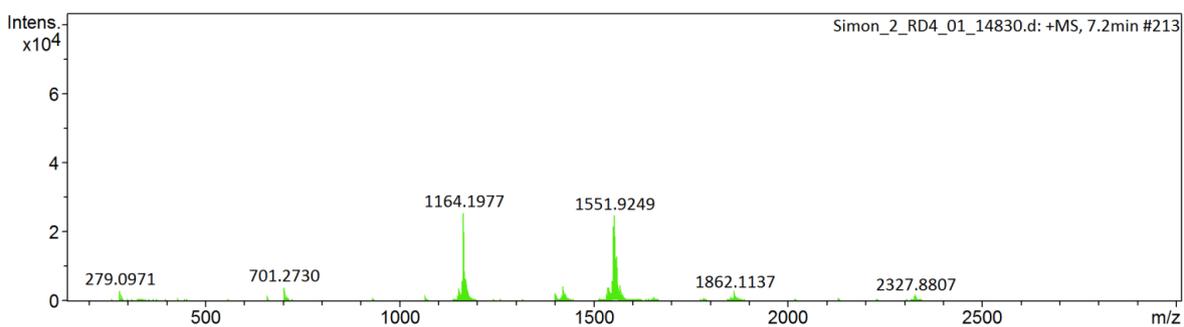
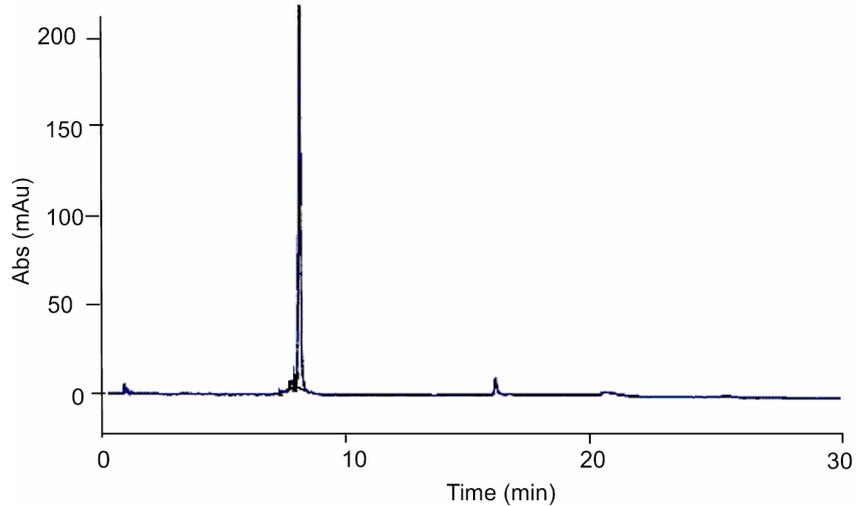


Top left: RP-HPLC profile of compound **1d** (Gradient: 10-100% B over 10 min, $\lambda = 254$ nm). Top right RP-HPLC profile of compound **1e**. The * annotation on the RP-HPLC profiles corresponds to the biotin-oxidized derivative. HRMS (ESI+) spectrum of foldamer **1d** (ESI-HRMS spectra for **1e** was omitted since they are enantiomers).

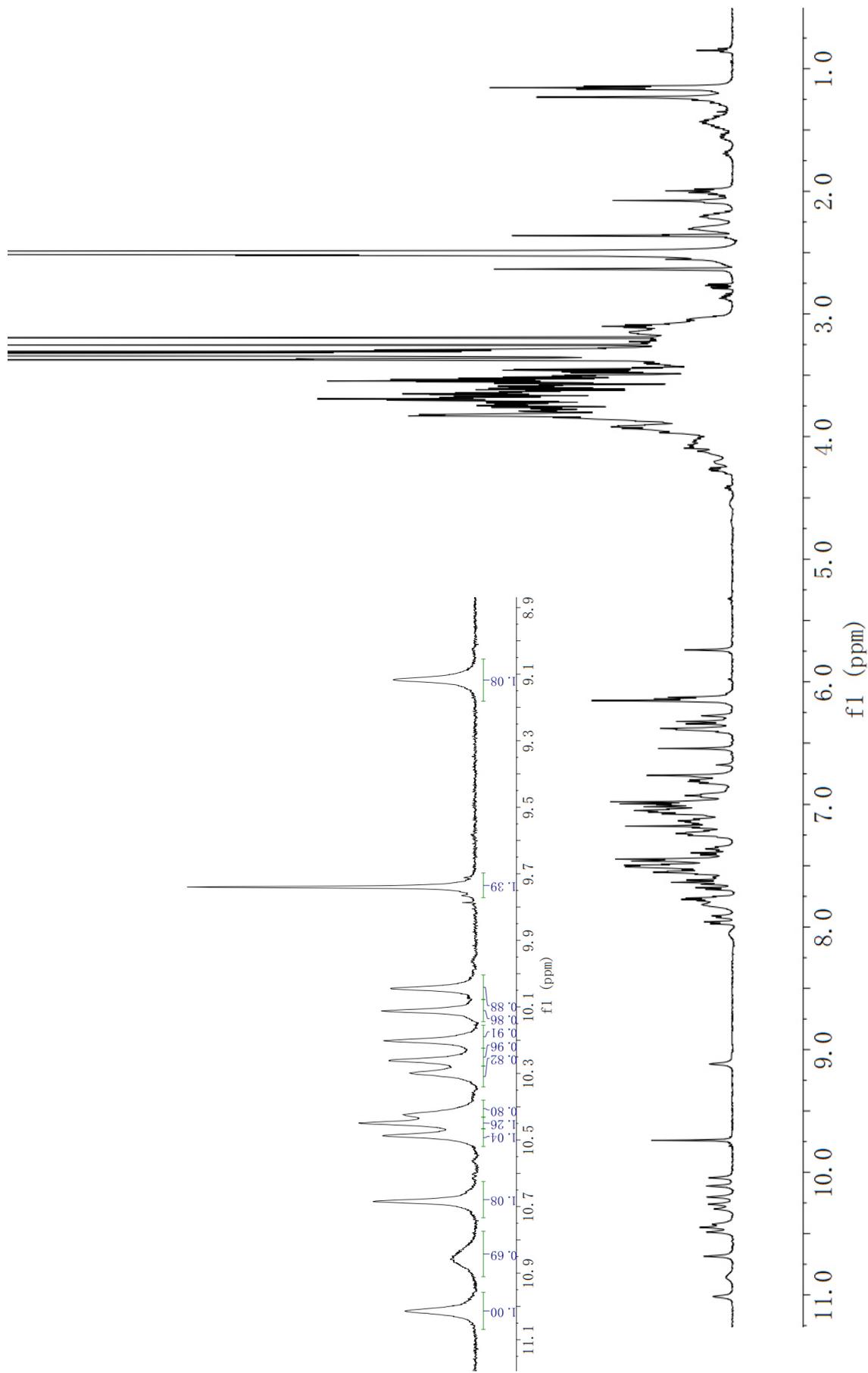




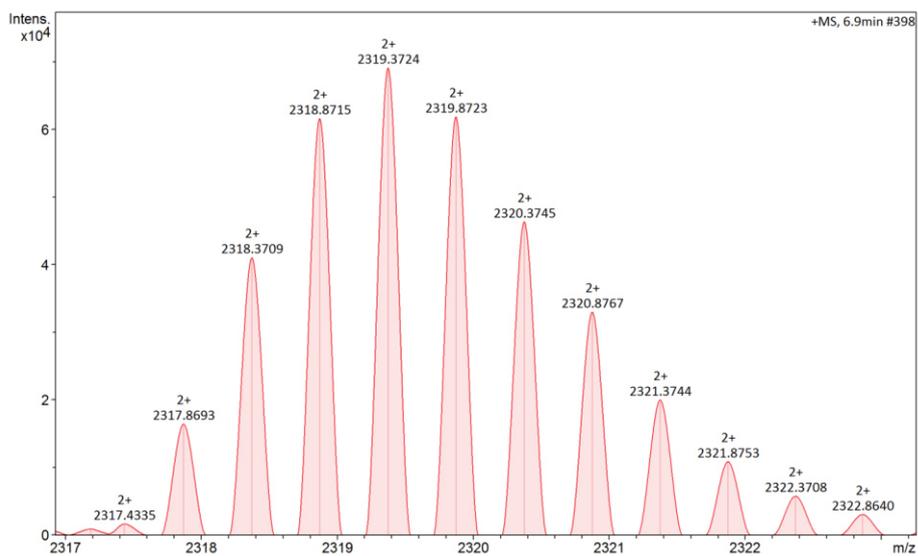
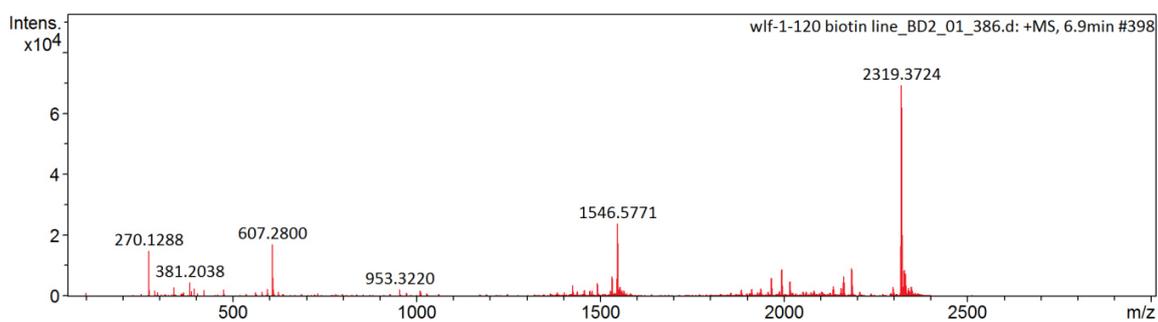
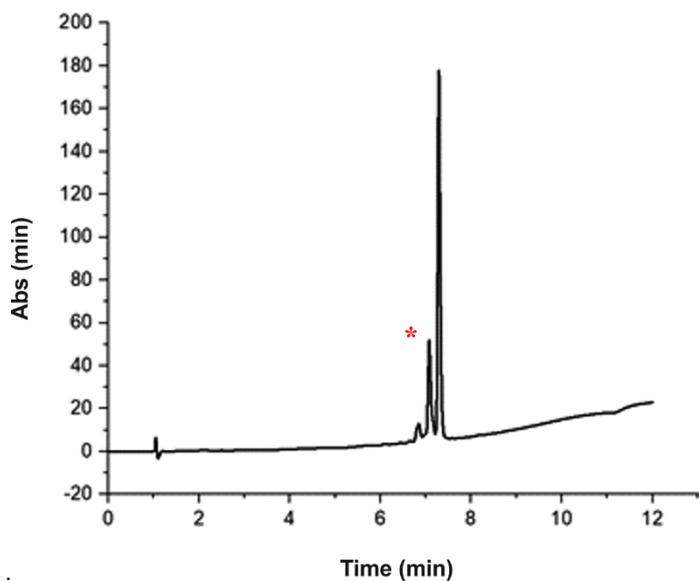
Full ¹H NMR spectra (500 MHz, DMF-*d*₇) of **1d** and zoomed region, monitoring at time point 0 h (blue), 18 h (green), 42 h (black), 25 °C. The red diamond indicates signal of the CH₃ group of chiral **B** unit from diastereomeric conformer. The same process was followed for **1e**.



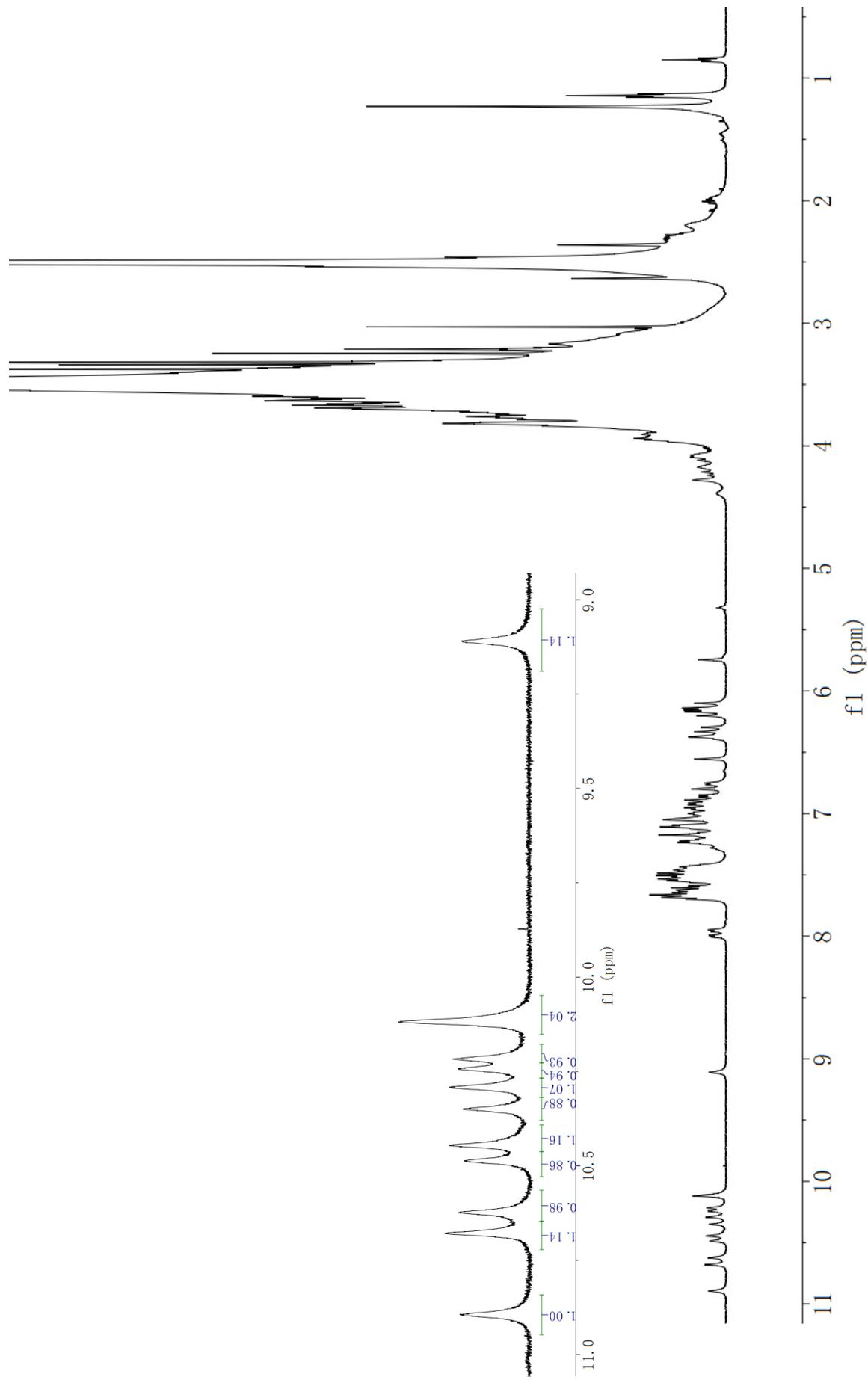
Analytical HPLC profile of pure foldamer **2a** (Gradient: 0 – 20% B over 17 min, then 20 – 100% B over 3 min, then 100% B for 5 min., $\lambda = 254$ nm) and ESI⁺-HRMS spectrum of pure foldamer **2a**



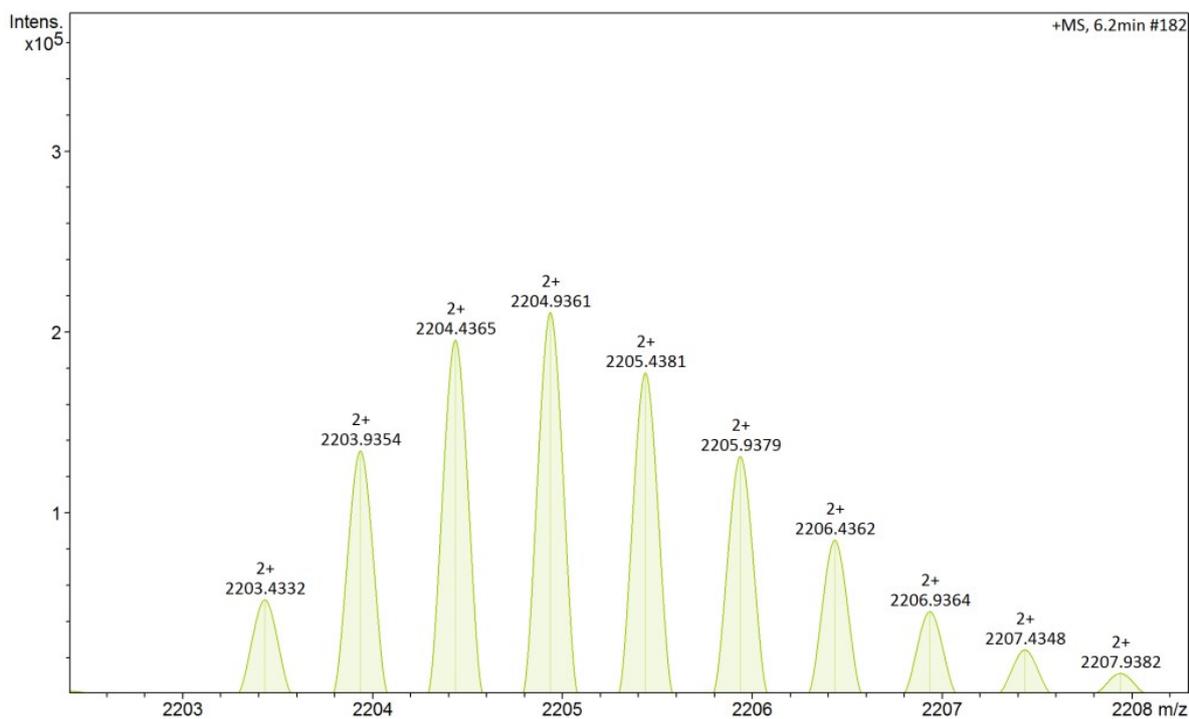
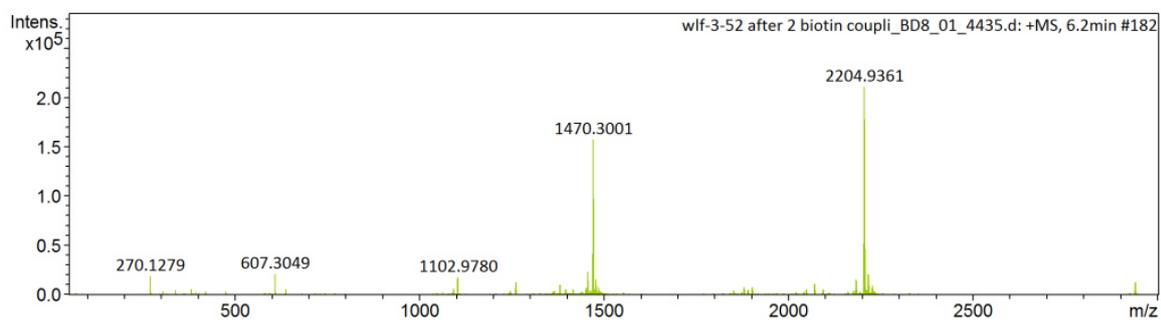
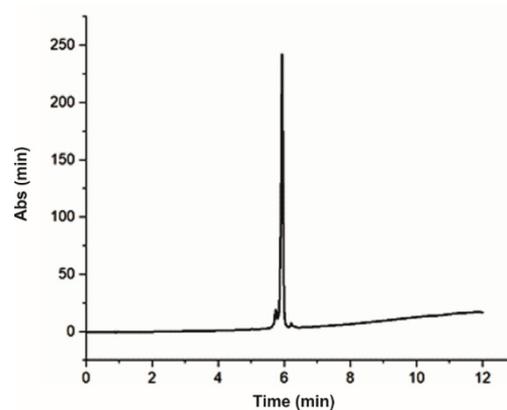
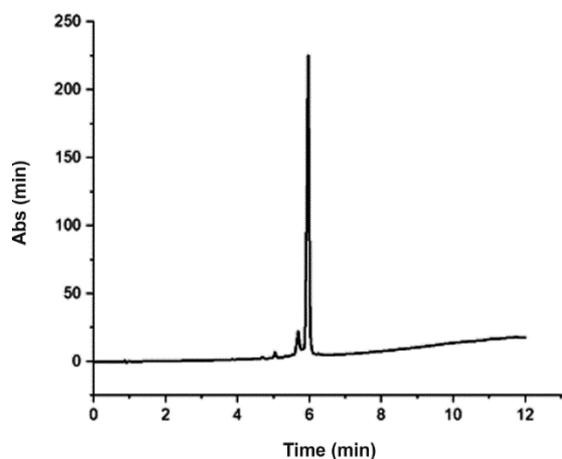
¹H NMR spectrum of foldamer **2a** and zoomed region (500 MHz, DMSO-*d*₆), 25 °C.



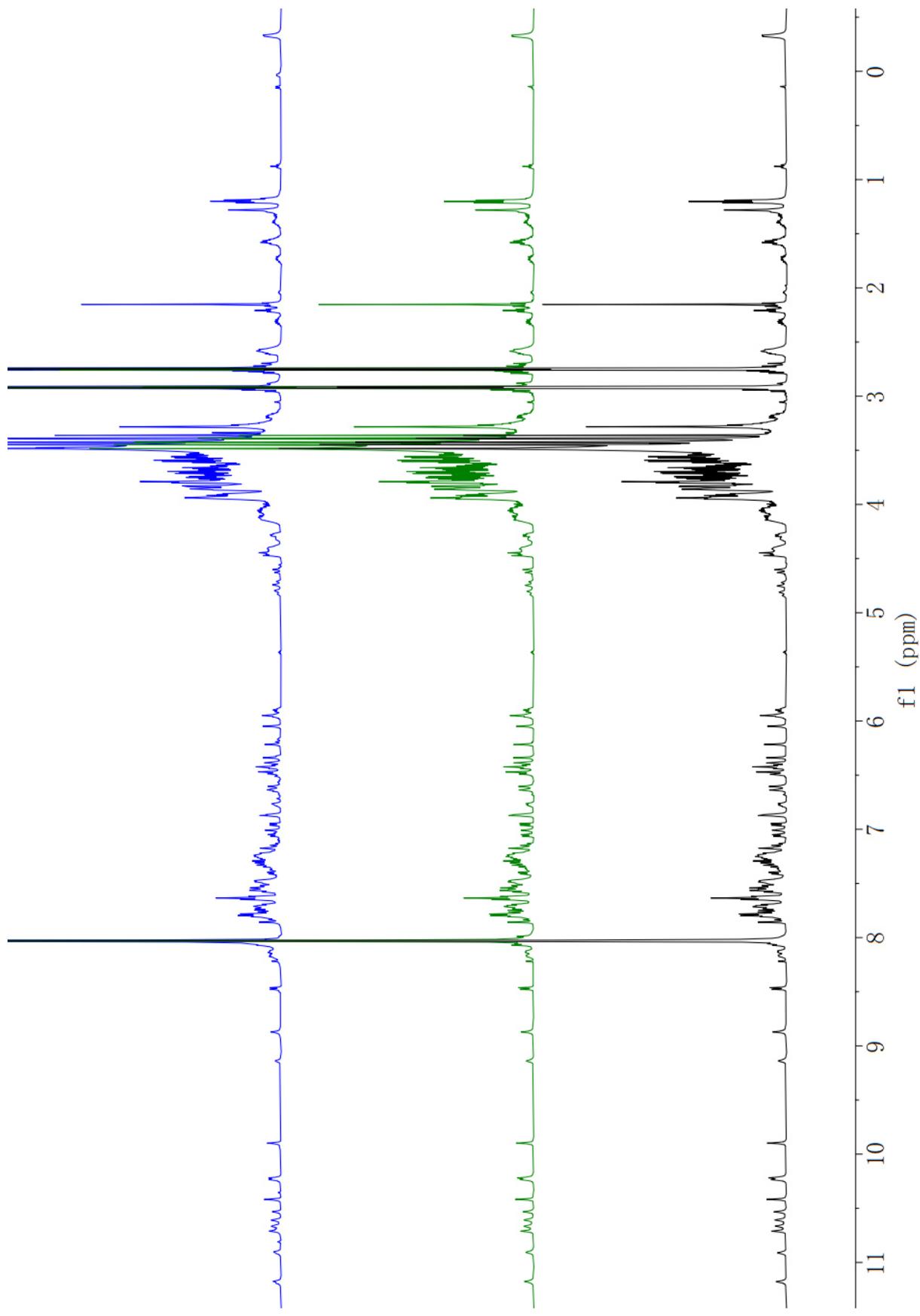
Analytical HPLC (10-100% B over 10 min, 254 nm) profile of foldamer **2b**. HRMS (ESI+) profile of foldamer **2b**. The * annotation corresponds to the biotin-oxidized derivative.

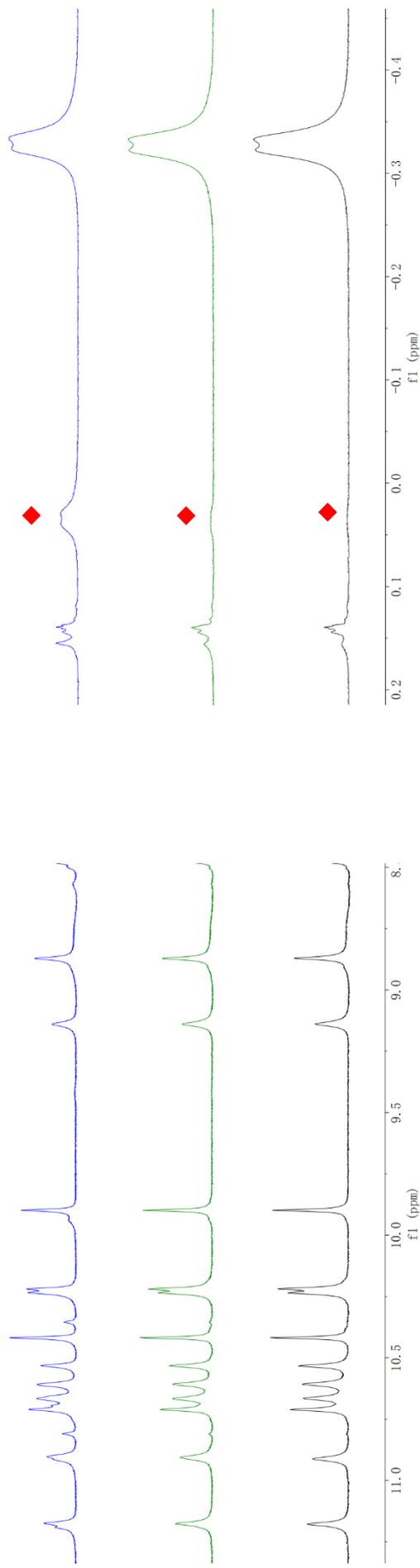


^1H NMR Spectrum of foldamer **2b** and zoomed region (500 MHz, $\text{DMSO-}d_6$, $25\text{ }^\circ\text{C}$).

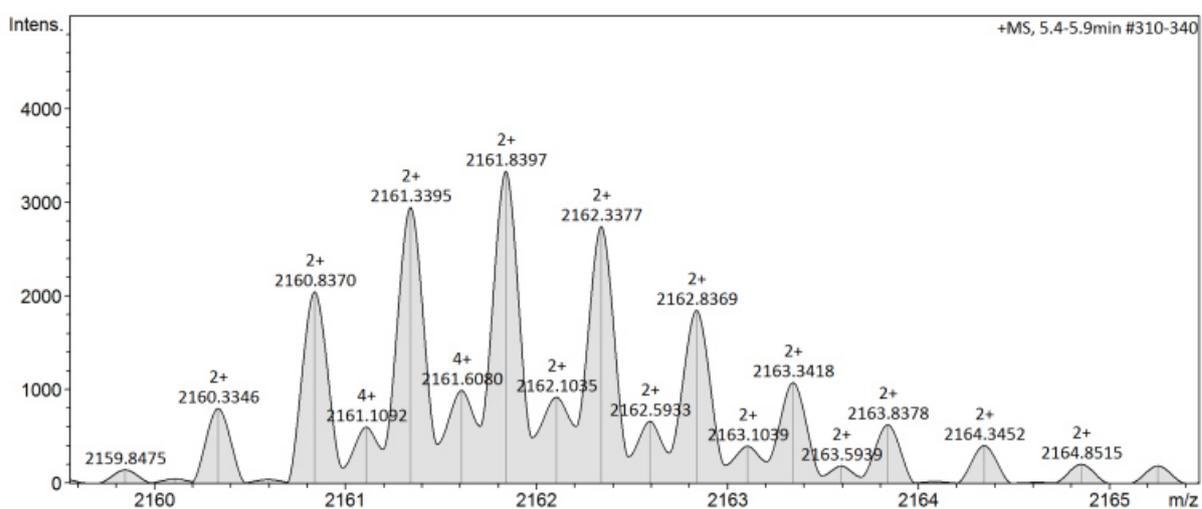
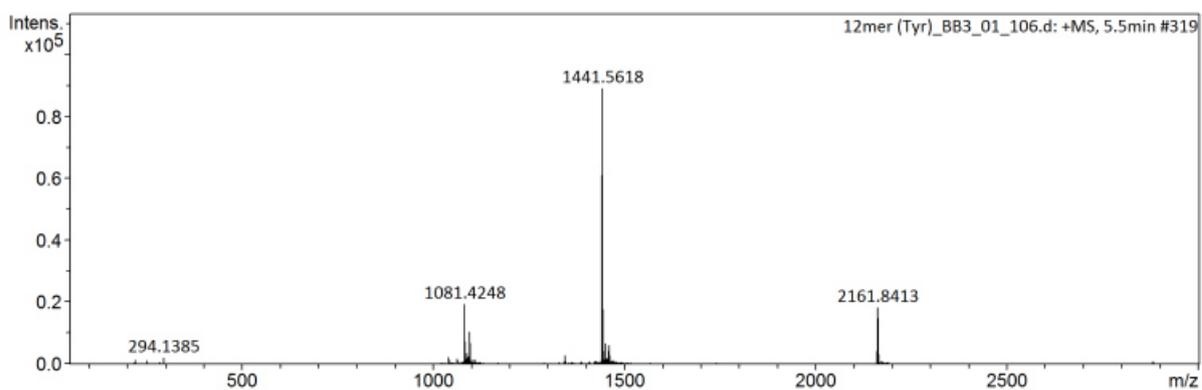
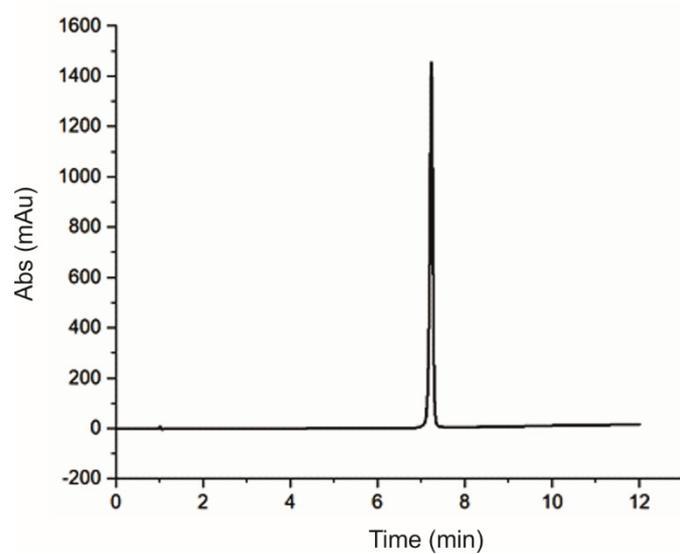


Analytical RP-HPLC profile of pure **2c** (top left) and **2d** (top right) (10-100% B over 10 min). ¹H NMR Spectrum (500 MHz, DMF-*d*₇, 25 °C), monitoring at time point 0 h (blue), 18 h (green), 42 h (black), 25 °C) of **2c**. HRMS (ESI+) profile of foldamer **2c** (ESI-HRMS spectra for **2d** was omitted since they are enantiomers)..

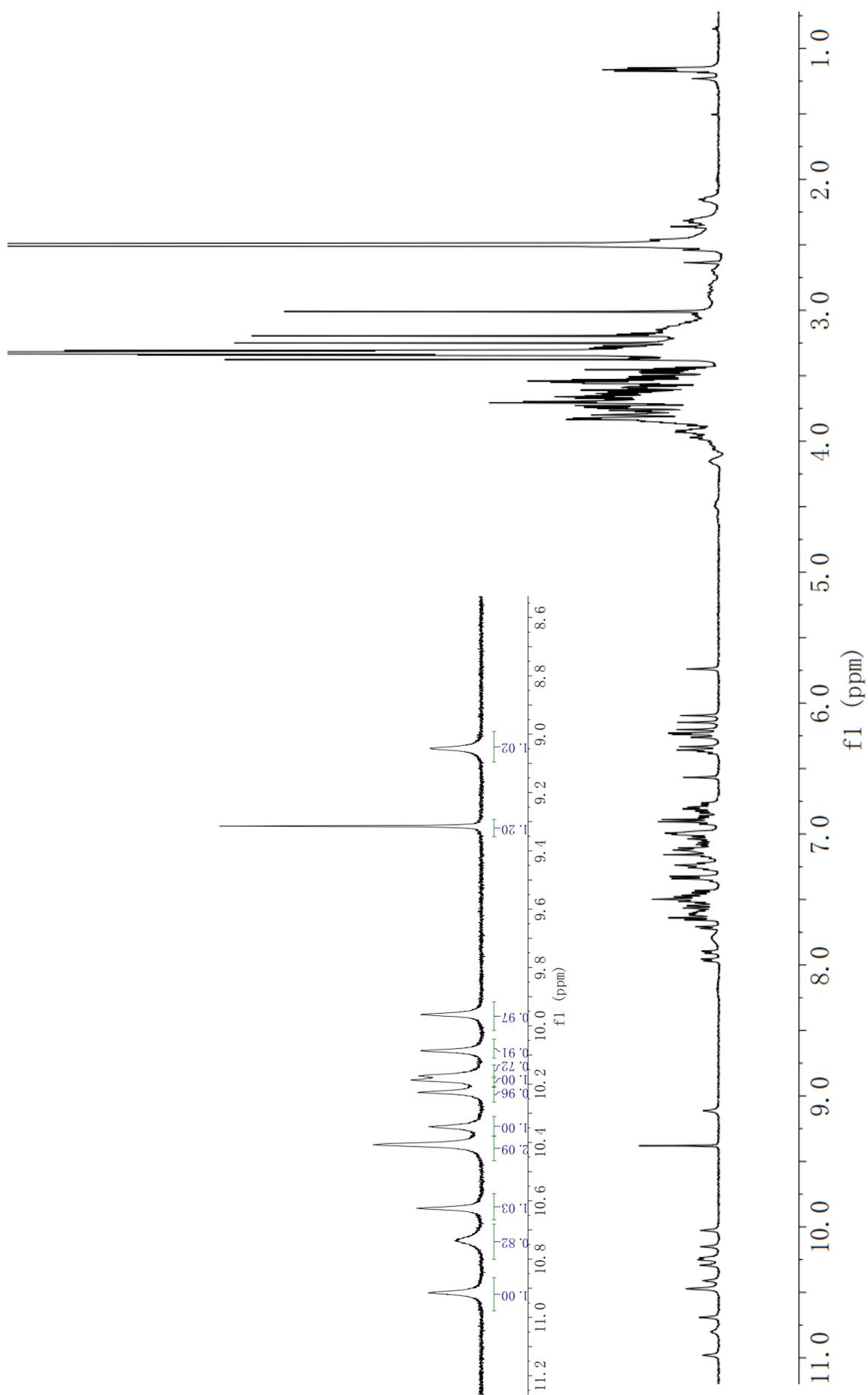




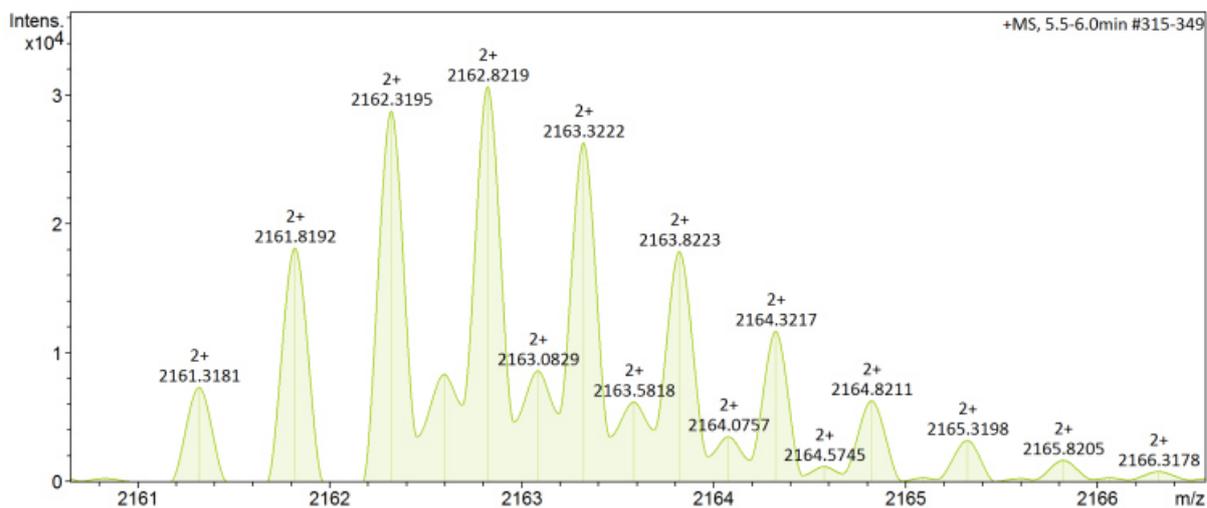
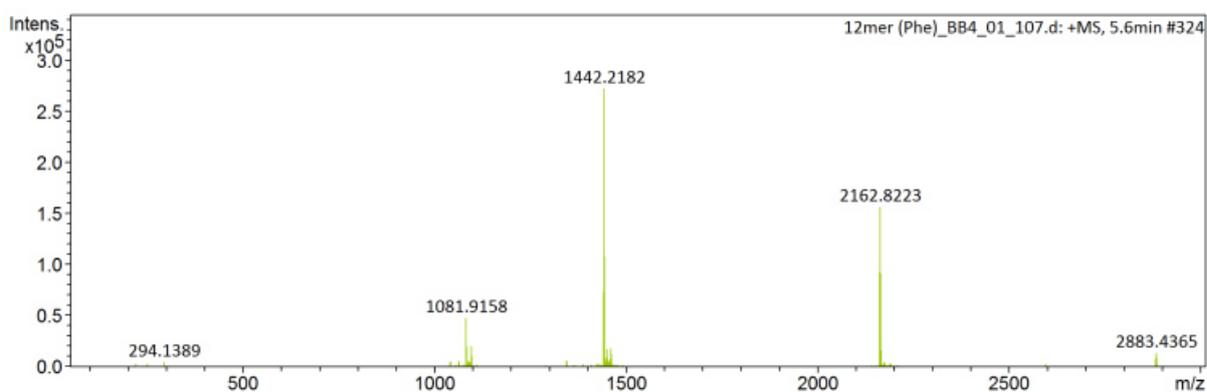
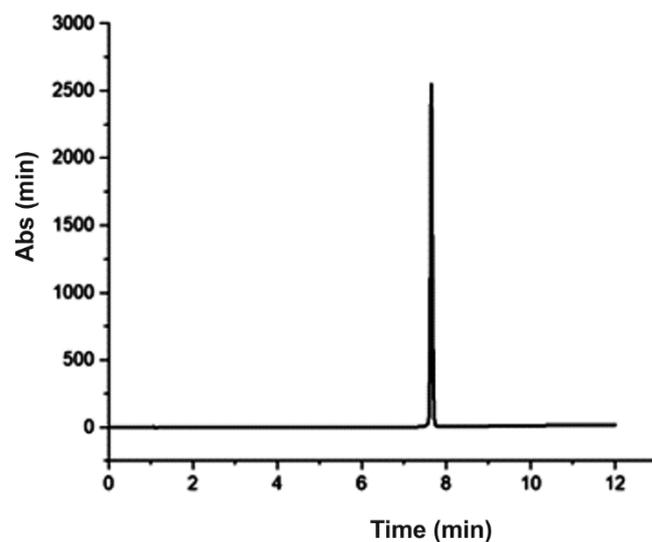
Full ¹H NMR spectra (500 MHz, DMF-d₇) of **2c** and zoomed region, monitoring at time point 0 h (blue), 18 h (green), 42 h (black), 25 °C. The red diamond indicates signal of the CH₃ group of chiral **B** unit from diastereomeric conformer. The same process was followed for **2d**.



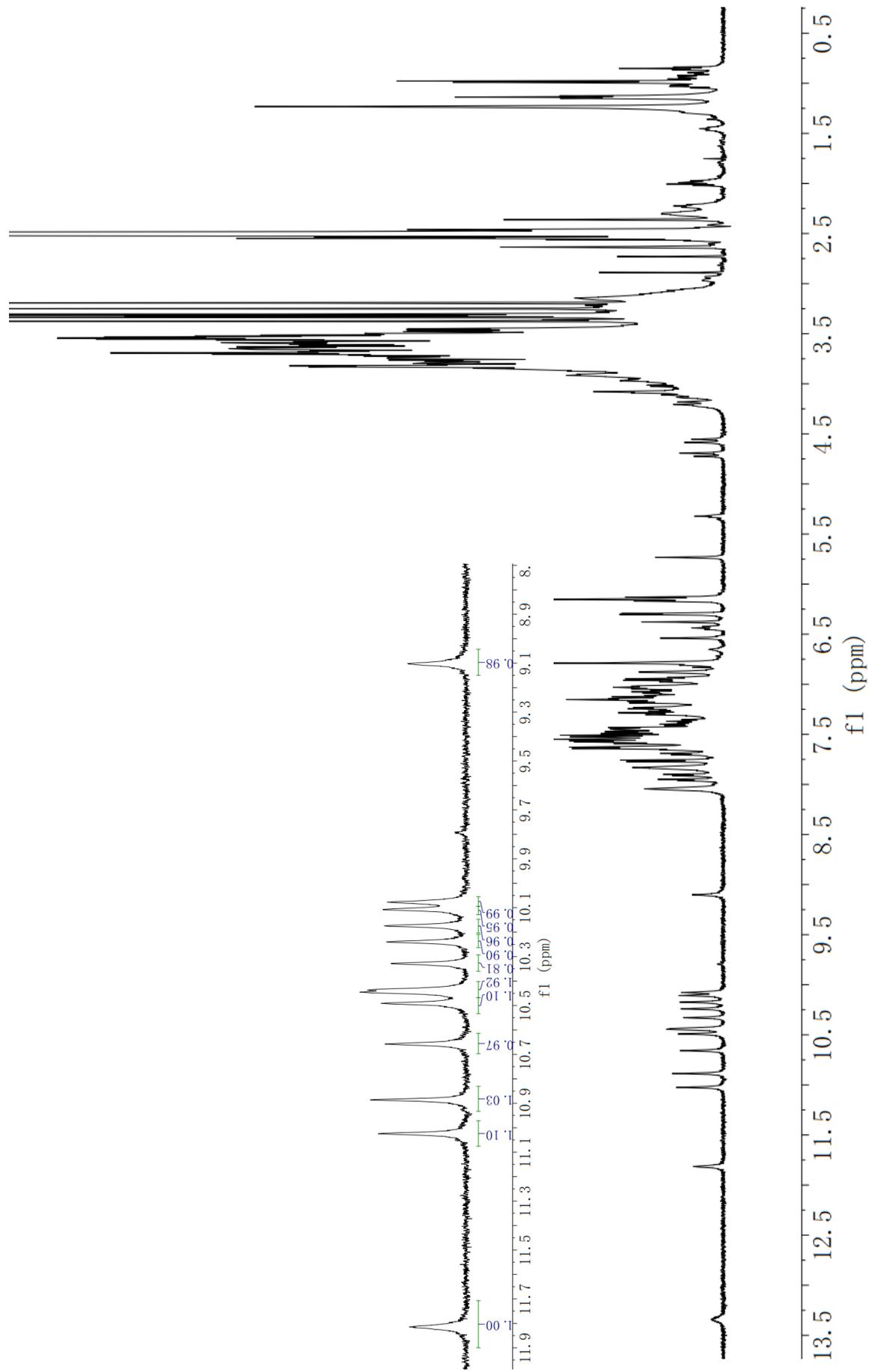
Analytical RP-HPLC (10-100% B over 10 min, 254 nm) profile of **2e**. HRMS (ESI⁺) spectrum of foldamer **2e**.



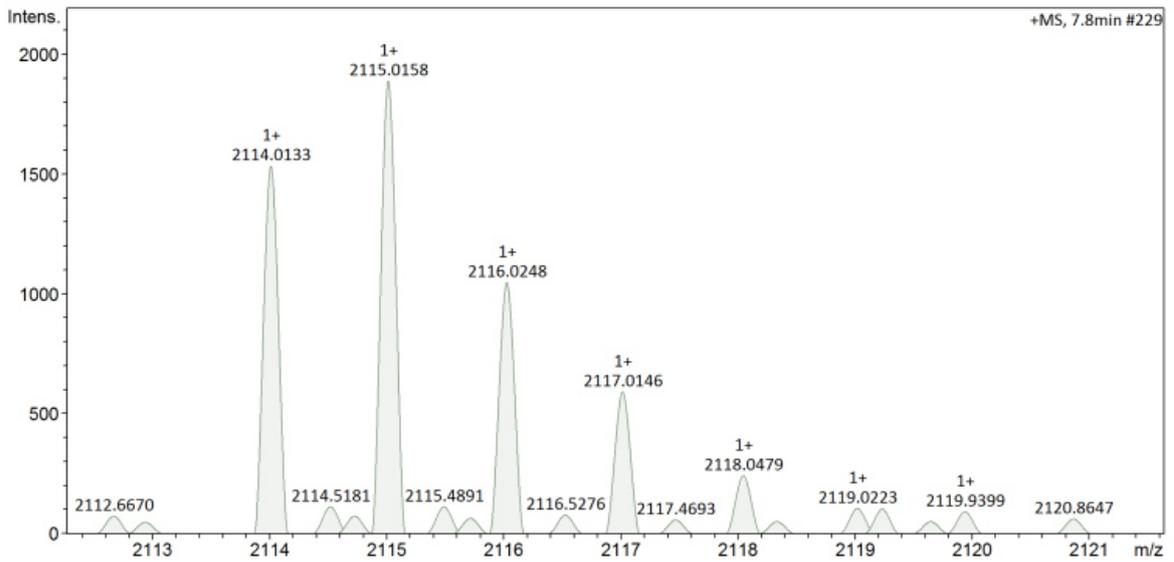
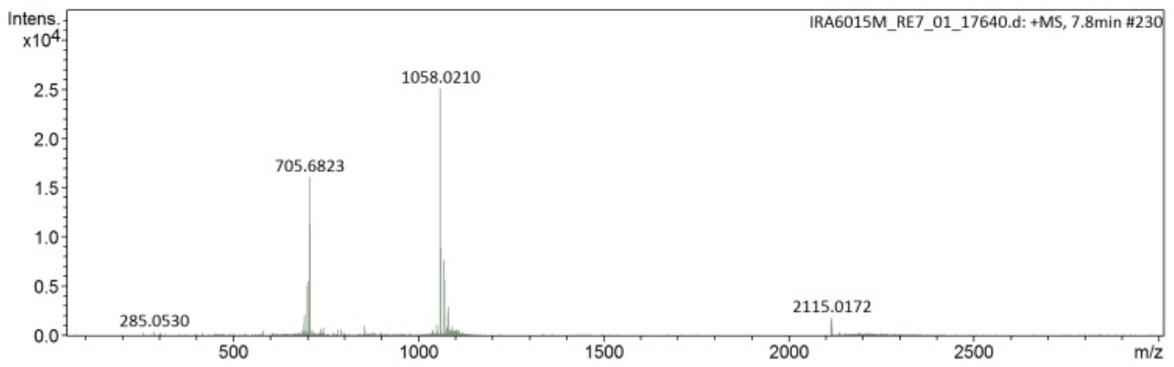
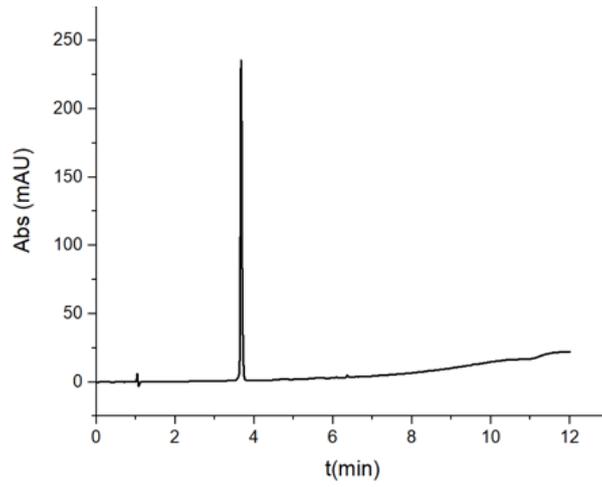
¹H NMR Spectrum of foldamer **2e** (500 MHz, DMSO-*d*₆) and zoomed region, 25 °C.



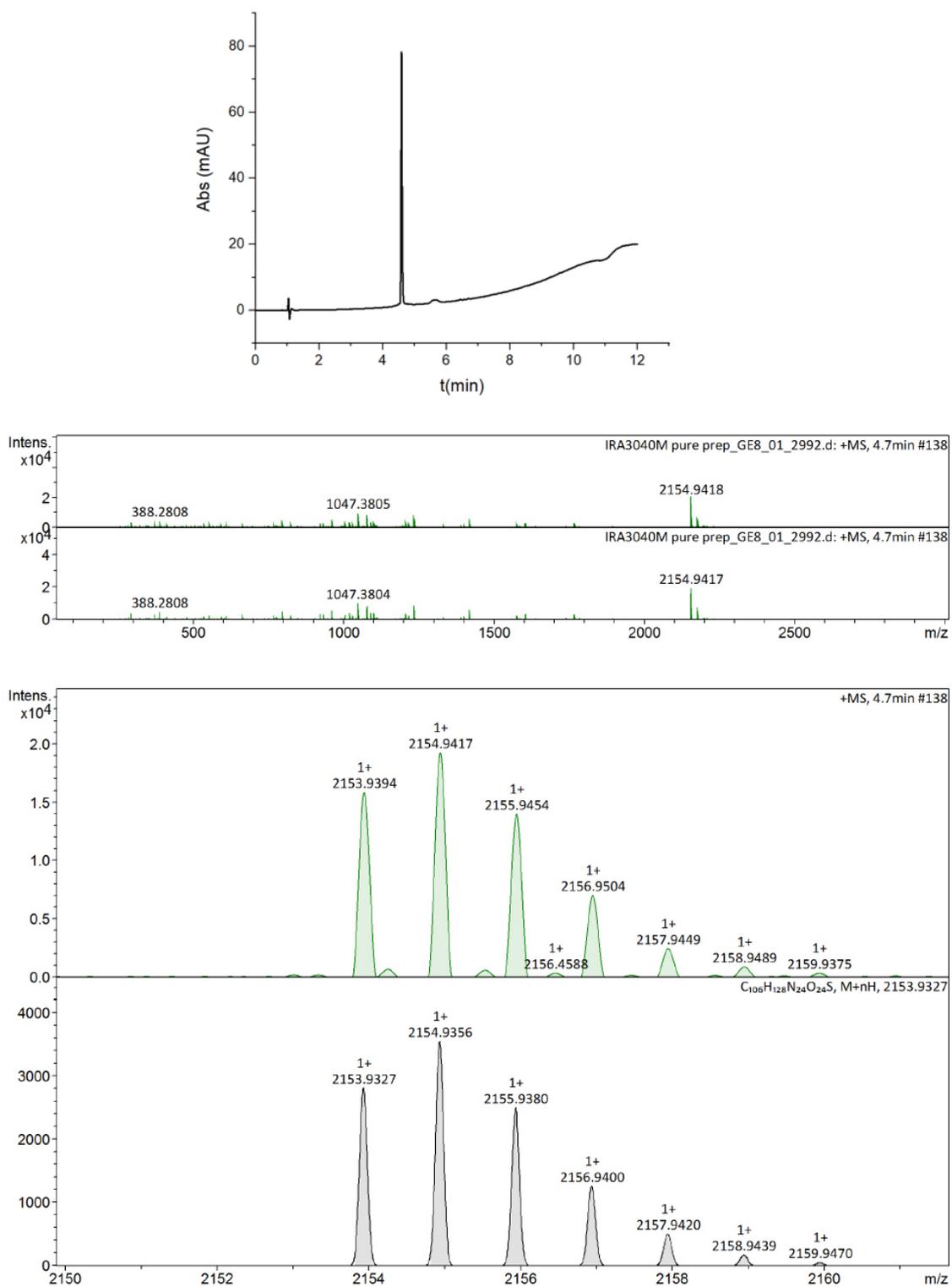
Analytical HPLC (10-100% B over 10 min, 254 nm) profile of foldamer **2f**. HRMS (ESI+) profile of foldamer **2f**.



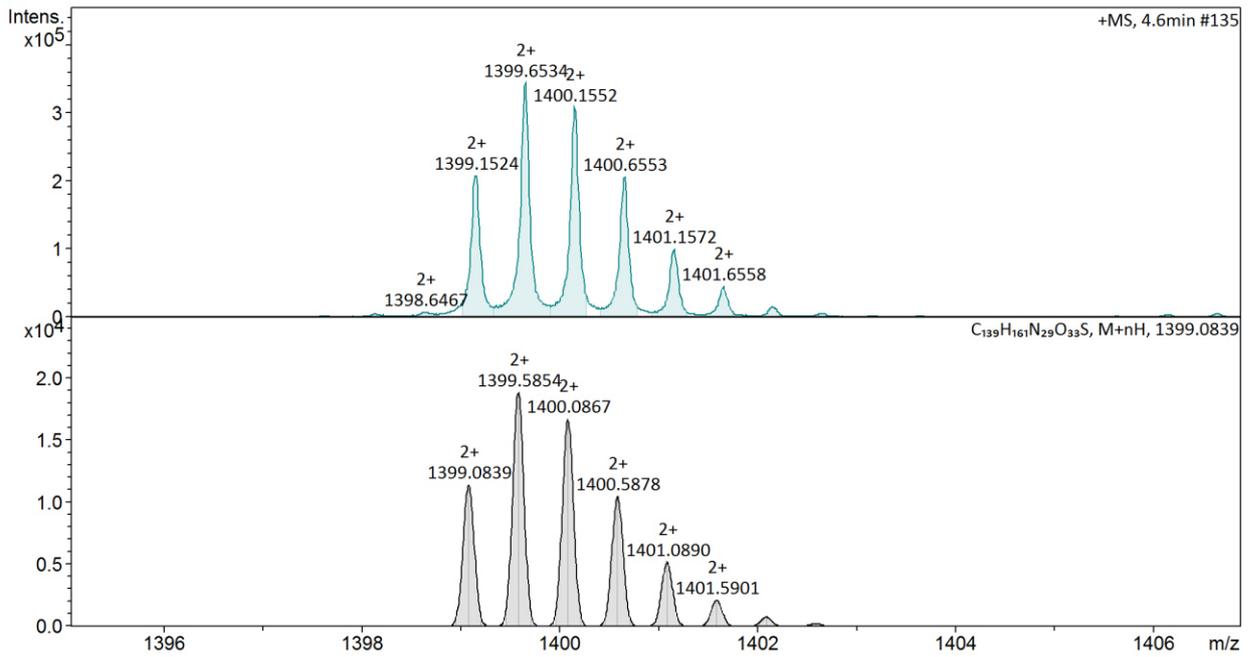
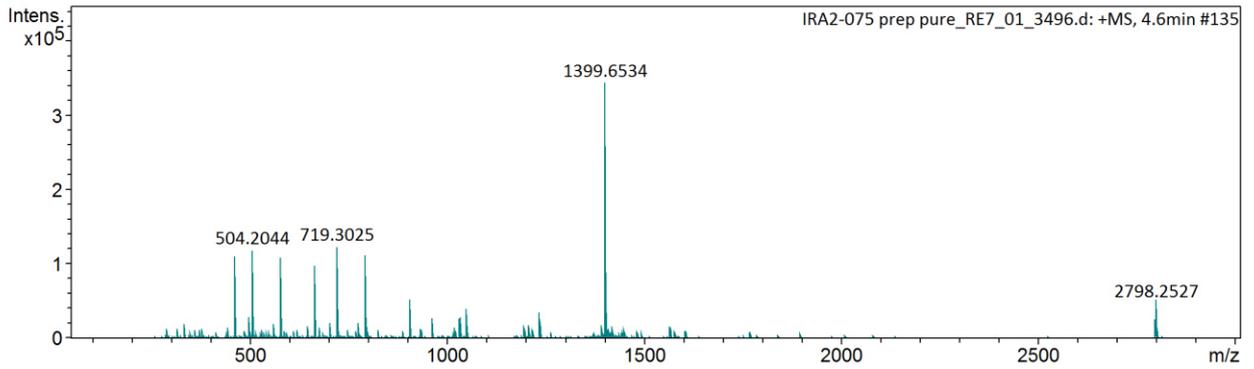
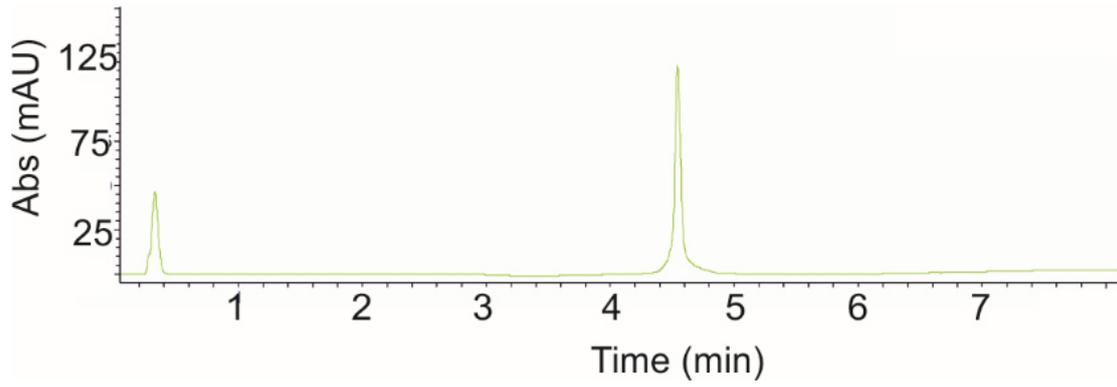
^1H NMR Spectrum of foldamer **2f** (500 MHz, $\text{DMSO-}d_6$) and zoomed region $25\text{ }^\circ\text{C}$.



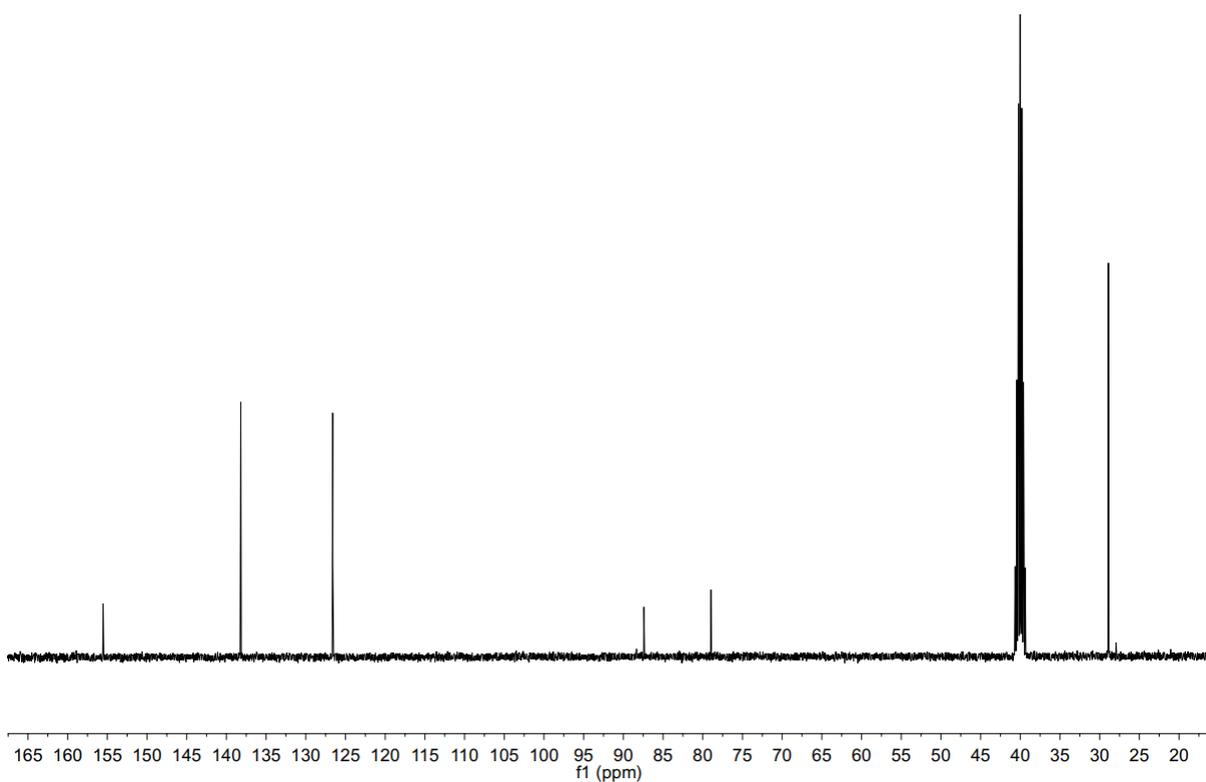
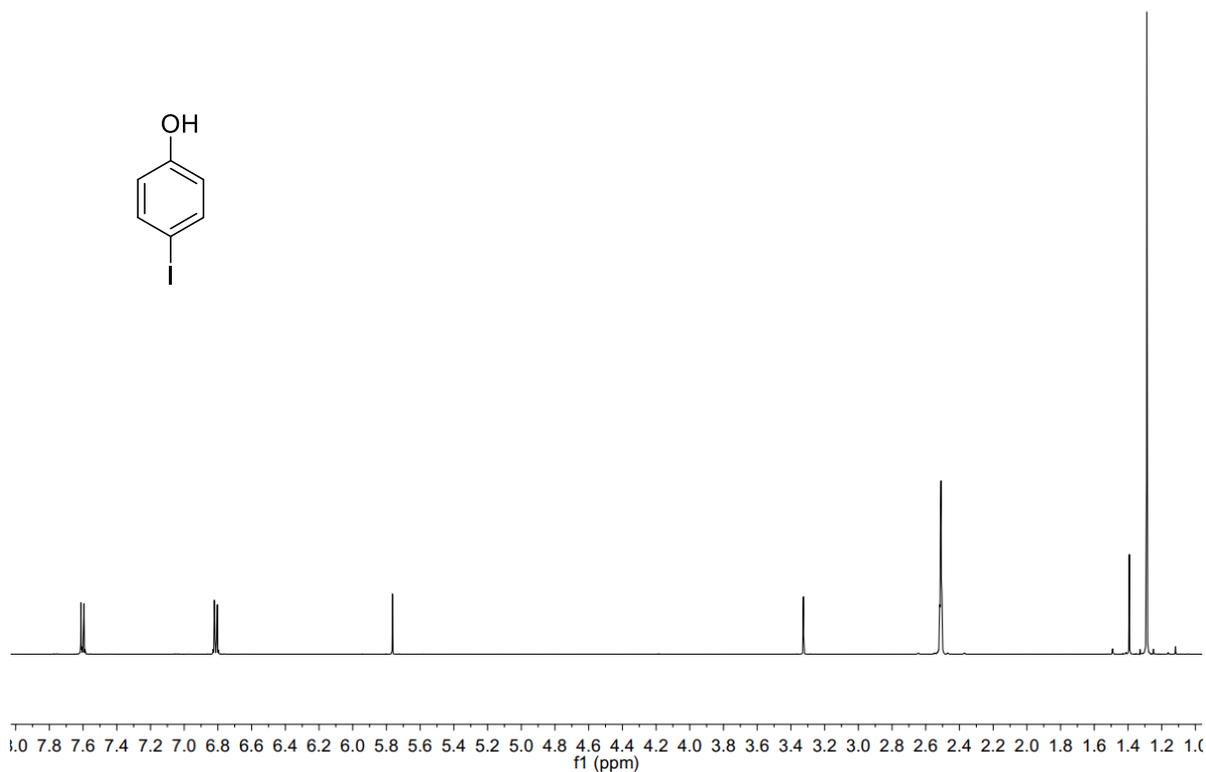
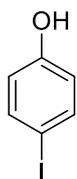
RP-HPLC profile and HRMS (ESI+) spectra of peptide **3a**



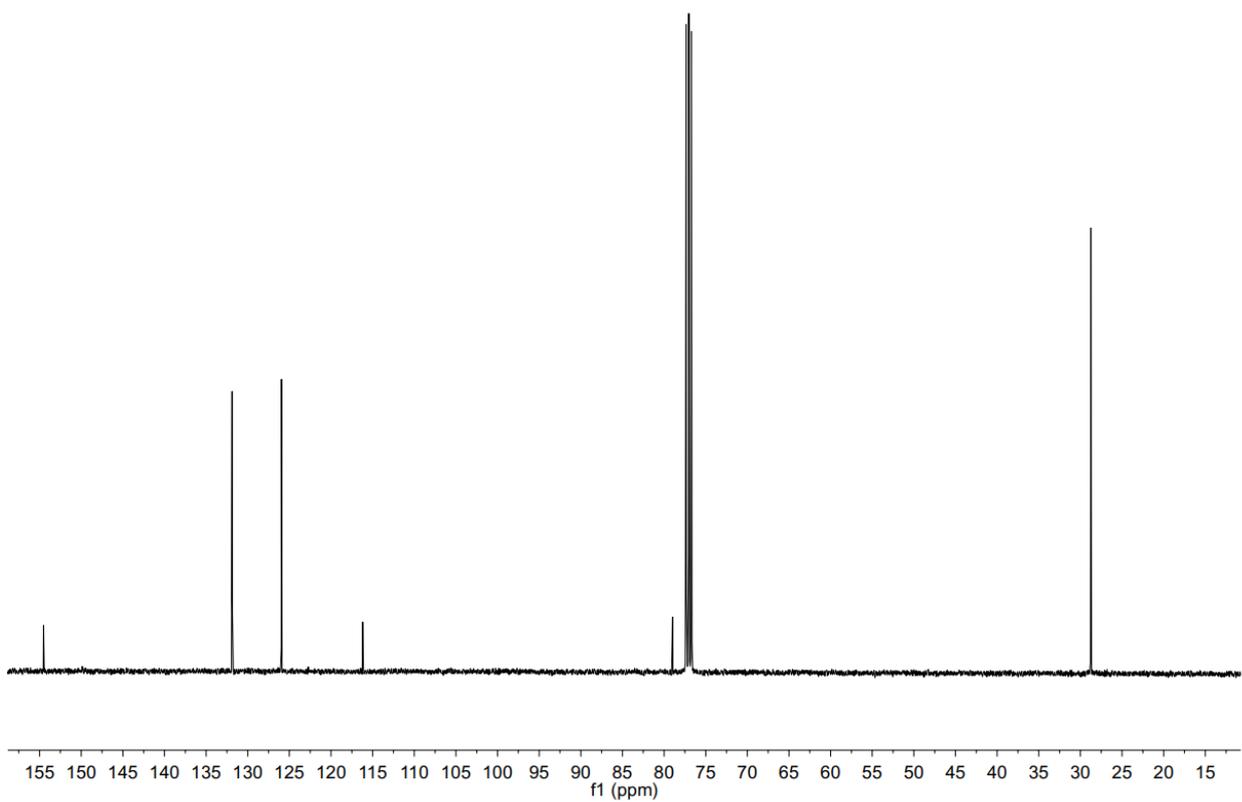
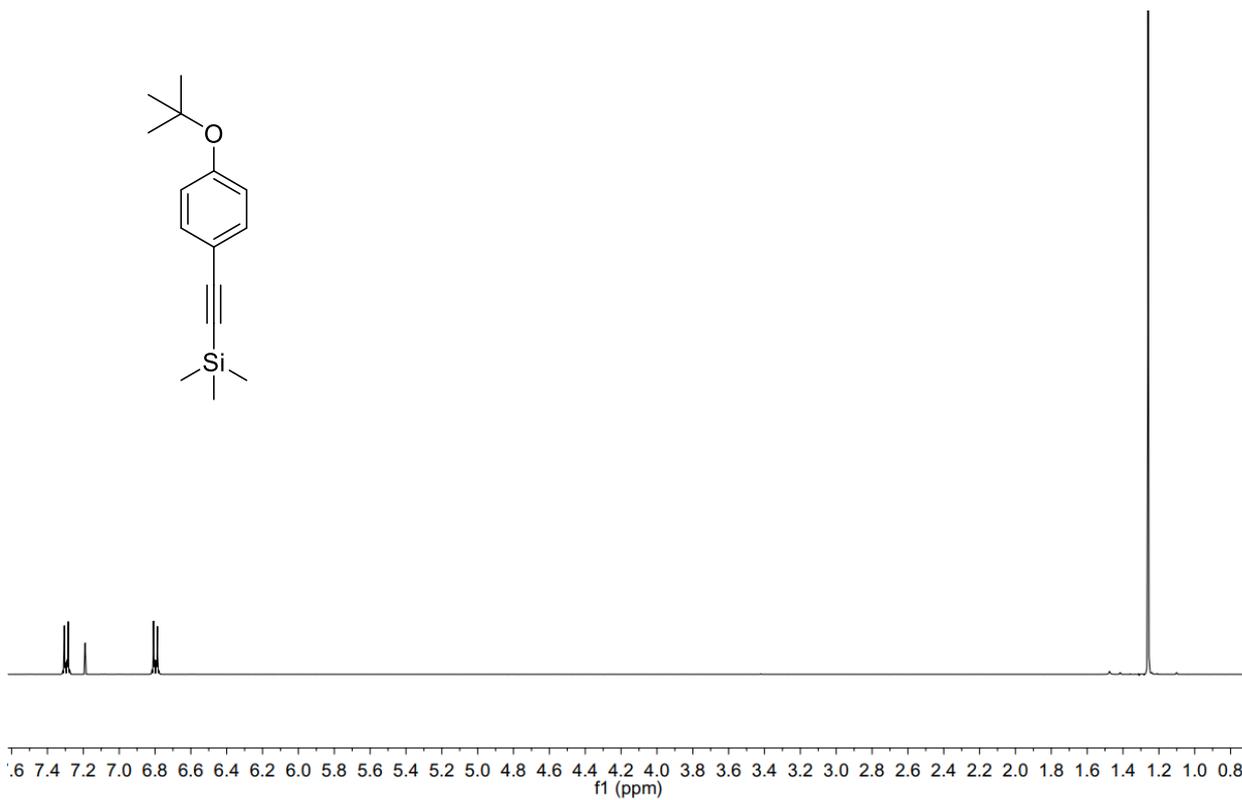
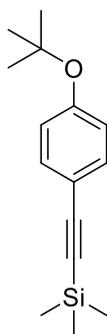
RP-HPLC profile and HRMS (ESI+) spectra of peptide **4a**



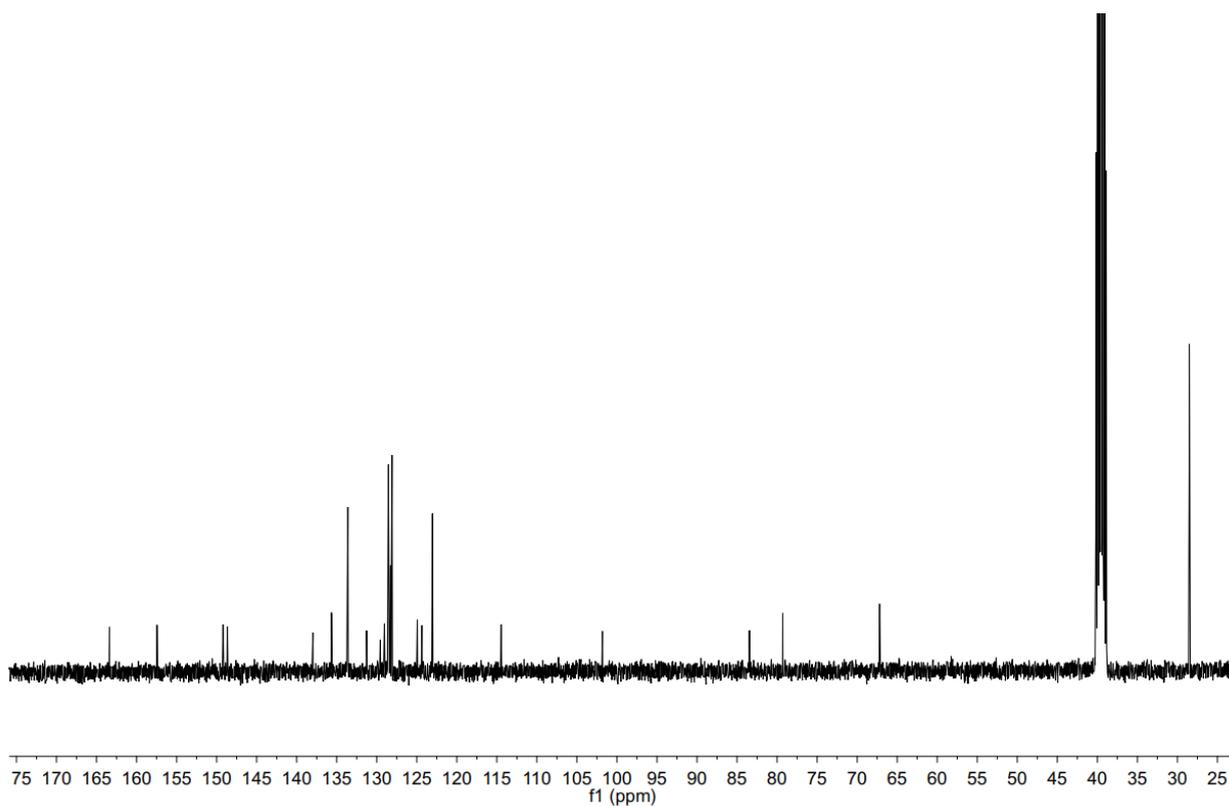
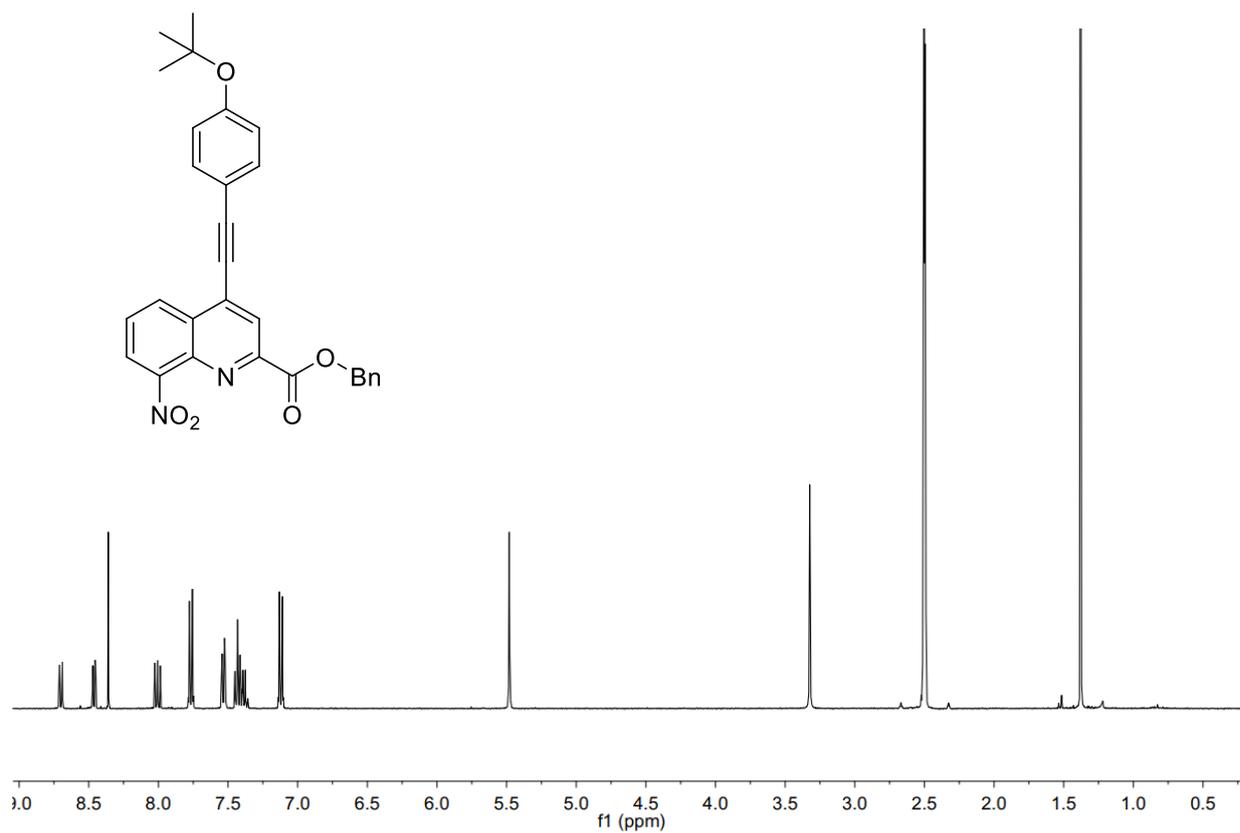
RP-HPLC profile and HRMS (ESI+) spectra of peptide 4c



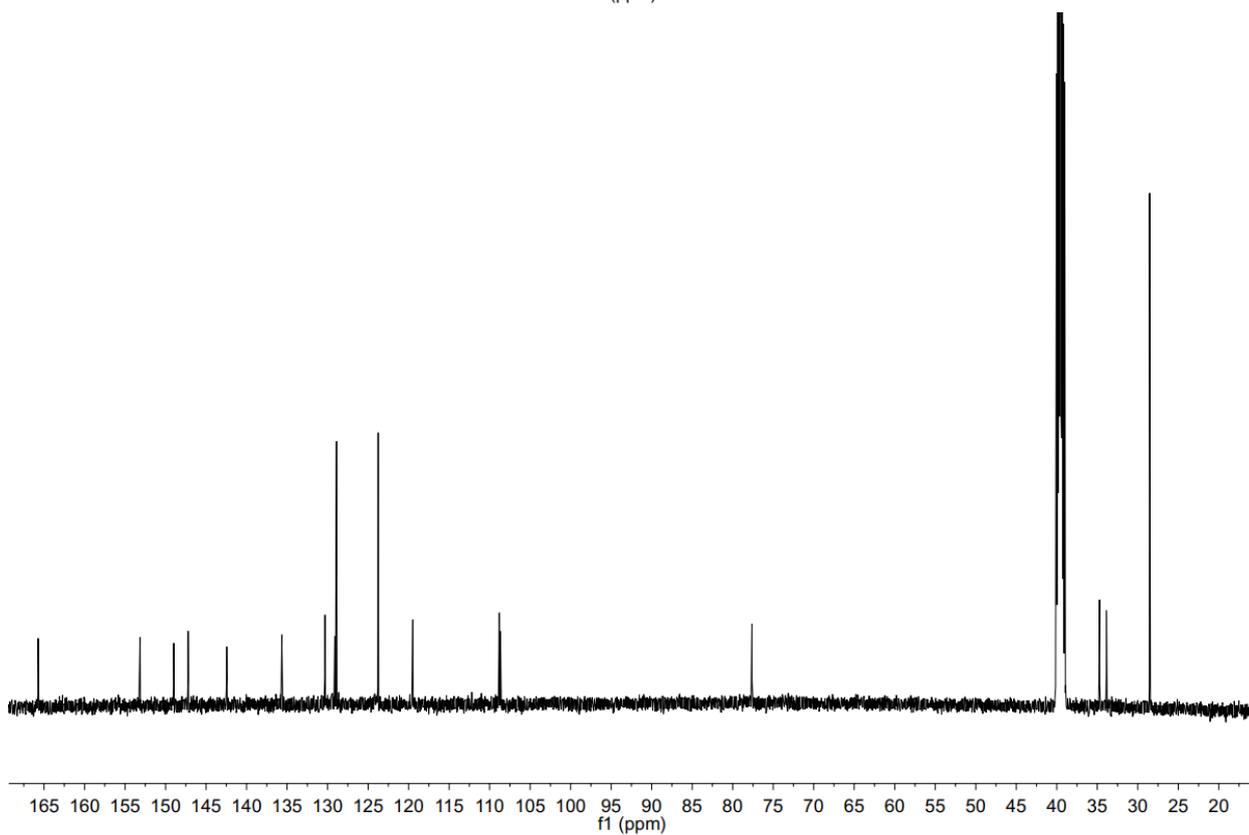
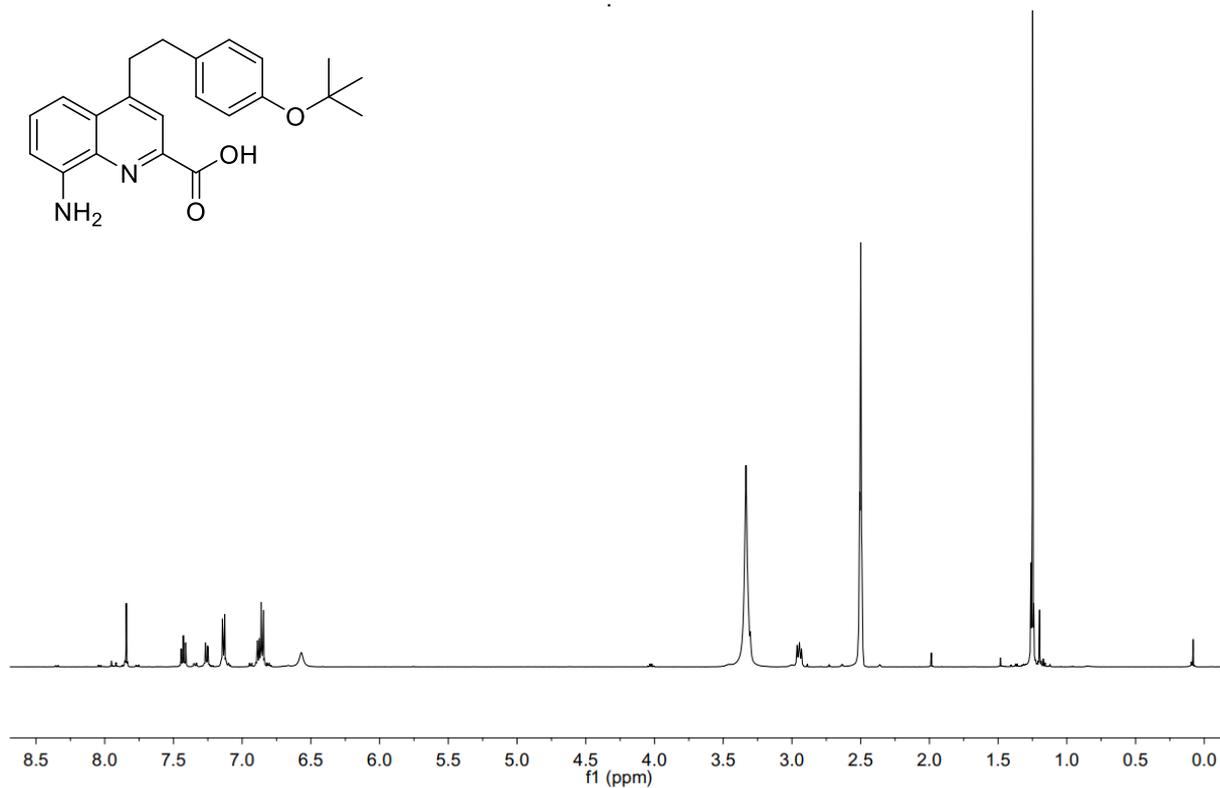
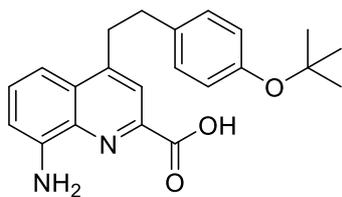
NMR spectra of compound 7: ^1H NMR (500 MHz, DMSO- d_6) and ^{13}C NMR Spectrum (126 MHz, DMSO- d_6)



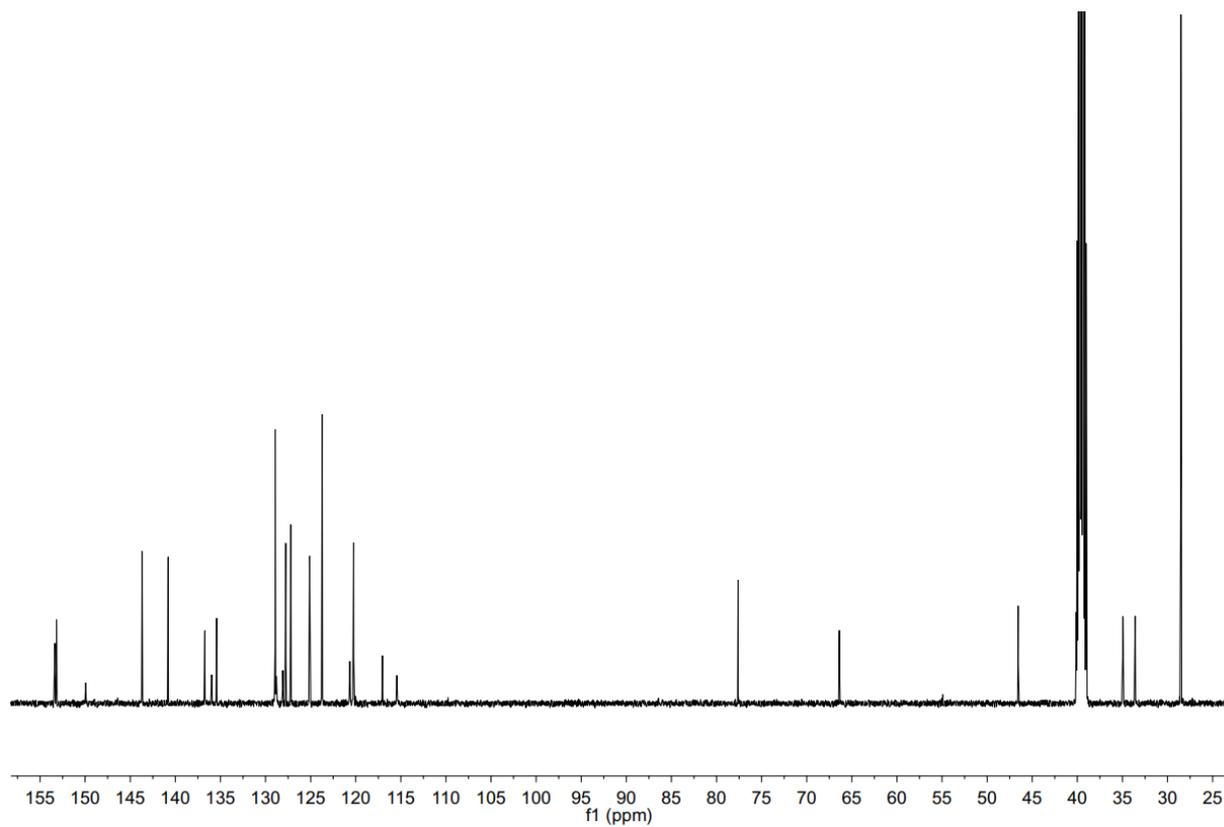
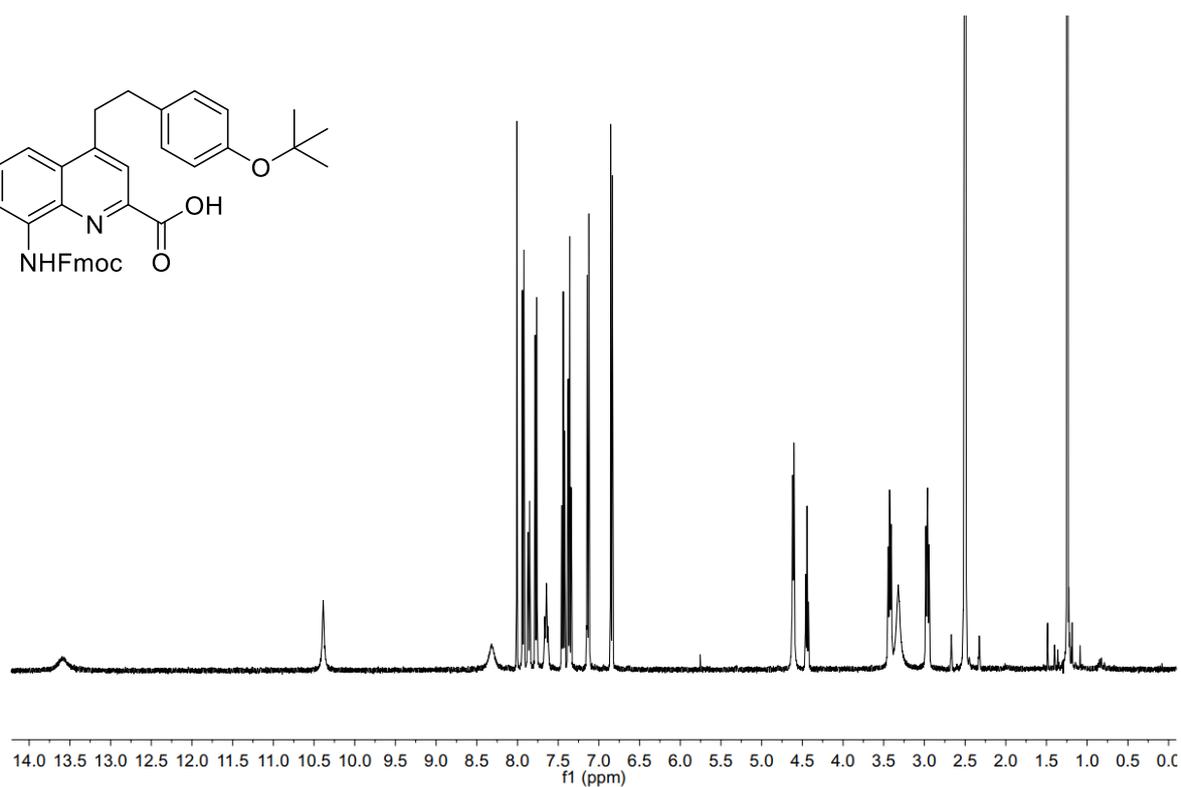
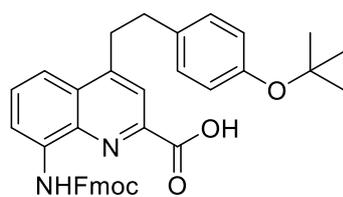
NMR spectra of compound **8**: ¹H NMR Spectrum (500 MHz, CDCl₃) and ¹³C NMR Spectrum (126 MHz, CDCl₃)



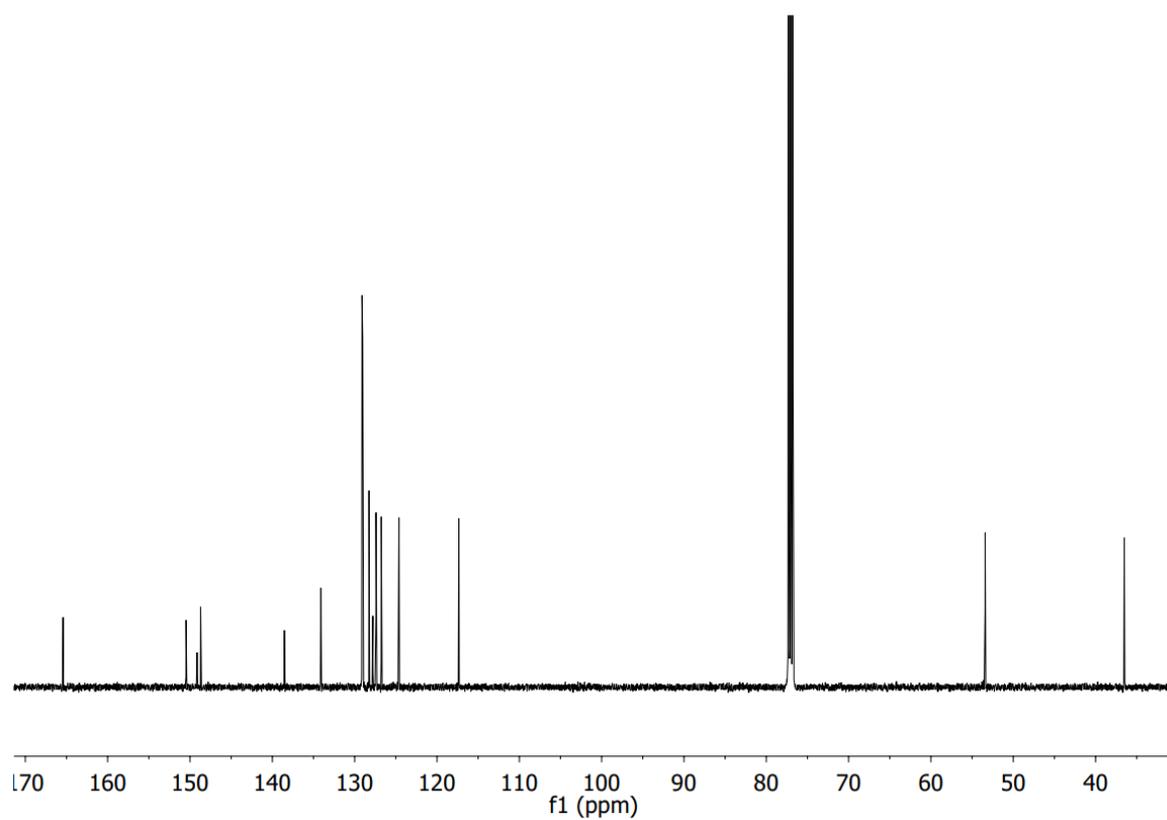
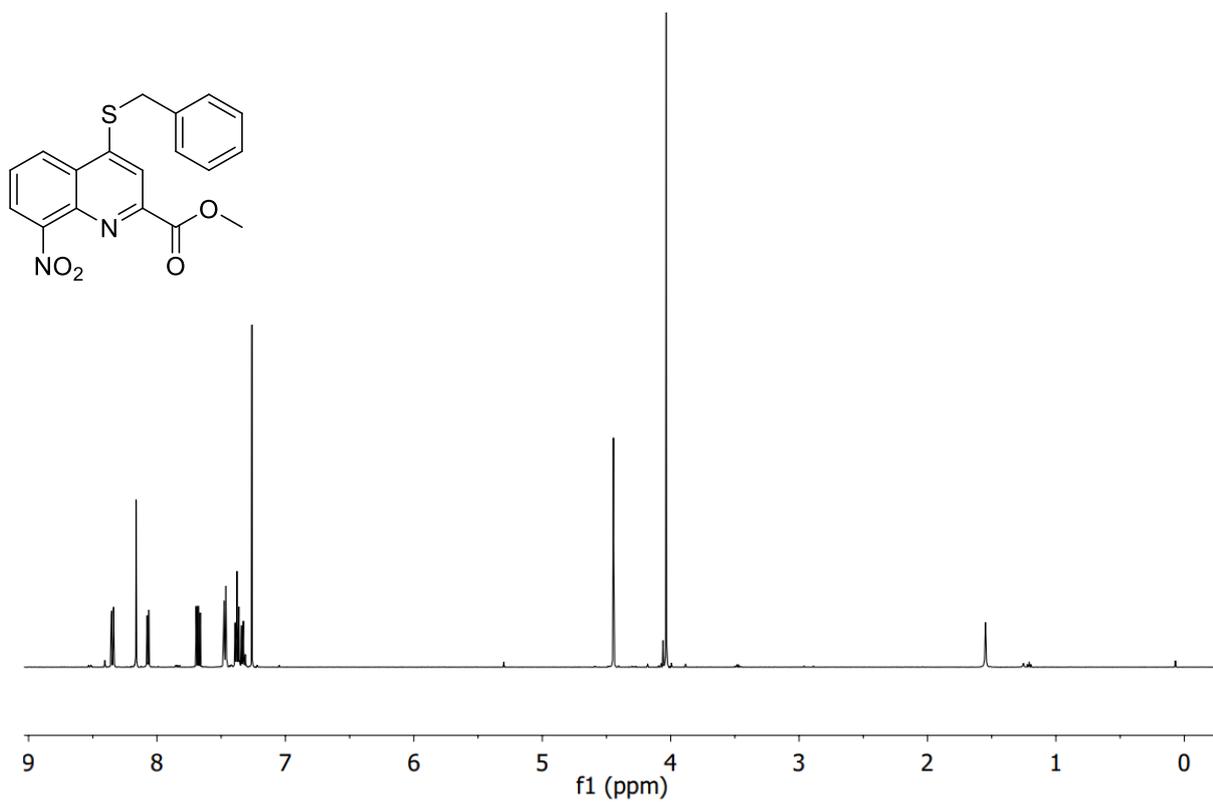
NMR spectra of compound 9: ^1H NMR (500 MHz, $\text{DMSO-}d_6$) and ^{13}C NMR Spectrum (126 MHz, $\text{DMSO-}d_6$)



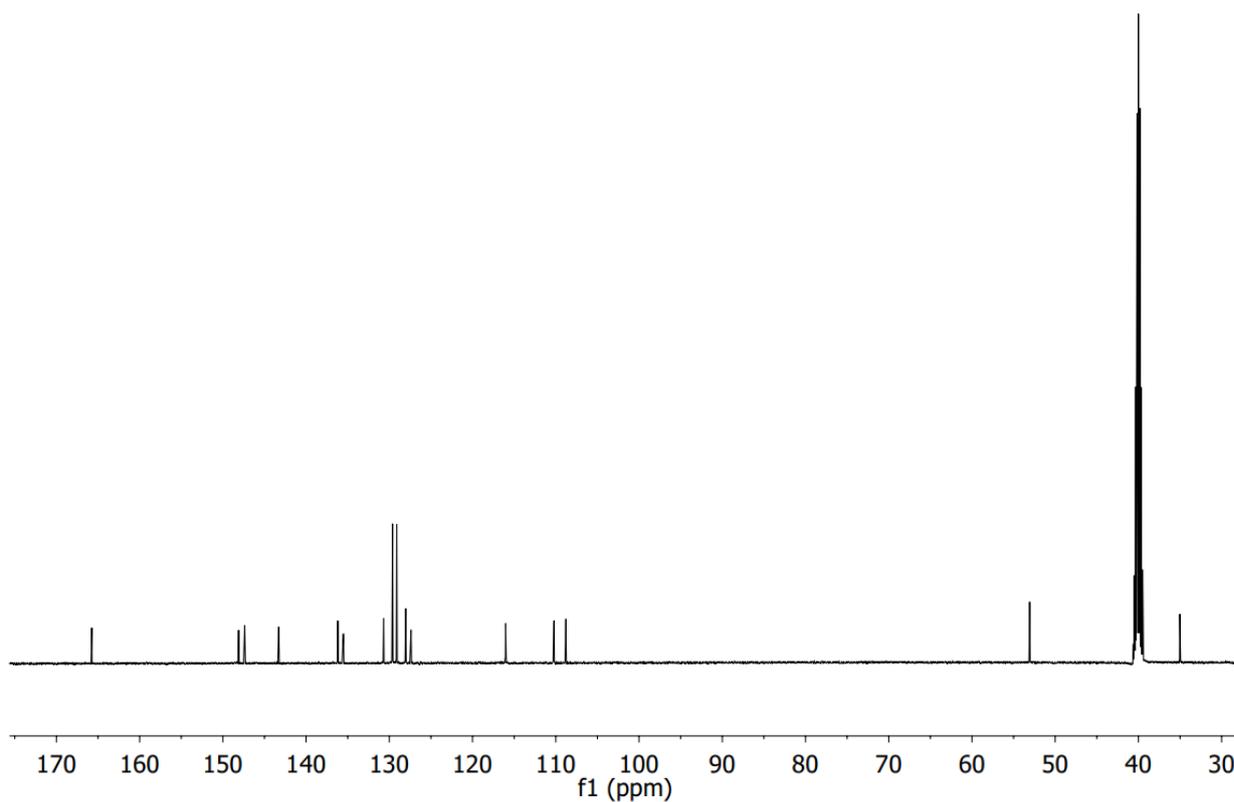
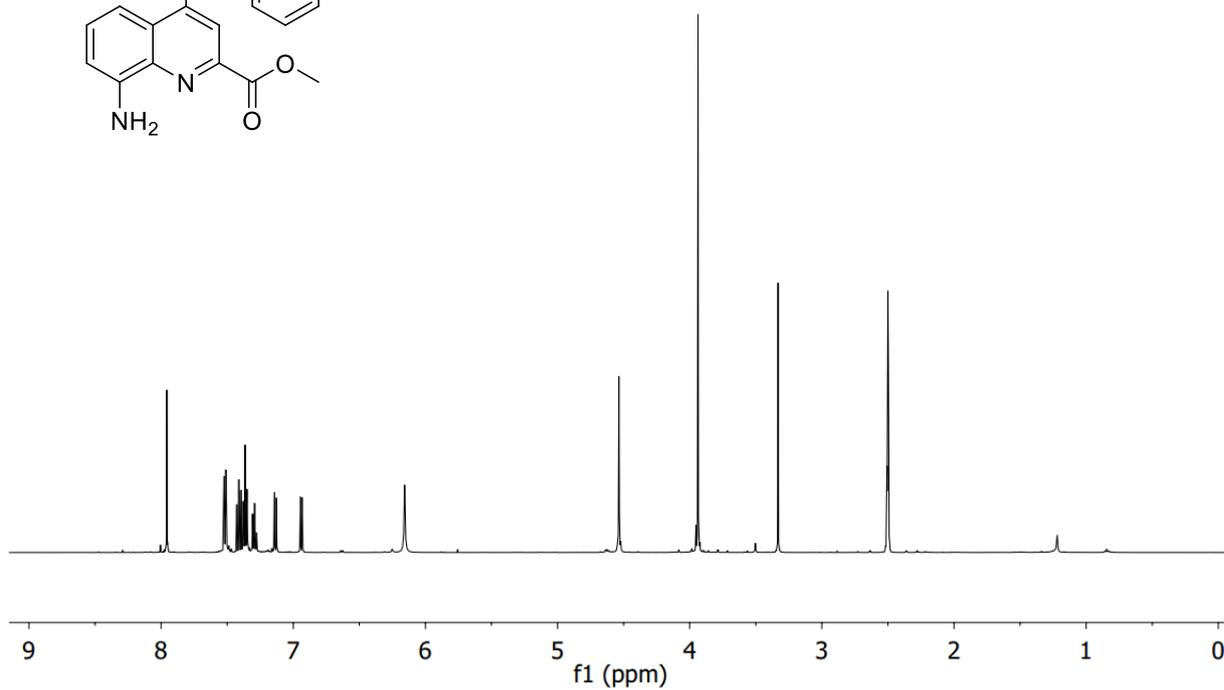
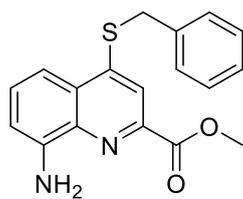
NMR spectra of compound 10: ¹H NMR (500 MHz, DMSO-*d*₆) and ¹³C NMR Spectrum (126 MHz, DMSO-*d*₆)



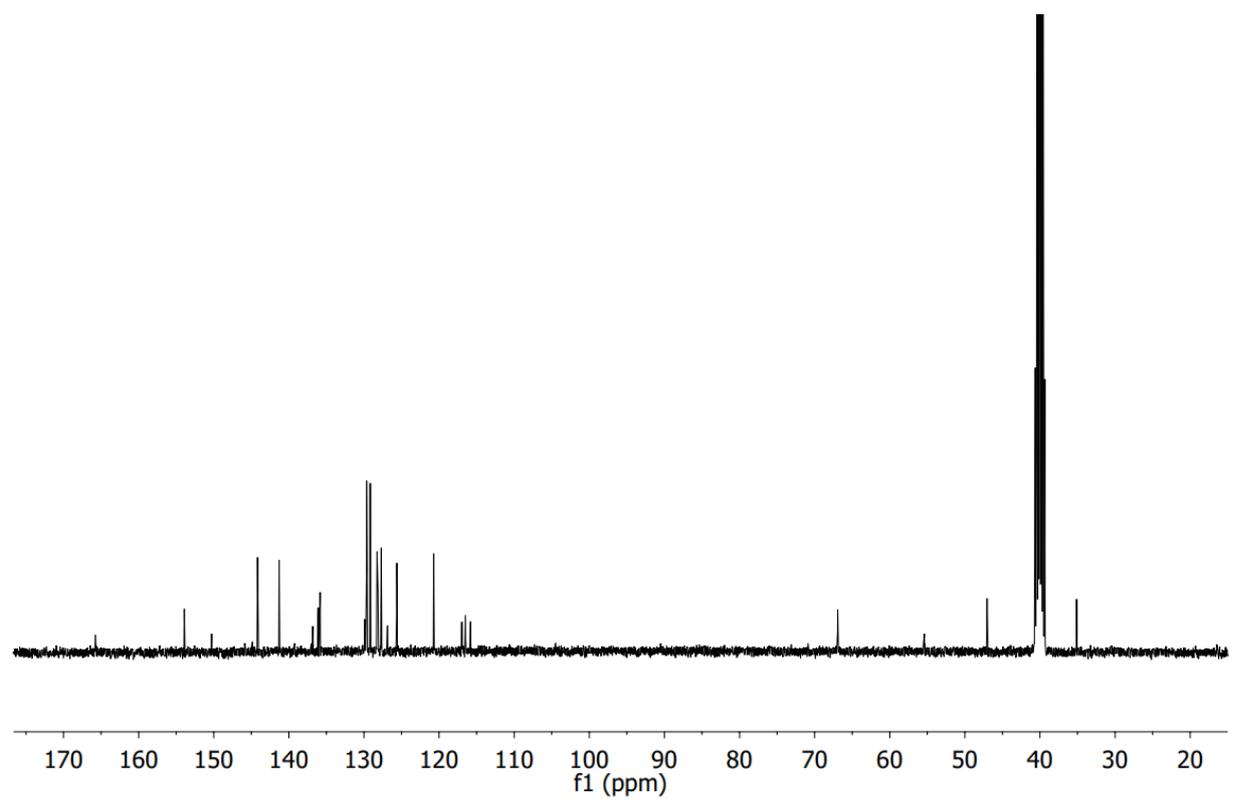
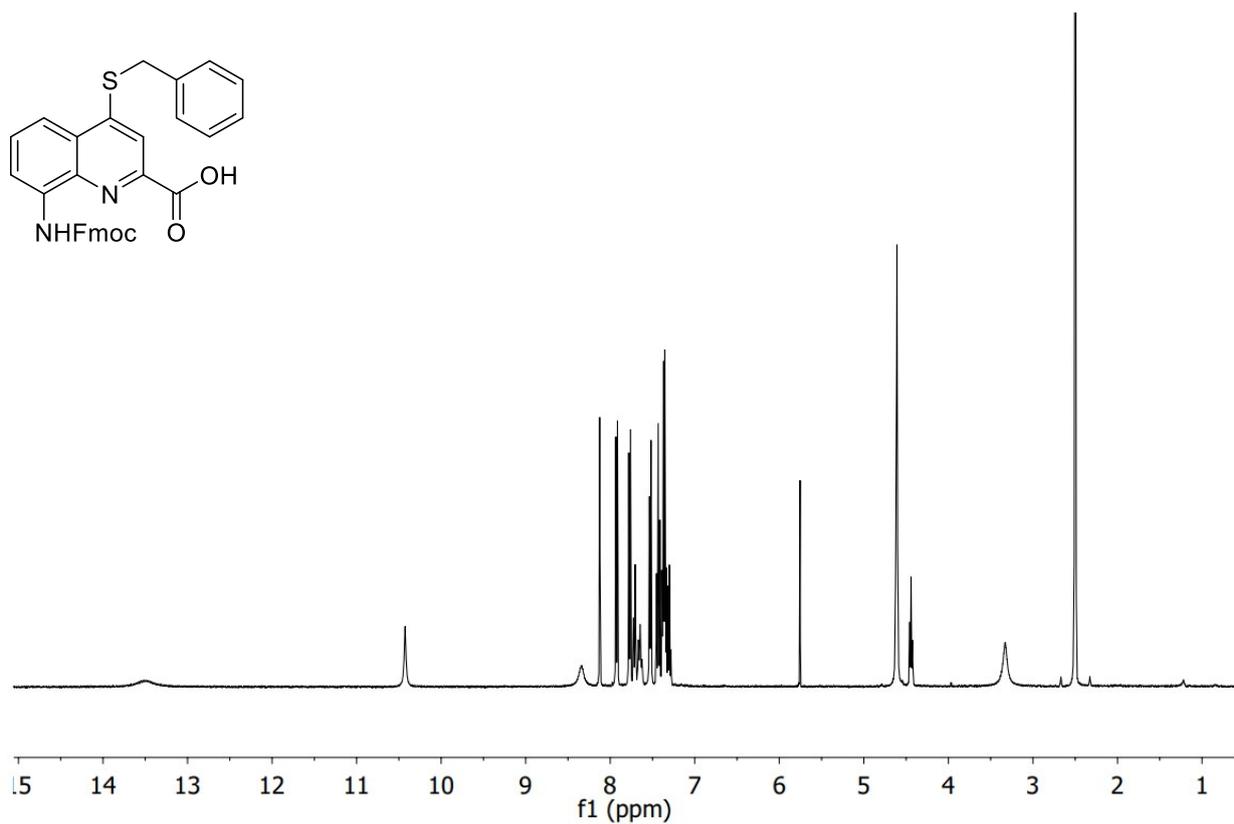
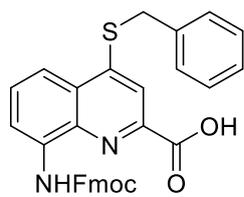
NMR spectra of compound 11: ¹H NMR (500 MHz, DMSO-*d*₆) and ¹³C NMR Spectrum (126 MHz, DMSO-*d*₆)



NMR spectra of compound **13**: ¹H NMR (500 MHz, CDCl₃) and ¹³C NMR Spectrum (126 MHz, CDCl₃)

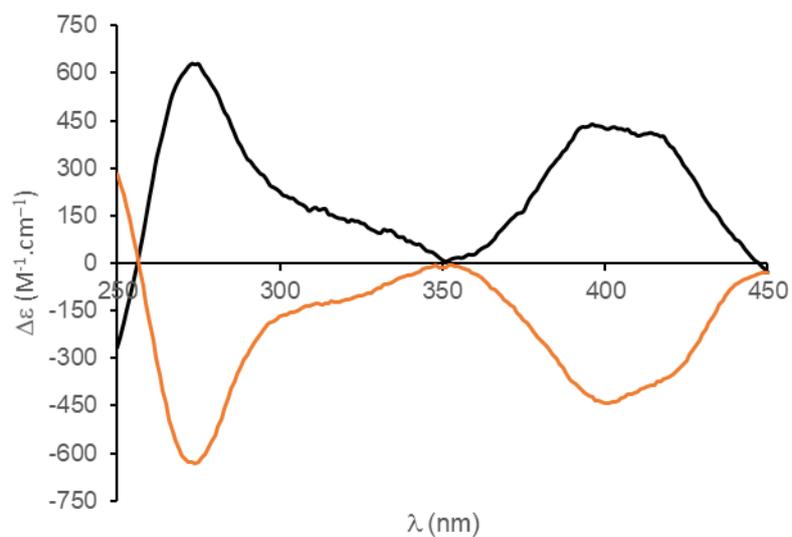


NMR spectra of compound 14: ^1H NMR (500 MHz, $\text{DMSO-}d_6$) and ^{13}C NMR Spectrum (126 MHz, $\text{DMSO-}d_6$)

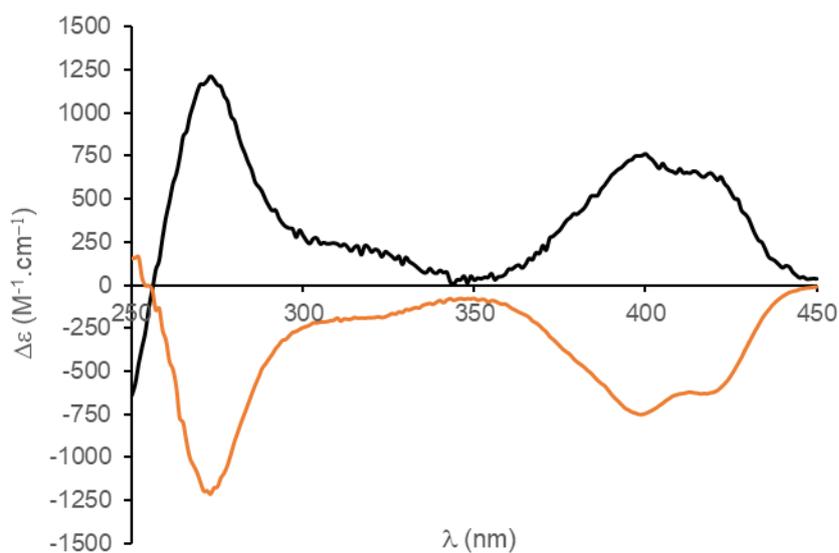


NMR spectra of compound **15**: ¹H NMR (500 MHz, DMSO-*d*₆) and ¹³C NMR Spectrum (126 MHz, DMSO-*d*₆)

8.6 CD spectra of foldamers 1d-1e, 2c-2d



CD spectra of **2c** (orange curve) and **2d** (black curve) recorded in TBST-D buffer at 25°C.



CD spectra of **1d** (orange curve) and **1e** (black curve) recorded in TBST-D buffer at 25°C.

8.7 References

- [1] Y. Hayashi, J. Morimoto, H. Suga, *ACS. Chem. Biol.*, 2012, **7**, 607-613.
- [2] D. J. Ford, N. M. Duggan, S. E. Fry, J. Ripoll-Rozada, S. M. Agten, W. Liu, L. Corcilius, T. M. Hackeng, R. van Oerle, H. M. H. Spronk, A. S. Ashhurst, V. Mini Sasi, J. A. Kaczmarek, C. J. Jackson, P. J. B. Pereira, T. Passioura, H. Suga, R. J. Payne, *J. Med. Chem.*, 2021, **64**, 7853-7876.
- [3] a) M. Nawatha, J. M. Rogers, S. M. Bonn, I. Livneh, B. Lemma, S. M. Mali, G. B. Vamisetti, H. Sun, B. Bercovich, Y. Huang, A. Ciechanover, D. Fushman, H. Suga, A. Brik, *Nat. Chem.*, 2019, **11**, 644-652; b) Y. Yamagishi, I. Shoji, S. Miyagawa, T. Kawakami, T. Katoh, Y. Goto, H. Suga, *Chem. Biol.*, 2011, **18**, 1562-1570.
- [4] a) B. Baptiste, C. Douat-Casassus, K. Laxmi-Reddy, F. Godde, I. Huc, *J. Org. Chem.*, 2010, **75**, 7175-7185; b) 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. Engl.*, 2014, **53**, 883-887; c) X. Hu, S. J. Dawson, P. K. Mandal, X. de Hatten, B. Baptiste, I. Huc, *Chem. Sci.*, 2017, **8**, 3741-3749; d) C. Tsiamantas, S. J. Dawson, I. Huc, *Comptes Rendus. Chimie*, 2016, **19**, 132-142.
- [5] S. Dengler, R. T. Howard, V. Morozov, C. Tsiamantas, W. E. Huang, Z. Liu, C. Dobrzanski, V. Pophristic, S. Brameyer, C. Douat, H. Suga, I. Huc, *Angew. Chem. Int. Ed. Engl.*, 2023, **62**, e202308408.
- [6] O. Al Musaimi, A. Basso, B. G. de la Torre, F. Albericio, *ACS. Comb. Sci.*, 2019, **21**, 717-721.
- [7] V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat, I. Huc, *Chem. –Eur. J.*, 2023, **29**, e202300898.
- [8] S. Dengler, P. K. Mandal, L. Allmendinger, C. Douat, I. Huc, *Chem. Sci.*, 2021, **12**, 11004-11012.
- [9] G. Bartoli, M. Bosco, M. Locatelli, E. Marcantoni, P. Melchiorre, L. Sambri, *Org. Lett.*, 2005, **7**, 427-430.
- [10] J. P. Davidson, O. Lubman, T. Rose, G. Waksman, S. F. Martin, *J. Am. Chem. Soc.*, 2002, **124**, 205-215.
- [11] M. Zwillinger, P. S. Reddy, B. Wicher, P. K. Mandal, M. Csekei, L. Fischer, A. Kotschy, I. Huc, *Chem. –Eur. J.*, 2020, **26**, 17366-17370.

9. Main text: Structure-based Design of an Aromatic Helical Foldamer-Protein Interface

Author contribution: The project was planned in collaboration with I. Huc. P. S. Reddy performed synthesis of Q^{Phe} and one sequence. B. L. d'Estaintot, J. Sigl, L. Fischer and T. Granier crystallized the foldamer-protein complexes and solved the structure. Y. Wei and Y. Zhang performed the Alphaspace 2.0 and purposed on side chain options. V. Pophristic and Z. Liu performed the molecular dynamic simulation setup and binding mode analysis for the proposed sequences. C. Douat performed the BLI test, co-supervised the project and assisted with chemical synthesis. The manuscript was written by me, C. Douat, and I. Huc. Most synthesis, analytical characterization, new B monomer synthesis, development of fluorescence competitive assays were performed by me.

Structure-based design of an aromatic helical foldamer-protein interface

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

DOI: 10.1039/x0xx00000x

Lingfei Wang,^a Céline Douat,^a Johannes Sigl,^a Post Sai Reddy,^b Lucile Fischer,^b Béatrice Langlois d'Estaintot,^b Zhiwei Liu,^c Vojislava Pophristic,^c Yuwei Yang,^d Yingkai Zhang^d and Ivan Huc^{*a}

The starting point of this study is the solid state structure of a complex between human carbonic anhydrase II (HCAII) and a helically folded tetradecaamide aromatic foldamer with a nanomolar HCAII ligand appended at the N terminus of the helix. In this complex, the foldamer is achiral but its handedness is biased by diastereoselective interaction with the protein. Computational analysis of the HCAII surface and inspection of the initial solid state structure led to the suggestion of main chain and side chain modifications of the foldamer helix that would result in an extension of the foldamer protein interface as well as in absolute helix handedness control. Molecular dynamics simulations validated several of these suggested modifications as potentially resulting in favorable foldamer-protein contacts. Five new Fmoc-protected amino acid building blocks bearing new biogenic-like side chains were synthesized. Nine new tetradecaamide sequences with or without the appended HCAII ligand were synthesized on solid phase and purified by RP-HPLC. The solid state structures of four of these sequences in complex with HCAII were obtained and validated the main design principles: (i) side chains can be predictably introduced at precise positions of the foldamer surface to create new contacts with the protein; (ii) side chains modifications do not alter main chain behavior and can be implemented independent from each other; (iii) some main chain units derived from quinoline-, pyridine-, or benzene-based δ -amino acids are largely interchangeable without altering the overall helix curvature in the context of a complex with a protein. An assessment of the K_D values required the adaptation of an existing fluorescence competition assay and suggested that the side chain and main chain modifications introduced in the new sequences did not result in significant improvement of the affinity of the foldamers to HCA

Introduction

Aromatic oligoamides represent a large class of compounds that can be used to recognize proteins and nucleic acids and that may interfere with their functions in multiple ways.¹ They comprise natural products such as distamycin,² cystobactamids,³ and albicidin,⁴ drug molecules such as suramin that has been crystallized bound to numerous proteins,⁵ rod-like oligomers many of which have been developed as α -helix mimetics,⁶ and oligomers that adopt helically folded conformations.⁷⁻¹⁰ We have been interested in the latter because their relatively large size offers the possibility to cover a large surface area of a protein target, which is relevant to protein-protein and protein-nucleic acid interactions, two types of interactions that are difficult to inhibit with small molecules.¹¹ Helical aromatic oligoamide foldamers (AOFs) and in particular those derived from 8-amino-2-quinolinecarboxylic

acid (Fig. 1) also possess the advantage that their conformations are very stable in particular in protic solvents,¹² and that synthetic methods exist to introduce various biogenic-like side chains at their periphery.¹³

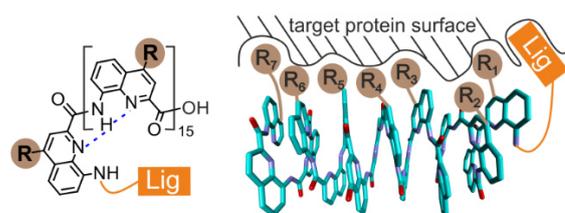


Fig. 1. Chemical formula (left) of a hexadecaamide of 8-amino-2-quinolinecarboxylic acid bearing a protein ligand (Lig) at the N terminus and biogenic-like side chains (R groups) in position 4. Schematic representation (right) of the helical structure of the hexadecaamide with some R groups interacting with a protein surface to which the ligand is also bound.

The potential of helical AOFs to interfere with protein function has been highlighted in the context of amyloid fibers⁷ and DNA-binding proteins.⁸ For the latter, AOFs that specifically mimic the shape of charge distribution of DNA have been developed. In contrast, methods are still missing to design *ab initio* a helical AOF protein binder that does not mimic an already known binding epitope, for example through the introduction of biogenic-like side chains complementary to the

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

^b CBMN (UMR5248), Univ. Bordeaux-CNRS-INP, Institut Européen de Chimie et Biologie, 2 rue Escarpiot, 33600 Pessac (France).

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

^d Department of Chemistry, New York University, 10003 New York, New York (USA)
Supplementary Information available: Supplementary figures and data, detailed experimental protocols and characterisation of new compounds. See DOI: 10.1039/x0xx00000x

protein surface. To assess the potential of some helical AOFs to interact with a given protein surface, we introduced a tethering approach where a covalent or non-covalent linkage confines the AOF at the surface of the protein (Fig. 1).^{8,9} Tethering between small molecules or between small molecules and their protein target is a common approach in the context of drug research to compensate for initially weak binding. It lies at the heart of linker design in fragment-based approaches,¹⁴ including in the context of template-assisted strategies,¹⁵ and of covalent ligands.¹⁶ Tethering of AOFs to a protein target was used to detect foldamer-protein interactions upon observing a preferred handedness in an achiral oligomer.^{8,9} Taken alone, the

achiral AOF exists as a racemic mixture of right-handed (*P*) and left-handed (*M*) enantiomeric conformers. If either helix (*M* or *P*) helix interacts better than the other with the protein surface, the resulting change of proportion leads to an induced circular dichroism (CD) signal. An AOF CD signal is easy to detect because AOFs absorb above 350 nm, in regions where proteins are transparent. Using human carbonic anhydrase II (HCAII) as a model system, strong helix handedness induction was observed with several helical AOFs linked to the protein via a nanomolar ligand. Subsequently, solid state structures of such complexes were obtained that confirmed the preferred *P* helix handedness and informed about foldamer protein contacts.^{8,9}

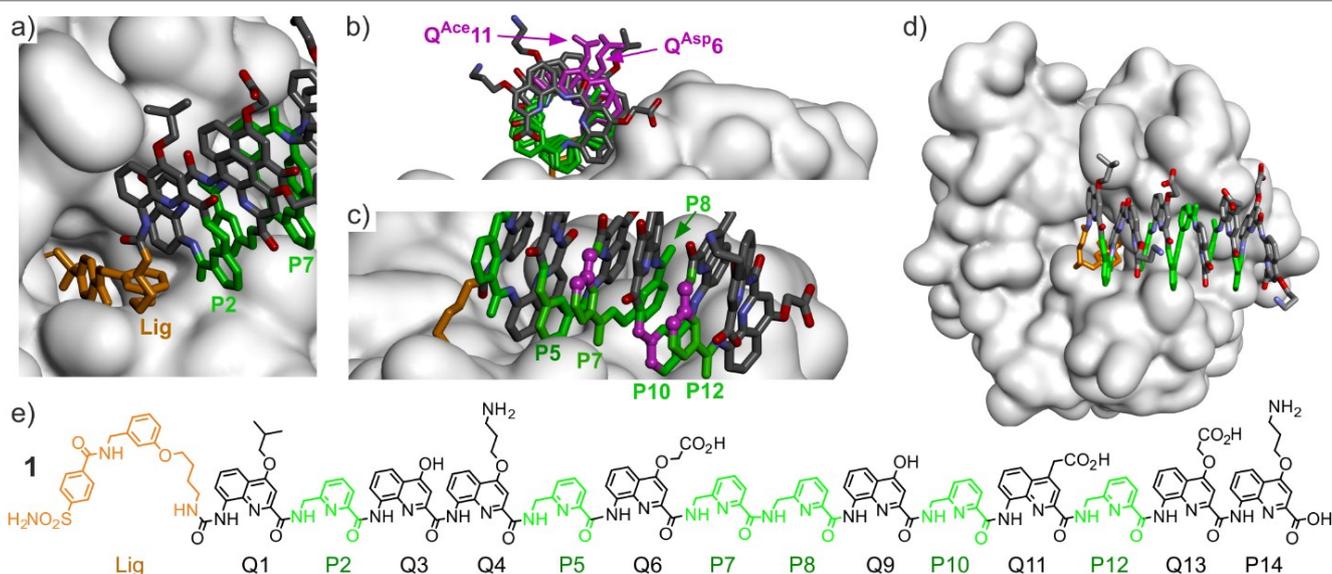


Fig. 2. a), b), c) and d) show different views of the solid state structure of the complex between HCAII and AOF 1.⁹ Only the surface of the protein is shown. The foldamer is shown in stick representation. Pyridine units (P) are shown in green and the HCAII ligand is shown in gold. In b), two residues of interest are shown in purple. In c) three carbon atoms of three P residues are shown in purple balls. These carbons would belong to the additional benzenic ring when implementing a P->Q mutation at these positions. Hydrogen atoms have been omitted for clarity. e) Structural formula of 1.

In the case of tetradecamide sequence **1**, the solid state structure showed a large contact surface area between the foldamer and HCAII (Fig. 2).⁹ However, it has not yet been shown whether such a structure could serve to further design the foldamer to extend its contact with the protein surface. Furthermore, although the AOFs are known to be rigid, it remained to be demonstrated whether side chain and main chain modifications could be implemented without altering their overall structure, a task difficult to achieve with *e.g.* a peptide or an aliphatic peptidic foldamer.¹⁷ Here, we show that, with the help of computational tools, the solid state structure of the complex **1**•HCAII can be used as a starting point to place side chains at defined positions in space to further elaborate the foldamer-protein interface. We validate that the foldamer structure remains independent of side chain variations and even some main chain variations. Although the changes implemented have not resulted in significant changes in the dissociation constant of the complexes, the results further validate the concept that AOF helices can serve as reliable scaffolds to display biogenic-like side chains at the surface of a protein.

Results and discussion

Synthesis and general design principles

Oligoamide sequence **1** consists of eight Q^{xxx} δ -amino acid monomers presenting different biogenic-like side chains in position 4, and of six P residues (Figs. 2d, 3). P residues bring the same contribution to helix curvatures as Q^{xxx} but they are more flexible. Their initial role was to make helix handedness dynamics fast enough to be practically monitored, *e.g.* in the course of minutes to hours,¹⁸ for example when helix handedness bias takes place upon binding of **1** to HCAII.⁹ Unexpectedly, P residues were found to be directly involved in foldamer-protein contacts in the solid structure of **1**•HCAII (Fig. 2b). Their role thus extends to that of interacting units despite the fact that they carry no biogenic-like side chains.

For the purpose of extending the foldamer-HCAII contacts a number of Q^{xxx} monomers were considered bearing side chains in position 4, 5 or 6 (Fig. 3). B^{xxx} monomers were also involved. B^{xxx} monomers are δ -amino acids as well and may thus bring a contribution to helix curvature similar to that of Q and P.¹⁹ In

addition, they carry a stereogenic center that has been shown to quantitatively bias helix handedness.²⁰ Sequences comprising B^{Xxx} monomers are thus designed to be one handed even when they are not bound to HCAII. The *S* configuration of B^{Xxx} monomers is intended to favor *P* helicity. Among the monomers used in this study (Fig. 3), some were previously described, some were used only in computations, and five – B^{Ind}, B^{Gpr}, Q^{Gly}, Q^{5Ph} and Q^{6gp} – were newly synthesized (Fig. S1). The preparation of these building block in a form suitable for solid phase synthesis is presented in detail in the Supplementary

Information. All monomers were produced with a free carboxylic acid and an Fmoc-protected main chain amine. In addition, the side chains of B^{Gpr} and Q^{6gp} were protected with Boc groups. Typically, side chain installation involved Sonogashira or Suzuki cross-coupling reactions on a bromoaryl precursor.¹³ Of note, Q^{5In} was also prepared but it was found to undergo oxidative degradation in air, especially after insertion in oligoamide sequences, and its preparation is not reported here.

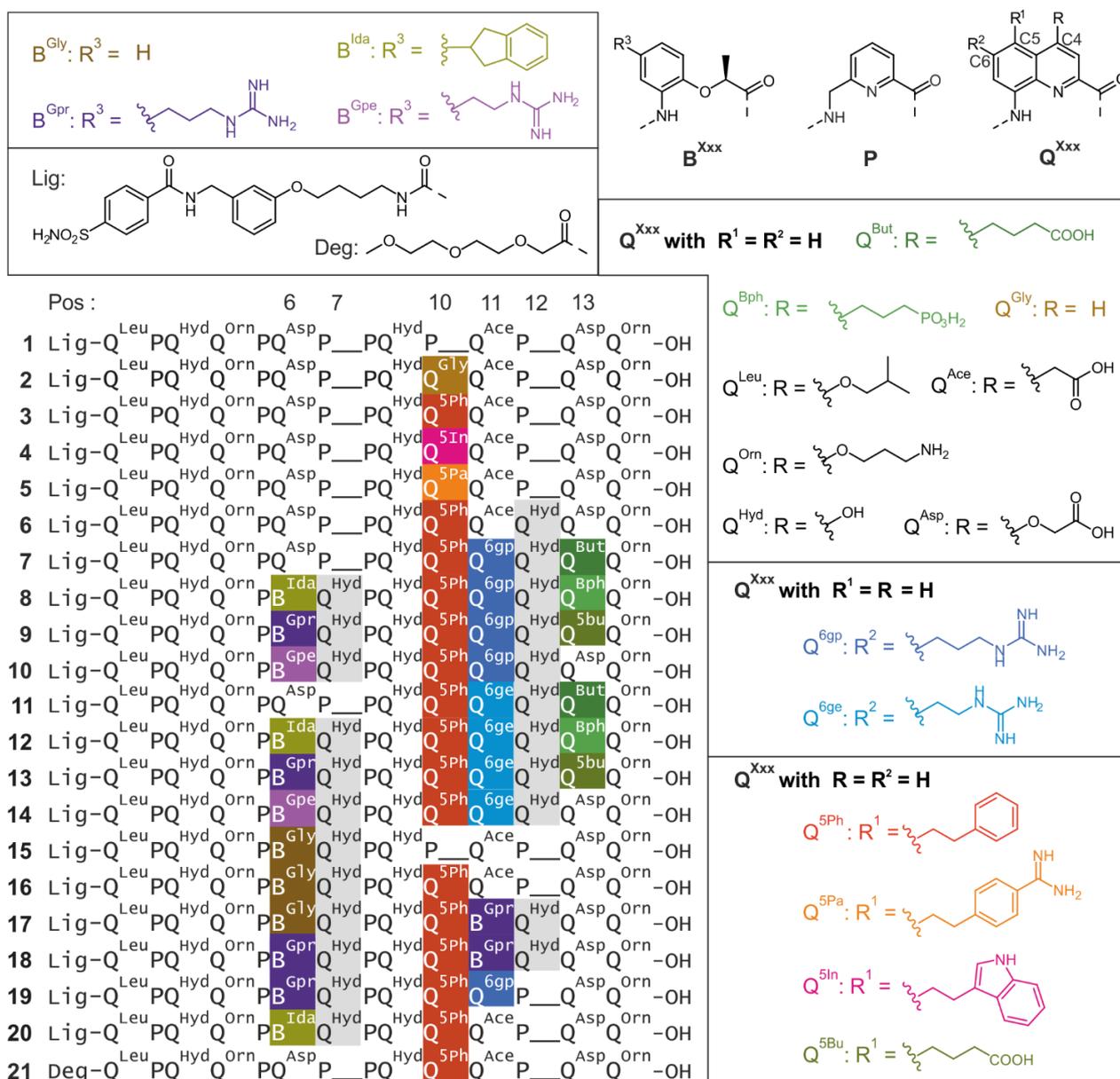


Fig. 3. Structural formulas of P, Q^{Xxx}, and B^{Xxx} δ -amino acids and of N-terminal Lig and Tail functional groups. Carbon atoms in position 4, 5, and 6 of Q^{Xxx} are indicated and carry R, R¹, and R² side chains, respectively. The Xxx three letter code used for the side chains is sometimes inspired by the three letter code of α -amino acids bearing similar side chains, even when they may not exactly match. The three letter code also indicates when the side chain is in position 5 or 6 of the quinoline ring. Sequences **1-21** are defined with the letter code used in this study. To facilitate residue identification, the colors of the highlighted residues match with the side chain colors in the adjacent boxes.

Sequences **1-3** and **15-21** (Fig. 3) were synthesized on solid phase using an established *in situ* acid chloride activation protocol for the coupling steps.²¹ An improved procedure for

the on-resin introduction of the HCAII ligand at the N-terminus of the helix using a urea linkage was also developed. Sequences

were purified by RP-HPLC after TFA-mediated resin cleavage and side chain deprotection.

Sequences **4-14** were investigated in computational studies but not synthesized. Sequences **2-5** were the focus of a first phase of our investigation. They are analogues of **1** in which P10 is replaced by different, more rigid Q10, residues. In the case of **3**, it could be verified experimentally that helix handedness bias upon binding to HCAII takes place despite the added rigidity, albeit significantly slower than with **1** (Fig. S2). This first phase led to the installation of a Q^{5Ph}10 residue in **3** instead of P10 in **1**. Q^{5Ph}10 was conserved in all subsequently synthesized sequences but **15**. In a second phase, residue variations in positions 6, 7, 11, 12 and 13 were assessed computationally in sequences **6-15** (see next section) and a selection of these variations was experimentally implemented in **15-20**. Sequence **21** is an analogue of **3** lacking the N-terminal ligand.

Including chiral B^{xxx} units to favor *P* helix handedness was desirable, for example to avoid conformational changes in the course of a *K_D* value determination. For this purpose, residues 6 and 11 were chosen as possible locations. This choice was based on the observation that the side chains in position 4 of the quinoline rings of Q^{Asp6} and Q^{Ace11} of **1** lie far from the HCAII surface (Fig. 2c). Removing these side chains and part of the pyridine ring of Q when performing a Q->B mutation should not alter the HCAII-foldamer interactions observed in the solid state. In contrast, the carbon atoms in position 5 and 6 of the quinoline rings of Q^{Asp6} and Q^{Ace11} in **1** seem better oriented to introduce a side chain that may interact with the HCAII surface and a B^{xxx} monomer may offer a similar side chain presentation.

Before implementing Q->B mutations, another design feature had to be considered. While chiral B units have been shown to quantitatively bias helix handedness in the context of (Q)_n oligomers,²⁰ this has not been validated when the helix also contains multiple more flexible P units, as in **1**. Indeed, partial handedness bias has occasionally been observed when Q monomers are mixed with other monomers.²² To mitigate the risk that the chiral B-containing sequences would not be quantitatively one handed, we replaced some P units by Q in the vicinity of the B6 and B11 monomers. With their additional fused benzenic ring, Q monomers are bulkier than P. The structure of the **1**•HCAII complex showed that this extra bulk could be accommodated without generating clashes in P7, P10 and P12, but not in P5 and P8 (Fig. 2c). Chiral B-containing sequences **15-20** therefore contain at least one and sometimes up to three Q monomers at residues 7, 10 or 12.

The one-handed nature of the new chiral foldamers could be verified by ¹H NMR spectroscopy through the observation of a single set of signals. On top of ensuring quantitative handedness bias, the additional Q residues also resulted in slow helix handedness inversion in water. In case handedness bias was incomplete when the foldamer was first dissolved in water, e.g. for RP-HPLC purification, it may no longer proceed to completion. This pitfall is easily detected by the observation of two distinct sets of signals on the ¹H NMR spectra, corresponding to *P* and *M* diastereomeric conformers. To solve this problem, one can dissolve and incubate the compound in an organic solvent such as DMF, where helix handedness

inversion takes places faster,¹² before evaporating and redissolving in water.

Computational design

Protein surface analysis. The potential of the HCAII surface for interacting with biogenic-like residues was assessed with AlphaSpace, a computational analysis tool designed for fragment-centric topographical mapping.²³ The assessment proceeded in two phases. In a first phase, the surface in the vicinity of the HCAII active site was analysed, leading to the identification of potential binding pockets Po1-Po5 (Fig. 4b). Po1 has the highest ligandability (highest Bscore) and corresponds to the HCAII active site where HCAII ligands usually bind.²⁴ In the **1**•HCAII complex, Po1, Po2, Po3 and Po5 are filled by the N-terminal ligand and the helix backbone, as indicated by the color patches in Fig. 4a, leaving essentially no space to add functionalities on the foldamer helix to further enhance contacts with the protein surface. In contrast, Po4 was identified as a sizeable (158 Å³) cavity nearby P10. Since a P10Q mutation appeared to be feasible without causing steric clashes (Fig. 2c), various side chains were docked in Po4 while being connected to the C5 carbon of the quinoline ring of Q10. All 274 side chains of the Swiss amino acid database²⁵ were tested. In each case, the amino acid was replaced by the quinoline residue and Autodock Vina²⁶ was used to score interactions between the side chain in position 5 and Po4 (Fig. S3). Q^{5Ph}, Q^{5Pa}, and Q^{5In} were selected as having a sufficiently low estimated Δ*G* and as being at the same time synthetically accessible. As presented in detail below, subsequent computational steps, synthesis and structural analysis eventually delivered the solid state structure of the **3**•HCAII complex where Po4 is indeed filled by the phenethyl side chains of Q^{5Ph}10.

The second surface analysis was performed on the **3**•HCAII complex in order to identify potentially ligandable sites in the vicinity of the foldamer helix where foldamer-protein contacts may be extended through the addition of foldamer side chains (Fig. 4c). This analysis led to the identification of pockets Po6-Po9. Po9 consists of the space left in Po4 that is not occupied by the side chain of Q^{Ph}10, hinting at the possibility that this side chain may be further elaborated (Fig. S4). However, this option was not explored as many of the suggested side chains were synthetically challenging. We focused instead on Po6-Po8 which all lie on the same side of the foldamer helix, and may potentially be reached with side chains on residues Q6, P8, Q11, and Q13 (Fig. 4a). As explained above for Q10, side chains that both had a reasonable docking score and appeared to be synthetically accessible were kept for subsequent investigations (Figs. S5, S6). For Q6 and Q11, we have mentioned above that side chains in position 4 of the quinoline do not establish contacts with the protein surface (Fig. 2b) and that these positions were considered for the introduction of chiral B residues to control helix handedness. Instead, the HCAII surface analysis suggested the side chains in position 6 of the quinoline ring might establish contacts with the protein. This eventually led to the mutation of Q^{Asp6} into guanidinium-containing B^{GPr6} or B^{GPe6} in sequences **9, 10, 13, 14, 18, 19**, as well as indane-containing residue B^{Ida6} in sequences **8, 12, 20**. Similarly,

mutation of Q^{Ace}11 to guanidinium-containing residues Q^{6BP}11, Q^{6Ge}11 or B^{GPR}11 was implemented in sequences **7-14** and **16-19**.

Possible modifications of Q^{ASP}13 were inspired by a salt bridge between this residue and Lys24 of HCAII observed in one of the solid state structures presented below. To better reach this Lys24, residues Q^{BUT}13, Q^{BPH}13 and Q^{5BU}13 were considered in sequences **7-9** and **11-13**. It should be pointed that, given the extensive HCAII surface that the foldamer helix covers, opportunities for mutations and for the creation of new foldamer-protein contacts were too numerous to be considered at the same time. For instance, pockets Po10-Po14 were not investigated (Fig. 4c).

Molecular Dynamics (MD) simulations. Prior to investing time and resources in the preparation of new monomers and new sequences suggested by the HCAII surface analysis, the effect of foldamer modifications in sequences **3-14** on interactions with HCAII were evaluated using MD simulations in explicit water

using the AMBER22 package.^{27,28} The initial HCAII and foldamer structures and positions were those of the solid state structure. The ff14SB force field²⁹ was used for α -amino acid residues. The general AMBER force field (GAFF),³⁰ with improved torsional parameters for arylamides,³¹ was used for the foldamer (see supplementary information for details). One additional simulation in the presence of 125 mM NaCl was performed on **3**•HCAII. It resulted in minor changes such as slightly larger fluctuations, deviations, and a reduction of the occurrence of salt bridges. To inspect the interaction between foldamer and HCAII, we carried out a combination of structure visualization, calculations of root mean square displacements (RMSD) of protein and foldamer backbone atoms with respect to the solid state structure, as well as analysis of specific residue-to-residue distances. With the exceptions of sequences **7** and **11**, all backbone RMSDs stayed within 3 Å of the solid state structure. Because of the larger deviations observed for **7** and **11**, not much could be concluded for these two sequences.[‡]

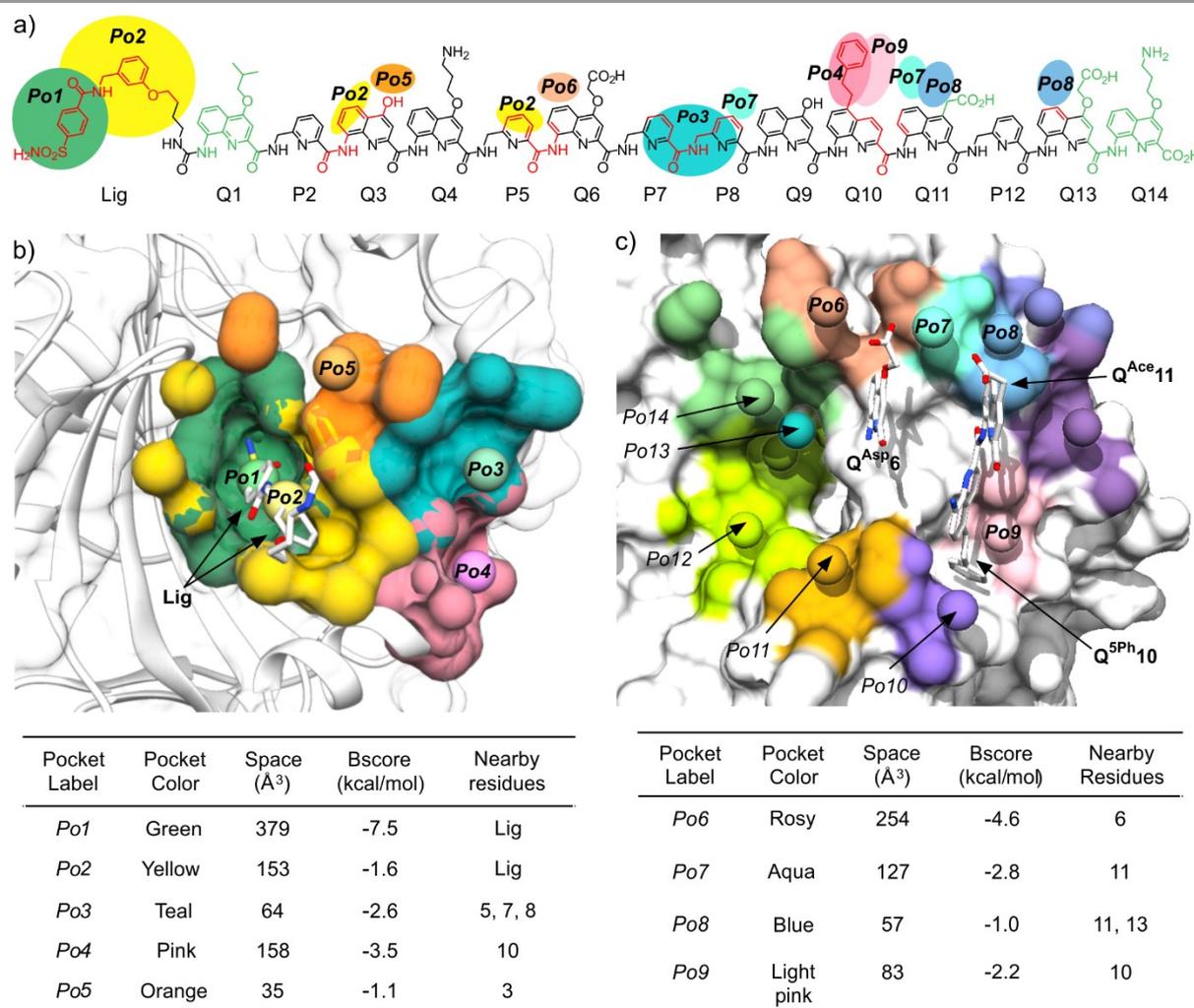


Fig. 4. a) Structural formula of **3** including Q and P residue numbering. The bonds shown in red highlight parts of the molecule in direct contact with the HCAII surface in the solid state structure of the **3**•HCAII complex (Fig. 5). The bonds shown in green highlight parts of the foldamer involved in intercomplex contacts in the crystal lattice of the solid state structure of the **3**•HCAII complex. The color patches indicate the pockets near the foldamer main chain or side chains in the solid state structure of the complex. Pockets are colored and numbered as in b) and c). b) Pocket analysis of the surface of HCAII restricted to the vicinity of the ligand binding site and the contact area with the foldamer helix in the solid structure of the **1**•HCAII complex (Fig. 2). c) Pocket analysis of the surface of HCAII restricted to the vicinity of the contact area with the foldamer helix in the solid structure of the **5**•HCAII complex (Fig. 5). In b) and c), pockets have been assessed in terms of their volume and their ligandability (BScore). Pockets Po10-Po14 were not included in this study.

A first aspect concerns the rationale that led to selecting sequences **3-14**. Because the number of possible single mutations was large, these were not investigated individually. Most sequences, carry two, three or four simultaneous side chain modifications with respect to **1**, at positions 6, 10, 11 and 13. This way, all but two side chain modifications were examined in at least two distinct MD simulations. The first essential result is that the various side chain and, sometimes, main chain modifications mainly depend on where they are implemented, and generally do not depend from one another. When a mutation is performed at a given position, the behavior of the new residue tends not to vary with mutations at other positions. This is a major advantage for making predictions and sharply contrasts with aliphatic peptides where local modifications may impact global behavior.¹⁷ The consistent behavior of each new residue regardless of other sequence modifications also tells that no mutation led to a major steric clash that would disturb the whole structure. In addition, owing to the independent behavior of the side chains, we could perform an analysis per interaction site/pocket, instead of an

analysis per sequence. The results are presented in Figs. S7-S11 and a representative example is shown in Fig. 5.

The MD simulations reflected the strong interaction between the ligand and HCAII (Figs. S7-S9). In all simulations, the bond between the HCAII-bound Zn²⁺ ion and the ligand sulfonamide group, tight contacts between the two aryl groups of the ligand and pockets Po1 and Po2, and contacts between Q3 and P5 with HCAII in Po2, all remained well in place (Figs. 2a, 4a,b). Concerning the mutations of P10 implemented to fill Po4, Q^{5Ph}10 was found to form stable hydrophobic contacts with Pro137, Leu203, Glu204 and Cys205, as highlighted by the histograms of distance shown in Fig. 5. Note that sequences **4** and **5** show some deviations in these histograms because their different Q^{5In}10 and Q^{5Pa}10 residues establish distinct contacts. The indole side chain of Q^{5In}10 in sequence **4** lies closer to Pro137. This residue was actually synthesized but had stability issues that hampered experimental investigations. The benzamidinium side chain of Q^{5Pa}10 in sequence **5** appeared to be too large for pocket Po4 and its position fluctuated more. It was not considered further and Q^{5Ph}10 was conserved in all subsequent experiments.

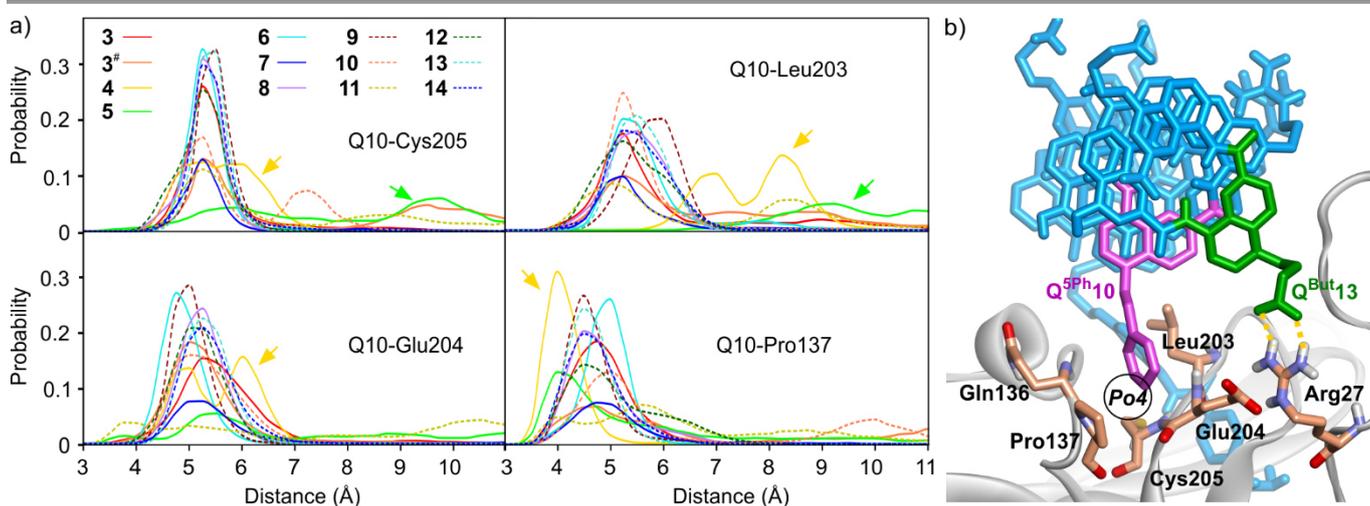


Fig. 5. a) MD Simulations of complexes between HCAII and foldamers **3-14** showing the occupancy of pocket Po4 by the side chain of Q10, that is, Q^{5Ph}10 for all sequences but **4** (Q^{5In}10) and **5** (Q^{5Pa}10). The histograms show the probability, through the entire simulation time, of the distance between the center of mass (COM) of the phenyl ring of Q^{5Ph}10 or Q^{5Pa}10 or of the pyrrole ring of Q^{5In}10, and HCAII residues Cys205 (position of C β), Glu204 (COM of C β and C γ), Leu203 (COM of all side chain C atoms), and Pro137 (COM of C β , C γ , and C δ). Two distinct simulations were performed with sequence **3**. The simulation marked with # included 125 mM NaCl. Arrows highlight the different positioning of Q^{5In}10 and Q^{5Pa}10 in the pocket. The histograms show weaker probabilities for sequences **7** and **11** due to strong deviations from the initial structure in these two cases (the Q10 residues are most of the time at distances >12 Å from Po4). b) Snapshot from the MD simulation of **7**•HCAII showing the relevant residues. In this snapshot, one can also spot a transient salt bridge between the carboxylate side chain of Q^{But}13 and the Arg27 (dashed yellow lines). This salt bridge does not occur frequently.

The interactions between side chains on residues 6, 11 and 13 and pockets Po6-Po8 (Fig. S10) can be summarized as follows. Overall, the distance histograms show larger variations than for the contacts in pockets Po1-Po4. The salt bridge between Q^{Asp}13 and Lys24 seen in one solid state structure (see below) was absent or present in small percentage (3% to 41%) of the time along the trajectories in aqueous solution. Using other negatively charged side residues Q^{But}13, Q^{Bph}13, and Q^{Bu}13 made little difference. These were therefore not tested experimentally. In the case of residue 6, the hydrophobic side chains of B^{Ida}6 or the cationic side chains of B^{Gpe}6 or B^{Gpr}6 could potentially form contacts with Phe20 and Asp19, respectively, within pocket 6. Some of these residues were subsequently

synthesized and implemented in sequences **19-20**. Finally, the benefit of cationic residues in position 11 to fill pocket Po8 was not clear. Salt bridges were established only during small fractions of simulation time. With sequence **10**, a possible exception to the independent role of the side chains was observed with an apparent positive cooperative effect of the guanidinium-containing side chains of B^{Gpe}6 and Q^{6Gp}11 (Fig. S10).

MD simulations revealed an additional, unplanned, favourable foldamer-protein contact between Q^{Hyd}7 and Gln135 (Fig. S11). As mentioned above, P7Q^{Hyd} and P12Q^{Hyd} mutations were introduced to rigidify the helix and mitigate the risk that chiral B residue may not quantitatively bias helix

handedness. Q^{Hyd} residues were selected for that purpose because their side relatively acidic yet small hydroxy side chain would not decrease foldamer water solubility. Nevertheless, MD simulations suggest that the side chain of Q^{Hyd}7 can also hydrogen bond to Gln135.

Structure elucidation

Crystallization was attempted for all foldamer-HCAII complexes. In the case of **2**•HCAII, **3**•HCAII, **16**•HCAII and **20**•HCAII, single crystals suitable for X-ray diffraction analysis were obtained (Fig. S12) and the solid state structures were elucidated in the $P2_12_12$ space group at a resolution of 1.4, 2.1, 1.6, and 2.1 Å, respectively (Fig. S13, S14). For the four structures, crystallization conditions were similar to that of **2**•HCAII, and so were the unit cells and packing arrangements. Some parts of the foldamer molecule are involved in intercomplex contacts in the crystal lattice, including the side chain of Q^{Ac}11 (Figs. 4a, S15),

these contacts are all conserved in foldamers **1**, **2**, **3**, **16** and **20**. In retrospect, we hypothesized that the unsuccessful crystallization of the complexes with **17**, **18** and **19** may be assigned to the mutation of Q^{Ac}11 into B^{Gpr}11 or Q^{Gpr}11 in these compounds.

The structure of **2**•HCAII validated that a P10Q^{Gly} mutation of the foldamer could be performed without any steric clash between the protein and Q^{Gly}10 or any alteration of the foldamer helix shape (Fig. 6a,b). This structure also revealed a salt bridge between the carboxylate of the Q^{Asp}13 and residue Lys24 (Fig. 7a). In the structure of **1**•HCAII, the side chain of Lys24 was only partly visible in the electron density map and the salt bridge was overlooked. As mentioned above, MD simulations suggested that this salt bridge is not stable in aqueous solution and is not convincingly stabilized when using anionic side chains longer than in Q^{Asp}13, or placed in position 5 of the quinoline ring, or having a dianionic phosphonate group, as in Q^{But}, Q^{Bph}, and Q^{5Bu}.

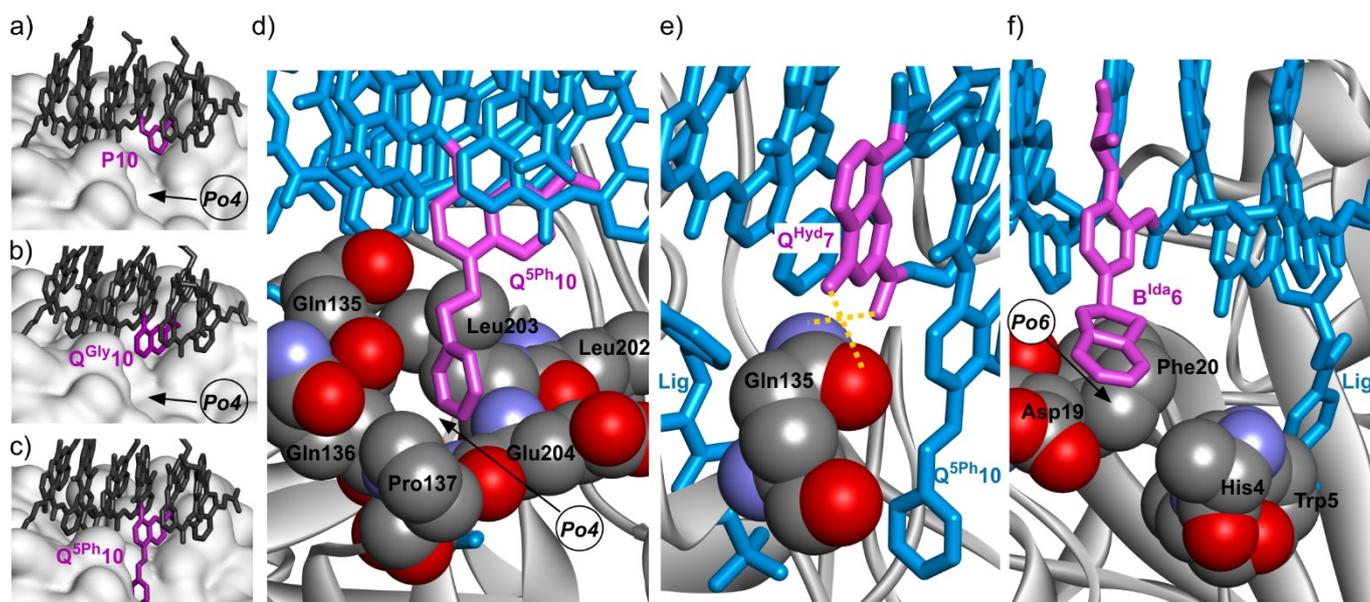


Fig. 6. Solid state structures of HCAII in complex with: a) **1**, b) **2**, c) **3**, d,e) **16**, f) **20**. In a)-c), the protein is shown as a white soft surface and the foldamer is shown in stick representation in gray except the residue in position 10 colored in purple. In d)-f), the protein is shown in gray ribbon representation except relevant amino acids which are in space filling representation. The foldamer is shown in blue stick representation except the residue in position 10 in d), the residue in position 7 in e) and the residue in position 6 in f), which are colored in purple. Hydrogen bonds are indicated as yellow dashed lines. Pockets Po4 and Po6 are defined in Fig. 4

The structure of **3**•HCAII then validated that the phenethyl side chain of Q^{5Ph}10 filled Po4 as predicted by computations (Fig. 6c, d). The methylene carbon atom linked to the quinoline ring was found at 3.9 Å from a methyl group of Leu203, and one carbon atom of the phenyl ring lies within 3.5 Å from the nitrogen atom of Pro137. The structure of **16**•HCAII confirmed the position of Q^{5Ph}10 found in **3**•HCAII and validated the double mutation Q^{Asp}6B^{Gly} and P7Q^{Hyd} intended to introduce helix handedness bias (through B^{Gly}6) and to make the helix more rigid (through Q^{Hyd}6). The proximity of Q^{Hyd}7 and Gln135 (Fig. 6e) makes the hydrogen bonding observed in MD simulations plausible. In the solid state, the Gln135 amide NH₂ hydrogen bonds to the main chain carbonyl of Q^{Hyd}7. In addition, the proximity between the

primary amide of Gln135 and the hydroxy side chain of Q^{Hyd}7 likely favors contacts with the latter as well, be it in a protonated or deprotonated state. Finally, the structure of **20**•HCAII validated that the indane side chain of B^{Ida}6 filled Po6 again as predicted by computations, establishing contacts with Phe20 (Fig. 6e). In this structure, Lys24 was again visible in the electron density map, but in a conformation where hydrogen bonding to Q^{Asp}13 is not established (Fig. 7b), different from the structure of **2**•HCAII. Of note, in all structures, Q^{Asp}13 is involved in a salt bridge with a lysine (Lys80) belonging to another HCAII molecule of the crystal lattice as part of the intercomplex contacts (Fig. S15). This probably influences, that is, competes with the formation of the salt bridge with Lys24 in the solid state.

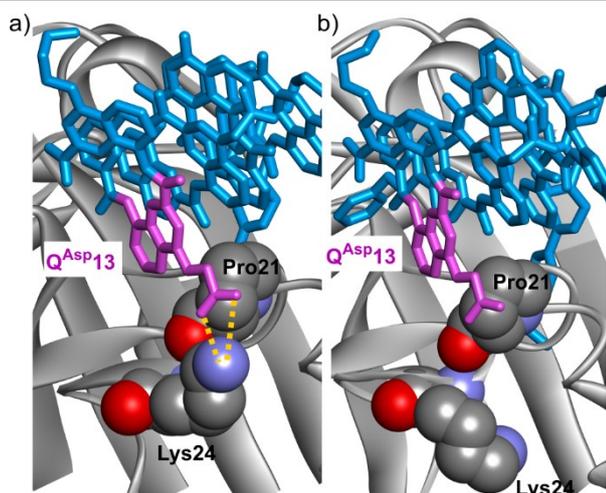


Fig. 7. Solid state structures of HCAII in complex with: a) **2**, b) **20**. The protein is shown in gray ribbon representation except P21 and Lys24, which are in space filling representation. The foldamer is shown in blue stick representation except the residue in position 13, which is colored in purple. Hydrogen bonds are indicated as yellow dashed lines.

Altogether, the solid state structures validate the predictions made by computations. They demonstrate the equivalent contribution to main chain helix curvature of P, Q and B monomers in the context of a foldamer-protein contact area. They also demonstrate that the positions of the foldamer side chains and the type of interactions they may engage with the protein are predictable.

Binding studies

We set out to measure the binding affinities of *P*-helical, chiral B-containing sequences **15–20** for HCAII to assess the extent to which they reflected the changes introduced in the foldamers. Sequences **1–3** are potentially problematic as they exist as a racemic mixture of *M* and *P* helical conformers that must have different K_D values and whose proportion evolve with time upon binding to HCAII, hence the focus on **15–20**. This assessment proved challenging. Simple ligands such as **22** (Fig. 8a) – the fragment of **15–20** that fills pockets Po1 and Po2 of HCAII – bind in the low nM range. Getting accurate K_D values to comment on potentially small effects for such strong binding is delicate. Furthermore, we have shown that appending a foldamer on **22** has one major consequence: both the association and dissociation kinetics are slowed down by almost two orders of magnitude.^{8a} Unlike with classical small molecule HCAII ligands, dissociation becomes so slow that techniques such as surface plasmon resonance (SPR) or biolayer interferometry (BLI) no longer deliver reliable results.

We turned to a recently published assay that exploits the quenching of the fluorescence of nanomolar ligand **23** upon binding to HCAII (Fig. 8a).³² The low K_D value of **23** makes it suitable to perform competition (displacement) assays with compounds binding in the same concentration range. However, some optimization of the assay was required to perform experiments with foldamer-containing ligands because the

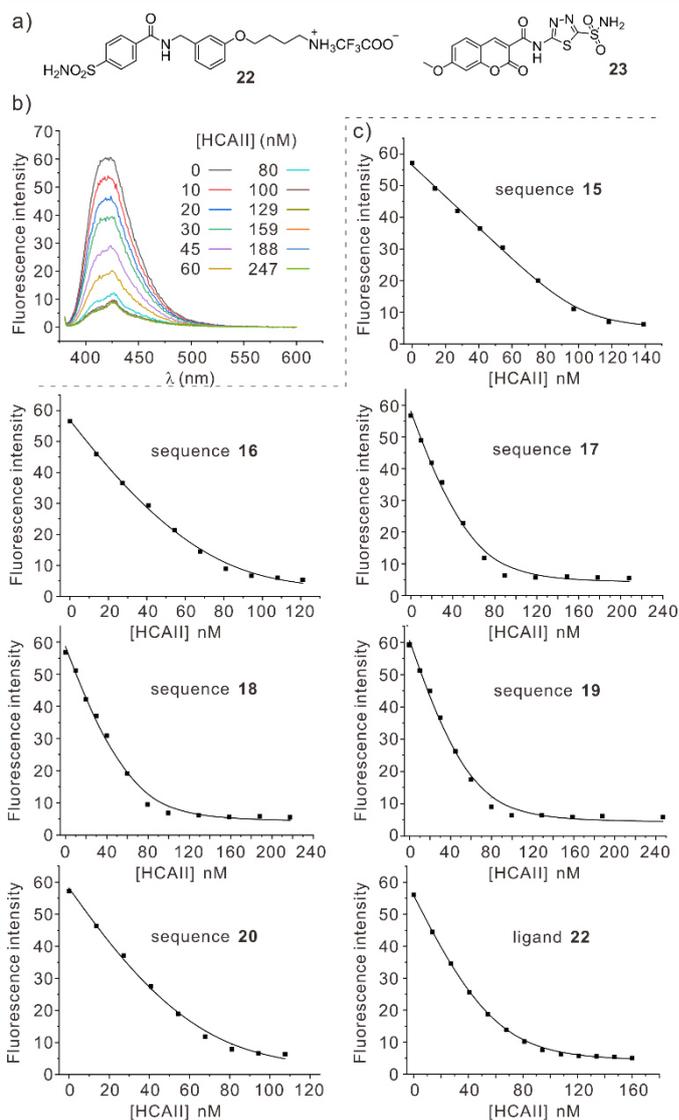


Fig. 8. a) Structural formula of HCAII ligand **22** and of fluorescence probe **23** used in the competition assay. b) Changes in fluorescence spectra (380 nm – 600 nm) upon HCAII titration of sequence **19**. [**19**] = [**23**] = 50 nM; c) Experimental (■) and calculated values using a 1:1 binding isotherm (-) of fluorescence intensity of sequences **15–20** and **22** titrated with HCAII. [competing species] = [**23**] = 50 nM. Note that the curve fitting is shown at one wavelength (419 nm) as an illustration but that the K_D values (Table 1) were calculated by simultaneously fitting data recorded in the 380–600 nm range.

foldamers absorb both at the excitation (373 nm) and emission (400–450 nm) wavelengths of **23**. Performing classical direct displacement titrations where a foldamer is added to a solution containing **23** and HCAII would be complicated by variable inner filter effects. Instead, we performed titrations in which aliquots of an HCAII solution, typically 1 μ M, were added to a solution already containing a foldamer (50 nM) and **23** (50 nM). The foldamer and **23** were also present at the same concentrations in the HCAII solution. This way, the concentrations of fluorophore and foldamer were kept constant and only the ratio of HCAII was varied. A representative titration is shown in Fig. 8b and the corresponding K_D values are shown in Table 1. With this assay, the K_D value for simple ligand **22** was 10 nM compared to 5 nM previously measured by SPR with HCAII

immobilized on the SPR chip under slightly different buffer conditions.^{5,8a}

Table 1. Dissociation constants of the complexes formed with HCAII determined by the fluorescence competition assay.

HCAII binder ^a	K_D (nM) ^b
22 (reference ligand)	10
15	1.5
16 (with Q ^{SPhe10})	5.2
17 (with Q ^{SPhe10} , B ^{Gpr11})	10
18 (with B ^{Gpr6} , Q ^{SPhe10} , G ^{pr11})	7.2
19 (with B ^{Gpr6} , Q ^{SPhe10} , G ^{pr11})	9.3
20 (with B ^{Id26} , Q ^{SPhe10})	7.4

^a Some remarkable features are indicated in parenthesis. ^b Values were found to be repeatable within $\pm 15\%$ in duplicate experiments.

In order to confirm these data, we tried to develop an alternate competition assay using BLI. A new biotinylated HCAII ligand **24** (Fig. 9a) was synthesized which, after immobilization on streptavidin sensors allowed for an accurate K_D determination of its association with HCAII (Fig. 9). Immobilized **24** may in principle act as a reporter of the concentrations of free HCAII in solution. However, in this case as well, the kinetics were slow, a steady state regime was not reached. A calibration curve could in principle be produced by intercepting a value on the sensorgrams after a fixed amount of time instead of waiting until a steady state is reached. However, this proved not to be accurate enough to reliably determine the free HCAII concentration in solution.

Coming back to the K_D values measured with the fluorescence competition assay, it appears that sequence **15**, with no added side chain in positions 6, 10 and 11, is the best binder and that all others bind similarly. These results should nevertheless be taken with caution. The foldamers suffer from low water solubility and a contribution from foldamer aggregation cannot be excluded. For instance, when the fluorescence titrations were performed at higher concentrations (*e.g.* [**23**] = 100 nM, [foldamer] = 200 nM, Fig. S16), the apparent K_D were higher than in Table 1, consistent with an effect of aggregation that reduces the effective foldamer concentration for binding to HCAII. The values in Table 1 could therefore also reflect that sequences **16-20** aggregate more than **15**, *e.g.* because of their hydrophobic Q^{SPhe10} residue.

It remains that none of the additional side chain combination of **16-20** appear to result in a strong enhancement of their affinity for HCAII. It should be pointed that the structure-based design is intended to stabilize the complex, that is, to slow down complex dissociation. The effect of the additional side chains on the kinetics of complex formation, *e.g.* potentially slowing it down, remains unknown and is not taken into account in the computations. Early studies on HCAII ligands had shown that higher affinity correlated with faster complex formation rather than slower complex dissociation.³³ Another early study also reported the lack of effect of extending an HCAII ligand in the search for secondary binding sites, a result comparable to ours in an approach conceptually similar, albeit

with much smaller molecules.³⁴ Finally, it may be that the very architecture of foldamers **15-20** makes it difficult for side chain modification to result in strong effects. These compounds consist of small nanomolar ligand to which is appended a much

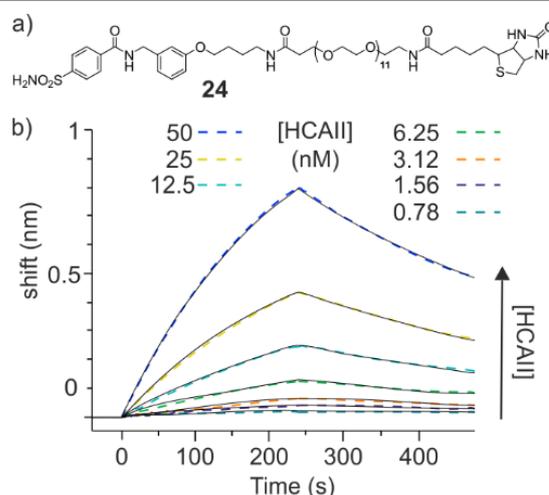


Fig. 9. a) Structural formula of biotinylated HCAII ligand **24**. b) BLI sensorgrams (black solid lines) of the titration of **24** immobilized on the streptavidin sensors at different HCAII concentrations. Calculated curves based on a 1:1 binding isotherm (colored dashed lines) fit with the measured values, yielding $K_D = 38.2$ nM.

larger foldamer that has inherently no affinity for HCAII even in the low micromolar range – no induced CD is observed at 35 μ M with **21** which lacks an HCAII ligand (Fig. S2). In other words, the starting affinity of the foldamers for HCAII is too low to hope that a few modifications will bring it to an interesting range of K_D value. In this respect, the choice of HCAII as a model system was perhaps not ideal. HCAII is a therapeutically relevant target and transmembrane isoforms HCAIX and HCAXII are overexpressed in some cancers and identified as potential targets as well.^{24,28} Nevertheless, we selected HCAII as a model system mainly for its robustness, easy overexpression, good crystal growth ability and the availability of simple nanomolar ligands that could act as tethers. To our knowledge, the vicinity of the HCAII active site is not involved in protein-protein interactions and deprived of any hotspot that may facilitate foldamer binding.

Conclusions

In summary, starting from the crystal structure of the complex between HCAII and tetradecaamide foldamer **1**, we have used computational tools to identify main chain and side chain modifications that may result in an extended foldamer-protein interface. New monomers and sequences incorporating these monomers were synthesized and several solid state structures of complexes with HCAII validated the design principles. We find that Q, B, and P main chain variations are interchangeable also in the context of a foldamer-protein complex. We also find that side chains may generally be introduced independently from one another, a result that was consistent in both MD simulations and solid state structures. This behavior is in sharp

contrast with that of peptides and peptidic foldamers in which a local change, *e.g.* a side chain modification, may also result in a different behavior of the main chain.¹⁷ The robustness of aromatic foldamer helices should therefore represent a good starting point for the structure-based design of protein binders that cover large protein surface areas. Nevertheless, the modifications explored in this study did not result into stronger associations nor did they deliver foldamers that would bind HCAII without a ligand or a covalent tether to mediate the interactions. The grand challenge of the *ab initio* design of an aromatic foldamer to bind a given protein surface remains unmet. In the meantime, other studies have revealed the potential of some aromatic helical foldamers to mimic α -helices^{13a} or DNA double helices.¹⁰ Solid state structures have been obtained of complexes between chromosomal protein Sac7d and a DNA-mimic foldamer^{10d} and between a fragment of ubiquitin ligase E6AP and a foldamer-peptide macrocycle.³⁵ In these complexes, no ligand or covalent tethering are involved. Furthermore, reliable K_D determination methods are available. These structure thus represent new candidates to apply the structure-based design principles validated here. Steps in these directions are being made and will be reported in due course.

Author contributions

Except for the first and last authors, the author list groups the authors by institution and does not reflect a ranking of their contributions. LW, PSR and CD performed solution phase and solid phase syntheses. LW performed protein expression. BLE, LW and JS performed crystal growth. LF, JS and TG carried out crystallographic structure elucidation. LW and CD performed fluorescence and BLI titrations. ZL and YY performed computational studies. VP, YZ and IH supervised the research. IH, LW, CD and ZL wrote the manuscript. All authors reviewed and edited the manuscript and approved its final version.

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) through project CRC1309-C7 (project ID 325871075) to I. H. It has benefited from the facilities and expertise of IECB Biophysical and Structural Chemistry platform (BPCS), CNRS UAR3033, Inserm US001, Univ. Bordeaux. We thank P. Mateus for valuable advice for the implementation of the fluorescence competition assay and Dr. V. Morozov for providing some HCAII. We also thank J. Buratto, M. Soler-Lopez (ESFR beamline ID30B), M. Bowler (ESRF beamline ID30A-1), and M. Savko (SOLEIL beamline Proxima-2A) for assistance during X-ray data collection.

Conflicts of interest

There are no conflicts to declare.

Data availability

A data availability statement (DAS) is required to be submitted alongside all articles. Please read our [full guidance on data availability statements](#) for more details and examples of suitable statements you can use.

Notes and references

‡ Sequences **7** and **11** are very similar – they differ only from one methylene group in the side chain of Q11 – so the fact that in both cases the MD trajectories significantly deviate from initial solid state coordinates is probably not a coincidence. The exact reason is not clear but it appears that the combination of an anionic side chain on Q6 and a cationic side chain on Q11 may lead to this behavior. These two side chains are five units apart, *i.e.* exactly two helix turns, and thus only about 7 Å from each other. Salt bridges between them are observed especially in the case of **7** (40% of the simulation time) which has a longer cationic side chain.

§ For SPR measurements: 2:98 DMSO/aqueous phosphate saline buffer PBS at pH 7.4 (vol/vol) at 25 °C, with PBS = 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄, 2.7 mM KCl and 137 mM NaCl. For fluorescence measurements: 50 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) at pH 7.2.

- 1 T. Seedorf, A. Kirschning and D. Solga, *Chem. – Eur. J.*, 2021, **27**, 7321.
- 2 P. G. Baraldi, M. Del Carmen Nunez, A. Espinosa and R. Romagnoli, *Curr. Top. Med. Chem.*, 2004, **4**, 231; Y Hiraku, S. Oikawa and S. Kawanishi, *Nucl. Acid Res.*, 2002, **2**, 95; A. Paul, P. Guo, D. W. Boykin and W. D. Wilson, *Molecules*, 2019, **24**, 1.
- 3 S. Baumann, J. Herrmann, R. Raju, H. Steinmetz, K. I. Mohr, S. Huttel, K. Harmrolfs, M. Stadler and R. Muller, *Angew. Chem., Int. Ed.*, 2014, **53**, 14605.
- 4 E. Michalczyk, K. Hommernick, I. Behroz, M. Kulike, Z. Pakosz-Stepien, L. Mazurek, M. Seidel, M. Kunert, K. Santos, H. von Moeller, B. Loll, J. B. Weston, A. Mainz, J. G. Heddle, R. D. Sussmuth and D. Ghilarov, *Nat. Catal.*, 2023, **6**, 52; S. M. Hashimi, *J. Antibiot.*, 2019, **72**, 785.
- 5 J. Zeelen, M. van Straaten, J. Verdi, A. Hempelmann, H. Hashemi, K. Perez, P. D. Jeffrey, S. Halg, N. Wiedemar, P. Maser, F. N. Papavasiliou and C. E. Stebbins, *Nat. Microbiol.*, 2021, **6**, 392; G. H. M. Salvador, T. R. Dreyer, A. A. S. Gomes, W. L. G. Cavalcante, J. I. Dos Santos, C. A. Gandin, M. de Oliveira Neto, M. Gallacci and M. R. M. Fontes, *Sci. Rep.*, 2018, **8**, 10317; L. Jiao, S. Ouyang, M. Liang, F. Niu, N. Shaw, W. Wu, W. Ding, C. Jin, Y. Peng, Y. Zhu, F. Zhang, T. Wang, C. Li, X. Zuo, C. H. Luan, D. Li and Z. J. Liu, *J. Virol.*, 2013, **87**, 6829; M. T. Murakami, E. Z. Arruda, P. A. Melo, A. B. Martinez, S. Calil-Elias, M. A. Tomaz, B. Lomonte, J. M. Gutierrez and R. K. Arni, *J. Mol. Biol.*, 2005, **350**, 416; G. H. M. Salvador, T. R. Dreyer, W. L. Cavalcante, F. F. Matioli, J. I. Dos Santos, A. Velazquez-Campoy, M. Gallacci and M. R. Fontes, *Acta. Cryst.*, 2015, **71**, 2066; E. Mastrangelo, M. Pezzullo, D. Tarantino, R. Petazzi, F. Germani, D. Kramer, I. Robel, J. Rohayem, M. Bolognesi and M. Milani, *J. Mol. Biol.*, 2012, **419**, 198.
- 6 I. Saraogi, J. A. Hebda, J. Becerril, L. A. Estroff, A. D. Miranker and A. D. Hamilton, *Angew. Chem., Int. Ed.*, 2010, **49**, 736; T. Flack, C. Romain, A. J. P. White, P. R. Haycock and A. Barnard, *Org. Lett.*, 2019, **21**, 4433; R. A. Dohoney, J. A. Joseph, C. Baysah, A. G. Thomas, A. Siwakoti, T. D. Ball and S. Kumar, *ACS Chem. Biol.*, 2023, **18**, 1510; I. Arrata, C. M. Grison, H. M. Coubrough, P. Prabhakaran, M. A. Little, D. C. Tomlinson, M. E. Webb and A. J. Wilson, *Org. Biomol. Chem.*, 2019, **17**, 3861; G. M. Burslem, H. F. Kyle, A. L. Breeze, T. A. Edwards, A. Nelson, S. L. Warriner and A. J. Wilson, *ChemBiochem*, 2014, **15**, 1083; S. Kumar and A. D. Hamilton, *J. Am. Chem. Soc.*, 2017, **139**,

- 5744; J. P. Plante, T. Burnley, B. Malkova, M. E. Webb, S. L. Warriner, T. A. Edwards and A. J. Wilson, *Chem. Commun.*, 2009, **34**, 5091; N. H. Stillman, J. A. Joseph, J. Ahmed, C. Zuwu Baysah, R. A. Dohoney, T. D. Ball, A. G. Thomas, T. C. Fitch, C. M. Donnelly and S. Kumar, *Nat. Commun.*, 2024, **15**, 3658.
- 7 S. Kumar, M. Birol, D. E. Schlamadinger, S. P. Wojcik, E. Rhoades and A. D. Miranker, *Nat. Commun.*, 2016, **7**, 11412; J. Ahmed, T. C. Fitch, C. M. Donnelly, J. A. Joseph, T. D. Ball, M. M. Bassil, A. Son, C. Zhang, A. Ledreux, S. Horowitz, Y. Qin, D. Paredes and S. Kumar, *Nat. Commun.*, 2022, **13**, 2273.
- 8 (a) J. Buratto, C. Colombo, M. Stupfel, S. J. Dawson, C. Dolain, B. Langlois d'Estaintot, L. Fischer, T. Granier, M. Laguerre, B. Gallois and I. Huc, *Angew. Chem., Int. Ed.*, 2014, **53**, 883; (b) M. Jewginski, T. Granier, B. Langlois d'Estaintot, L. Fischer, C. D. Mackereth and I. Huc, *J. Am. Chem. Soc.*, 2017, **139**, 2928; (c) M. Vallade, M. Jewginski, L. Fischer, J. Buratto, K. Bathany, J.-M. Schmitter, M. Stupfel, F. Godde, C. D. Mackereth and I. Huc, *Bioconjugate Chem.*, 2019, **30**, 54.
- 9 P. S. Reddy, B. Langlois d'Estaintot, T. Granier, C. D. Mackereth, L. Fischer and I. Huc, *Chem. – Eur. J.*, 2019, **25**, 11042.
- 10 (a) K. Ziach, C. Chollet, V. Parissi, P. Prabhakaran, M. Marchivie, V. Corvaglia, P. P. Bose, K. Laxmi-Reddy, F. Godde, J.-M. Schmitter, S. Chaignepain, P. Pourquier and I. Huc, *Nat. Chem.*, 2018, **10**, 511; (b) V. Corvaglia, D. Carbajo, P. Prabhakaran, K. Ziach, P. K. Mandal, V. D. Santos, C. Legeay, R. Vogel, V. Parissi, P. Pourquier and I. Huc, *Nucleic. Acids. Res.*, 2019, **47**, 5511; (c) V. Kleene, V. Corvaglia, E. Chacin, I. Forne, D. B. Konrad, P. Khosravani, C. Douat, C. F. Kurat, I. Huc and A. Imhof, *Nucleic. Acids. Res.*, 2023, **51**, 9629; (d) D. Deepak, J. Wu, V. Corvaglia, L. Allmendinger, M. Scheckenbach, P. Tinnefeld and I. Huc, *Angew. Chem., Int. Ed.*, 2025, **64**, e202422958.
- 11 L.-G. Milroy, T. N. Grossmann, S. Hennig, L. Brunsveld and C. Ottmann, *Chem. Rev.* 2014, **114**, 4695. H. Lu, Q. Zhou, J. He, Z. Jiang, C. Peng, R. Tong and J. Shi, *Sig. Transduct. Target. Ther.* 2020, **5**, 213; J. M. Ostrem, U. Peters, M. L. Sos, J. A. Wells and K. M. Shokat, *Nature*, 2013, **503**, 548; C. V. Pagba, A. K. Gupta, A. K. Naji, D. van der Hoeven, K. Churion, X. Liang, J. Jakubec, M. Hook, Y. Zuo, M. Martinez de Kraatz, J. A. Frost and A. A. Gorfe, *ACS. Bio. Med. Chem. Au.*, 2022, **2**, 617.
- 12 T. Qi, V. Maurizot, H. Noguchi, T. Charoenraks, B. Kauffmann, M. Takafuji, H. Ihara and I. Huc, *Chem. Commun.*, 2012, **48**, 6337.
- 13 (a) M. Zwillinger, P. S. Reddy, B. Wicher, P. K. Mandal, M. Csékei, L. Fischer, A. Kotschy and I. Huc, *Chem. – Eur. J.*, 2020, **26**, 17366. (b) M. Zwillinger, P. Sőregi, F. Sanchez, C. Douat, M. Csékei, I. Huc and A. Kotschy, *J. Org. Chem.*, 2025, doi.org/10.1021/acs.joc.4c02900; (c) X. Hu, S. J. Dawson, P. K. Mandal, X. de Hatten, B. Baptiste and I. Huc, *Chem. Sci.*, 2017, **8**, 3741.
- 14 O. Ichihara, J. Barker, R. J. Law and M. Whittaker, *Mol. Inform.*, 2011, **30**, 298; J. M. Mattheisen, C. Limberakis, R. B. Ruggeri, M. S. Dowling, C. W. am Ende, E. Ceraudo, T. Huber, C. L. McClendon and T. P. Sakmar, *J. Am. Chem. Soc.*, 2023, **145**, 11173; A. Heikal, Y. Nakatani, W. Jiao, C. Wilson, D. Rennison, M. R. Weimar, E. J. Parker, M. A. Brimble and G. M. Cook, *Bioorg. Med. Chem. Lett.*, 2018, **28**, 2239; X. Wu, Y. Zhang, S. Liu, C. Liu, G. Tang, X. Cao, X. Lei and J. Peng, *Bioorg. Chem.*, 2022, **127**, 105921; B. C. Doak, R. S. Norton and M. J. Scanlon, *Pharmacol. Ther.*, 2016, **167**, 28.
- 15 M. F. Schmidt and J. Rademann, *Trends Biotechnol.*, 2009, **27**, 512; M. Mondal and A. K. H. Hirsch, *Chem. Soc. Rev.*, 2015, **44**, 2455; P. Frei, R. Hevey and B. Ernst, *Chem. – Eur. J.*, 2019, **25**, 60.
- 16 W. Lu, M. Kostic, T. Zhang, J. Che, M. P. Patricelli, L. H. Jones, E. T. Chouchani and N. S. Gray, *RSC Chem. Biol.*, 2021, **2**, 354.
- 17 M. E. Perrin, B. Li, J. Mbianda, M. Bakail, C. André, G. Moal, P. Legrand, V. Ropars, C. Douat, F. Ochsenein and G. Guichard, *Chem. Commun.*, 2023, **59**, 8696; J. Mbianda, M. Bakail, C. André, G. Moal, M. E. Perrin, G. Pinna, R. Guerois, F. Becher, P. Legrand, S. Traoré, C. Douat, G. Guichard and F. Ochsenein, *Sci. Adv.*, 2021, **7**, eabd9153. H. Wang, R. S. Dawber, P. Zhang, M. Walko, A. J. Wilson and X. Wang, *Chem. Sci.*, 2021, **17**, 5977.
- 18 M. Vallade, P. Sai Reddy, L. Fischer and I. Huc, *Eur. J. Org. Chem.* 2018, 5489.
- 19 D. Bindl, P. K. Mandal and I. Huc, *Chem. – Eur. J.*, 2022, **28**, e202200538
- 20 D. Bindl, E. Heinemann, P. K. Mandal and I. Huc, *Chem. Commun.*, 2021, **57**, 5662.
- 21 S. Dengler, P. K. Mandal, L. Allmendinger, C. Douat and I. Huc, *Chem. Sci.*, 2021, **12**, 11004; V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat and I. Huc, *Chem. –Eur. J.*, 2023, **29**, e202300898.
- 22 V. Corvaglia, J. Wu, D. Deepak, M. Loos and I. Huc, *Chem. – Eur. J.*, 2024, **28**, e202303650
- 23 D. Rooklin, C. Wang, J. Katigbak, P. S. Arora and Y. Zhang, *J. Chem. Inf. Model.*, 2015, **55**, 1585; J. Katigbak, H. Li, D. Rooklin and Y. Zhang, *J. Chem. Inf. Model.*, 2020, **60**, 1494
- 24 T. Gokcen, I. Gulcin, T. Ozturk and A. C. Goren, *J. Enzyme Inhib. Med. Chem.*, 2016, **31**, 180; A. Maresca, C. Temperini, L. Pochet, B. Masereel, A. Scozzafava and C. T. Supuran, *J. Med. Chem.*, 2010, **53**, 335; C. T. Supuran, *Nat. Rev. Drug Discov.*, 2008, **7**, 168.
- 25 D. Gfeller, O. Michielin and V. Zoete, *Nucl. Acids Res.*, 2013, **41**, D327.
- 26 O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 455-461
- 27 D.A. Case, H.M. Aktulga, K. Belfon, D.S. Cerutti, G.A. Cisneros, V.W.D. Cruz eiro, N. Forouzes, T.J. Giese, A.W. Götz, H. Gohlke, S. Izadi, K. Kasavajhala, M.C. Kaymak, E. King, T. Kurtzman, T.-S. Lee, P. Li, J. Liu, T. Luchko, R. Luo, M. Manathunga, M.R. Machado, H.M. Nguyen, K.A. O’Hearn, A.V. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, A. Risheh, S. Schott-Verdugo, A. Shajan, J. Swails, J. Wang, H. Wei, X. Wu, Y. Wu, S. Zhang, S. Zhao, Q. Zhu, T.E. Cheatham III, D.R. Roe, A. Roitberg, C. Simmerling, D.M. York, M.C. Nagan and K.M. Merz Jr., *J. Chem. Inf. Model.*, 2023, **63**, 6183.
- 28 T. O. Wambo, L. Y. Chen, S. F. McHardy and A. T. Tsin, *Biophys. Chem.*, 2016, **214-215**, 54.
- 29 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696.
- 30 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157.
- 31 Z. Liu, A. M. Abramyan, V. Pophristic, *New J. Chem.*, 2015, **39**, 3229
- 32 P. Koutnik, E. G. Shcherbakova, S. Gozem, M. G. Caglayan, T. Minami, P. Anzenbacher, *Chem* 2017, **2**, 27.
- 33 R. W. King and A. S. V. Burgen, *Proc. R. Soc. Lond. B Biol. Sci.*, 1976, **193**, 107
- 34 A. Jain, S. G. Huang and G. M. Whitesides, *J. Am. Chem. Soc.*, 1994, **116**, 5057.
- 35 S. Dengler, R. T. Howard, V. Morozov, C. Tsiamantas, W. Huang, Z. Liu, C. Dobrzanski, V. Pophristic, S. Brameyer, C. Douat, H. Suga and I. Huc, *Angew. Chem. Int. Ed.*, 2023, **62**, e202308408

10. Supplementary Information: Structure-based Design of an Aromatic Helical Foldamer-Protein Interface

Supplementary Information

for

Structure-based design of an aromatic helical foldamer-protein interface

Lingfei Wang,^a Céline Douat,^a Johannes Sigl,^a Post Sai Reddy,^b Lucile Fischer,^b Béatrice Langlois d'Estaintot,^b Zhiwei Liu,^c Vojislava Pophristic,^c Yuwei Yang,^d Yingkai Zhang^d and Ivan Huc^{*a}

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

b. CBMN (UMR5248), Univ. Bordeaux-CNRS-INP, Institut Européen de Chimie et Biologie, 2 rue Escarpit, 33600 Pessac (France).

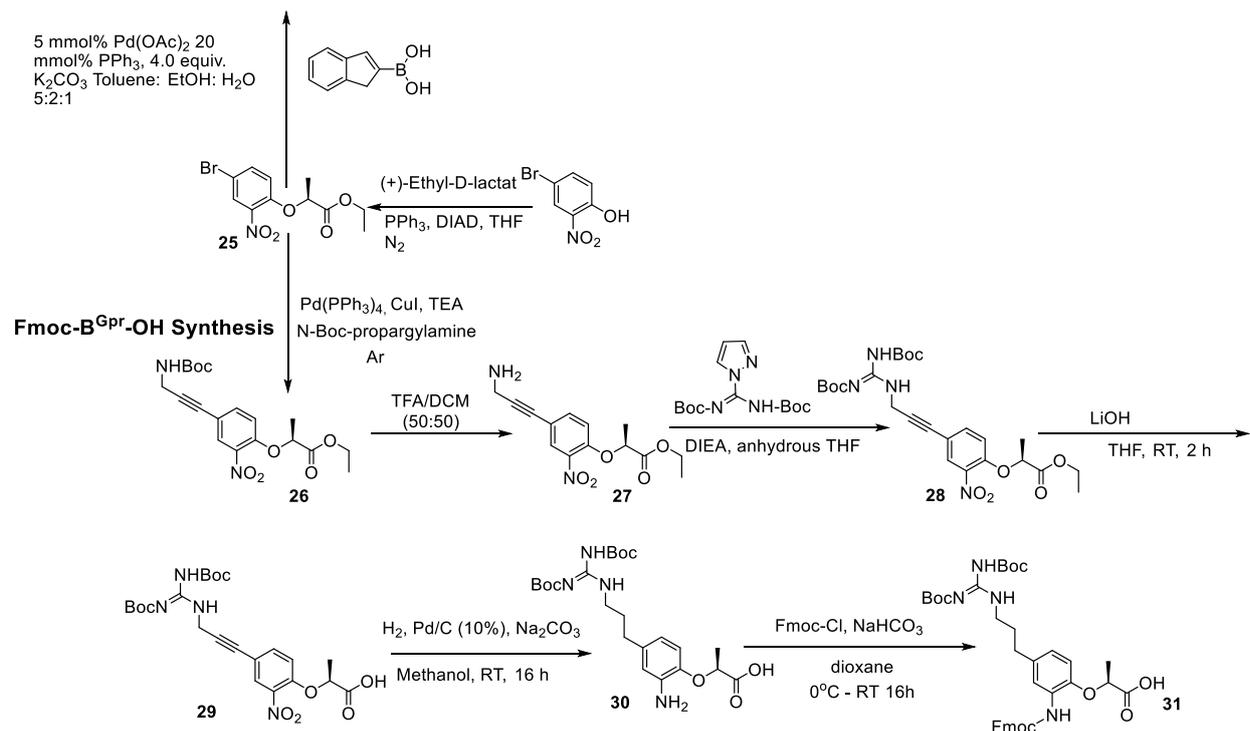
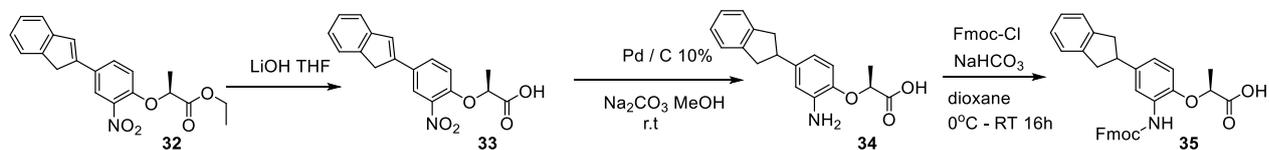
c. Department of Chemistry & Biochemistry, Rowan University, 201 Mullica Hill Road, 08028 Glassboro, New Jersey (USA).

d. Department of Chemistry, New York University, 10003 New York, New York (USA)

10.1 Supplementary Figures and Tables

10.1.1 Supplementary figures

Fmoc-B^{Ida}-OH Synthesis



Fmoc-Q^{5Ph}-OH Synthesis

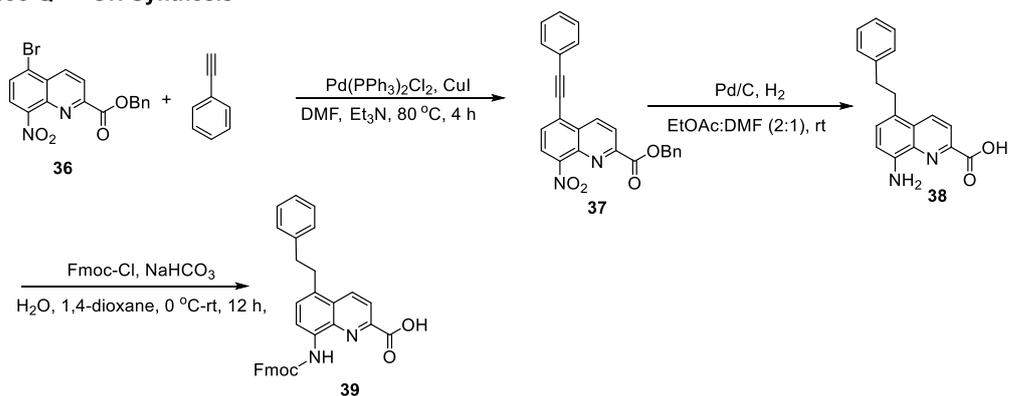
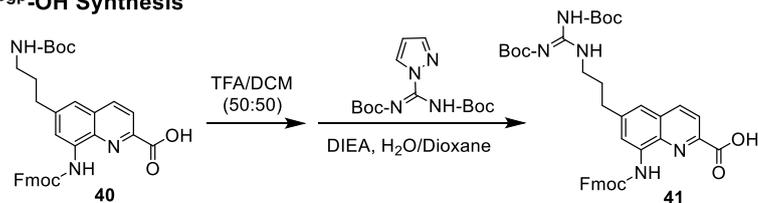
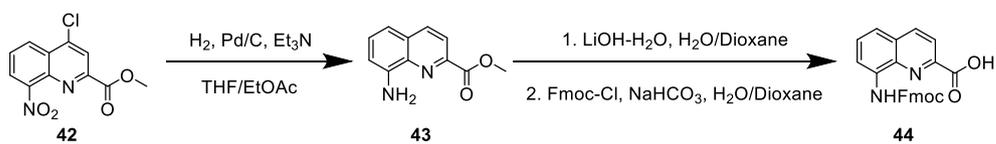


Figure S1 (part 1). Synthetic route of new monomers and biotinylated HCAII ligand.

Fmoc-Q^{6gp}-OH Synthesis



Fmoc-Q^{Gly}-OH Synthesis



Biotinylated HCA II Ligand Synthesis

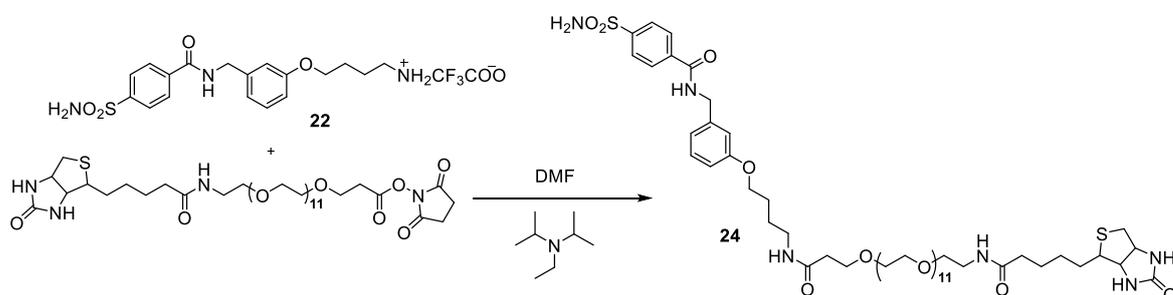


Figure S1 (part 2). Synthetic route of new monomers and biotinylated HCAII ligand.

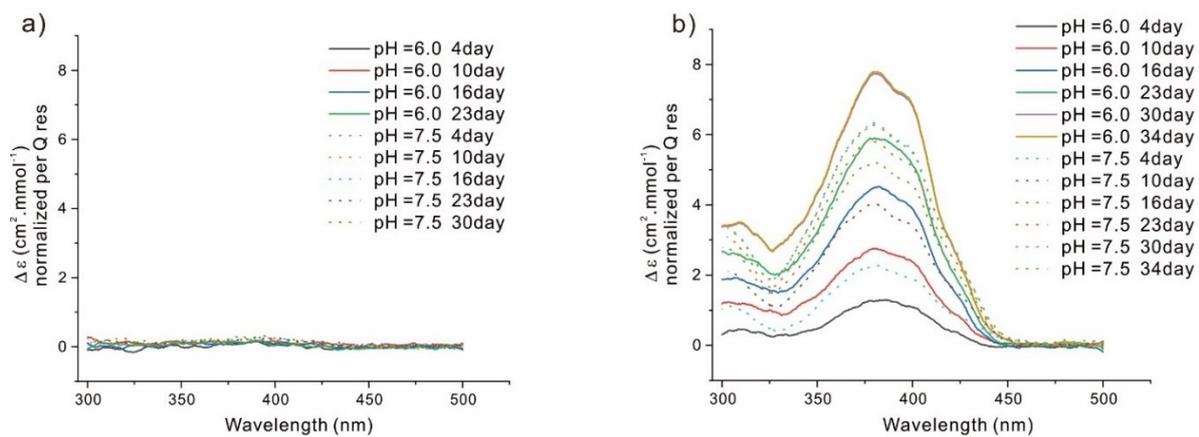
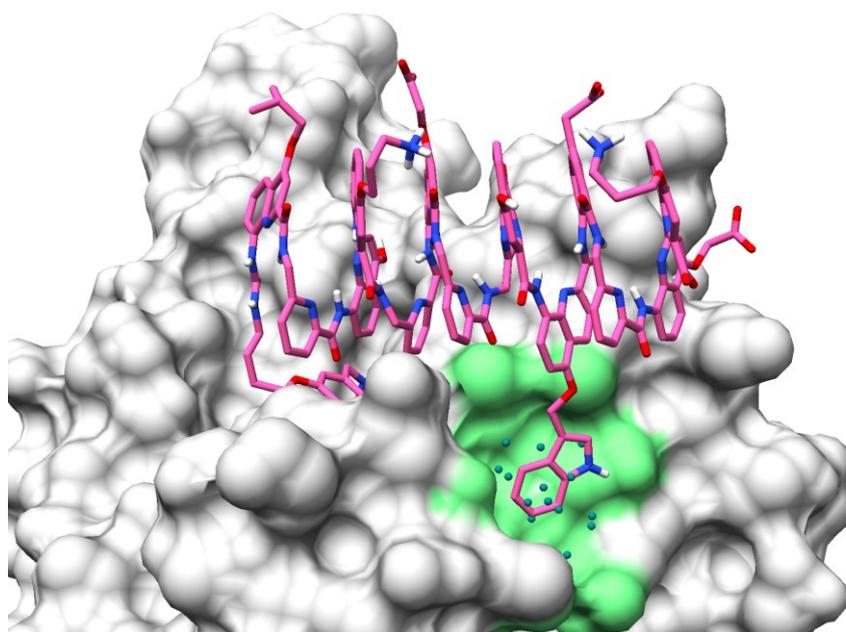
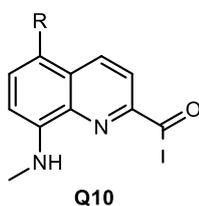
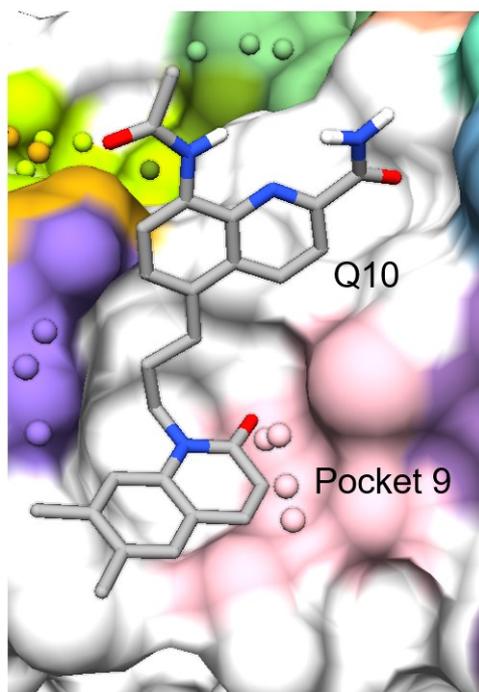


Figure S2. CD spectra of **21** (a) and **3** (b) in the presence of HCAII in phosphonate buffer at pH 6.0 and 7.5 at 20 °C. The concentrations of HCAII and foldamers are both 34.5 μ M.



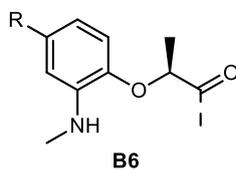
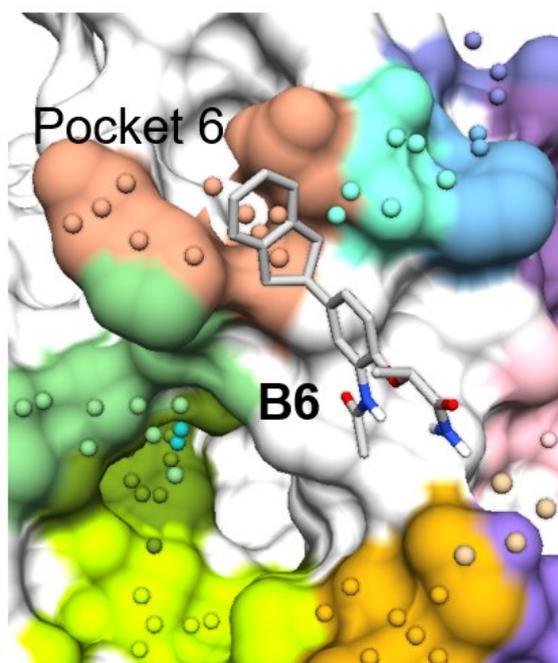
Side chain	ΔG (kcal/mol)
 Q^{5Ph}	-1.4
	-1.7
 Q^{5Pa}	-2.2
	-1.6
	-1.7
 Q^{5In}	-2.0
	-2.1
	-2.5

Figure S3. Docking and scoring of side chains in position 5 of residue Q10 for their interaction with Po4. Note that side chains in position 4 or 6 may be too far from, or too close to, pocket 4. Hence, proposed side chains were placed in position 5. The top scorers are shown.



Side chain	ΔG (kcal/mol)
	-2.6
	-2.5
	-2.5
	-2.4
	-2.3

Figure S4. Docking and scoring of side chains in position 5 of residue Q10 for their interaction with Po4/Po9. The side chains were not derived from the Swiss side chain database but were manually generated. These side chains were not further investigated.



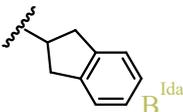
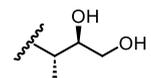
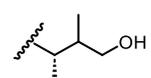
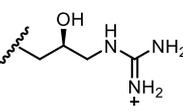
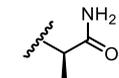
Side chain	ΔG (kcal/mol)
 B^{Ida}	-1.11
	-0.93
	-0.77
	-0.77
	-0.71

Figure S5. Docking and scoring of side chains of residue B6 for their interaction with Po6. The indane side chain and analogues of the guanidinium-containing side chain (lacking the hydroxy group) were further considered.

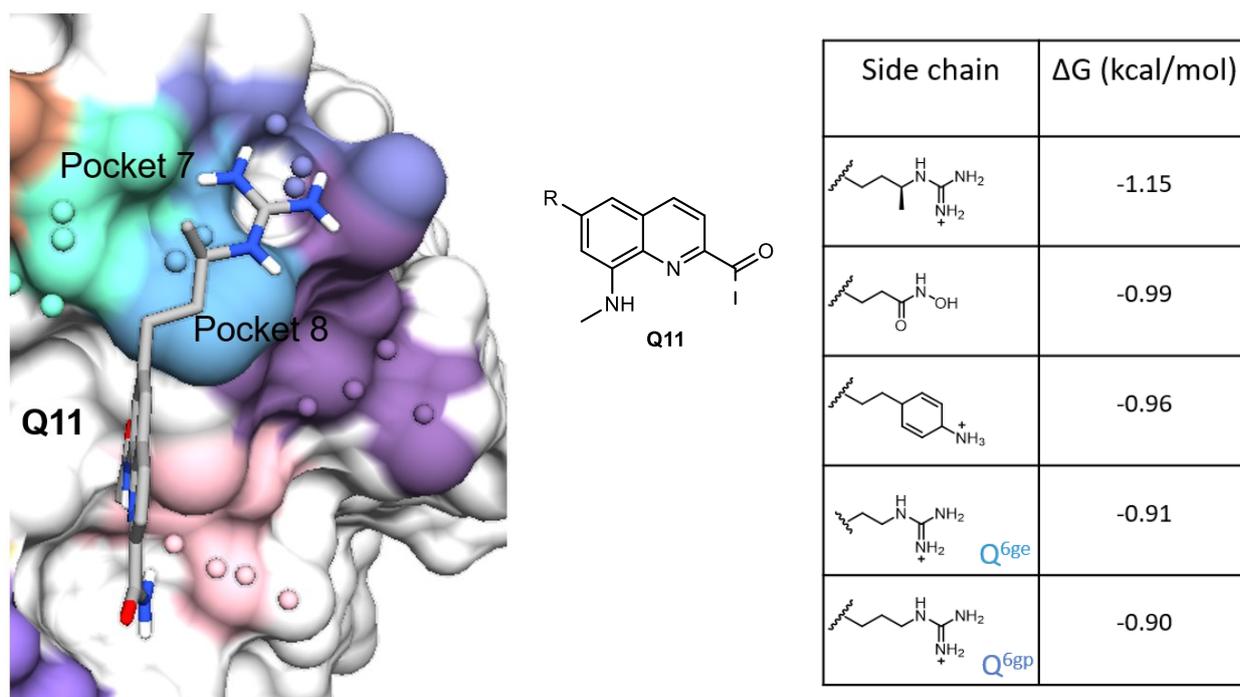


Figure S6. Docking and scoring of side chains ion position 6 of residue Q11 for their interaction with Po7 and Po8. Guanidinium-containing side chains of different lengths were further considered.

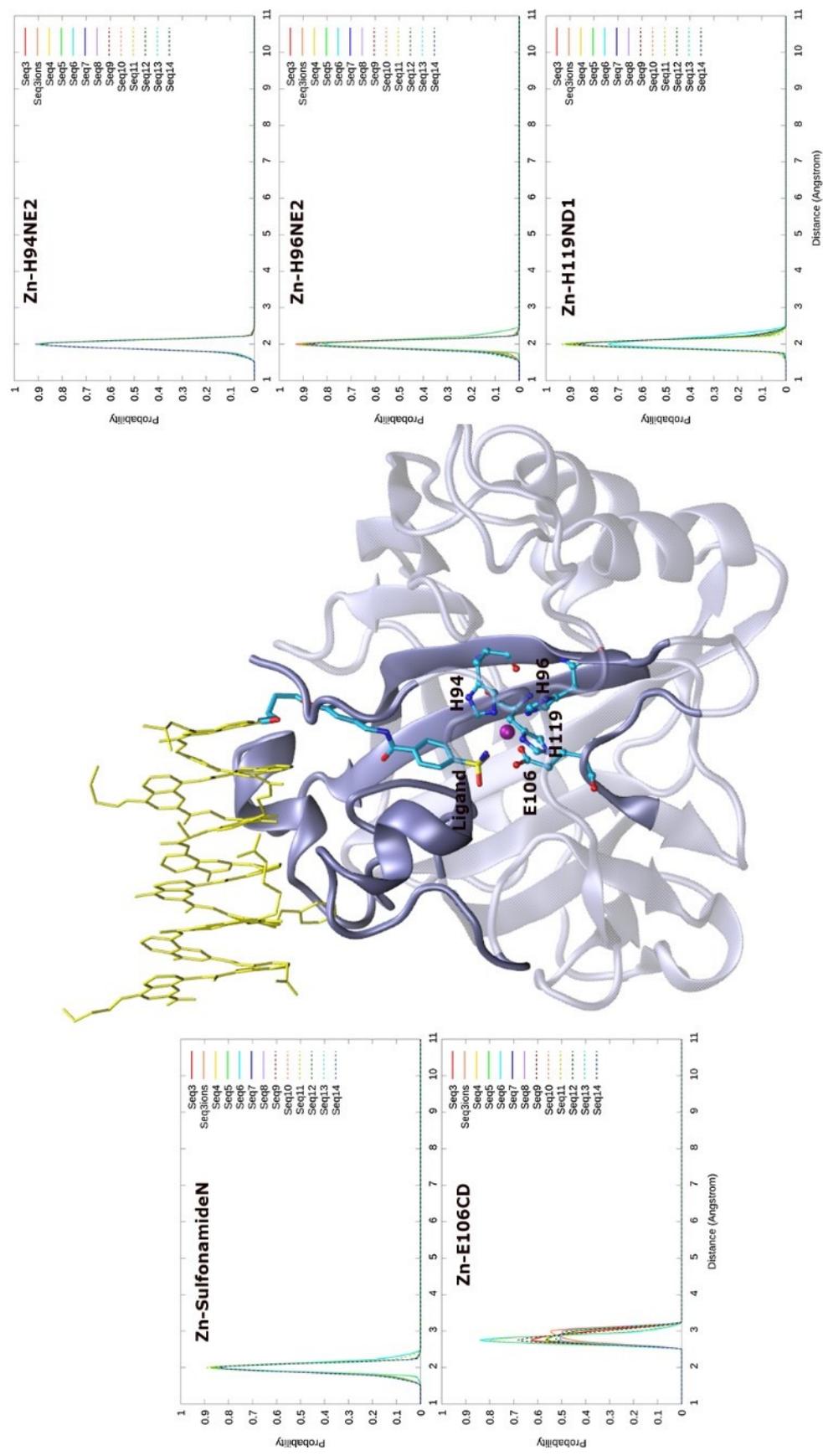


Figure S7. Binding at the ligand binding site. Histograms of distances between Zn²⁺ to HIS94 tele-N, HIS96 tele-N, HIS119 pro-N (right), ligand sulfonamide N and GLU106 δC are calculated based on 500 ns trajectory. The first four distances are constrained during simulation using a weak force constant to ensure proper ligand binding.

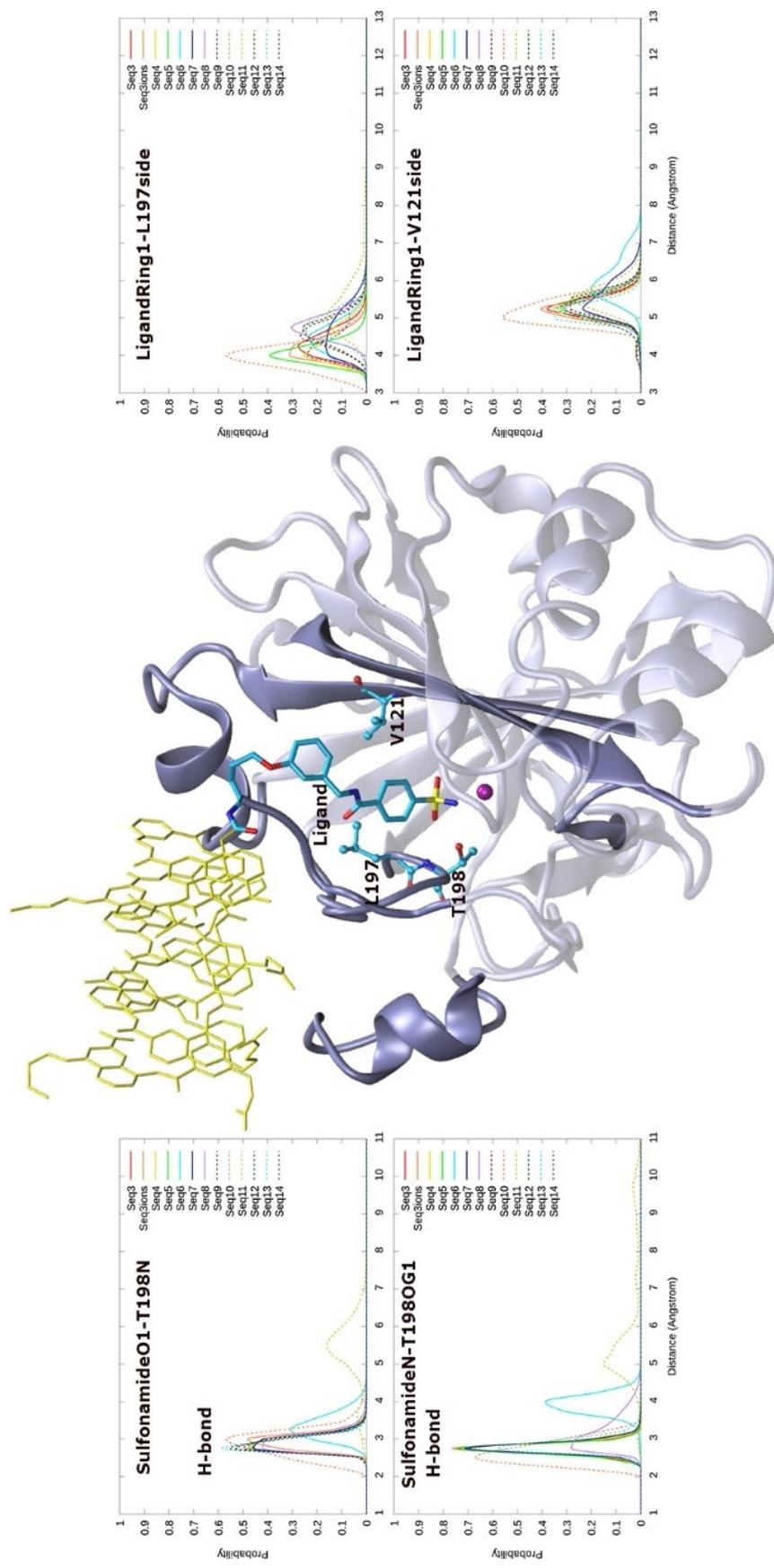


Figure S8. Interaction at pocket 1. Histograms of distances between ligand groups and protein residues indicating either hydrogen bond (H-bond) or hydrophobic interaction are calculated based on 500 ns trajectory.

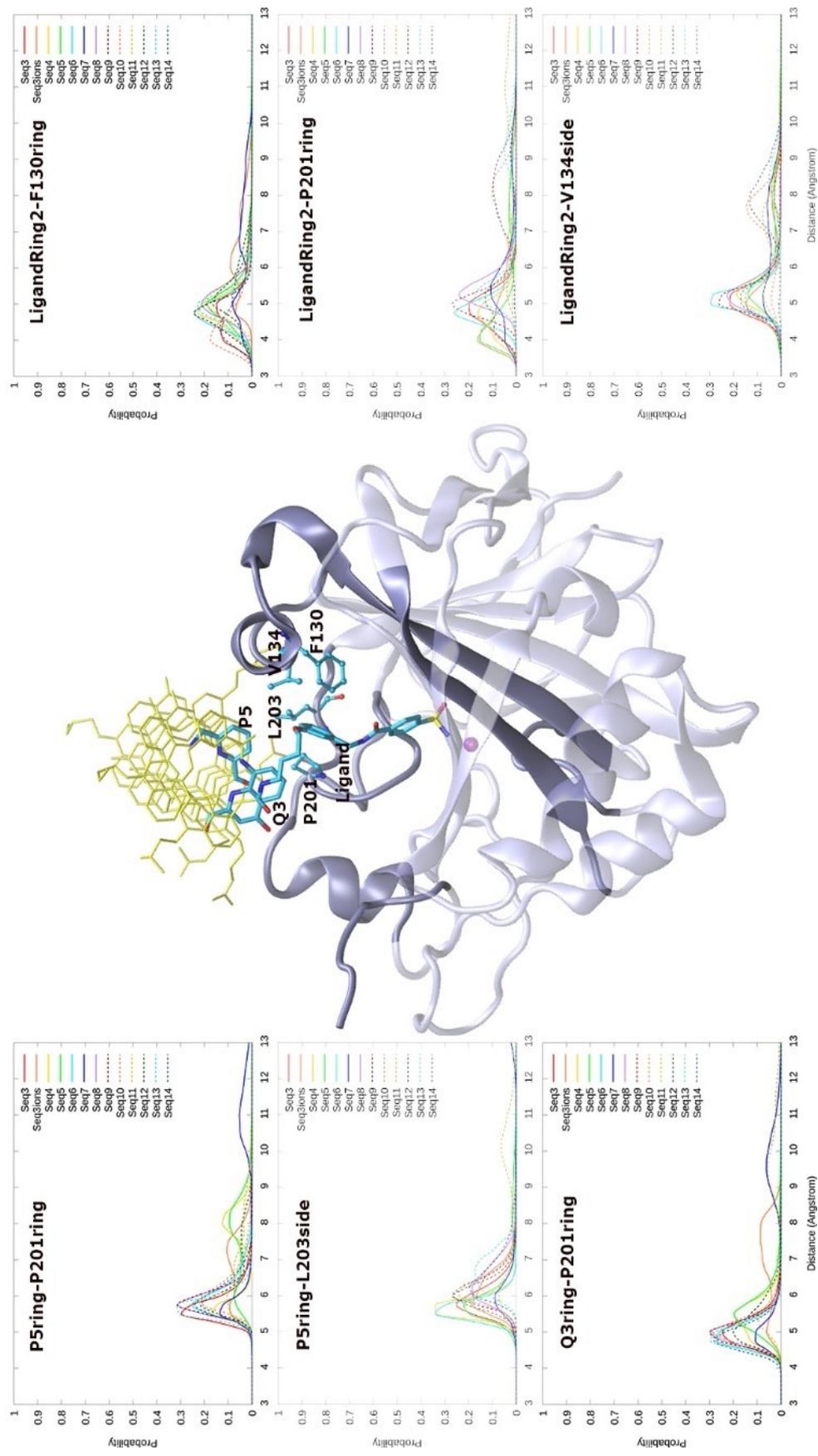


Figure S9. Interaction at pocket 2. Histograms of distances between ligand or foldamer groups and protein residues indicating hydrophobic interaction are calculated based on 500 ns trajectory. (Q3ringN indicates the aromatic ring on the N side of Q3).

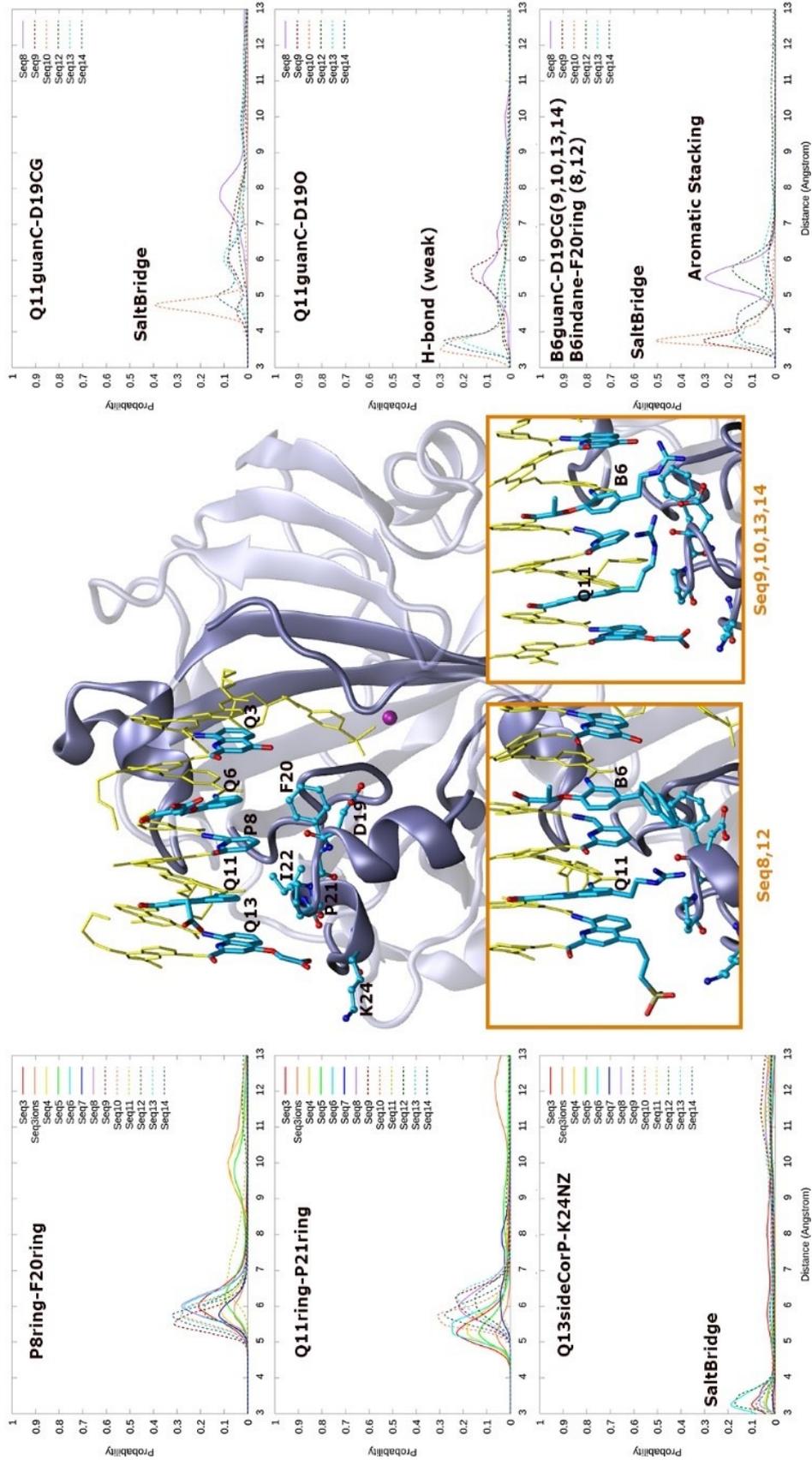


Figure S10. Interaction at pocket 3 or 6 or 7 or 8. Insertions with orange boarders are from different sequences with side chains on unit 6, 11, 13. Histograms of distances between foldamer groups and protein residues indicating salt bridge, H-bond or hydrophobic interactions are calculated based on 500 ns trajectory data.

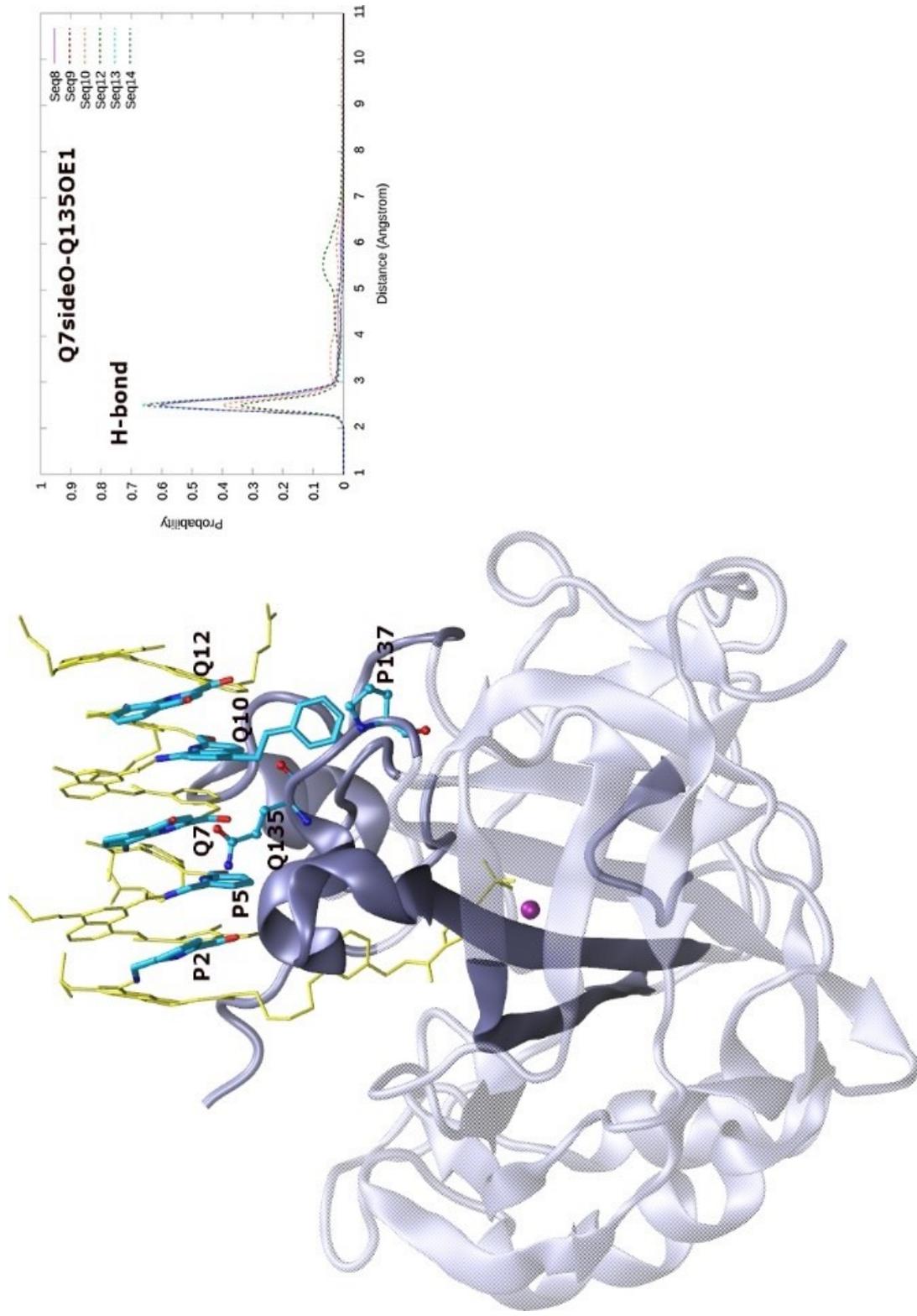


Figure S11. Interaction at pocket 10. Histograms of distances between Q7 sidechain O (of OH group on position 4) and GLN135OE1 indicating H-bond are calculated based on 500 ns trajectory data.

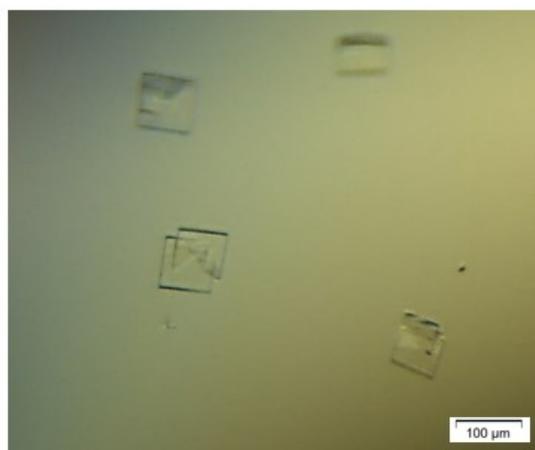
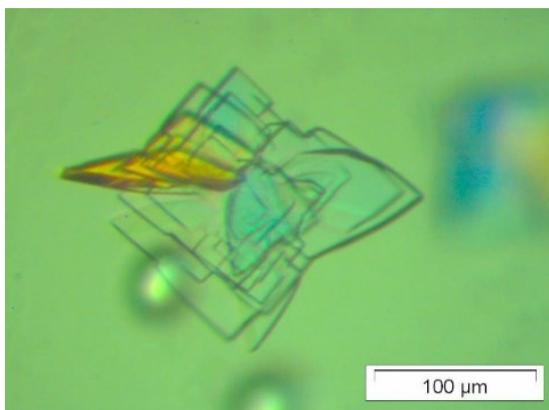
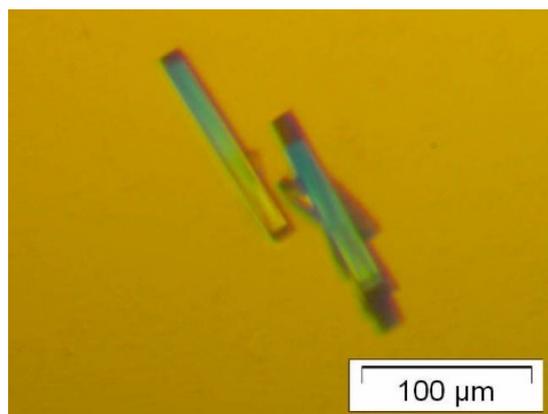
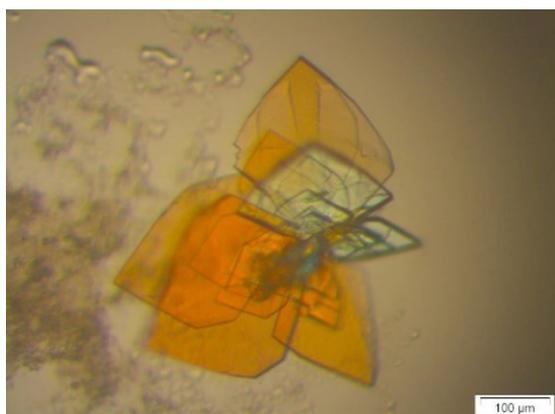


Figure S12. Crystals of complexes of HCAII and **2** (top left), **3** (top right), **16** (bottom left), **20** (bottom right)

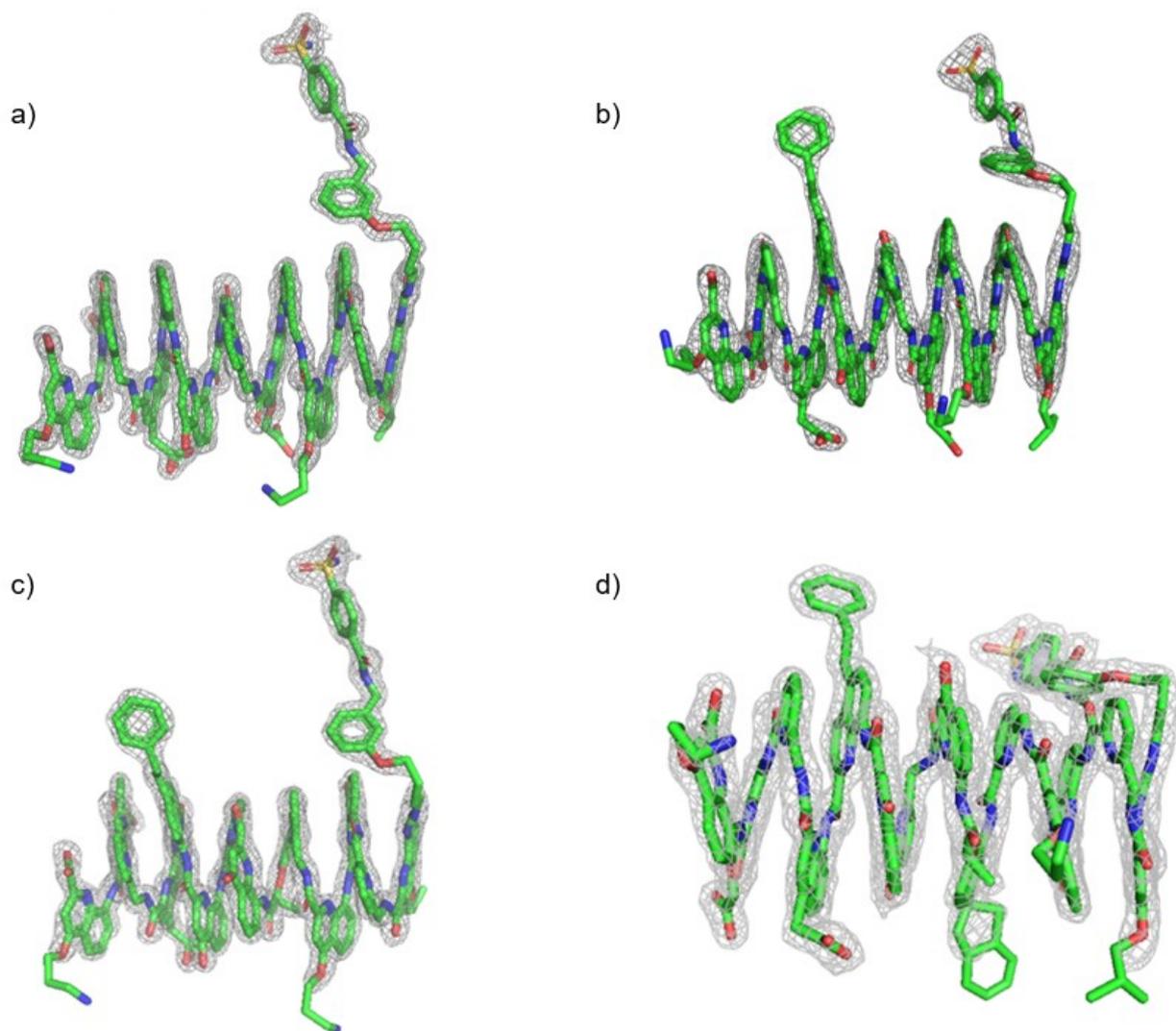


Figure S13. Electron density maps at 1.5σ cut off of four foldamer helices when bound to HCA11 in the solid-state, a): 2, b): 3, c): 16, d): 20.

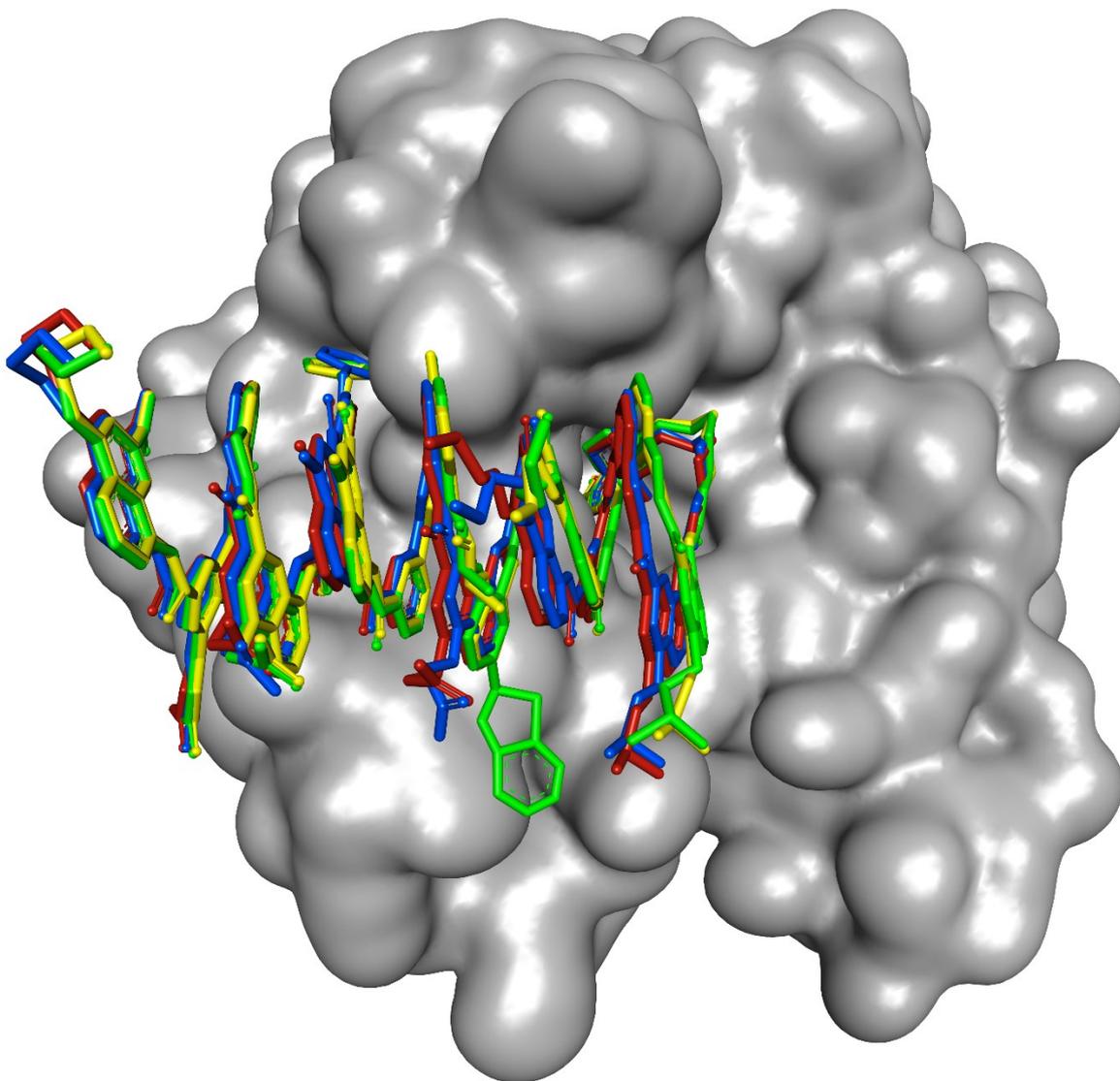


Figure S14: Overlay of foldamer sequences bound to HCAII as they are found in the solid state. The HCAII molecules have been superimposed. Sequence **2** is shown in red, sequence **3** in blue, sequence **16** in yellow, sequence **20** in green. The HCAII surface is colored in grey.

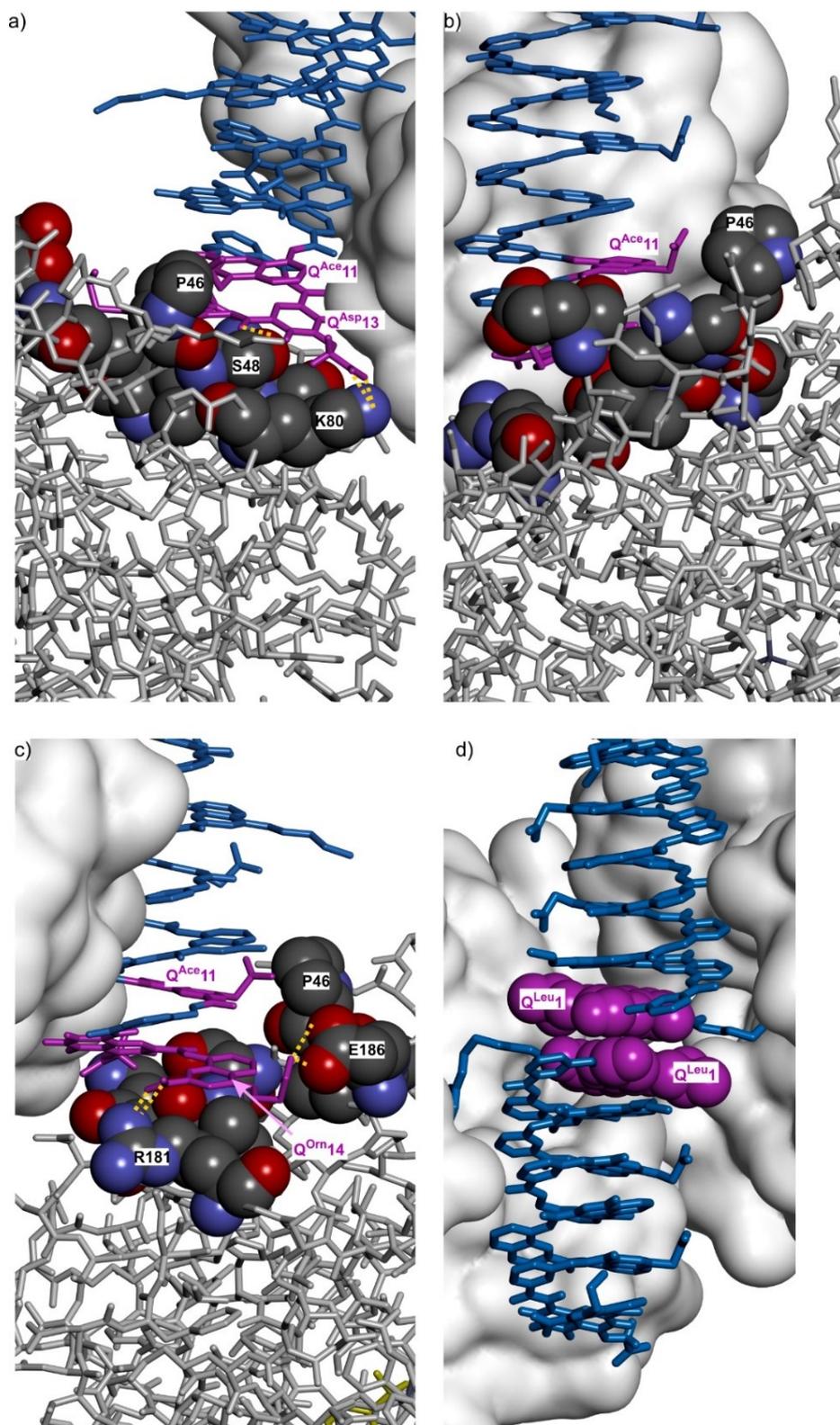


Figure S15. Intercomplex contacts involving the foldamer in the solid state structures

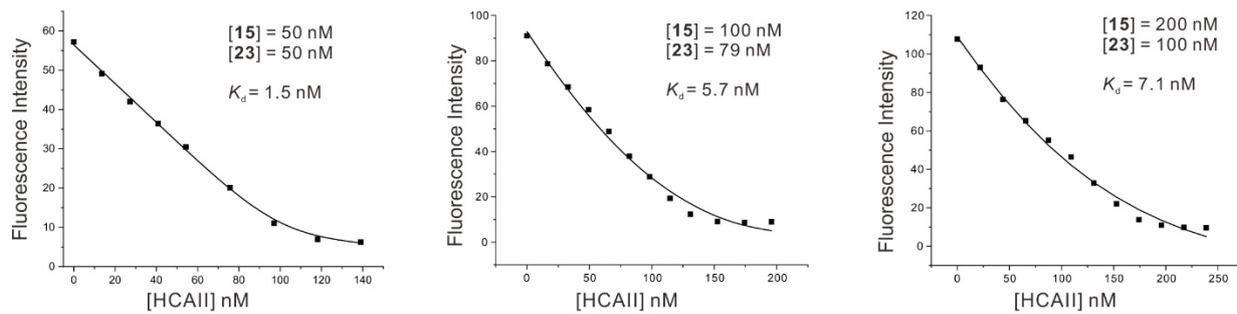


Figure S16. Experimental (■) and calculated (black line -) values for fluorescence intensity of HCAII titration to sequences **15** in different foldamer **15** and fluorescence probe **23** concentration combinations. Note that the curve fitting is shown at one wavelength (419 nm) as an illustration but that the K_D values were calculated by simultaneously fitting 380-600 nm wavelength measured.

10.1.2 Supplementary tables

Table S1. Screening results of unit 6

SMILES	Score (kcal/mol)
<chem>CC(=O)Nc1cc(C2Cc3ccccc3C2)ccc1OC(C)C(N)=O</chem>	-1.1
<chem>CC(=O)Nc1cc([C@@H](C)[C@@H](O)CO)ccc1OC(C)C(N)=O</chem>	-0.9
<chem>CC(=O)Nc1cc([C@@H](C)[C@@H](C)O)ccc1OC(C)C(N)=O</chem>	-0.8
<chem>CC(=O)Nc1cc(C[C@@H](O)CNC(N)=[NH2+])ccc1OC(C)C(N)=O</chem>	-0.8
<chem>CC(=O)Nc1cc([C@H](O)C(N)=O)ccc1OC(C)C(N)=O</chem>	-0.7
<chem>CC(=O)Nc1cc(CC(=O)c2ccccc2N)ccc1OC(C)C(N)=O</chem>	-0.7
<chem>CC(=O)Nc1cc(CCNC(N)=[NH2+])ccc1OC(C)C(N)=O</chem>	-0.7
<chem>CC(=O)Nc1cc(C[C@H](O)C(=O)O)ccc1OC(C)C(N)=O</chem>	-0.6
<chem>CC(=O)Nc1cc(C[C@H](O)C(N)=O)ccc1OC(C)C(N)=O</chem>	-0.6
<chem>CC(=O)Nc1cc(C[C@@H]2CN[C@H](F)N2)ccc1OC(C)C(N)=O</chem>	-0.6
<chem>CC(=O)Nc1cc([C@@H](C)CC(N)=O)ccc1OC(C)C(N)=O</chem>	-0.6
<chem>CC(=O)Nc1cc([C@@H](C)CC(=O)O)ccc1OC(C)C(N)=O</chem>	-0.6
<chem>CC(=O)Nc1cc([C@H](C)CC(=O)O)ccc1OC(C)C(N)=O</chem>	-0.6
<chem>CC(=O)Nc1cc(CCCNC(N)=[NH2+])ccc1OC(C)C(N)=O</chem>	-0.6
<chem>CC(=O)Nc1cc(C[C@H](F)C(=O)O)ccc1OC(C)C(N)=O</chem>	-0.6
<chem>CC(=O)Nc1cc([C@H](O)CO)ccc1OC(C)C(N)=O</chem>	-0.5
<chem>CC(=O)Nc1cc(CCO)ccc1OC(C)C(N)=O</chem>	-0.5
<chem>CC(=O)Nc1cc([C@@H](C)CCBr)ccc1OC(C)C(N)=O</chem>	-0.5
<chem>CC[C@H](C)c1ccc(OC(C)C(N)=O)c(NC(C)=O)c1</chem>	-0.5
<chem>CC[C@@H](C)c1ccc(OC(C)C(N)=O)c(NC(C)=O)c1</chem>	-0.5
<chem>CC(=O)Nc1cc(CCc2ccccc2)ccc1OC(C)C(N)=O</chem>	-0.5
<chem>CC(=O)Nc1cc(CC(F)F)ccc1OC(C)C(N)=O</chem>	-0.5
<chem>CC(=O)Nc1cc(CCN)ccc1OC(C)C(N)=O</chem>	-0.5
<chem>CC(=O)Nc1cc([C@H](C)C(=O)O)ccc1OC(C)C(N)=O</chem>	-0.5
<chem>CC(=O)Nc1cc(CCC(=O)NO)ccc1OC(C)C(N)=O</chem>	-0.4
<chem>CNC(=O)CCc1ccc(OC(C)C(N)=O)c(NC(C)=O)c1</chem>	-0.4
<chem>CC(=O)Nc1cc([C@H](C)C=O)ccc1OC(C)C(N)=O</chem>	-0.4
<chem>CC(=O)Nc1cc(CCC[C@@H](N)C(=O)O)ccc1OC(C)C(N)=O</chem>	-0.4
<chem>CC(=O)Nc1cc(CCON)ccc1OC(C)C(N)=O</chem>	-0.4
<chem>CC(=O)Nc1cc(C(C)C)ccc1OC(C)C(N)=O</chem>	-0.4
<chem>CC(=O)Nc1cc(CCC(=O)O)ccc1OC(C)C(N)=O</chem>	-0.4

CC(=O)Nc1cc(CCC(N)=O)ccc1OC(C)C(N)=O	-0.4
CC(=O)Nc1cc(C(O)O)ccc1OC(C)C(N)=O	-0.4
CC(=O)CCc1ccc(OC(C)C(N)=O)c(NC(C)=O)c1	-0.4
COC(=O)CCc1ccc(OC(C)C(N)=O)c(NC(C)=O)c1	-0.4
CC(=O)Nc1cc(CCON=C(N)N)ccc1OC(C)C(N)=O	-0.4
CC(=O)Nc1cc(CCCC(=O)C(=O)O)ccc1OC(C)C(N)=O	-0.4
CC(=O)Nc1cc(CCCC(=O)O)ccc1OC(C)C(N)=O	-0.4
CC(=O)Nc1cc(CCCNC(N)=S)ccc1OC(C)C(N)=O	-0.4
CC(=O)Nc1cc([C@H](O)C(C)C)ccc1OC(C)C(N)=O	-0.4
CC(=O)Nc1cc([C@@H](O)C(C)C)ccc1OC(C)C(N)=O	-0.4
CC(=O)Nc1cc(CCCCO)ccc1OC(C)C(N)=O	-0.4
CC(=O)Nc1cc(CCCCN)ccc1OC(C)C(N)=O	-0.4
C=Cc1ccc(OC(C)C(N)=O)c(NC(C)=O)c1	-0.3
CC(=O)Nc1cc([C@@H](C)O)ccc1OC(C)C(N)=O	-0.3
CC(=O)Nc1cc([C@H](C)O)ccc1OC(C)C(N)=O	-0.3
C=C(CCCc1ccc(OC(C)C(N)=O)c(NC(C)=O)c1)C(=O)O	-0.3
CC(=O)Nc1cc(CCCNC(N)=[NH2+])ccc1OC(C)C(N)=O	-0.3
CC(=O)Nc1cc(CCCNC(N)=O)ccc1OC(C)C(N)=O	-0.3
C=CCc1ccc(OC(C)C(N)=O)c(NC(C)=O)c1	-0.3
CC(=O)Nc1cc(CCCCCN)ccc1OC(C)C(N)=O	-0.3
CCc1ccc(OC(C)C(N)=O)c(NC(C)=O)c1	-0.3
CC(=O)Nc1cc(C[C@@H](C)C=O)ccc1OC(C)C(N)=O	-0.3
CCC(CC)c1ccc(OC(C)C(N)=O)c(NC(C)=O)c1	-0.3
CC(=O)Nc1cc(CC=O)ccc1OC(C)C(N)=O	-0.3
CC[C@@H](O)c1ccc(OC(C)C(N)=O)c(NC(C)=O)c1	-0.3
CC(=O)Nc1cc(C[C@@H](C)C(=O)O)ccc1OC(C)C(N)=O	-0.3
CC(=O)Nc1cc(C[C@H](C)C(F)(F)F)ccc1OC(C)C(N)=O	-0.3
CCCCC1ccc(OC(C)C(N)=O)c(NC(C)=O)c1	-0.3
CC(=O)Nc1cc(CNC(N)=[NH2+])ccc1OC(C)C(N)=O	-0.3
CC(=O)Nc1cc(CCCO)ccc1OC(C)C(N)=O	-0.3
CC(=O)Nc1cc(CCSC(F)F)ccc1OC(C)C(N)=O	-0.3
CC(=O)Nc1cc(CO)ccc1OC(C)C(N)=O	-0.2
CCc1ccc(OC(C)C(N)=O)c(NC(C)=O)c1	-0.2
CCCCc1ccc(OC(C)C(N)=O)c(NC(C)=O)c1	-0.2
CC(=O)Nc1cc(CCCN)ccc1OC(C)C(N)=O	-0.2

<chem>CC(=O)Nc1cc(CCSCO)ccc1OC(C)C(N)=O</chem>	-0.2
<chem>CC(=O)Nc1cc(C[C@H](C)CF)ccc1OC(C)C(N)=O</chem>	-0.2
<chem>CC(=O)Nc1cc(CN)ccc1OC(C)C(N)=O</chem>	-0.2
<chem>CC(=O)Nc1cc(CCSC#N)ccc1OC(C)C(N)=O</chem>	-0.2
<chem>CC(=O)Nc1cc(CC2CCCC2)ccc1OC(C)C(N)=O</chem>	-0.2
<chem>CC(=O)Nc1cc([C@H](O)C(Cl))ccc1OC(C)C(N)=O</chem>	-0.2
<chem>CC(=O)Nc1cc([C@H](O)C(=O)O)ccc1OC(C)C(N)=O</chem>	-0.2
<chem>CCSCCc1ccc(OC(C)C(N)=O)c(NC(C)=O)c1</chem>	-0.2
<chem>CC(=O)Nc1cc(C(F)(F)F)ccc1OC(C)C(N)=O</chem>	-0.2
<chem>CC(=O)Nc1cc(CC(C)C)ccc1OC(C)C(N)=O</chem>	-0.2
<chem>CC(=O)Nc1cc(CC(=O)O)ccc1OC(C)C(N)=O</chem>	-0.1
<chem>CC(=O)Nc1cc(C)ccc1OC(C)C(N)=O</chem>	-0.1
<chem>CC(=O)Nc1cc(CC[C@H]2C=C[C@@H](N)C=C2)ccc1OC(C)C(N)=O</chem>	-0.1
<chem>CC(=O)Nc1cc(CC2CCCCC2)ccc1OC(C)C(N)=O</chem>	-0.1
<chem>CSCCc1ccc(OC(C)C(N)=O)c(NC(C)=O)c1</chem>	-0.1
<chem>CC(=O)Nc1cc(CCl)ccc1OC(C)C(N)=O</chem>	-0.1
<chem>CO[C@H](C)c1ccc(OC(C)C(N)=O)c(NC(C)=O)c1</chem>	-0.1
<chem>CC(=O)Nc1cc(CCS)ccc1OC(C)C(N)=O</chem>	-0.1

Table S2. Screening results of unit 10

SMILES	Score (kcal/mol)
<chem>CC(=O)Nc2ccc(Cc1cccc1)c3ccc(C(N)=O)nc23</chem>	-1.4
<chem>CC(=O)Nc2ccc(Cc1ccc(Cl)cc1)c3ccc(C(N)=O)nc23</chem>	-1.7
<chem>CC(=O)Nc2ccc(Cc1ccc(C(N)=[NH2+])cc1)c3ccc(C(N)=O)nc23</chem>	-2.2
<chem>CC(=O)Nc2ccc(Cc1ccc(O)cc1)c3ccc(C(N)=O)nc23</chem>	-1.6
<chem>COc3ccc(Cc1ccc(NC(C)=O)c2nc(C(N)=O)ccc12)cc3</chem>	-1.7
<chem>CC(=O)Nc3ccc(Cc1c[nH]c2cccc12)c4ccc(C(N)=O)nc34</chem>	-2.0
<chem>CC(=O)Nc3ccc(Cc1c[nH]c2cc(Cl)ccc12)c4ccc(C(N)=O)nc34</chem>	-2.1
<chem>CC(=O)Nc3ccc(Cc1c[nH]c2c(O)c(N)ccc12)c4ccc(C(N)=O)nc34</chem>	-2.5

Table S3: Screening results of unit 11

SMILES	Score (kcal/mol)
<chem>CC(=O)Nc1cc(CC[C@H](C)NC(N)=[NH2+])cc2ccc(C(N)=O)nc12</chem>	-1.1
<chem>CC(=O)Nc1cc(CCC(=O)NO)cc2ccc(C(N)=O)nc12</chem>	-1.0
<chem>CC(=O)Nc1cc(CC[C@H]2C=C[C@@H](N)C=C2)cc2ccc(C(N)=O)nc12</chem>	-1.0
<chem>C=C(CCCc1cc(NC(C)=O)c2nc(C(N)=O)ccc2c1)C(=O)O</chem>	-0.9
<chem>CC(=O)Nc1cc(CC=Cc2ccccc2)cc2ccc(C(N)=O)nc12</chem>	-0.9
<chem>CC(=O)Nc1cc(CCNC(N)=[NH2+])cc2ccc(C(N)=O)nc12</chem>	-0.9
<chem>CC(=O)Nc1cc(CCCNC(N)=[NH2+])cc2ccc(C(N)=O)nc12</chem>	-0.9
<chem>CC(=O)Nc1cc(CCC[C@@H](N)C(=O)O)cc2ccc(C(N)=O)nc12</chem>	-0.9
<chem>CC(=O)Nc1cc(CCCNC(N)=O)cc2ccc(C(N)=O)nc12</chem>	-0.8
<chem>CC(=O)Nc1cc(CCCNC(N)=[NH2+])cc2ccc(C(N)=O)nc12</chem>	-0.8
<chem>CC(=O)Nc1cc(CCCNC(N)=S)cc2ccc(C(N)=O)nc12</chem>	-0.8
<chem>CNC(=O)CCc1cc(NC(C)=O)c2nc(C(N)=O)ccc2c1</chem>	-0.8
<chem>CC(=O)Nc1cc(CCCN)cc2ccc(C(N)=O)nc12</chem>	-0.7
<chem>CC(=O)Nc1cc(CCCC(=O)C(=O)O)cc2ccc(C(N)=O)nc12</chem>	-0.7
<chem>CC(=O)Nc1cc(CCCO)cc2ccc(C(N)=O)nc12</chem>	-0.7
<chem>CC(=O)Nc1cc(CCCCN)cc2ccc(C(N)=O)nc12</chem>	-0.7
<chem>CC(=O)Nc1cc(CCCCCN)cc2ccc(C(N)=O)nc12</chem>	-0.7
<chem>CC(=O)Nc1cc(CCON=C(N)N)cc2ccc(C(N)=O)nc12</chem>	-0.7
<chem>COC(=O)CCc1cc(NC(C)=O)c2nc(C(N)=O)ccc2c1</chem>	-0.7
<chem>CC(=O)Nc1cc(CCC(C)C)cc2ccc(C(N)=O)nc12</chem>	-0.7
<chem>CC(=O)CCc1cc(NC(C)=O)c2nc(C(N)=O)ccc2c1</chem>	-0.6
<chem>CC(=O)Nc1cc(CCCC(=O)O)cc2ccc(C(N)=O)nc12</chem>	-0.6
<chem>CC(=O)Nc1cc(CCC(=O)O)cc2ccc(C(N)=O)nc12</chem>	-0.6
<chem>CC(=O)Nc1cc(CCCCO)cc2ccc(C(N)=O)nc12</chem>	-0.6
<chem>CCCCC1c1cc(NC(C)=O)c2nc(C(N)=O)ccc2c1</chem>	-0.6
<chem>CC(=O)Nc1cc(CCC(N)=O)cc2ccc(C(N)=O)nc12</chem>	-0.6
<chem>CC(=O)Nc1cc(CCON)cc2ccc(C(N)=O)nc12</chem>	-0.6
<chem>CC(=O)Nc1cc(CCS(C)(=O)=O)cc2ccc(C(N)=O)nc12</chem>	-0.6
<chem>CC(=O)Nc1cc(CCSC(F)F)cc2ccc(C(N)=O)nc12</chem>	-0.6
<chem>CCCCc1cc(NC(C)=O)c2nc(C(N)=O)ccc2c1</chem>	-0.6
<chem>CC(=O)Nc1cc(CCSCO)cc2ccc(C(N)=O)nc12</chem>	-0.5
<chem>CC(=O)Nc1cc(CC[S@@](C)=O)cc2ccc(C(N)=O)nc12</chem>	-0.5

CC(=O)Nc1cc(CC[S@@](C)=O)cc2ccc(C(N)=O)nc12	-0.5
CC(=O)Nc1cc(CCSC#N)cc2ccc(C(N)=O)nc12	-0.5
CC(=O)Nc1cc(CCN)cc2ccc(C(N)=O)nc12	-0.5
CC(=O)Nc1cc(CCO)cc2ccc(C(N)=O)nc12	-0.5
CCSCCc1cc(NC(C)=O)c2nc(C(N)=O)ccc2c1	-0.4
CCCc1cc(NC(C)=O)c2nc(C(N)=O)ccc2c1	-0.4
C=CCc1cc(NC(C)=O)c2nc(C(N)=O)ccc2c1	-0.4
CSCCc1cc(NC(C)=O)c2nc(C(N)=O)ccc2c1	-0.4
CC(=O)Nc1cc(CCe2ccccc2)cc2ccc(C(N)=O)nc12	-0.4
CC(=O)Nc1cc(CCS)cc2ccc(C(N)=O)nc12	-0.4
CCc1cc(NC(C)=O)c2nc(C(N)=O)ccc2c1	-0.3
CC(=O)Nc1cc(CC=O)cc2ccc(C(N)=O)nc12	-0.3

Table S4 Data-collection and refinement statistics (Statistics for the highest-resolution shell are shown in parentheses.)

	HCAII-3	HCAII-2	HCAII-16	HCAII-20
Data collection				
X-ray source	SOLEIL-Proxima-2	SOLEIL-Proxima-2	ESRF- ID30A-1	ESRF- ID30B
Wavelength (Å)	0.979995	0.979999	0.965459	0.87313
Resolution (Å) (last shell)	50.00-2.11 (2.24-2.11)	65.00-1.40 (1.48-1.40)	45.89-1.64 (1.74-1.64)	17.17-2.05 (2.12-20.5)
Space Group	P 2 ₁ 2 ₁ 2			
<i>Cell parameters</i>				
a, b, c (Å)	80.25, 81.29, 46.12	79.596, 81.601, 46.72	78.512, 81.515, 45.893	81.6816, 78.9698, 46.116
α, β, γ (°)	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00
Asymmetric unit	1 complex	1 complex	1 complex	1 complex
Unique reflexions	17876 (2756)	60722 (9648)	67930 (10792)	19166 (3555)?
Multiplicity	6.50 (6.57)	13.27 (13.19)	2.43 (2.31)	4.4 (4.60)
Completeness (%)	99.8 (97.1)	99.9 (99.6)	97.0 (95.7)	99.04 (97.20)
I/σ(I)	10.79 (1.22)	18.31 (1.17)	7.34 (0.51)	7.20 (1.88)
R _{meas} (%)	11.7 (145.9)	7.3 (230.2)	9.2 (212.6)	16.8 (55.31)
CC _{1/2} (%)	99.8 (55.7)	100 (58.8)	99.8 (22.9)	96 (80.1)
Refinement				
R _{work}	0.2003 (0.366)	0.1555 (0.391)	0.1836 (0.412)	0.1708 (0.2056)
R _{free}	0.2524 (0.402)	0.1825 (0.395)	0.2108 (0.385)	0.2309 (0.2937)
r.m.s. bonds (Å)	0.007	0.009	0.009	0.008
r.m.s. angles (°)	1.035	1.425	1.025	2.018
<i>No. of atoms</i>				
Total	2410	2649	2591	2649
protein	2020	2056	2073	2036
ligand	235	227	233	242
water	102	341	278	362
Overall B-factor (Å ²)	50.084	22.393	31.218	25.15
protein	53.522	25.045	34.907	23.02
ligand	43.355	20.683	29.174	22.81

water	55.946	37.165	42.376	31.91
Ramachandran favored (%)	96.47	96.86	96.50	96.47
Molprobit score	1.26	0.91	1.23	1.57
PDB code	9GAK	9GAM	9GAJ	9HGB

10.2 Computational studies

10.2.1 Pocket analysis and pocket-centric side chain screening

Pocket analysis is performed using AlphaSpace 2.0,¹ which detects and quantitatively evaluates concave space on the surface of a protein or protein-ligand complex. In the analysis results, pocket space is a geometric feature related to the size and shape of a pocket, and Bscore measures the optimal amount of free energy that can be gained by occupying a pocket. The procedure of side chain screening is as follows: to increase pocket occupancy, pocket-centric screening is carried out to select suitable side chains. Side chains are either manually designed or selected from SwissSideChain (<https://www.swisssidechain.ch/>). Side chains are attached to the foldamer backbone unit which is located near the pocket of interest. Conformers are generated using RDKit (<https://www.swisssidechain.ch/>) by keeping the backbone atoms fixed. AutoDock Vina² minimization procedure is performed to calculate the binding affinity of each conformer and the highest binding affinity value is used as the score of the side chain.

Side chain screening is performed for pocket 4, pocket 6, and pocket 7 and 8 with side chain alternatives attached to unit 6, unit 10 and unit 11 respectively. For unit 10, the side chain candidates are manually designed; while for unit 6 and unit 11, the candidates are from SwissSideChain.

10.2.2 Molecular dynamics simulations

We carried out MD simulations on sequence **3** to **14** bound to HCAII (with Zn²⁺) in explicit water using the AMBER22 package.³ We built the initial structures of all systems based on the HCAII-**3** crystal structure. Basically, the initial structure of HCAII was directly adopted from the crystal structure and then the foldamer helix was positioned by aligning its backbone with that of HCAII-**3** crystal structure. The foldamer structure was constructed by connecting structurally pre-tuned arylamide building blocks and residues for the ligand. All building blocks/residues are created using a multi-conformational RESP fitting protocol.^{4,5} Each system was then solvated by explicit TIP3P water molecules in a periodic box measuring about 82 Å along each side. The ff14SB force field⁶ was used for α -amino acid residues. The general AMBER force field (GAFF),⁷ with improved torsional parameters for arylamide,⁴ was used for the foldamer. All systems were equilibrated using the same procedure involving solvent minimization, heating and NPT simulation at 1 atm and 300K. Production runs using the NVT ensemble at 300K were then carried out for 500 ns per system. Weak constraints were put in place to constrain the distance between Zn²⁺ to the tele-N, tele-N and pros-N of His residues 94, 96, 119, respectively and the nitrogen of the sulfonamide group of the ligand, to make sure the ligand binds properly.

10.3 Crystallography

The recombinant HCAII enzyme was expressed and purified according to references.⁸ Prior to crystallization, compounds **2**, **3**, **16** and **20** were solubilized in pure DMSO. HCAII (0.3 mM) was preincubated with 1.05 equiv. of foldamer **3**, with 1.05 equiv. of foldamer **2**, with 1.1 equiv. of foldamer **16** and with 1.05 equiv. of foldamer **20** in 50 mM Tris buffer (pH 7.8) containing 3 mM NaN₃.

HCAII-2: For the binary complex of HCAII with foldamer **2**, drops consisted of 0.5 μL of complex solution and 0.5 μL of the precipitant solution containing 0.2 M lithium sulfate, 0.1 M Tris (pH 8.5) PEG 4000 20%, and NaN₃ 3 mM. The drops were equilibrated by vapor diffusion against the precipitant solution at room temperature, and platelets appeared after 2 to 4 weeks and grew to their final size (250 x 75 x 20 μm) within 1 to 2 months. They were cryo-protected in the precipitant solution supplemented by 33% glycerol.

HCAII-3: For the binary complex of HCAII with foldamer **3**, drops consisted of 0.4 μL of complex solution and 0.4 μL of the precipitant solution containing lithium sulfate 0.2 M, Tris 0.1 M pH 8.8, 18% PEG 4000, and NaN₃ 3 mM. The drops were equilibrated by vapor diffusion against the precipitant solution at room temperature, and bladed crystals appeared after two days (125 × 75 × 17 μm). They were cryo-protected in the precipitant solution supplemented by 33% glycerol.

HCAII-16: For the binary complex of HCAII with foldamer **16**, drops consisted of 0.4 μL of complex solution and 0.4 μL of the precipitant solution containing 0.2 M lithium sulfate, 0.1 M Tris (pH 8.0) 24% PEG 4000, and 3 mM NaN₃. The drops were equilibrated by vapor diffusion against the precipitant solution at room temperature, and platelets appeared after 3 days (100 × 50 × 20 μm). They were cryo-protected in the precipitant after addition of one drop of LV CryoOil™.

HCAII-20: For the binary complex of HCAII with foldamer **20** drops consisted of 0.8 μL of complex solution and 0.8 μL of the precipitant solution containing 0.2 M lithium sulfate, 0.1 M Tris (pH 8.5), 30% PEG 4000, and 3 mM NaN₃. The drops were equilibrated by vapor diffusion against the precipitant solution at room temperature, and plate-shaped crystals appeared after 14 days (200 × 120 × 25 μm). They were cryo-protected in the precipitant after addition of 20% glycerol prior to flash freezing in liquid nitrogen.

Data were collected on microfocus beamline Proxima-2A at synchrotron SOLEIL for complexes HCAII-3 and HCAII-2, at synchrotron ESRF on beamline ID30A-1 for complex HCAII-16 and at ESRF beamline ID30B at 100 K for complex HCAII-20.

All data were reduced with XDS or CrysAlisPro (Rigaku Oxford Diffraction, (2024), CrysAlisPro Software system, version 1.171.43.130a, Rigaku Corporation, Wroclaw, Poland)⁹ X-ray structures of HCAII-2, HCAII-3, HCAII-16 were solved by molecular replacement using the program Phaser¹⁰ and

atomic coordinates of a previous protein/foldamer complex (PDB code 6QT9)¹¹ as a search model. The X-ray structure of HCAII-**20** was solved by molecular replacement using programs Phaser¹⁰ and atomic coordinates of an apo-protein structure (PDB: 5EHV)¹² as a search model. Refinement was carried out using Refmac¹³ and Phenix¹⁴ and manual model building using Coot.¹⁵ The topology files used to build and refine the modified inhibitors have been generated using Prodrgr and Phenix eLBOW.¹⁶ The X-ray structures were validated using Molprobity¹⁷ prior to deposition in the RCSB Protein Data Bank (entry codes 9GAK, 9GAM, 9GAJ and 9HGB).

10.4 Biophysical measurements

10.4.1 CD and UV-Vis spectroscopy

Circular Dichroism (CD) spectra were recorded on a Jasco J-815 Circular Dichroism spectrometer using quartz cells of 2 mm optical path length. Scans were measured at 20°C, over a wavelength range of 300–500 nm, with a response time of 0.5 sec and a scanning speed of 50 nm/min. The CD data represents an average of two scans. All CD were baseline-corrected for signal contributions due to the buffer containing HCAII (HCAII 34.5 μM in 50 mM NaH_2PO_4 buffer at pH 7.4 or 6.0). Samples were prepared by adding 1 equiv. of the foldamer (10 mM solution in pure DMSO) to a solution containing HCAII (34.5 μM , in a 50 mM NaH_2PO_4 buffer at pH 7.4 or 6.0).

All ultraviolet–visible (UV/Vis) absorbance measurements were done with a Jasco V-750 spectrophotometer instrument using a 1 cm quartz cuvette. Measurements were performed at 20 °C if not stated otherwise.

A series of solutions of compound **23** (from 0.3 μM to 3.0 μM) were prepared in aqueous 50 mM HEPES buffer (pH 7.2), without degassing as previously described from a DMSO stock solution. The UV absorbance of **23** at 359 nm was shown in **Figure S17** and the whole UV spectrum was shown in **Figure S20**. From 2.0 μM to 3.0 μM (labelled in pink color), the absorbance decreases, which indicates a potential aggregation of **23**. The molar extinction coefficient ($\epsilon_{359\text{nm}}$) was therefore calculated from Beer-Lambert's law and Abs values measured from a range of concentration of 0.3 μM to 2.0 μM . The linear regression gave a value of the $\epsilon_{359\text{nm}}$ of $26613 \text{ M}^{-1} \text{ cm}^{-1}$, which is in good agreement with the reported value $27800 \text{ M}^{-1} \text{ cm}^{-1}$.¹⁸

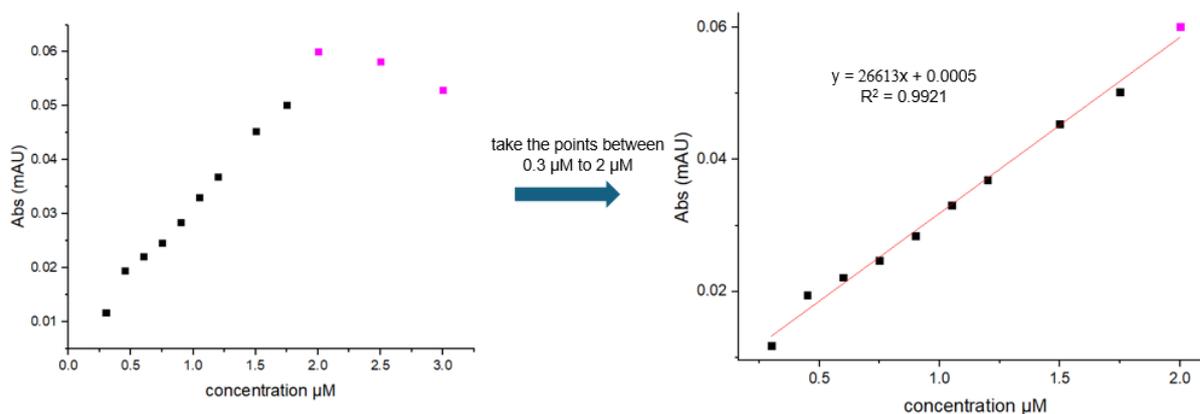


Figure S17. UV absorbance of **23** at 359 nm at different concentrations and UV calibration curve of **23** to compare with the value from literature (after removal of two Abs values above 2 μM).

10.4.2 Fluorescence binding assay

This fluorescence assay was adapted from the paper of Anzenbacher.¹⁸ All fluorescence titrations were done in aqueous 50 mM HEPES buffer (pH = 7.2) at room temperature, without degassing of the samples and using a quartz fluorescence cuvette with 1 cm path length on Varian Cary Eclipse Spectrophotometer 05813. The titration assays were carried out in two parts: 1) sensor-protein titration (**Figure S18**); 2) sensor-foldamer competition assay (**Figure S21**). Several experiments were repeated and validated a reproducibility within 15% error.

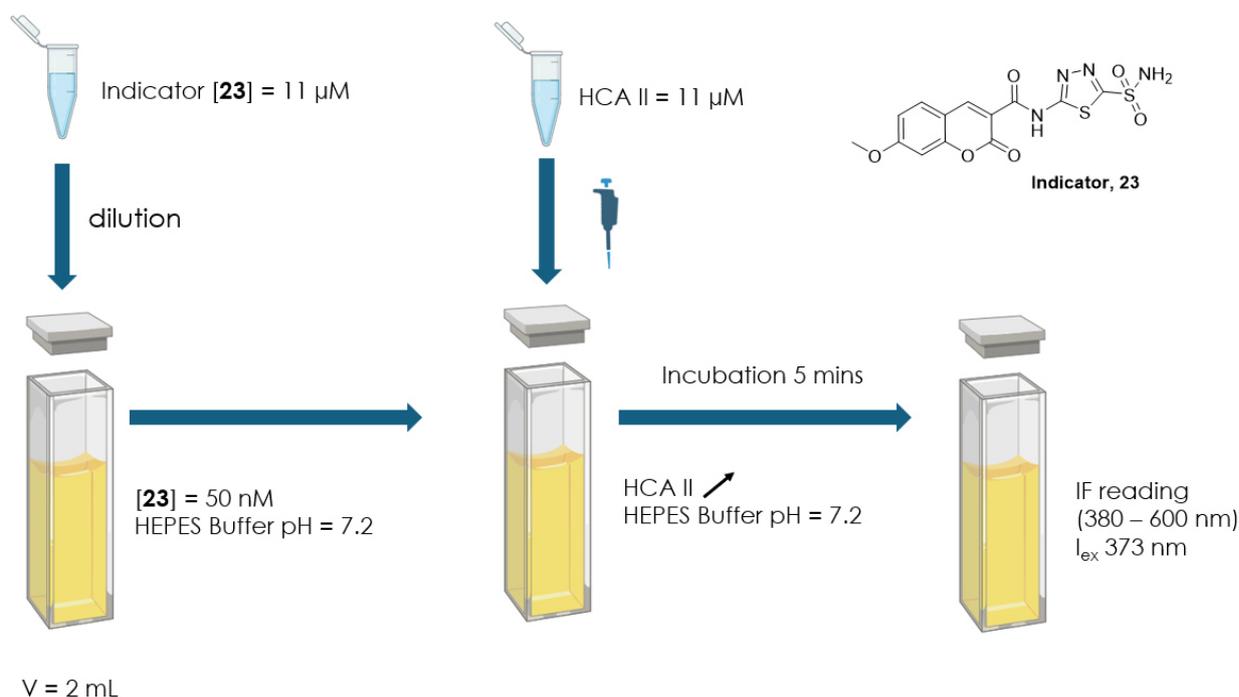


Figure S18. Schematic procedure of the titration experiment between indicator 23 and HCAII (adapted from literature¹⁸).

10.4.2.1 K_D measurement of the 23•HCAII complex

1. Fluorophore **23** was synthesized according to literature¹⁸ and next dissolved in pure DMSO to get a stock solution at 3 mM.
2. To the solution of **23** (50 nM) in HEPES buffer (50 mM, pH 7.2) in the fluorescence cuvette (2 mL), incremental volumes of a stock solution of HCAII (11 μM) in HEPES buffer were added. After each addition and gentle agitation, the resulting mixture was incubated for 5 min at RT. The intensity of fluorescence (IF) was then recorded from 380 nm to 600 nm ($\lambda_{\text{ex}} = 373$ nm).
3. In total around 2.3 equiv. of protein was added (final concentration of protein in the cuvette was 114 nM) (see **Figure S19**). The data were plotted to calculate the K_D value using Hypspec software.¹⁹ For curve fitting, the binding model was set to 1:1. The spectra of all emitting species were recorded from 380 nm to 600 nm ($\lambda_{\text{ex}} = 373$ nm) in separate experiments and set as “known spectrum”. The K_D

value was determined considering all the wavelengths recorded. Errors quoted are standard deviations of the overall constants given directly by the program for the input data.

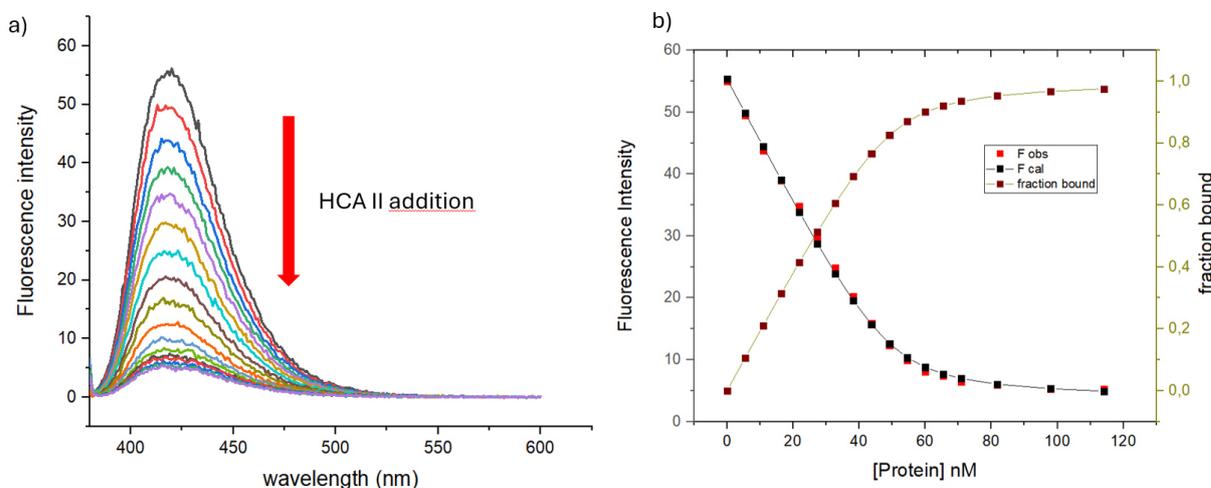


Figure S19. a) Fluorescence spectrum change upon HCAII addition; b) The intensity of fluorescence (IF) values read at $\lambda_{em} = 419$ nm were plotted. The black squares are the read IF values and the red squares are the calculated IF. The overall fitting gives a K_D value for **23** of 1.69 nM, which is in good agreement with the K_D value reported and will next be used for the calculation of foldamer K_D values.

10.4.2.2 Sensor aggregation and reasons for protocol optimization

For foldamer binding affinity determination, we sought to implement a competitive fluorescence titration but performed the experiment after protocol optimization. In the published protocol, the competitors were added to a solution of HCAII-**23** mixture, and the IF change was recorded with increasing concentration of competitors. Titration curves could thus be obtained, and the K_D value of competitors could be calculated. In our experiment, we used foldamers as competitors and we opted not to add the foldamers to a solution of **23**-HCAII complex, but instead, we **added the protein to a solution of foldamer and **23**** by preparing two solutions:

Solution 1 contained the foldamer and **23**

Solution 2 contained the protein, foldamer and **23** mixture (foldamer and **23** concentrations were the same as in **solution 1**).

The **solution 1** was then titrated by **solution 2**, and the read emission spectra were used to calculate the K_D value of foldamers (competitors) by HypSpec software. This experimental design was necessary to address the inner filter effect, which arises from two key factors. Firstly, since quinoline-based foldamers absorb at 373 nm (*i.e.* at the excitation wavelength of **23**), if high concentrations of foldamer were added to the solution of **23**-HCAII complex, the intensity of excitation light would be partly

absorbed by the foldamers, leading to a decrease of emission light intensity of the **23** (primary inner filter effect).

Secondly, quinoline-based foldamers absorb in the wavelength range between 400-450 nm (UV spectrum of foldamer **3** is exemplified in **Figure S20**), overlapping with the emission spectrum of **23**. This inherent aromatic foldamer absorption might diminish the detected fluorescence intensity (secondary inner filter effect). By maintaining the foldamer and **23** concentrations constant during titration, this inner filter effect was kept constant, allowing changes in fluorescence to be attributed solely to the competitive interactions between the protein, **23**, and foldamers. We also worked at a low concentration of **23** and foldamers in the measurement and tested different concentration combinations to make sure the obtained values were reliable.

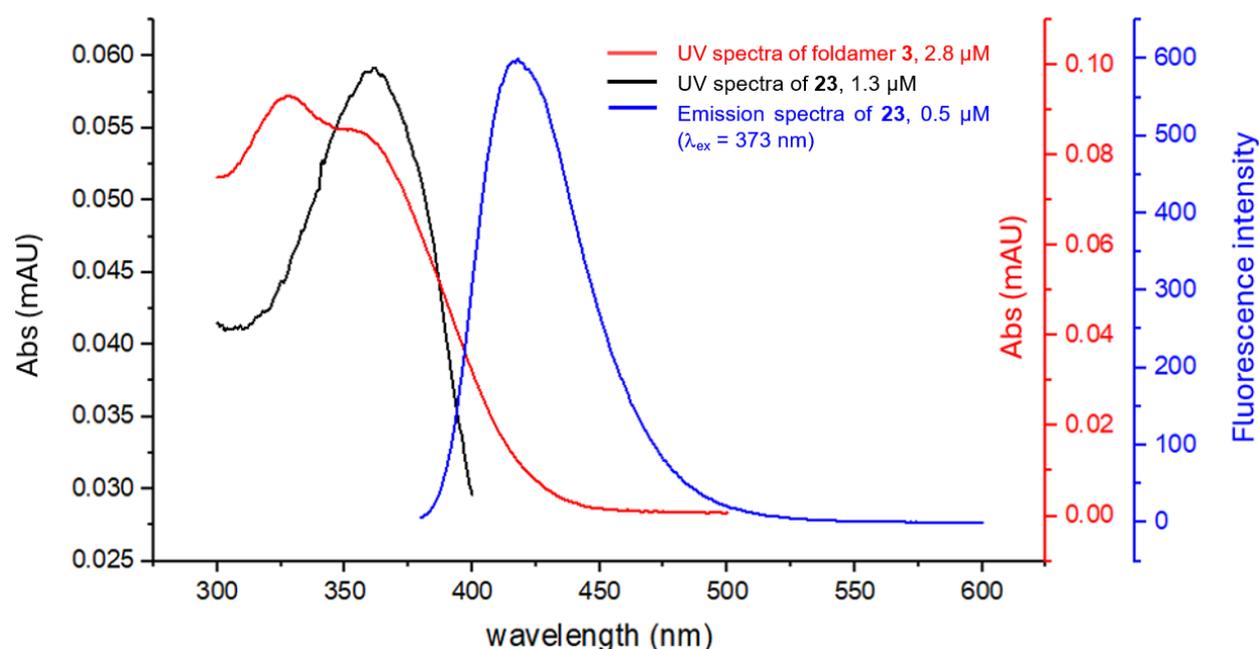
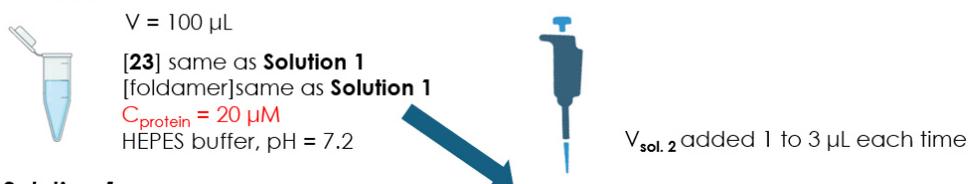


Figure S20. Overlay of UV-vis absorption spectra of foldamer **3** and probe **23**, fluorescence emission spectrum of **23** ($\lambda_{\text{ex}} = 373$ nm).

10.4.2.3 K_D measurement of HCAII complexes with AOFs 15-20 and ligand 22

Solution 2:



Solution 1:

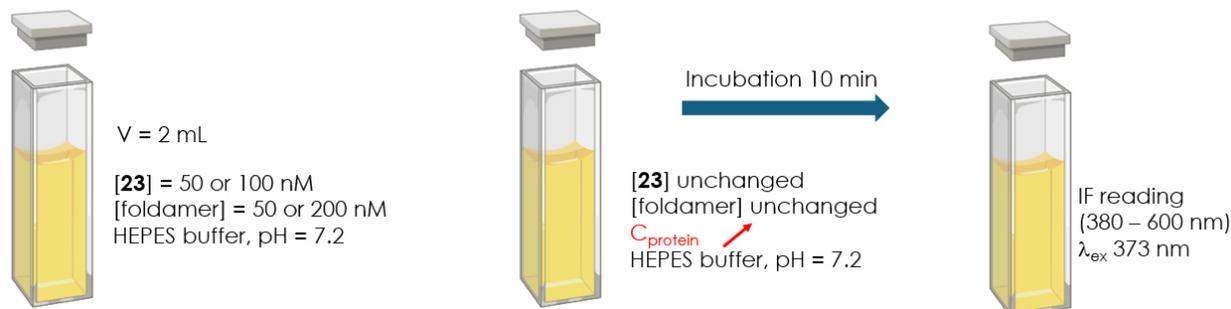
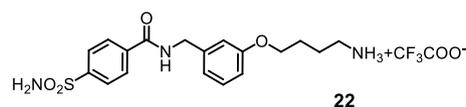


Figure S21. Schematic procedure of a competition experiment between **23** and foldamer (or ligand **22**).

1. **Solution 1:** **23** indicator stock solution and foldamer (or ligand **22**) stock solution were diluted in 50 mM HEPES buffer in the cuvette. The total volume was set at 2 mL.



2. **Solution 2:** Foldamer (or ligand **22**), **23** and protein (100 μM in HEPES buffer) stock solutions were diluted using 50 mM HEPES buffer. Concentration of foldamer and **23** were kept at 50 nM, (to remain identical to reference *solution 1*), final protein concentration was 20 μM ; total volume was 100 μL . Using solution 2 allowed that the concentrations of foldamer and **23** were kept constant (no dilution) upon protein solution addition ($\lambda_{\text{ex}} = 373 \text{ nm}$).

3. Aliquots of *solution 2* (1 to 3 μL) were added to *solution 1* and the IF was recorded from 380 nm to 600 nm ($\lambda_{\text{ex}} = 373 \text{ nm}$).

4. As for curve fitting, a 1:1 binding model was applied. The K_D value of **23** (1.69 nM) was inserted as a constant. The spectra of all emitting species were recorded from 380 nm to 600 nm ($\lambda_{\text{ex}} = 373 \text{ nm}$) in separate experiments and set as “known spectrum”. The K_D values foldamer were calculated considering all the wavelengths and the intensity of fluorescence at 419 nm is depicted.

10.4.3 BioLayer Interferometry

BLI measurements were carried out at 25 °C on an Octet R8 BLI Sartorius instrument, using streptavidin biosensors (SA). The buffer was 100 mM HEPES, 150 mM NaCl, 0.05% Tween, 0.1% DMSO, pH 7.4 (HEPES-D). At first and after a baseline in buffer for 120 sec, the biotinylated arylsulfonamide ligand **24** was loaded on the sensors (8 sensors, full column) with a loading at 4 µg/mL over 30 sec. The sensors were then washed with the buffer, a second baseline was recorded for 120 sec and then the association was performed with a range of HCAII concentrations from 50 nM to 0.78 nM for 240 sec before recording the dissociation over the same time in HEPES-D buffer. The K_D value was obtained after global curve fitting with the software embedded with the Octet R8 instrument.

10.5 Chemical Synthesis

10.5.1 General

Commercial reagents (suppliers: Abcr, Fisher Scientific, Merck, Sigma-Aldrich, TCI, BLDpharm or VWR) were used without further purification unless otherwise stated. LL Wang resin (100–200 mesh) was purchased from Sigma-Aldrich. CI-MPA protide resin® was purchased from CEM-Germany. 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 CaH₂ prior to use. Anhydrous tetrahydrofuran (THF) and dichloromethane (DCM) were obtained via an MBRAUN SPS-800 solvent purification system. Ultrapure water was collected on a Sartorius arium® pro VF ultrapure water system. Reactions were monitored by thin layer chromatography (TLC) on Merck silica gel 60-F254 plates and observed under UV light. Column chromatography purifications were carried out on Merck GEDURAN Si60 (40–63 μm). Nuclear magnetic resonance (NMR) spectra were recorded on an Avance III HD 400 MHz Bruker BioSpin spectrometer or an Avance III HD 500 MHz Bruker BioSpin spectrometer equipped with a broad band observe 5-mm BB-H&FD CryoProbe™ Prodigy. ¹H NMR measurements were performed at 25 °C unless stated otherwise. Water suppression was performed with excitation sculpting method. Processing was done with MestReNova (v.12.0.0-20080) NMR processing software from Mestrelab Research. Chemical shifts (δ) are reported in ppm and calibrated via residual solvent signals. Signal multiplicities are abbreviated as *s*, singlet; *d*, doublet; *t*, triplet; *q*, quartet, and *m*, multiplet. LC-MS spectra were recorded on a Bruker microTOF II in positive ionization mode. The instrument was calibrated in positive mode by direct infusion of a calibration solution (Agilent Technologies ESI-L Low Concentration Tuning Mix). The HPLC line was an Ultimate 3000 RP-HPLC system (ThermoFisher Scientific) equipped with a Nucleodur C18 gravity column (2 × 50 mm, 1.8 μm) at a flow rate of 0.33 mL/min. 0.2% formic acid and 0.02% TFA were added to the aqueous mobile phase (solvent A) and to acetonitrile (solvent B). The gradient was: 0-10 min, 10% to 100% solvent B at 50°C. The column eluent was monitored by UV detection at 214, 254, and 300 nm with a diode array detector. Analytical and semi-preparative reversed-phase (RP) high performance liquid chromatography (HPLC) were performed on a Thermo Fisher Scientific Ultimate 3000 HPLC System using MachereyNagel Nucleodur C18 Gravity columns (4 × 100 mm, 5 μm and 10 × 250 mm, 5 μm) or Macherey-Nagel Nucleodur C8 Gravity columns (4 × 50 mm, 5 μm and 10 × 100 mm, 5 μm) with different gradients composed of solvent A (0.1% TFA water) and B (0.1% TFA acetonitrile). Microwave-assisted solid phase foldamer synthesis (SPFS) was performed with a CEM® Discover Bio manual microwave apparatus. The temperature within the reactor vessel was monitored with an optical fiber probe. Automated SPFS was done on a PurePep® Chorus synthesizer (Gyros Protein Technologies) by applying induction heating.

10.5.2 Experimental procedures for chemical synthesis and purification

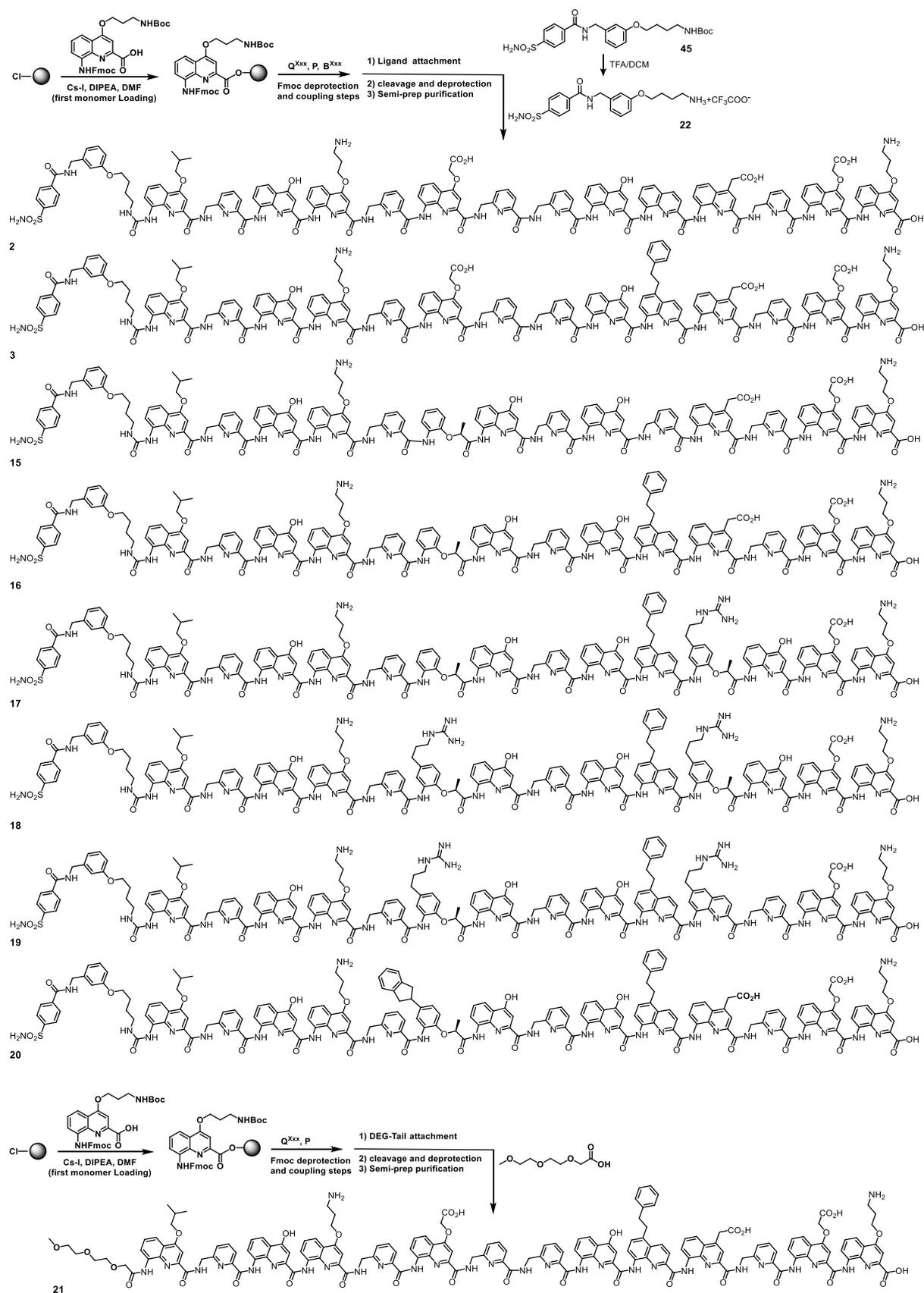
10.5.2.1 Solid phase synthesis and purification of aromatic oligoamide foldamers (AOFs)

The loading of the first Fmoc-Q-OH monomer and loading determination were done according to reported protocols using Cl-MPA protide resin.²⁰ The Q, P, B monomers^{21,22} were iteratively coupled on solid support as recently reported²³ using the PurePep® Chorus synthesizer. For ligand coupling see chapter 5.1.2. Cleavage from the resin and sidechain deprotection were performed simultaneously with a freshly prepared TFA solution containing triisopropylsilane (TIS) and water (TFA/TIS/H₂O, 95:2.5:2.5, v/v/v). After cleavage, the crude material was lyophilized and purified by semi-preparative RP-HPLC using a solvent mixture of A and B to furnish target foldamer sequences with purity over 95%.

10.5.2.2 Attachment of the Ligand and the DEG tail on resin-bound AOFs

This coupling step was optimized from the previously reported protocol²⁴: The Fmoc-protected resin-bound 14mer (10 μmol scale) was first subjected to 20% piperidine in NMP for 5 min, washed with NMP (3 × 3 mL), and this step was repeated once. Then the resin was suspended in anhydrous THF (0.75 mL), followed by dry DIPEA (17 μL, 100 μmol, 10 equiv.). Triphosgene (15 mg, 50 μmol, 5 equiv.) was dissolved in anhydrous THF (0.75 mL), poured into the reaction vessel (RV), the RV was placed in the microwave and the resin was subsequently heated under microwave irradiations (50°C, 25 W, 5 min). The resin was next filtered off, washed with dry THF (5 × 3 mL) to remove any trace of triphosgene.

On the day preceding the ligand installation on solid phase, compound **45**²⁴(scheme 1) (24 mg, 50 μmol, 5 equiv.) was dissolved in a DCM/TFA mixture (1:1, v/v) to remove the Boc group and after solvent evaporation placed overnight under the vacuum line to remove traces of TFA. Then, the next day the resulting TFA salt **42** was dissolved in dry NMP (0.75 mL), followed by the addition of freshly distilled DIPEA (17 μL, 100 μmol, 10 equiv.), and the solution was poured in the RV containing the freshly activated isocyanate resin. The resin was heated up under microwave irradiations (50°C, 25 W, 15 min). The resin was next washed with anhydrous THF (5 × 3 mL), and this step was repeated once in the presence **42** and DIPEA.



Scheme 1. Overview of the SPFS and chemical structures of all the synthesized AOF sequences (The final DEG tail group was installed via same coupling method of Q, B and P monomers).

10.5.3 AOF monomer and biotinylated HCAII ligand synthesis

Compound 25: (+)-Ethyl-D-lactate (1.1ml, 8.8 mmol, 1.0 equiv., enantiomer purity over 99%), Triphenylphosphine (2.87 g, 10.56 mmol, 1.2 eq) and 4-bromo-2-nitrophenol (1.83 g, 8.8 mmol, 1.0 equiv.) were dissolved in 80 mL THF under N₂ protection. The mixture was cooled down to 0°C then DIAD (2.1 mL, 10.6 mmol, 1.2 eq) was added to the mixture slowly under the N₂ protection. The reaction mixture was stirred at RT for three hours. The solvent was removed under vacuum and the crude was purified by silica gel column chromatography (CyHex/EtOAc, 9:1, v/v) to yield pale yellow crystalline powder (2.4 g, 7.92 mmol, 91%)

¹H-NMR (500MHz, CDCl₃): δ = 7.96 (d, *J*₁ = 4 Hz 1H), 7.58 (dd, *J*₁ = 8Hz, *J*₂ = 4Hz, 1H), 6.86 (d, *J* = 8 Hz, 1H), 4.80 (q, 1H), 4.21 (m, 2H), 1.68 (d, 2H), 1.25 (t, 3H). **¹³C NMR** (126 MHz, DMSO-*d*₆) δ 170.67, 150.31, 141.16, 136.65, 128.53, 117.60, 113.29, 74.95, 61.94, 18.41, 14.19. **HRMS** (ESI-) calcd. for C₁₁H₁₃BrNO₅ [M-H]⁻ 316.9898, found: 316.9881.

Compound 26: Compound **25** (2.4 g, 7.92 mmol, 1.0 equiv.), N-Boc-propargylamine (1.84 g, 11.8 mmol, 1.5 equiv.) and TEA (2.76 mL, 2.5 equiv.) were dissolved in anhydrous THF (32 mL), and the solution was degassed 3 times by freeze pumping. Then CuI (114 mg, 396 μmol, 5 mol%) and Pd(PPh₃)₄ (396 μmol, 5 mol%) were added under Ar atmosphere and the reaction mixture was stirred at 60 °C for 18 hours. After reaction mixture dilution with H₂O, the mixture was extracted with DCM (3 ×), dried over MgSO₄ and solvents were removed under reduced pressure. The crude was purified further by silica gel column chromatography (CyHex/EtOAc 8:2 → 6:4, v/v), yielding **26** (2.5 g, 6.4 mmol, 80%) as a yellow oil.

¹H-NMR (500 MHz, CDCl₃) δ 7.87 (d, *J* = 2.1 Hz, 1H), 7.49 (dd, *J* = 8.7, 2.2 Hz, 1H), 6.87 (d, *J* = 8.7 Hz, 1H), 4.83 (q, *J* = 6.8 Hz, 1H), 4.76 (s, 1H), 4.21 (qq, *J* = 7.4, 3.7 Hz, 2H), 4.13 (p, *J* = 4.9, 4.1 Hz, 2H), 1.68 (d, *J* = 6.8 Hz, 3H), 1.60 (s, 2H), 1.47 (s, 9H), 1.23 (t, *J* = 7.1 Hz, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 170.57, 155.27, 150.72, 140.27, 136.72, 128.84, 116.42, 115.58, 86.74, 80.32, 74.60, 61.79, 28.36, 18.27, 14.05. **HRMS** (ESI+) calcd. for C₁₉H₂₄N₂O₇Na [M+Na]⁺ 415.1481, found 415.1474

Compound 28: Compound **26** (2.4 g, 6.1 mmol, 1.0 equiv.) was suspended in DCM (5 mL) and TFA (5 mL) was added to the suspension. The solution was stirred for 60 min. Then the solvent was removed under reduced pressure and traces of TFA were finally removed by lyophilization. Compound **27** was directly used for the following process without further purification. The obtained TFA salts was dissolved in THF (82 mL) and DIPEA (4.25 mL, 24.4 mmol, 4 equiv.) was added to the mixture and the solution was cooled down to 0 °C. N,N'-Di-Boc-1H-pyrazole-1-carboxamidine (2.8 g, 9.1 mmol, 1.5 equiv.) was added and the resulting mixture was stirred overnight. The solvents were removed under reduced pressure, and the crude was dissolved in EtOAc and the organic phase was successively washed with 10% citric acid, saturated NaHCO₃, saturated NaCl and dried over Na₂SO₄. The crude was purified

by silica gel column chromatography (CyHex/EtOAc 5:1, v/v) yielding **28** (2.0 g, 3.74 mmol, 60%) as a green solid.

¹H-NMR (500 MHz, CDCl₃): δ = 11.5 (s, 1H), 8.53 (t, 1H), 7.89 (d, *J*₁ = 4 Hz, 1H), 7.51 (dd, *J*₁ = 8.0 Hz, *J*₂ = 4.0 Hz, 1H), 6.87 (d, *J*₁ = 8.0 Hz, 1H), 4.84 (q, 1H), 4.45 (d, 2H), 4.21 (q, 2H), 1.68 (d, 3H), 1.51 (s, 9H), 1.50 (s, 9H), 1.23 (t, 3H). **¹³C NMR** (101 MHz, CDCl₃) δ 170.85, 151.32, 139.88, 137.31, 129.14, 115.68, 114.88, 84.52, 80.79, 74.40, 62.11, 30.26, 18.23, 14.06. **HRMS** (ESI+) calcd. for C₂₅H₃₅N₄O₉ [M+H]⁺ 535.2404, found 535.2398.

Compound 29: Compound **28** (2.0 g, 3.74 mmol, 1.0 equiv.) was dissolved in THF (56 mL). After addition of a solution of LiOH (180 mg, 7.5 mmol, 2 equiv.) in H₂O (20 mL), the reaction mixture was stirred for 30 min at RT. Then, the mixture was acidified to approximately pH = 2 using 1 M HCl in H₂O. The resulting aqueous phase was extracted with DCM (3 ×) and dried over MgSO₄. After removing the solvents under reduced pressure, compound **29** was recovered quantitatively.

¹H-NMR (400 MHz, CDCl₃) δ 11.45 (s, 1H), 8.59 (s, 1H), 7.91 (d, *J* = 2.1 Hz, 1H), 7.54 (dd, *J* = 8.7, 2.1 Hz, 1H), 6.95 (d, *J* = 8.8 Hz, 1H), 5.06 (s, 3H), 4.88 (q, *J* = 6.8 Hz, 1H), 4.35 (s, 2H), 1.72 (d, *J* = 6.8 Hz, 3H), 1.50 (d, *J* = 2.6 Hz, 18H). **¹³C-NMR**: (101 MHz, CDCl₃) δ 172.76, 162.76, 155.70, 153.04, 150.75, 139.83, 137.14, 129.21, 115.31, 85.20, 83.78, 81.30, 80.08, 74.55, 67.96, 31.49, 28.19, 28.05, 25.60, 18.22. **HRMS** (ESI+) calcd. for C₂₃H₃₁N₄O₉ [M+H]⁺ 507.2091, found 507.2085.

Compound 30: Compound **29** (2.0 g, 3.74 mmol, 1.0 equiv.) and Na₂CO₃ (480 mg, 4.53 mmol, 1.2 equiv.) were dissolved in MeOH (56 mL). The solution was purged with N₂ for three times before adding Pd/C (200 mg, 10% w/w) and the N₂ was replaced by H₂ atmosphere. The reaction mixture was stirred at room temperature for 17 h, filtered over celite and washed with MeOH. Solvents were evaporated under reduced pressure yielding compound **30** (1.8 g, 3.74 mmol) quantitatively.

¹H-NMR (400 MHz, DMSO-*d*₆) δ 11.48 (s, 1H), 8.29 (s, 1H), 6.57 (d, *J* = 8.0 Hz, 1H), 6.42 (d, *J* = 2.1 Hz, 1H), 6.24 (dd, *J* = 8.1, 2.1 Hz, 1H), 4.89 (s, 2H), 4.04 (q, *J* = 6.8 Hz, 1H), 3.24 (m, 2H), 2.38 (t, *J* = 7.7 Hz, 2H), 1.72 (m, 2H), 1.47 (s, 9H), 1.38 (s, 9H), 1.35 (d, 3H). **¹³C-NMR** (101 MHz, DMSO-*d*₆) δ 175.76, 145.46, 139.72, 134.00, 116.02, 115.78, 114.50, 78.47, 49.06, 32.48, 30.80, 28.49, 28.12, 20.19. **HRMS** (ESI+) calcd. for C₂₃H₃₇N₄O₇ [M+H]⁺ 481.2662, found 481.2657.

Compound 31: Compound **30** (1.8 g, 3.74 mmol, 1.0 equiv.) and NaHCO₃ (1.8 g, 18.7 mmol, 5.0 equiv.) were dissolved in H₂O (83 mL). Then, Fmoc-Cl (1.5 g, 4.8 mmol, 1.3 equiv.) dissolved in dioxane (83 mL) was added at 0 °C over 1 h. The reaction mixture was stirred at 0 °C for one additional hour and then at RT for 18 h. After the reaction mixture was acidified to approximately pH = 2 using 1 M HCl in H₂O, the aqueous phase was extracted with DCM (3×), dried over MgSO₄ and the solvents were removed under reduced pressure. The residue was purified by puriFlash[®]xs 520Plus purification system

(line A: water, line B: ACN; 30% - 100% B 15min then 100%B 10min) to yield compound **31** (1.5 g, 2.13 mmol, 57%) as a white solid.

¹H-NMR (400 MHz, DMSO-*d*₆) δ 13.15 (s, 1H), 11.49 (s, 1H), 8.69 (s, 1H), 8.32 (t, *J* = 5.7 Hz, 1H), 7.92 (d, *J* = 7.5 Hz, 2H), 7.74 (dd, *J* = 7.3, 3.4 Hz, 2H), 7.51 (s, 1H), 7.43 (t, *J* = 7.4 Hz, 2H), 7.35 (td, *J* = 7.4, 1.2 Hz, 2H), 6.89 (d, *J* = 1.2 Hz, 2H), 4.76 (q, *J* = 6.8 Hz, 1H), 4.42 (dd, *J* = 7.3, 3.8 Hz, 2H), 4.32 (t, *J* = 7.0 Hz, 1H), 3.29 (m, 4H), 1.76 (p, *J* = 7.5 Hz, 2H), 1.56 (d, *J* = 6.8 Hz, 3H), 1.43 (d, 18H). **¹³C-NMR** (101 MHz, DMSO-*d*₆) δ 174.12, 163.56, 155.72, 153.94, 152.55, 144.19, 141.20, 135.18, 128.17, 127.61, 125.72, 124.09, 120.65, 83.32, 78.57, 66.59, 47.03, 40.68, 40.47, 32.29, 30.67, 28.46, 28.08, 18.99. **HRMS** (ESI+) calcd. for C₃₈H₄₇N₄O₉ [M+H]⁺ 703.3343, found 703.3339

Compound 32: The protocol was based on reported literature²⁵ and slightly modified. To a 250 mL two-neck flask flushed with a positive pressure of N₂, Pd(OAc)₂ (55 mg, 0.25 mmol, 5% mol), (1*H*-inden-2-yl)boronic acid (800 mg, 5 mmol, 1 equiv.), K₂CO₃ (2.8 g, 20 mmol, 4 equiv.) were added and the flask was again flushed with N₂. Compound **29** (1.58 g, 5 mmol, 1 equiv.) was dissolved in a mixture of toluene/ ethanol/ H₂O (37.5 mL: 15 mL: 7.5 mL, v/v/v) and degassed three times. The degassed solution was transferred to the two-neck flask and the mixture was heated to 80 °C overnight. The reaction mixture was diluted with water and extracted with DCM (3×) and dried over Na₂SO₄. The crude was purified first by silica gel chromatography and then with a puriFlash[®]xs 520Plus purification system (line A: water, line B: ACN; 30% - 100% B 15min then 100%B 10min) to yield compound **32** (1.1 g, 3.15 mmol, 63%).

¹H-NMR (500 MHz, CDCl₃) δ 8.05 (d, *J* = 2.3 Hz, 1H), 7.72 (dd, *J* = 8.7, 2.3 Hz, 1H), 7.48 (dq, *J* = 7.3, 0.9 Hz, 1H), 7.41 (dt, *J* = 7.5, 0.9 Hz, 1H), 7.29 (td, *J* = 7.5, 1.2 Hz, 1H), 7.23 – 7.19 (m, 2H), 6.98 (d, *J* = 8.7 Hz, 1H), 4.86 (q, *J* = 6.8 Hz, 1H), 4.23 (qd, *J* = 7.1, 1.9 Hz, 2H), 3.76 (dd, *J* = 1.6, 0.8 Hz, 2H), 1.71 (d, *J* = 6.7 Hz, 3H), 1.26 (t, *J* = 7.1 Hz, 3H). **¹³C-NMR** (126 MHz, CDCl₃) δ 171.05, 150.18, 144.95, 143.46, 143.01, 141.04, 130.61, 130.40, 127.83, 127.02, 125.48, 123.92, 122.64, 121.46, 116.40, 74.96, 61.85, 39.11, 18.52, 14.23. **HRMS** (ESI+) calcd. for C₂₀H₂₀NO₅ [M+H]⁺ 354.1342, found 354.1291

Compound 33: Compound **32** (1.1 g, 3.1 mmol, 1 equiv.) was dissolved in THF (45 mL) and a LiOH (144 mg, 6.2 mmol, 2 equiv.) solution in water (15 mL) was added. After 1 h, the reaction mixture was acidified with 5% citric acid in water and the aqueous phase was extracted with DCM (3×). After drying with Na₂SO₄, the solvent was evaporated to yield compound **33** without further purification (0.9 g, 2.8 mmol, 90%).

¹H NMR (500 MHz, DMSO-*d*₆) δ 8.11 (d, *J* = 2.3 Hz, 1H), 7.91 (dd, *J* = 8.8, 2.3 Hz, 1H), 7.47 (d, *J* = 7.4 Hz, 1H), 7.40 (d, *J* = 7.7 Hz, 2H), 7.26 (td, *J* = 7.5, 1.2 Hz, 1H), 7.21 – 7.14 (m, 2H), 4.97 (d, *J* = 7.5 Hz, 1H), 3.84 (s, 2H), 1.50 (d, *J* = 6.8 Hz, 3H), **¹³C NMR** (126 MHz, DMSO-*d*₆) δ 172.72, 150.06,

145.24, 144.50, 143.48, 140.62, 131.26, 129.04, 127.41, 127.04, 125.35, 124.18, 122.02, 121.49, 116.38, 74.12, 39.07, 18.60. **HRMS** (ESI+) calcd. for $C_{18}H_{15}NO_5$ $[M+Na]^+$ 348.0842, found 348.0896

Compound 34: Compound **33** (900 mg, 2.8 mmol, 1 equiv.) and Na_2CO_3 (300 mg, 1.0 equiv.) were dissolved in MeOH (40 mL), and the solution was purged three times with N_2 positive pressure. Pd/C (90 mg, 10%, w/w) was added and the N_2 was replaced by H_2 atmosphere. The reaction mixture was stirred at RT for 17 h, filtered over a celite pad and washed with MeOH. Solvents were evaporated under reduced pressure to furnish compound **34** (0.8 g) in quantitative yield without further purification.

Compound 35: Compound **34** (0.8 g 2.8 mmol, 1.0 equiv.) and $NaHCO_3$ (1.3 g, 14 mmol, 5.0 equiv.) were dissolved in H_2O (62 mL). Then, Fmoc-Cl (1.04 g, 3.64 mmol, 1.3 equiv.) in dioxane (62 ml) was added at 0 °C over 1 h. The reaction mixture was stirred at 0 °C for one additional hour and then at RT for 18 h. The reaction mixture was next acidified to approximately pH = 2 using 1 M HCl in H_2O , the aqueous phase was extracted with DCM (3x), dried over $MgSO_4$ and solvents were removed under reduced pressure. The residue was purified by puriFlash[®]xs 520Plus purification system (line A: water, line B: ACN; 50% - 100% B 15min then 100%B 10min) to yield compound **30** (0.65 g, 2.13 mmol, 45%) a white solid.

¹H-NMR (400 MHz, $DMSO-d_6$) δ 13.15 (s, 1H), 8.68 (s, 1H), 7.90 (d, J = 7.5 Hz, 2H), 7.73 (d, J = 7.0 Hz, 2H), 7.66 (s, 1H), 7.42 (t, J = 7.4 Hz, 2H), 7.38 – 7.27 (t, J = 7.5 Hz, 2H), 7.23 (dd, J = 5.4, 3.3 Hz, 2H), 7.19 – 7.07 (m, 2H), 6.99 (d, J = 8.5, 1H), 6.90 (d, J = 8.5 Hz, 1H), 4.79 (q, J = 6.8 Hz, 1H), 4.47 – 4.37 (m, 2H), 4.32 (q, J = 8.6, 7.1 Hz, 1H), 3.55 (p, J = 8.6 Hz, 1H), 3.24 (dd, J = 15.5, 8.1 Hz, 2H), 2.91 (dd, J = 15.5, 9.1 Hz, 2H), 1.56 (d, J = 6.8 Hz, 3H). **¹³C-NMR** (101 MHz, $DMSO-d_6$) δ 206.99, 174.07, 154.02, 144.21, 143.00, 141.19, 138.80, 128.19, 127.61, 126.82, 125.76, 124.62, 122.84, 120.65, 114.98, 74.39, 66.60, 47.01, 44.88, 40.83, 31.18, 18.94. **HRMS** (ESI+) calcd. for $C_{33}H_{29}NO_5$ $[M+H]^+$ 519.2048, found 519.2008

Compound 37: Compound **36**²¹ (2.1 g, 5.42 mmol, 1 equiv.) was dissolved in DMF (30 mL) and TEA (30 mL) was added. The reaction mixture was degassed by using freeze-thaw techniques. Then $Pd(PPh_3)Cl_2$ (76 mg, 0.11 mmol, 0.02 equiv.) and CuI (41 mg, 0.22 mmol, 0.04 equiv.) were added and the reaction mixture was again degassed twice using freeze-thaw techniques. Phenylacetylene (0.89 mL, 8.14 mmol, 1.5 equiv.) was added and the reaction mixture was heated up to 80 °C for 4 h. The reaction mixture was poured into a mixture of water and DCM, layers were then separated. Aqueous layer was extracted one more time with DCM. Combined organic layers were washed with water, brine (2 times), dried over $MgSO_4$, then the solvent was removed under reduced pressure. The crude was purified by silica gel chromatography in pure DCM to yield the target compound **37** (1.7 g, 4.17 mmol, 77%) as a yellow solid.

¹H-NMR (400 MHz, CDCl₃) δ 8.92 (d, J = 8.7 Hz, 1H), 8.38 (d, J = 8.7 Hz, 1H), 8.14 (d, J = 7.8 Hz, 1H), 7.92 (d, J = 7.8 Hz, 1H), 7.71 – 7.63 (m, 2H), 7.58 – 7.51 (m, 2H), 7.50 – 7.31 (m, 7H), 5.52 (s, 2H). **¹³C-NMR** (101 MHz, CDCl₃) δ 164.46, 150.50, 147.83, 139.16, 136.45, 135.52, 132.07, 130.71, 130.16, 129.88, 128.85, 128.80, 128.52, 128.28, 126.24, 124.71, 123.18, 121.81, 99.22, 84.59, 67.94. **HRMS** (ESI+) calcd. for C₂₅H₁₆N₂O₄ [M+H]⁺ 409.1188, found 409.1185

Compound 38: Pd/C (165 mg, 10% m/m) was added to compound **37** (1.65 g, 4.04 mmol, 1 equiv.) in EtOAc and DMF solvent mixture (18 mL/9 mL, 2:1) under N₂. N₂ was bubbled, followed by H₂ bubbling, then the H₂ balloon was placed and stirred at RT for 4 h. The reaction mixture was filtered through a celite pad, washing with EtOAc until the eluent was colorless. The solvent was removed to yield the target compound as brown solid without further purification.

¹H-NMR (400 MHz, DMSO-*d*₆) δ 8.55 (d, J = 8.8 Hz, 1H), 8.06 (d, J = 8.8 Hz, 1H), 7.29 – 7.22 (m, 5H), 7.21 – 7.13 (m, 1H), 6.81 (d, J = 7.8 Hz, 1H), 6.40 (s, 2H), 3.16 (dd, J = 9.5, 6.5 Hz, 2H), 2.89 (s, 2H). **HRMS** calculated for C₁₈H₁₅N₂O₂: 291.1134. (M-H)⁻; Found :291.1137

Compound 39: 10% NaHCO₃ solution (71 mL) was added to the compound **38** in 1,4-dioxane (36 mL). Fmoc-Cl (1.36 g, 5.23 mmol, 1.3 equiv.) in 71 mL was added to the mixture dropwise in 0 °C. After adding, the solution was stirred in RT overnight. The reaction mixture was acidified using 5% citric acid, extracted with DCM (2 times). The combined organic layers were washed with water, brine (2 times), dried with Na₂SO₄. The solvent was evaporated, and crude was purified by silica gel chromatography with 5-10% MeOH in DCM mixture to yield 1.62 g target compound (4.07 mmol, 78%).

¹H-NMR (400 MHz, DMSO-*d*₆) δ 10.36 (s, 1H), 8.76 (d, J = 8.8 Hz, 1H), 8.20 (d, J = 8.8 Hz, 1H), 7.93 (d, J = 7.5 Hz, 2H), 7.77 (d, J = 7.4 Hz, 2H), 7.44 (td, J = 7.4, 1.2 Hz, 2H), 7.36 (td, J = 7.4, 1.2 Hz, 2H), 7.31 – 7.23 (m, 5H), 7.23 – 7.14 (m, 1H), 4.60 (d, J = 6.8 Hz, 2H), 4.44 (t, J = 6.8 Hz, 1H), 3.31 (t, J = 7.9 Hz, 4H), 2.94 (m, 2H). **¹³C-NMR** (126 MHz, DMSO-*d*₆) δ 165.58, 153.46, 145.85 – 144.80 (m), 143.71, 141.19, 140.80, 134.93, 134.03, 131.26, 128.95, 128.46, 128.23, 127.76, 127.63, 127.20, 125.94, 125.14, 120.43, 120.23, 115.81, 66.34, 46.58, 36.53, 32.95. **HRMS** (ESI-) calcd. for C₃₃H₂₅N₂O₄ [M-H]⁻ 513.1814, found 513.1813

Compound 41: Compound **40** was synthesized according to the previous protocol.²⁶ Compound **36** (400 mg, 0.72 mmol, 1 equiv.) was dissolved in TFA/DCM mixture (20 mL, 1:1, v/v). After Boc-group removal, the sample was lyophilized without further purification. The resulting TFA salt was dissolved in a water / dioxane mixture (20 mL, 1:1, v/v) followed by the addition of DIPEA (500 μL, 2.88 mmol, 4 equiv.). The reaction mixture became cloudy. *tert*-Butyl (((*tert*-butoxycarbonyl)amino)(1H-pyrazol-1-yl)methylene)carbamate (336 mg, .0.8 mmol, 1.5 equiv.) was then added and after 1 h the reaction mixture became clear. The reaction mixture was stirred overnight and after solvent evaporation the

crude was purified by silica gel chromatography with 5-10% MeOH in DCM mixture to yield compound **41** (300 mg, 0.42 mmol, 58%).

¹H-NMR (500 MHz, DMSO-*d*₆) δ 13.53 (s, 1H), 11.45 (s, 1H), 10.47 (s, 1H), 8.47 (d, *J* = 8.6 Hz, 1H), 8.27 (t, *J* = 5.7 Hz, 1H), 8.17 (d, *J* = 8.5 Hz, 1H), 7.93 (d, *J* = 7.5 Hz, 2H), 7.77 (dd, *J* = 7.4, 1.1 Hz, 2H), 7.54 (d, *J* = 1.8 Hz, 1H), 7.44 (td, *J* = 7.5, 1.1 Hz, 2H), 7.36 (td, *J* = 7.4, 1.2 Hz, 2H), 4.60 (d, *J* = 6.9 Hz, 2H), 4.45 (t, *J* = 6.9 Hz, 1H), 3.37 – 3.36 (m, 2H), 2.78 (t, *J* = 7.5 Hz, 2H), 1.92 (p, *J* = 7.3 Hz, 2H), 1.41 (s, 9H), 1.36 (s, 9H). **¹³C-NMR** (126 MHz, DMSO-*d*₆) δ 165.89, 155.70, 153.99, 152.43, 145.14, 144.15, 143.71, 141.27, 138.40, 136.03, 129.84, 128.26, 127.69, 125.62, 121.20, 120.75, 119.83, 117.80, 83.25, 78.56, 66.87, 47.07, 33.75, 29.96, 28.45, 28.22, 28.12, 28.03. **HRMS** (ESI+) calcd. for C₁₁H₁₀N₂O₂ [M+H]⁺ 203.0820, found 203.0766

Compound 43: Compound **42** (2 g, 7.5 mmol, 1 equiv) was dissolved in a mixture of ethyl acetate and THF (92 mL, 1:1, v/v). TEA (41 mL) was added, and the reaction mixture was flushed with N₂. Pd/C (200 mg, 10% w/w) was next added, and the reaction mixture was stirred under H₂ pressure (1 bar) for 24 hours. The reaction mixture was filtered through a celite pad, washed with ethyl acetate and the solvent was removed under vacuum to yield compound **43** (1.42 g, 7 mmol, 94%) as orange oil.

¹H NMR (500 MHz, DMSO-*d*₆) δ 8.35 (d, *J* = 8.5 Hz, 1H), 8.03 (d, *J* = 8.5 Hz, 1H), 7.43 (t, *J* = 7.8 Hz, 1H), 7.13 (dd, *J* = 8.1, 1.3 Hz, 1H), 6.94 (dd, *J* = 7.6, 1.3 Hz, 1H), 6.08 (s, 2H), 3.95 (s, 3H). **¹³C NMR** (126 MHz, DMSO-*d*₆) δ 165.35, 146.14, 143.92, 137.16, 136.40, 130.25, 129.74, 120.81, 113.28, 109.28, 52.53. **HRMS** calculated for C₁₁H₁₀N₂O₂: 203.0820 (M+H)⁺ Found: 203.0766

Compound 44: Compound **43** (1.42 g, 7 mmol, 1 equiv.) was dissolved in 30 mL dioxane and mixed with LiOH (441 mg, 10.5 mmol, 1.5 equiv.) in 15 mL H₂O. The reaction was monitored by TLC. After reaction completion, 1 M aqueous HCl (10.7 mL) was added to quench the reaction. The mixture was then cooled down to 0 °C and NaHCO₃ (2.94 g, 35 mmol, 5 equiv.) was added. A solution of Fmoc-Cl (2.2 g, 8.4 mmol, 1.2 equiv.) in dioxane (93 mL) was added dropwise over 1 h. After reaction completion, 10% aqueous citric acid was added until pH = 4. The organic layer was separated by adding DCM and dried over MgSO₄. Solvent was removed under vacuum and the compound **44** was precipitated in MeOH to yield a yellow powder (2.17 g, 5.3 mmol, 76%)

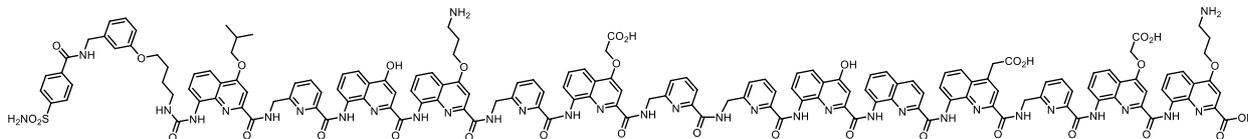
¹H NMR (500 MHz, DMSO-*d*₆) δ 13.57 (s, 1H), 10.46 (s, 1H), 8.60 (d, *J* = 8.5 Hz, 1H), 8.35 (s, 1H), 8.21 (d, *J* = 8.5 Hz, 1H), 7.93 (d, *J* = 7.5 Hz, 2H), 7.78 (d, *J* = 7.5 Hz, 2H), 7.72 (dd, *J* = 8.2, 1.4 Hz, 1H), 7.67 (t, *J* = 7.9 Hz, 1H), 7.44 (t, *J* = 7.4 Hz, 2H), 7.36 (td, *J* = 7.4, 1.1 Hz, 2H), 4.63 (d, *J* = 6.8 Hz, 2H), 4.45 (t, *J* = 6.8 Hz, 1H). **¹³C NMR** (126 MHz, DMSO-*d*₆) δ 165.74, 153.92, 145.75, 144.08, 141.21, 139.01, 136.95, 136.07, 129.89, 129.71, 128.19, 127.63, 125.54, 121.41, 121.05, 120.65, 116.60, 66.79, 46.97. **HRMS** (ESI+) calcd. for C₂₅H₁₈N₂O₄ [M+H]⁺ 411.1344, found 411.1334.

Compound 24: Compound **22**²⁴ (12.0 mg, 32 μ mol, 1.2 equiv.) was dissolved in a 5 mL round bottle flask with 500 μ L DMF, followed by DIPEA (9.3 μ L, 53 μ mol, 2 equiv.). Biotin-PEG₁₂-NHS ester (25.0 mg, 27 μ mol, 1.0 equiv., purchased from Iris) was dissolved in 500 μ L DMF and added to the reaction mixture. The reaction was monitored by HPLC and further purified by semi-prep HPLC to yield target compound as white powder (15.0 mg, 12.5 μ mol, 47%)

¹H NMR (500 MHz, DMSO-*d*₆) δ 9.20 (t, J = 6.0 Hz, 1H), 8.07 – 8.01 (m, 2H), 7.94 – 7.87 (m, 2H), 7.84 (q, J = 5.7 Hz, 2H), 7.49 (s, 2H), 7.26 – 7.19 (m, 1H), 6.91 – 6.85 (m, 2H), 6.83 – 6.77 (m, 1H), 6.42 (s, 1H), 6.36 (s, 1H), 4.46 (d, J = 5.9 Hz, 2H), 4.33 – 4.27 (m, 1H), 4.12 (dd, J = 7.8, 4.4 Hz, 1H), 3.94 (t, J = 6.4 Hz, 2H), 3.58 (t, J = 6.5 Hz, 2H), 3.52 – 3.43 (m, 43H), 3.38 (t, J = 5.9 Hz, 2H), 3.18 (q, J = 5.9 Hz, 2H), 3.08 (qd, J = 6.6, 5.8, 3.6 Hz, 3H), 2.81 (dd, J = 12.4, 5.1 Hz, 1H), 2.57 (d, J = 12.4 Hz, 1H), 2.29 (t, J = 6.4 Hz, 2H), 2.06 (t, J = 7.8 Hz, 2H), 1.74 – 1.63 (m, 2H), 1.65 – 1.55 (m, 1H), 1.63 – 1.56 (m, 1H), 1.55 – 1.39 (m, 2H), 1.36 – 1.21 (m, 2H). **HRMS** (ESI+) calcd. for C₅₅H₉₀N₆O₁₉S₂ [M+2H]²⁺ 602.2924; found 602.2994

10.5.4 Foldamer synthesis on solid support

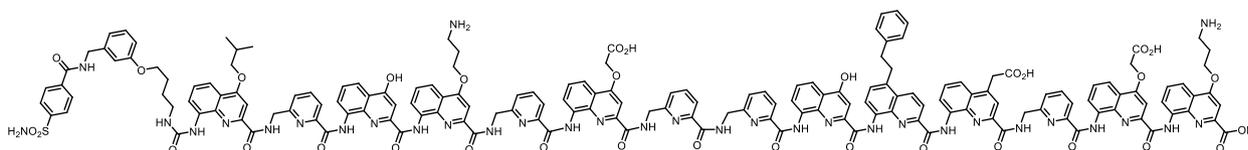
Compound 2



The scale was 22 μmol after first monomer loading determination. Target AOF was assembled on Cl-MPA protide resin using SPFS (Method 5.1.1, 5.1.2). 7 μmol of resin was used for ligand installation, after TFA cleavage, compound 2 was purified by semi-prep HPLC and recovered as yellow solid (4 mg, 1.3 μmol , 19%)

^1H NMR (300 MHz, $\text{DMSO-}d_6$) δ 12.10 (s, 1H), 11.73 (s, 1H), 11.23 (s, 1H), 10.99 (s, 1H), 10.93 (s, 1H), 10.74 (s, 1H), 10.63 (s, 1H), 10.22 (s, 1H), 9.97 (s, 1H), 9.12 (t, $J = 6.1$ Hz, 1H), 8.89 (s, 1H), 8.66 (s, 1H), 8.40 (s, 1H), 8.25 (dd, $J = 24.6, 8.0$ Hz, 3H), 8.04 – 7.95 (m, 2H), 7.95 (s, 1H), 7.88 (dd, $J = 8.5, 6.7$ Hz, 3H), 7.77 – 7.67 (m, 4H), 7.71 – 7.61 (m, 1H), 7.64 – 7.45 (m, 6H), 7.38 (s, 4H), 7.34 (d, $J = 7.6$ Hz, 1H), 7.26 – 7.15 (m, 2H), 7.12 (s, 8H), 7.13 – 6.96 (m, 4H), 6.96 – 6.57 (m, 7H), 6.53 (s, 1H), 6.44 (s, 1H), 6.34 (s, 1H), 6.21 (s, 1H), 5.39 – 5.32 (m, 1H), 4.82 (d, $J = 16.0$ Hz, 1H), 4.68 (d, $J = 15.9$ Hz, 1H), 4.45 – 4.33 (m, 7H), 4.21 (s, 1H), 3.97 (d, $J = 16.2$ Hz, 3H), 3.78 – 3.68 (m, 3H), 3.05 (s, 4H), 2.67 (s, 5H), 2.13 (s, 2H), 1.74 (d, $J = 16.1$ Hz, 1H), 1.42 (s, 3H), 1.26 (s, 6H), 1.13 – 1.03 (m, 6H). HRMS (ESI⁺) calcd. for $\text{C}_{160}\text{H}_{137}\text{N}_{33}\text{O}_{33}\text{S}$ $[\text{M}+2\text{H}]^{2+}$ 1539.9883, found 1540.1506

Compound 3

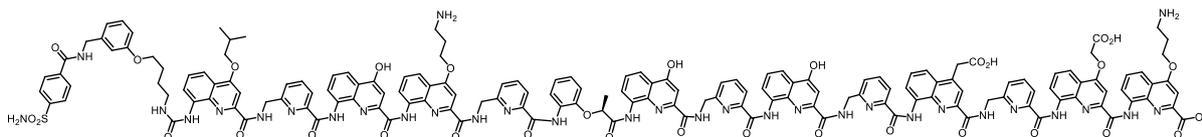


The scale was 15 μmol after first monomer loading determination. Target AOF was assembled on Cl-MPA protide resin using SPFS (Method 5.1.1, 5.1.2). 30 mg of crude was obtained after TFA cleavage, and the target compound was purified by semi-prep HPLC and recovered as a yellow solid (3.4 mg, 1.06 μmol , 7.1%).

^1H NMR (500 MHz, $\text{DMSO-}d_6$) δ 12.03 (s, 1H), 11.73 (s, 1H), 11.23 (s, 1H), 10.95 (s, 2H), 10.90 (s, 1H), 10.72 (s, 1H), 10.56 (s, 1H), 10.19 (s, 1H), 9.95 (s, 1H), 9.10 (t, $J = 6.0$ Hz, 1H), 8.89 (s, 1H), 8.66 (s, 1H), 8.49 (d, $J = 8.7$ Hz, 1H), 8.39 (s, 1H), 8.21 (dd, $J = 17.2, 7.5$ Hz, 2H), 7.98 (s, 1H), 7.98 (d, $J = 8.5$ Hz, 2H), 7.92 (s, 1H), 7.90 – 7.82 (m, 3H), 7.76 – 7.67 (m, 3H), 7.67 – 7.57 (m, 3H), 7.57 (s, 1H), 7.52 (m, 2H), 7.46 (m, 4H), 7.42 – 7.29 (m, 3H), 7.22 (m, 1H), 7.17 – 7.04 (m, 5H), 7.03 – 6.96 (m, 2H), 6.83 (d, $J = 7.6$ Hz, 1H), 6.78 (d, $J = 7.6$ Hz, 1H), 6.70 (dd, $J = 19.5, 5.3$ Hz, 3H), 6.60 (dd, $J = 8.2, 2.5$ Hz, 1H), 6.53 (s, 1H), 6.41 (s, 1H), 6.31 (s, 1H), 6.25 (s, 1H), 4.78 (d, $J = 15.9$ Hz, 1H), 4.65

(d, $J = 16.0$ Hz, 1H), 4.36 (d, $J = 5.9$ Hz, 2H), 4.32 (s, 4H), 4.19 (s, 1H), 3.96 (s, 3H), 3.87 (s, 1H), 3.72 (t, $J = 6.5$ Hz, 2H), 3.22 (s, 12H), 3.08 (t, $J = 7.3$ Hz, 1H), 3.03 (s, 5H), 3.01 (d, $J = 5.5$ Hz, 1H), 2.54 (s, 34H), 2.46 (s, 2H), 1.88 (s, 1H), 1.73 (d, $J = 15.3$ Hz, 1H), 1.63 (q, $J = 5.8$ Hz, 1H), 1.42 (dq, $J = 12.8, 6.7, 5.3$ Hz, 2H), 1.24 (d, $J = 9.3$ Hz, 8H), 1.17 (t, $J = 7.3$ Hz, 3H), 1.04 (dd, $J = 6.7, 2.3$ Hz, 5H), 0.96 (s, 3H), 0.85 (t, $J = 6.7$ Hz, 1H). HRMS (ESI⁺) calcd. for C₁₆₈H₁₄₃N₃₃O₃₃S [M+2H]²⁺ 1593.0225, found 1592.9842

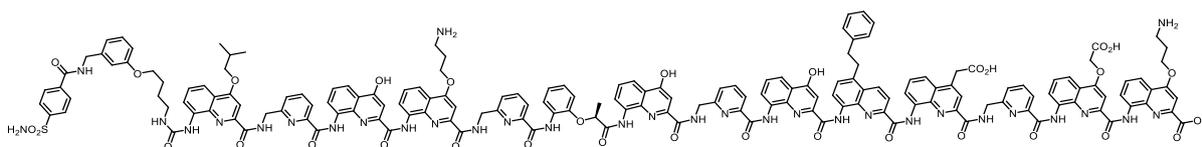
Compound 15



The scale was 13 μ mol after first monomer loading determination. Target AOF was assembled on Cl-MPA protide resin using SPFS (Method 5.1.1, 5.1.2). 25 mg crude was obtained after TFA cleavage, and target compound was purified by semi-prep HPLC and recovered as a yellow solid (1.1 mg, 0.37 μ mol, 2.8%).

¹H NMR (500 MHz, DMSO-*d*₆) δ 13.32 (s, 1H), 12.66 (s, 1H), 12.16 (s, 1H), 11.67 (s, 1H), 11.48 (s, 1H), 11.35 (s, 1H), 11.26 (s, 1H), 11.12 (s, 1H), 10.97 (s, 1H), 10.80 (s, 1H), 10.48 (s, 1H), 10.33 (s, 1H), 9.05 (t, $J = 6.0$ Hz, 1H), 8.93 (s, 1H), 8.71 (s, 1H), 8.18 (d, $J = 7.6$ Hz, 1H), 8.13 (s, 1H), 7.91 (d, $J = 8.4$ Hz, 3H), 7.86 – 7.77 (m, 12H), 7.73 (d, $J = 8.3$ Hz, 3H), 7.68 (d, $J = 9.1$ Hz, 6H), 7.62 (d, $J = 8.0$ Hz, 1H), 7.56 (d, $J = 7.7$ Hz, 2H), 7.50 – 7.39 (m, 6H), 7.34 – 7.23 (m, 4H), 7.18 (s, 1H), 7.08 (d, $J = 15.6$ Hz, 4H), 7.02 (t, $J = 7.8$ Hz, 1H), 6.72 (d, $J = 7.6$ Hz, 1H), 6.67 (s, 1H), 6.56 (dd, $J = 8.2, 2.5$ Hz, 1H), 6.40 (s, 1H), 6.35 (s, 1H), 6.28 (s, 2H), 6.16 (s, 1H), 6.07 (s, 1H), 5.86 (s, 1H), 4.79 (d, $J = 15.9$ Hz, 1H), 4.68 (d, $J = 16.4$ Hz, 1H), 4.38 (s, 2H), 4.29 (d, $J = 5.9$ Hz, 2H), 4.15 (s, 1H), 3.97 (s, 2H), 3.92 (d, $J = 14.9$ Hz, 1H), 3.68 (t, $J = 6.4$ Hz, 3H), 3.44 (s, 1H), 3.07 (s, 6H), 2.99 – 2.94 (m, 3H), 2.75 (s, 1H), 2.70 (s, 2H), 2.18 (s, 3H), 2.15 – 2.04 (m, 1H), 2.05 (s, 2H), 1.97 – 1.89 (m, 1H), 1.39 (t, $J = 7.5$ Hz, 3H), 1.22 (d, $J = 7.8$ Hz, 3H), 1.18 (p, $J = 5.7, 5.1$ Hz, 4H), 1.11 (t, $J = 7.3$ Hz, 1H), 1.04 – 0.98 (m, 7H), 0.78 (t, $J = 6.8$ Hz, 1H), -0.27 (s, 3H). HRMS (ESI⁺) calcd. for C₁₅₇H₁₃₆N₃₂O₃₂S [M+2H]²⁺ 1507.4932, found 1507.5003

Compound 16

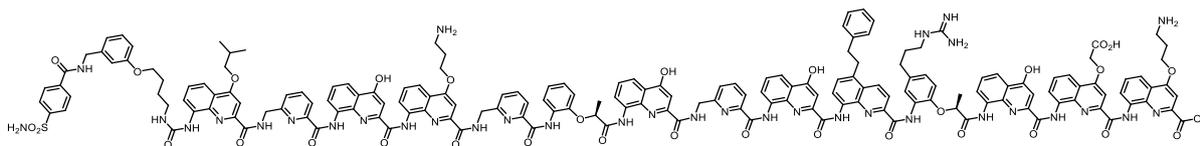


The scale was 15 μ mol after first monomer loading determination. Target AOF was assembled on Cl-MPA protide resin using SPFS (Method 5.1.1, 5.1.2). 33 mg crude was obtained after TFA cleavage,

Target compound was purified by semi-prep HPLC and recovered as yellow solid (3.7 mg, 1.17 μmol , 7.8%)

^1H NMR (500 MHz, $\text{DMSO-}d_6$) δ 11.46 (s, 1H), 11.36 (s, 1H), 11.20 (s, 1H), 11.14 (s, 1H), 10.92 (s, 1H), 10.87 (s, 1H), 10.69 (s, 1H), 10.61 (s, 1H), 10.44 (s, 1H), 9.12 (t, $J = 6.3$ Hz, 1H), 8.88 (s, 1H), 8.60 (s, 1H), 8.48 (d, $J = 10.5$ Hz, 1H), 8.18 (d, $J = 9.0$ Hz, 1H), 8.07 (d, $J = 8.9$ Hz, 2H), 7.97 – 7.90 (m, 5H), 7.84 (d, $J = 8.3$ Hz, 3H), 7.77 (s, 6H), 7.74 (d, $J = 9.3$ Hz, 1H), 7.73 – 7.66 (m, 2H), 7.60 (m, 3H), 7.49 (m, 2H), 7.47 (s, 5H), 7.43 (s, 2H), 7.37 (m, 3H), 7.30 (s, 1H), 7.21 (d, $J = 7.0$ Hz, 2H), 7.14 (dd, $J = 15.9, 8.3$ Hz, 3H), 7.07 (d, $J = 8.9$ Hz, 4H), 6.99 (d, $J = 12.6$ Hz, 3H), 6.88 (dd, $J = 15.8, 8.8$ Hz, 2H), 6.75 (d, $J = 8.9$ Hz, 3H), 6.68 (s, 2H), 6.59 (t, $J = 11.6$ Hz, 2H), 6.52 (d, $J = 9.6$ Hz, 2H), 6.43 (s, 1H), 6.33 (d, $J = 6.9$ Hz, 2H), 6.27 – 6.18 (m, 3H), 6.06 (t, $J = 8.7$ Hz, 1H), 5.72 – 5.66 (m, 1H), 4.84 (m, 1H), 4.69 (d, $J = 17.0$ Hz, 1H), 4.40 (s, 1H), 4.32 (d, $J = 6.6$ Hz, 1H), 1.37 (q, $J = 10.5, 9.0$ Hz, 3H), 1.21 (s, 6H), 1.16 (t, $J = 7.2$ Hz, 1H), 1.05 (d, $J = 7.8$ Hz, 8H), 0.81 (d, $J = 8.9$ Hz, 1H), -0.52 (d, 3H). HRMS (ESI+) calcd. for $\text{C}_{168}\text{H}_{144}\text{N}_{32}\text{O}_{32}\text{S}$ $[\text{M}+2\text{H}]^{2+}$ 1578.5274, found 1578.5383

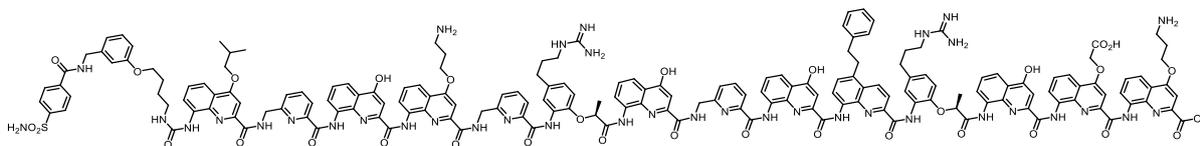
Compound 17



The scale was 15 μmol after first loading monomer determination. Target AOF was assembled on C1-MPA protide resin using SPFS (Method 5.1.1, 5.1.2). 25 mg crude was obtained after TFA cleavage. AOF was purified by semi-prep HPLC and recovered as a yellow solid (1.36 mg, 0.42 μmol , 2.8%).

^1H NMR (500 MHz, $\text{DMSO-}d_6$) δ 11.55 (s, 1H), 11.44 (s, 1H), 11.26 (s, 2H), 10.90 (s, 1H), 10.64 (s, 1H), 10.26 (s, 1H), 9.66 (s, 1H), 9.35 (s, 1H), 9.10 (s, 2H), 9.01 (s, 3H), 8.80 (s, 1H), 8.67 (s, 1H), 8.26 (s, 3H), 7.98 (d, $J = 8.2$ Hz, 3H), 7.87 (d, $J = 8.4$ Hz, 4H), 7.82 (s, 3H), 7.58 – 7.48 (m, 6H), 7.45 (s, 5H), 7.34 (s, 5H), 7.19 (s, 6H), 7.08 (d, $J = 7.3$ Hz, 3H), 6.78 (d, $J = 7.6$ Hz, 3H), 6.62 (d, $J = 11.6$ Hz, 6H), 6.26 (s, 4H), 6.12 (s, 3H), 6.03 (s, 3H), 5.32 (t, $J = 5.0$ Hz, 4H), 4.49 (s, 3H), 4.36 (d, $J = 6.1$ Hz, 3H), 3.30 (s, 21H), 2.54 (s, 19H), 2.02 – 1.96 (m, 16H), 1.46 (d, $J = 7.2$ Hz, 10H), 1.37 (d, $J = 16.3$ Hz, 3H), 1.28 (d, $J = 4.1$ Hz, 3H), 1.17 (t, $J = 7.4$ Hz, 8H), 1.08 (d, $J = 6.7$ Hz, 9H), 0.85 (t, $J = 6.8$ Hz, 12H), -0.59 (s, 3H). HRMS (ESI+) calcd. for $\text{C}_{172}\text{H}_{154}\text{N}_{34}\text{O}_{32}\text{S}$ $[\text{M}+2\text{H}]^{2+}$ 1621.5645, found 1621.5794

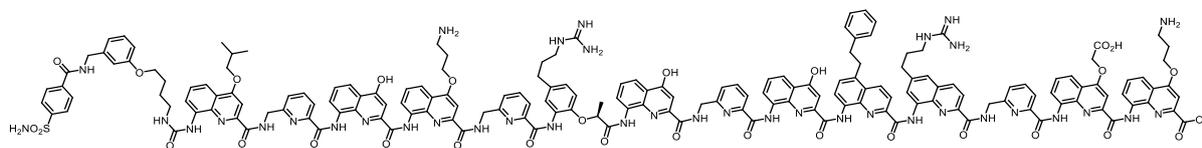
Compound 18



The scale was 12.5 μmol after first monomer loading determination. Target AOF was assembled on Cl-MPA protide resin using SPFS (Method 5.1.1, 5.1.2). 31 mg crude was obtained after TFA cleavage AOF was purified by semi-prep HPLC and recovered as yellow solid (1.43 mg, 0.43 μmol , 3.4%).

^1H NMR (500 MHz, $\text{DMSO-}d_6$) δ 11.75 (s, 1H), 11.64 (s, 1H), 11.48 (s, 1H), 11.42 (s, 1H), 11.36 (s, 1H), 11.29 (s, 1H), 11.20 (s, 1H), 10.91 (s, 1H), 10.80 (s, 1H), 10.21 (s, 1H), 9.26 (s, 1H), 9.06 (t, $J = 5.9$ Hz, 1H), 8.93 (s, 1H), 8.71 (s, 1H), 8.60 (s, 1H), 8.48 (d, $J = 7.6$ Hz, 1H), 8.31 (d, $J = 7.6$ Hz, 1H), 8.20 (m, 2H), 8.13 (d, $J = 6.9$ Hz, 1H), 8.07 (d, $J = 8.5$ Hz, 1H), 8.00 (s, 1H), 7.94 – 7.88 (m, 2H), 7.81 (d, $J = 2.1$ Hz, 2H), 7.82 – 7.70 (m, 3H), 7.67 (s, 1H), 7.62 – 7.53 (m, 2H), 7.45 (m, 13H), 7.29 (m, 6H), 7.16 (s, 1H), 7.14 – 6.96 (m, 5H), 6.91 (d, $J = 7.5$ Hz, 1H), 6.85 (d, $J = 7.7$ Hz, 2H), 6.82 – 6.69 (m, 6H), 6.67 (s, 1H), 6.56 (d, $J = 8.4$ Hz, 1H), 6.53 (d, $J = 7.7$ Hz, 2H), 6.49 (s, 1H), 6.37 (s, 1H), 6.27 (s, 2H), 6.10 (s, 1H), 6.05 (s, 2H), 5.94 (s, 1H), 5.85 (d, $J = 8.2$ Hz, 1H), 5.26 (t, $J = 5.0$ Hz, 1H), 5.07 (s, 4H), 4.41 (s, 3H), 4.29 (d, $J = 5.9$ Hz, 2H), 3.98 (s, 4H), 3.69 (s, 1H), 3.44 (s, 3H), 2.92 (s, 6H), 2.63 (s, 1H), 2.47 (s, 1H), 2.18 (s, 4H), 2.11 (dd, $J = 13.9, 6.8$ Hz, 1H), 2.06 (s, 3H), 2.01 (s, 1H), 1.97 (s, 5H), 1.96 – 1.87 (m, 2H), 1.51 (d, $J = 7.5$ Hz, 1H), 1.40 (s, 4H), 1.19 (d, $J = 23.2$ Hz, 12H), 1.11 (t, $J = 7.3$ Hz, 1H), 1.03 (dd, $J = 6.6, 1.5$ Hz, 7H), 0.82 – 0.75 (m, 2H), -0.58 (d, $J = 5.9$ Hz, 6H). HRMS (ESI+) calcd. for $\text{C}_{176}\text{H}_{163}\text{N}_{37}\text{O}_{32}\text{S}$ $[\text{M}+2\text{H}]^{2+}$ 1671.1094, found 1671.1189

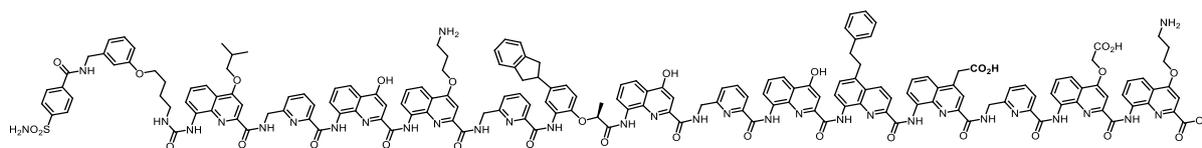
Compound 19



The scale was 12.5 μmol after first monomer loading determination. Target AOF was assembled on Cl-MPA protide resin using SPFS Method 5.1.1, 5.1.2). 36 mg crude was obtained after TFA cleavage. AOF was purified by semi-prep HPLC and recovered as yellow solid (3 mg, 0.91 μmol , 7.3%).

^1H NMR (500 MHz, $\text{Acetonitrile-}d_3$) δ 11.14 (s, 1H), 11.02 (s, 1H), 10.50 (s, 1H), 10.30 (s, 1H), 8.81 (s, 1H), 8.52 (s, 2H), 8.41 (s, 3H), 7.80 (dd, $J = 15.8, 6.9$ Hz, 10H), 7.68 (d, $J = 13.0$ Hz, 4H), 7.60 – 7.51 (m, 4H), 7.46 (s, 7H), 7.46 (t, $J = 7.7$ Hz, 4H), 7.42 – 7.31 (m, 4H), 7.14 – 6.91 (m, 8H), 6.86 – 6.75 (m, 4H), 6.65 (s, 2H), 6.58 (q, $J = 9.3, 8.8$ Hz, 4H), 5.97 (d, $J = 15.9$ Hz, 3H), 5.51 (d, $J = 10.3$ Hz, 1H), 5.22 (s, 1H), 3.58 (s, 2H), 3.09 (qd, $J = 7.9, 5.3$ Hz, 6H), 2.30 (s, 3H), 2.19 (s, 6H), 1.52 (s, 5H), 1.20 (q, $J = 8.8, 7.9$ Hz, 10H), 1.08 (d, $J = 6.2$ Hz, 4H), 1.05 (q, $J = 6.5$ Hz, 11H), -0.85 (s, 3H). HRMS (ESI+) calcd. for $\text{C}_{174}\text{H}_{160}\text{N}_{38}\text{O}_{30}\text{S}$ $[\text{M}+2\text{H}]^{2+}$ 1648.6043, found 1648.6123

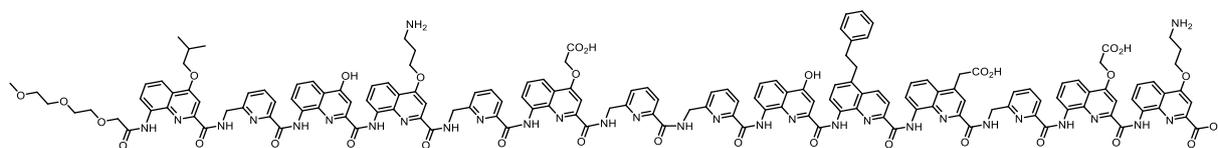
Compound 20



The scale was 13 μmol after first monomer loading determination. Target AOF was assembled on Cl-MPA protide resin using SPFS (Method 5.1.1, 5.1.2). 32 mg crude was obtained after TFA cleavage. AOF was purified by semi-prep HPLC and recovered as yellow solid (2.5 mg, 0.76 μmol , 6%).

^1H NMR (500 MHz, $\text{DMSO-}d_6$) δ 12.02 (s, 1H), 11.72 (s, 1H), 11.64 (s, 1H), 11.34 (s, 1H), 11.30 (s, 1H), 11.07 (s, 1H), 10.94 (d, $J = 16.0$ Hz, 2H), 10.73 (d, $J = 16.2$ Hz, 2H), 10.57 (s, 1H), 9.53 (s, 1H), 9.17 (t, $J = 6.0$ Hz, 1H), 9.06 (s, 1H), 8.87 (s, 1H), 8.55 (d, $J = 8.7$ Hz, 1H), 8.38 (d, $J = 7.9$ Hz, 1H), 8.30 (d, $J = 7.3$ Hz, 1H), 8.23 (d, $J = 7.4$ Hz, 1H), 8.04 – 7.98 (m, 3H), 7.92 (d, $J = 13.8$ Hz, 4H), 7.92 – 7.86 (m, 3H), 7.81 (s, 7H), 7.77 (d, $J = 16.6$ Hz, 1H), 7.73 (d, $J = 8.9$ Hz, 3H), 7.64 (ddd, $J = 26.4$, 14.3, 7.4 Hz, 6H), 7.55 (d, $J = 8.0$ Hz, 1H), 7.55 – 7.50 (m, 2H), 7.50 – 7.42 (m, 7H), 7.38 (s, 7H), 7.42 – 6.99 (m, 15H), 6.99 – 6.91 (m, 2H), 6.87 (d, $J = 7.3$ Hz, 1H), 6.83 (s, 1H), 6.80 (d, $J = 14.4$ Hz, 5H), 6.69 (dd, $J = 25.8$, 8.9 Hz, 2H), 6.60 (d, $J = 7.7$ Hz, 1H), 6.50 (s, 1H), 6.44 (s, 1H), 6.40 (d, $J = 7.8$ Hz, 1H), 6.36 (d, $J = 7.0$ Hz, 2H), 5.81 (s, 1H), 4.85 (d, $J = 16.1$ Hz, 1H), 4.73 (d, $J = 15.9$ Hz, 1H), 4.61 (s, 1H), 4.40 (d, $J = 5.9$ Hz, 3H), 4.32 (s, 1H), 4.17 (s, 1H), 4.01 (d, $J = 15.0$ Hz, 1H), 3.95 (s, 2H), 3.87 (d, $J = 17.2$ Hz, 3H), 3.50 (s, 1H), 3.24 (d, $J = 14.9$ Hz, 1H), 3.19 – 3.13 (m, 2H), 3.09 (dd, $J = 7.3$, 4.7 Hz, 1H), 3.01 (s, 2H), 2.95 (s, 2H), 2.86 – 2.78 (m, 2H), 2.69 (s, 1H), 2.21 – 2.07 (m, 2H), 2.07 (s, 5H), 2.04 – 1.95 (m, 1H), 1.60 (s, 2H), 1.44 (d, $J = 7.4$ Hz, 1H), 1.42 (s, 3H), 1.23 (s, 7H), 1.18 (t, $J = 7.3$ Hz, 1H), 1.06 (d, $J = 6.6$ Hz, 1H), 1.01 (d, $J = 6.8$ Hz, 6H), 0.88 – 0.82 (m, 1H), -0.20 (s, 3H). HRMS (ESI+) calcd. for $\text{C}_{177}\text{H}_{152}\text{N}_{32}\text{O}_{32}\text{S}$ $[\text{M}+2\text{H}]^{2+}$ 1636.558, found 1636.5659

Compound 21



The scale was 15 μmol after first monomer loading determination. Target compound was prepared on Cl-protide resin using SPPS (Method 5.1.1, 5.1.2). 39 mg crude was obtained after TFA cleavage, 7 mg target compound was obtained purified by semi-prep HPLC as yellow solid (2.4 μmol , 16%)

^1H NMR (500 MHz, $\text{DMSO-}d_6$) δ 13.33 (s, 1H), 12.03 (s, 1H), 11.39 (s, 1H), 11.33 (s, 1H), 11.14 (s, 1H), 10.94 (s, 1H), 10.87 (s, 1H), 10.67 (s, 1H), 10.59 (s, 1H), 10.51 (s, 1H), 10.08 (s, 1H), 9.85 (s, 1H), 9.30 (s, 1H), 8.47 (d, $J = 8.6$ Hz, 1H), 8.37 (s, 1H), 8.27 – 8.19 (m, 2H), 7.95 – 7.87 (m, 3H), 7.89 – 7.82 (m, 4H), 7.79 (q, $J = 6.7$, 6.1 Hz, 1H), 7.72 (d, $J = 11.7$ Hz, 4H), 7.71 – 7.62 (m, 2H), 7.65 – 7.52

(m, 4H), 7.55 – 7.47 (m, 4H), 7.47 (t, J = 7.5 Hz, 2H), 7.41 (t, J = 7.3 Hz, 1H), 7.38 – 7.23 (m, 7H), 7.17 (dd, J = 14.8, 7.9 Hz, 5H), 7.10 (t, J = 6.9 Hz, 1H), 7.10 – 7.01 (m, 6H), 7.01 – 6.89 (m, 5H), 6.81 (q, J = 8.0, 7.2 Hz, 4H), 6.68 – 6.63 (m, 2H), 6.54 (dd, J = 18.0, 7.3 Hz, 2H), 6.40 (d, J = 12.2 Hz, 1H), 6.28 (s, 1H), 6.17 (s, 1H), 5.16 (s, 1H), 4.77 (d, J = 15.8 Hz, 1H), 4.65 (d, J = 15.8 Hz, 1H), 4.37 (d, J = 13.1 Hz, 2H), 4.31 (d, J = 7.5 Hz, 2H), 4.22 (d, J = 14.8 Hz, 1H), 3.98 (d, J = 7.9 Hz, 1H), 3.96 – 3.86 (m, 2H), 3.65 (d, J = 15.0 Hz, 1H), 3.50 (s, 2H), 3.19 – 3.13 (m, 2H), 3.05 – 2.99 (m, 1H), 2.80 – 2.63 (m, 4H), 2.70 (s, 3H), 2.22 (dq, J = 20.2, 6.9 Hz, 2H), 2.16 – 2.04 (m, 2H), 1.78 (d, J = 24.4 Hz, 1H), 1.69 (d, J = 15.5 Hz, 1H), 1.24 (d, J = 12.1 Hz, 1H), 1.23 (s, 2H), 1.17 (t, J = 7.3 Hz, 1H), 1.08 (t, J = 6.9 Hz, 6H). HRMS (ESI+) calcd. for C₁₅₆H₁₃₄N₃₀O₃₂ [M+2H]²⁺ 1470.9948, found 1471.0129

10.5.5 NMR spectra and RP-HPLC chromatograms of new compounds

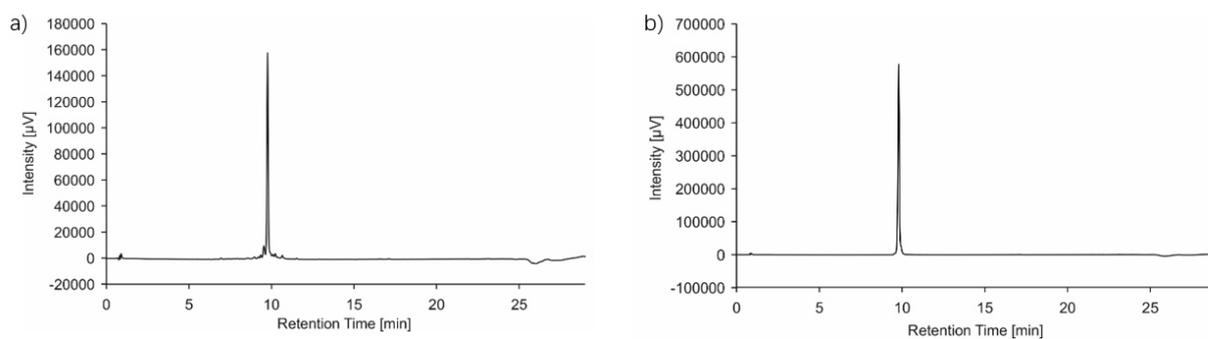
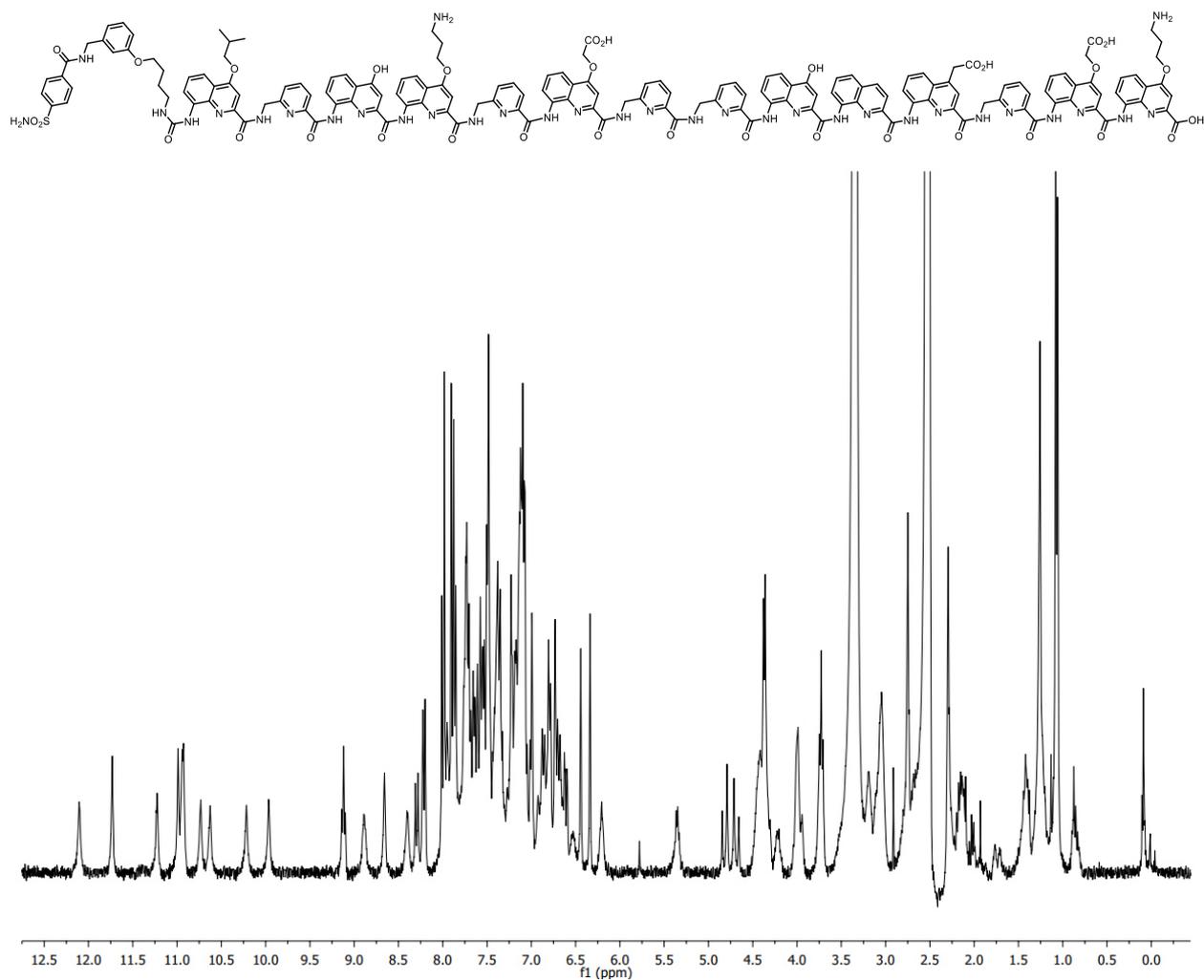


Figure S22. Analytical data of compound **2**. RP-HPLC chromatograms (a) after cleavage from the resin (C18, 5 to 100 B% over 23 min, 50 °C, $\lambda = 254$ nm) and (b) after purification (C18, 10 to 100 B% over 10 min, 50 °C, $\lambda = 254$ nm); A: 0.1% TFA water, B: 0.1% TFA acetonitrile



^1H NMR spectrum (300 MHz, $\text{DMSO-}d_6$, 25 °C)

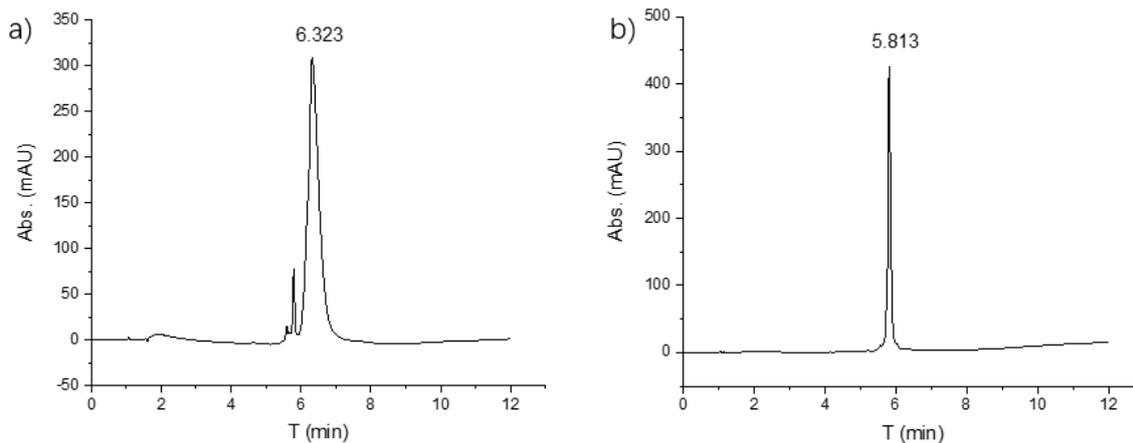
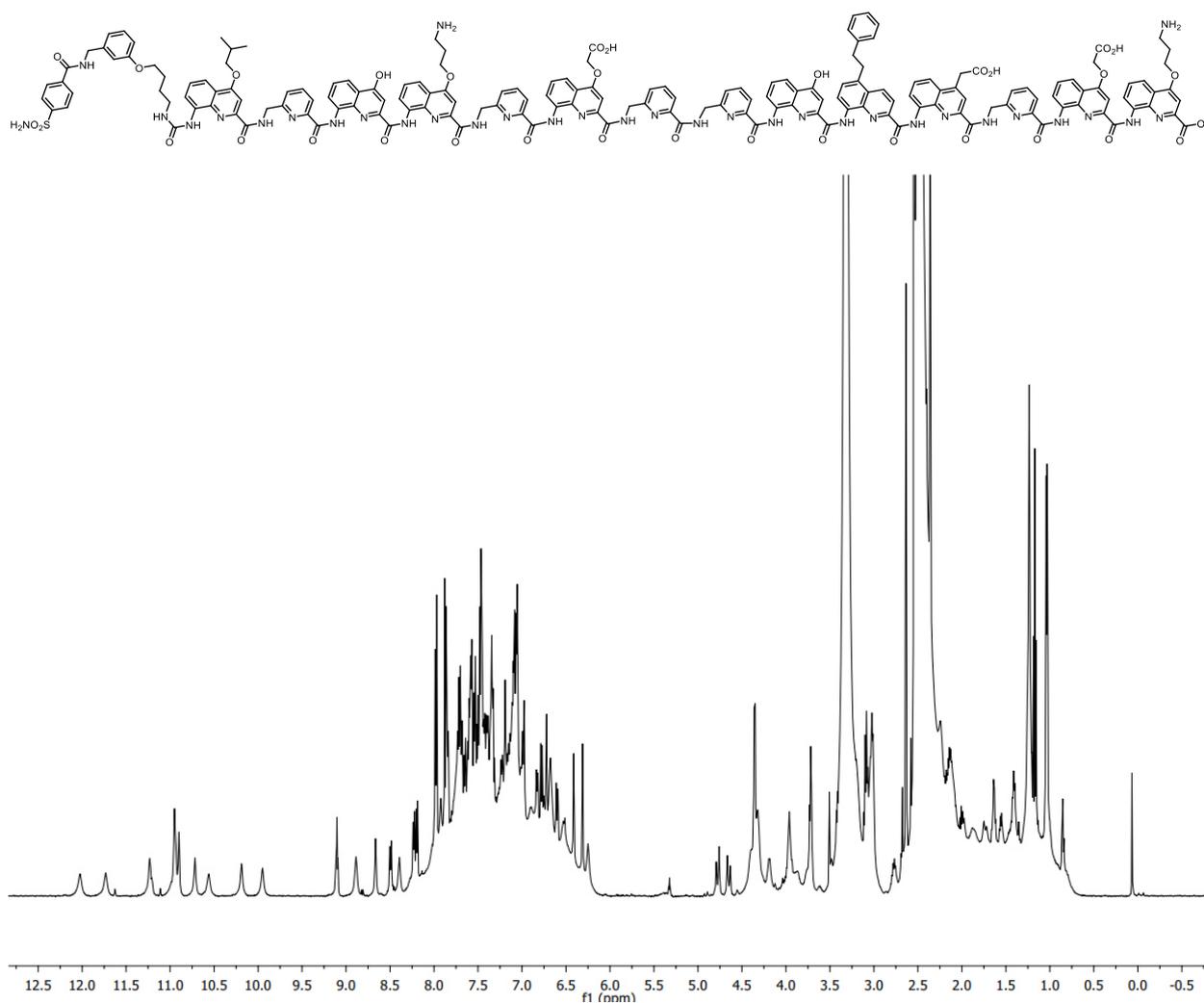


Figure S23. Analytical data of compound **3**. RP-HPLC chromatograms (a) after cleavage from the resin (C18, 10 to 100 B% over 10 min, 50 °C, $\lambda = 300$ nm) and (b) after purification (C18, 10 to 100 B% over 10 min, 50 °C, $\lambda = 254$ nm); A: 0.1% TFA water, B: 0.1% TFA acetonitrile



^1H NMR spectrum (500 MHz, $\text{DMSO-}d_6$, 25 °C)

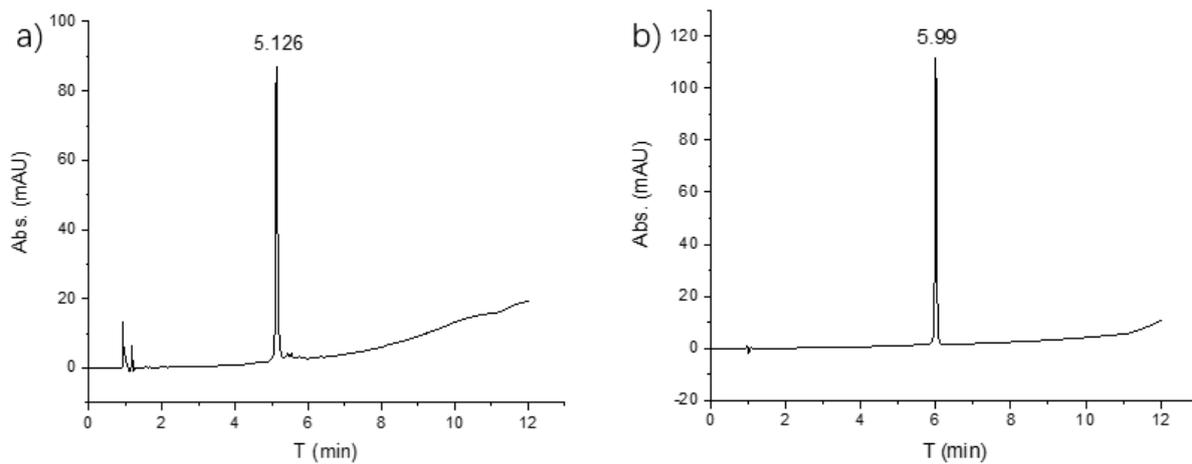
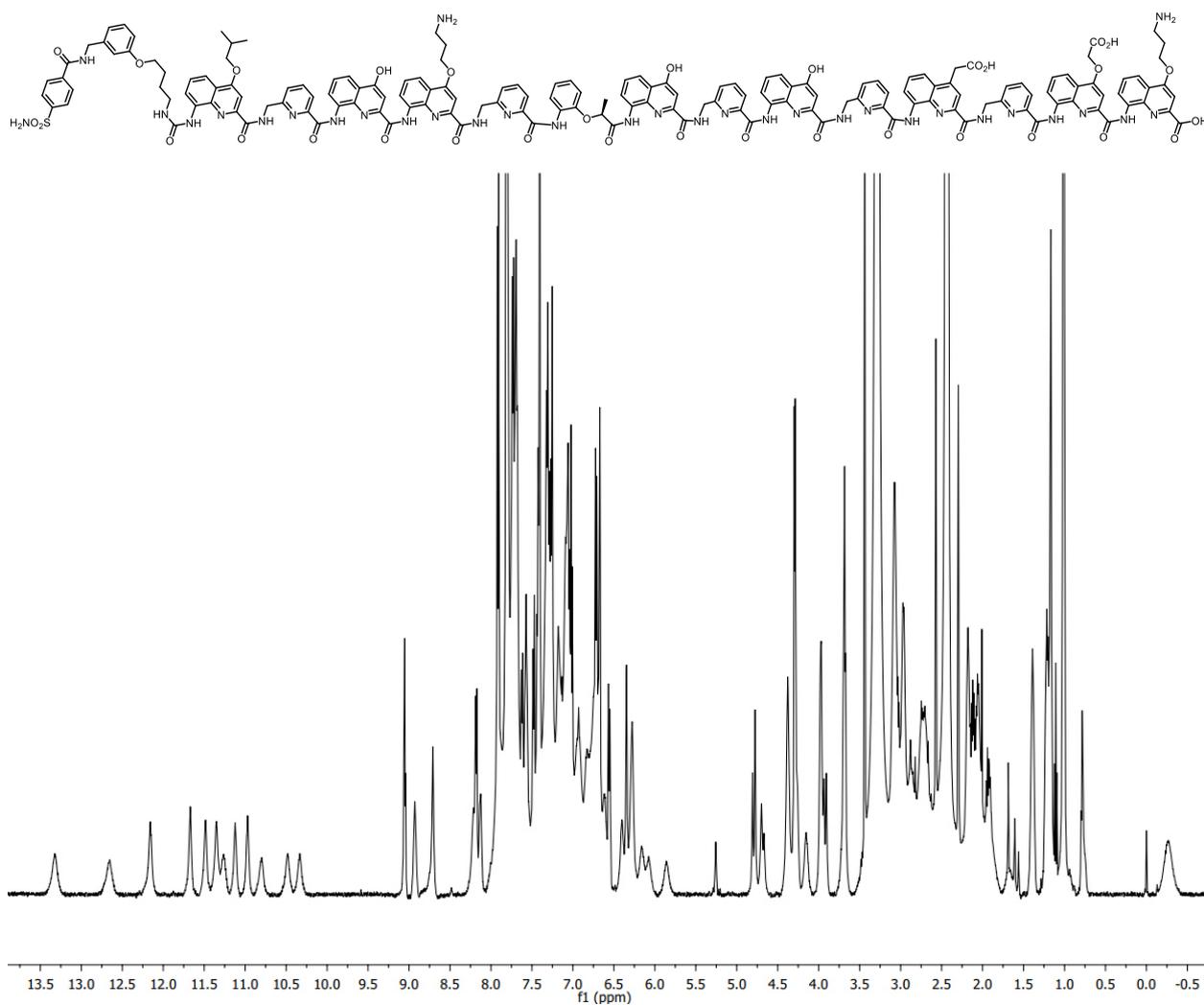


Figure S24. Analytical data of compound **15**. RP-HPLC chromatograms (a) after cleavage from the resin (C18, 10 to 100 B% over 10 min, 50 °C, $\lambda = 254$ nm) and (b) after purification (C18, 20 to 70 B% over 10 min, 50 °C, $\lambda = 254$ nm); A: 0.1% TFA water, B: 0.1% TFA acetonitrile.



^1H NMR spectrum (500 MHz, $\text{DMSO-}d_6$, 25 °C)

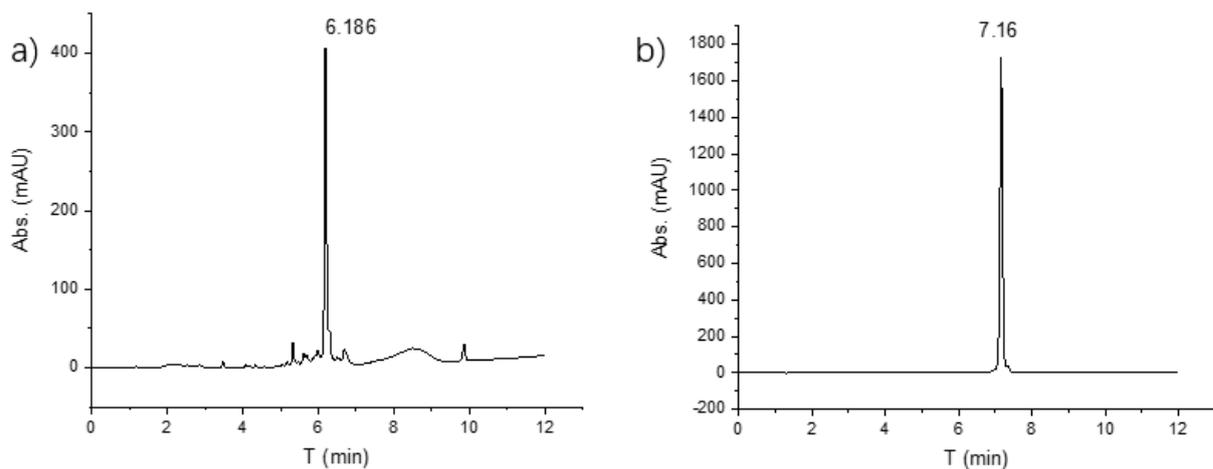
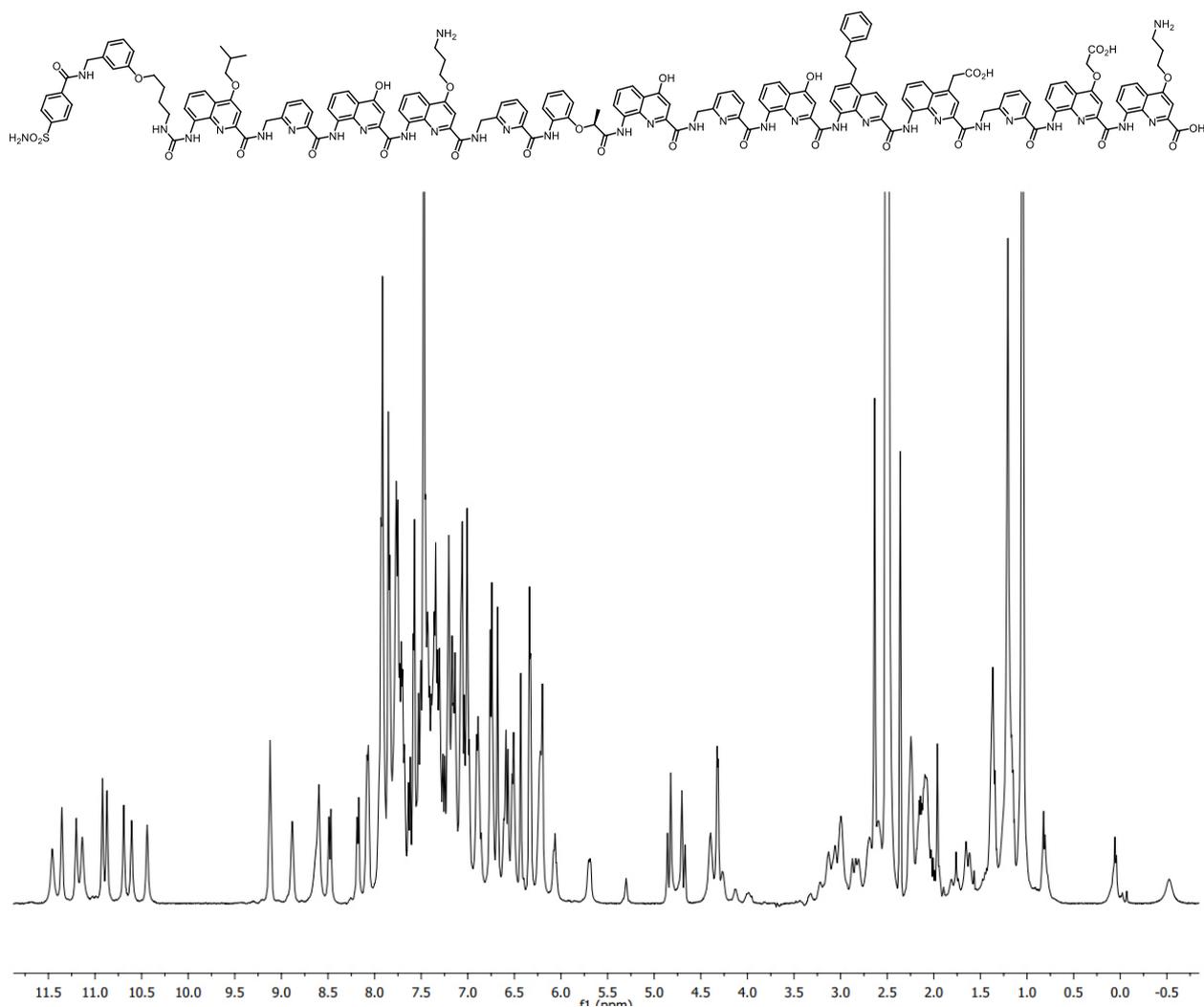


Figure S25. Analytical data of compound **16**. RP-HPLC chromatograms (a) after cleavage from the resin (C18, 10 to 100 B% over 10 min, 50 °C, $\lambda = 254$ nm) and (b) after purification (C18, 20 to 70 B% over 10 min, 50 °C, $\lambda = 254$ nm); A: 0.1% TFA water, B: 0.1% TFA acetonitrile.



^1H NMR spectrum (500 MHz, $\text{DMSO-}d_6/\text{H}_2\text{O}$ (9:1), 25 °C)

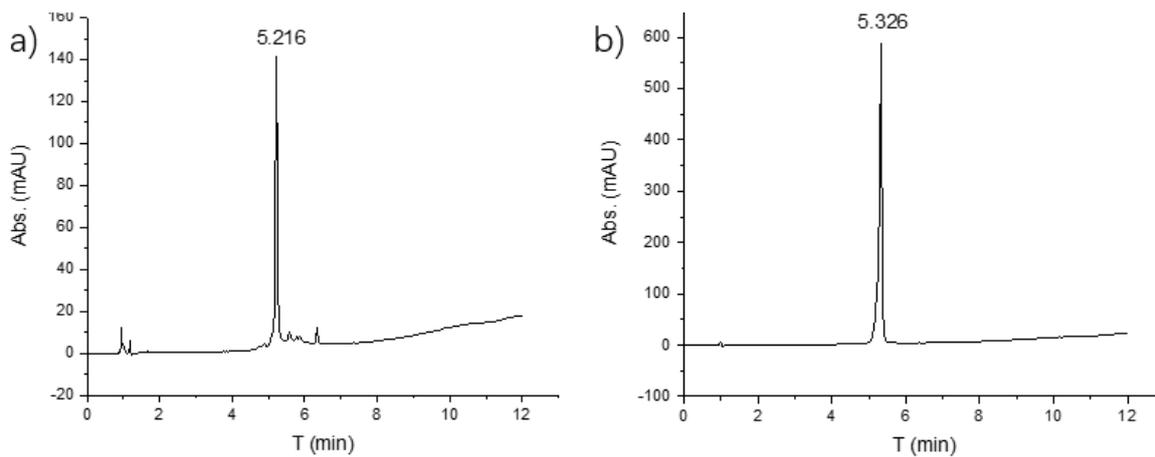
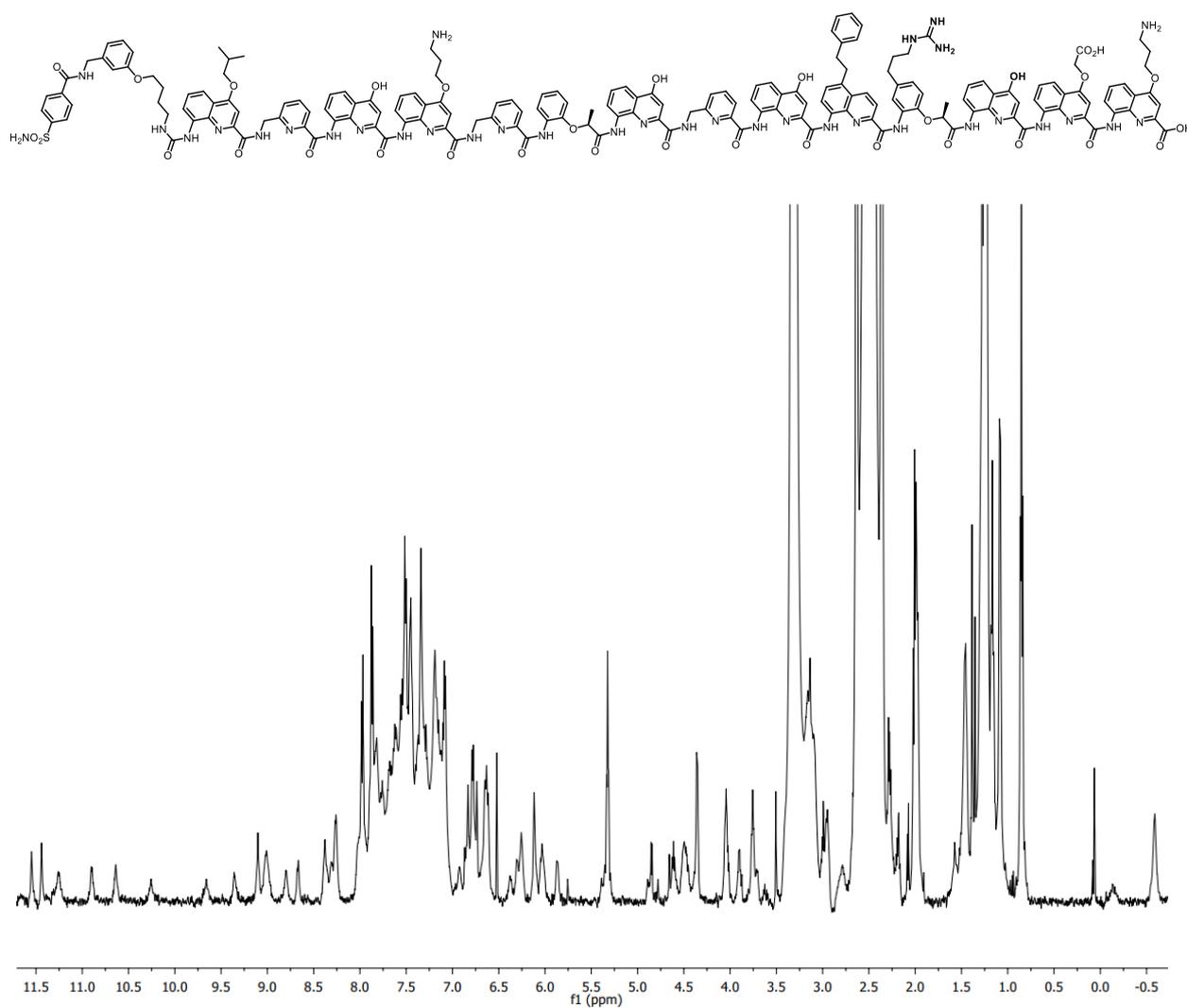


Figure S26. Analytical data of compound **17**. RP-HPLC chromatograms (a) after cleavage from the resin (C18, 10 to 100 B% over 10 min, 50 °C, $\lambda = 254$ nm) and (b) after purification (C18, 10 to 100 B% over 10 min, 50 °C, $\lambda = 254$ nm); A: 0.1% TFA water, B: 0.1% TFA acetonitrile.



^1H NMR spectrum (500 MHz, $\text{DMSO-}d_6$, 25 °C)

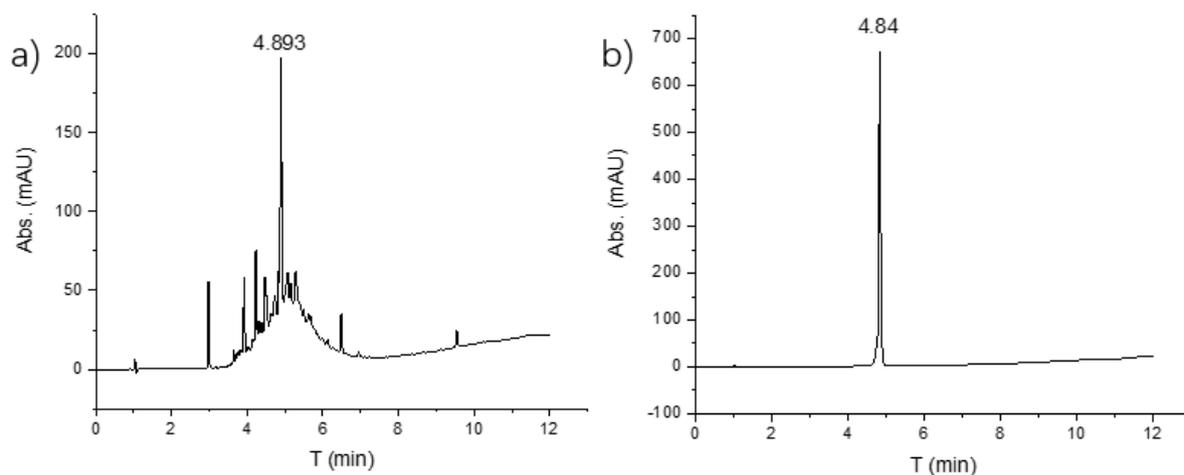
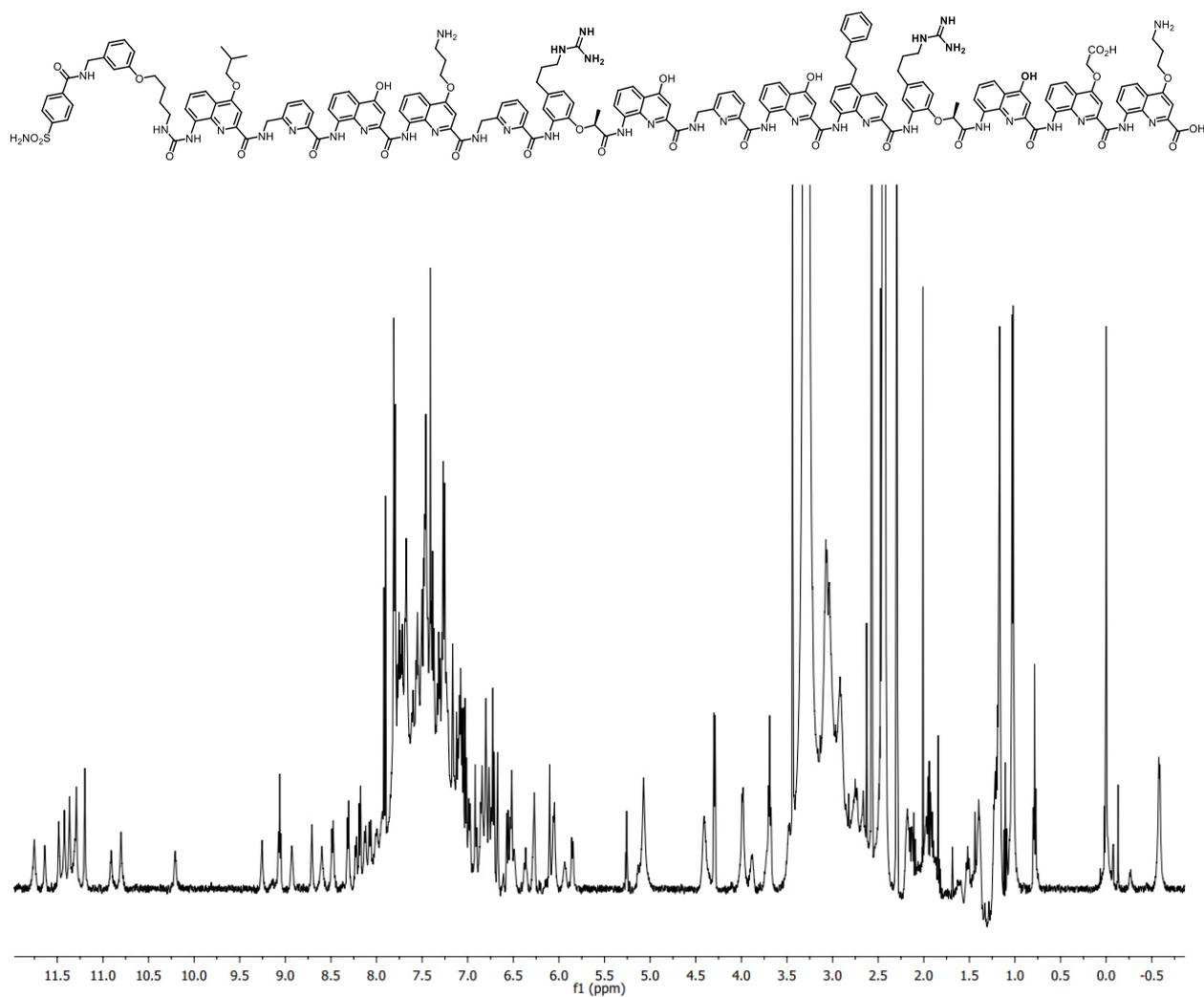


Figure S27. Analytical data of compound **18**. RP-HPLC chromatograms (a) after cleavage from the resin (C18, 10 to 100 B% over 10 min, 50 °C, $\lambda = 254$ nm) and (b) after purification (C18, 10 to 100 B% over 10 min, 50 °C, $\lambda = 254$ nm); A: 0.1% TFA water, B: 0.1% TFA acetonitrile.



^1H NMR spectrum (500 MHz, $\text{DMSO-}d_6$, 25 °C)

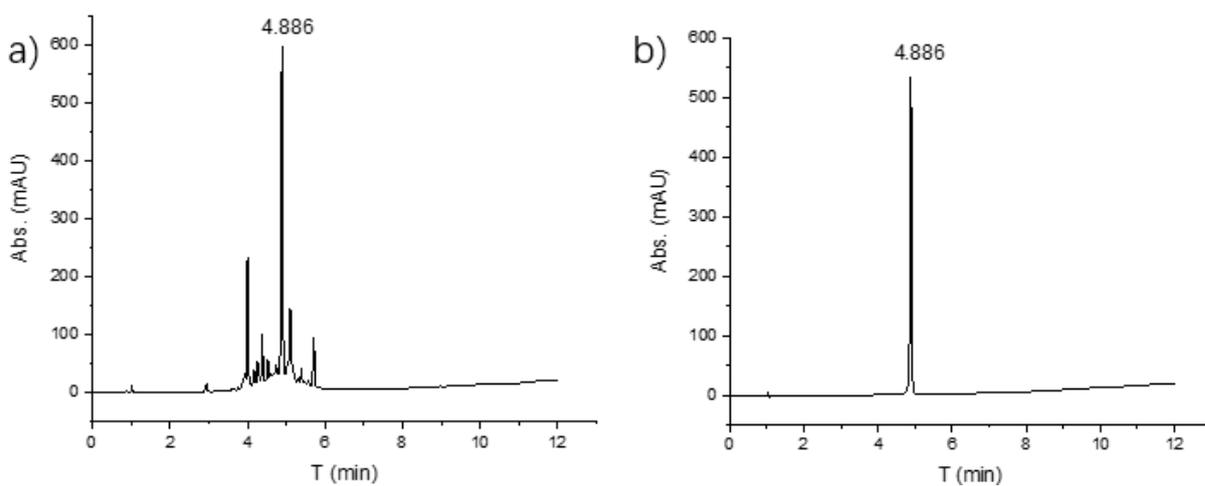
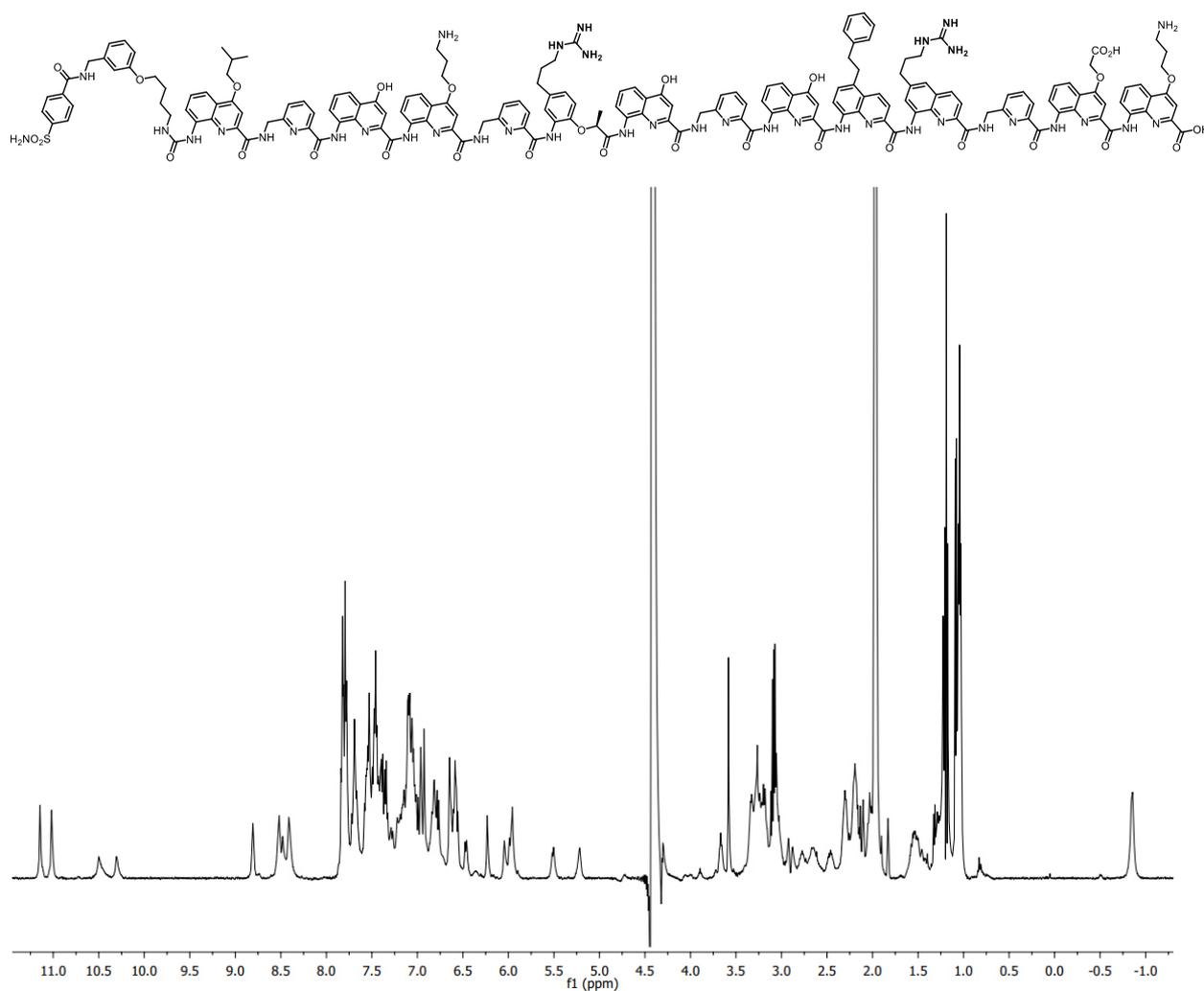


Figure S28. Analytical data of compound **19**. RP-HPLC chromatograms (a) after cleavage from the resin (C18, 10 to 100 B% over 10 min, 50 °C, $\lambda = 254$ nm) and (b) after purification (C18, 10 to 100 B% over 10 min, 50 °C, $\lambda = 254$ nm); A: 0.1% TFA water, B: 0.1% TFA acetonitrile.



^1H NMR spectrum with water suppression (500 MHz, $\text{H}_2\text{O}/\text{CD}_3\text{CN}$ (1:3, v/v), 25 °C).

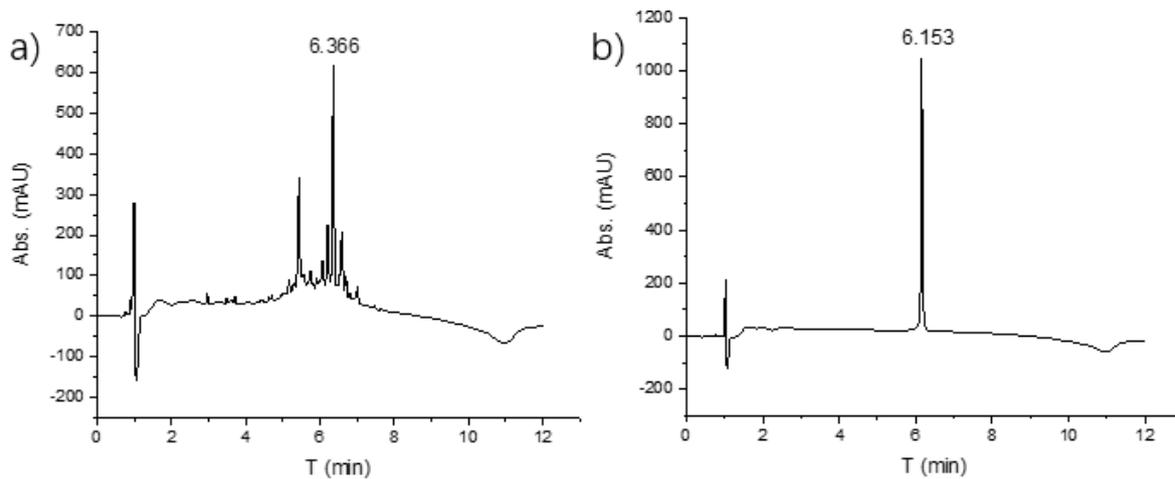
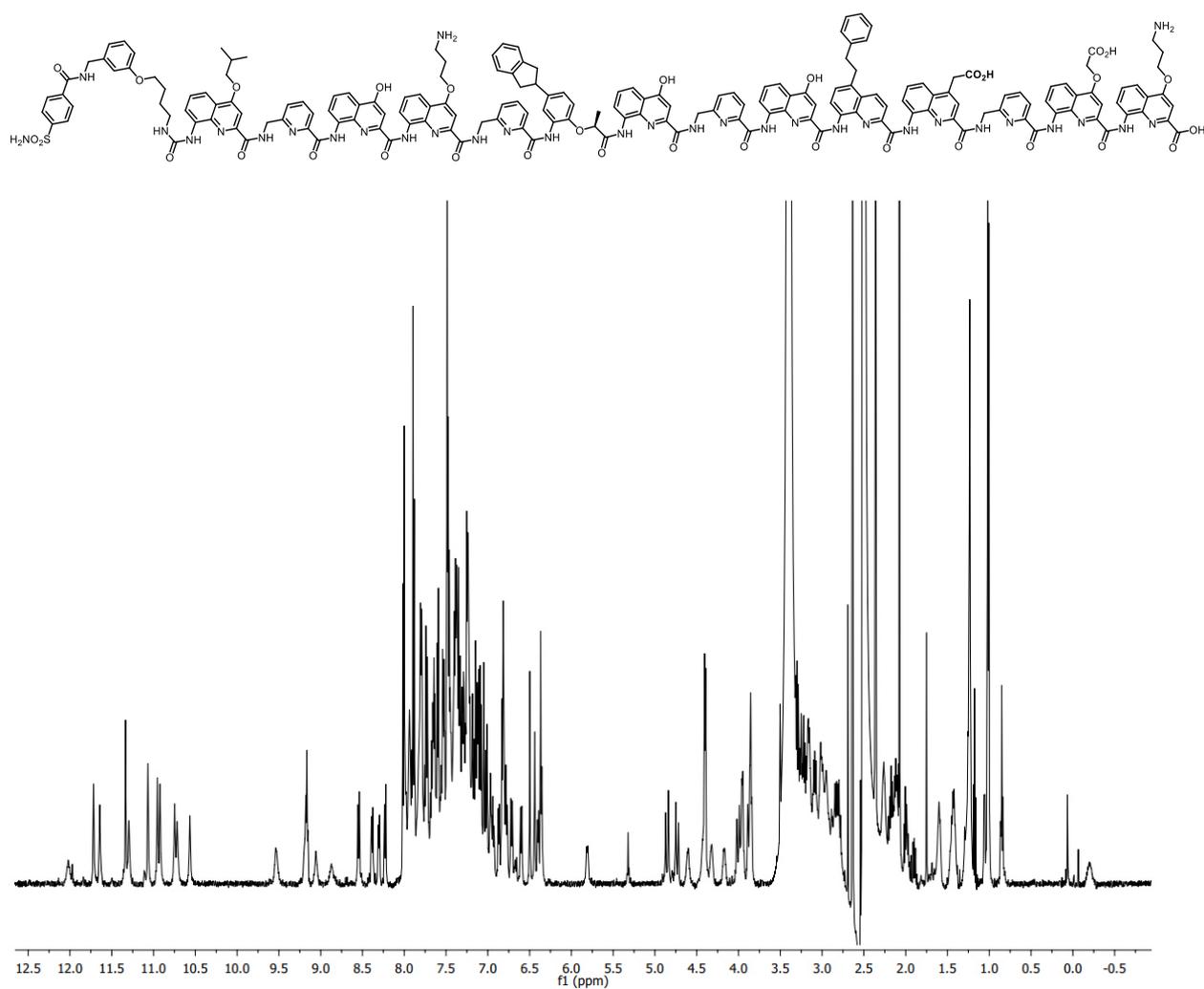


Figure S29. Analytical data of compound **20**. RP-HPLC chromatograms (a) after cleavage from the resin (C18, 10 to 100 B% over 10 min, 50 °C, $\lambda = 254$ nm) and (b) after purification (C18, 10 to 100 B% over 10 min, $\lambda = 254$ nm); A: 0.1% TFA water, B: 0.1% TFA acetonitrile.



^1H NMR spectrum (500 MHz, $\text{DMSO-}d_6$, 25 °C)

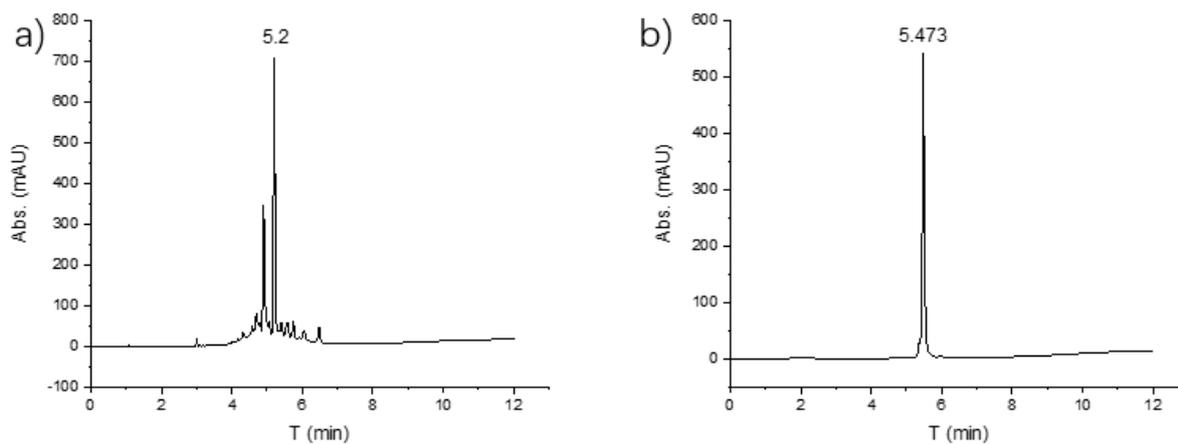
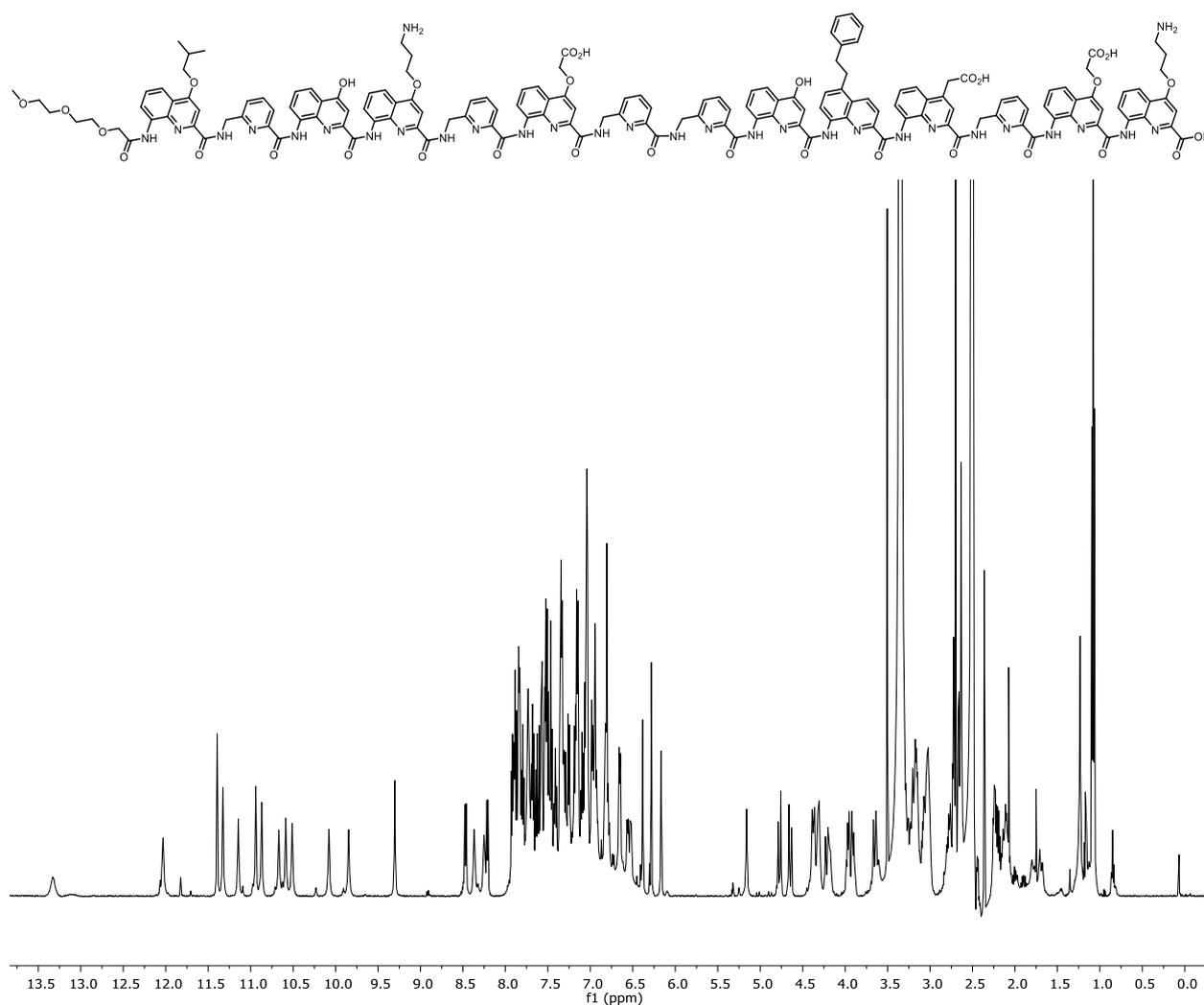


Figure S30. Analytical data of compound **21**. RP-HPLC chromatograms (a) after cleavage from the resin (C18, 10 to 100 B% over 10 min, 50 °C, $\lambda = 254$ nm) and (b) after purification (C18, 10 to 100 B% over 10 min, $\lambda = 254$ nm); A: 0.1% TFA water, B: 0.1% TFA acetonitrile



^1H NMR spectrum (500 MHz, $\text{DMSO}-d_6$, 25 °C)

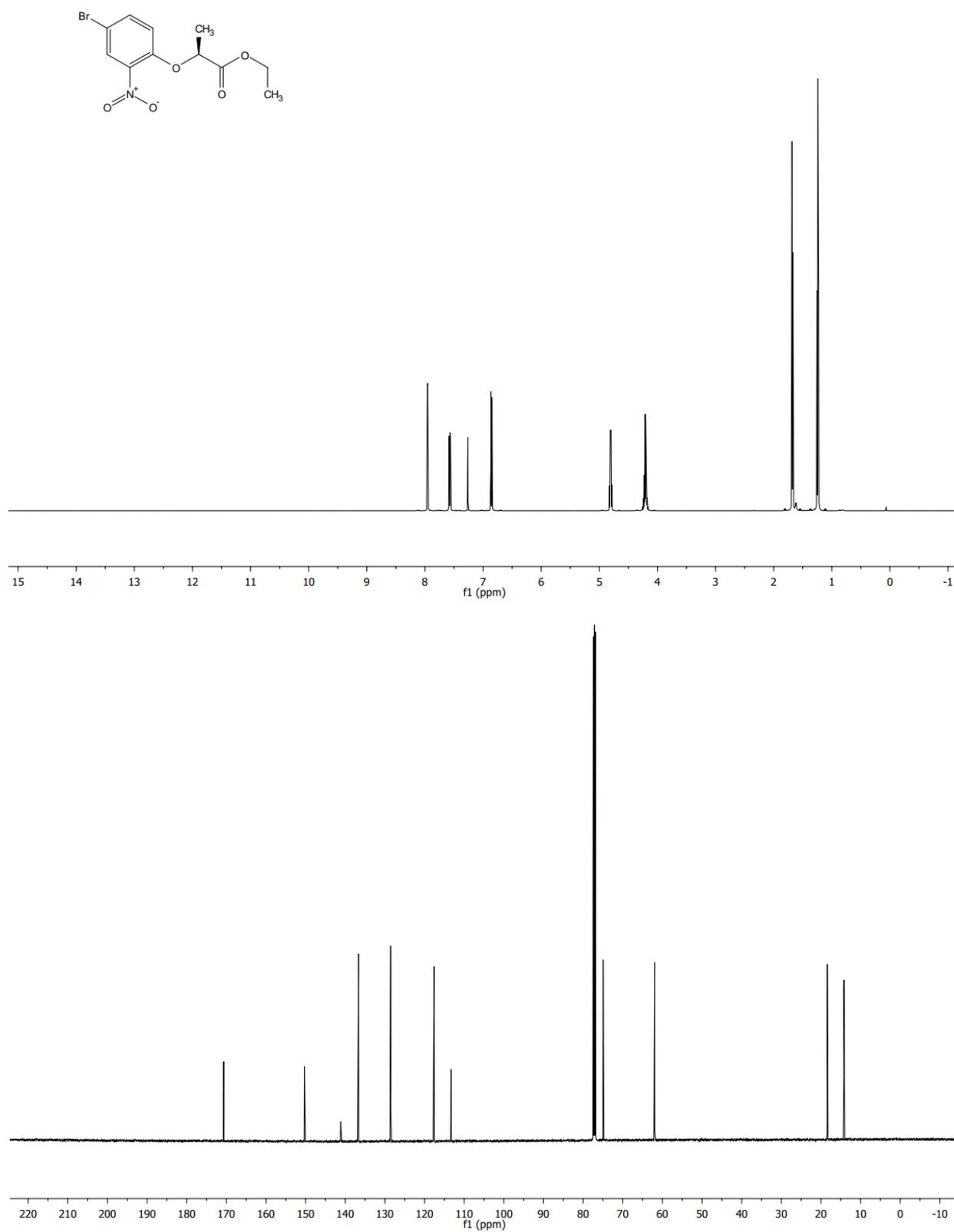


Figure S31. NMR spectra of **25**: ¹H NMR (500 MHz, DMSO-*d*₆) and ¹³C NMR (126 MHz, DMSO-*d*₆).

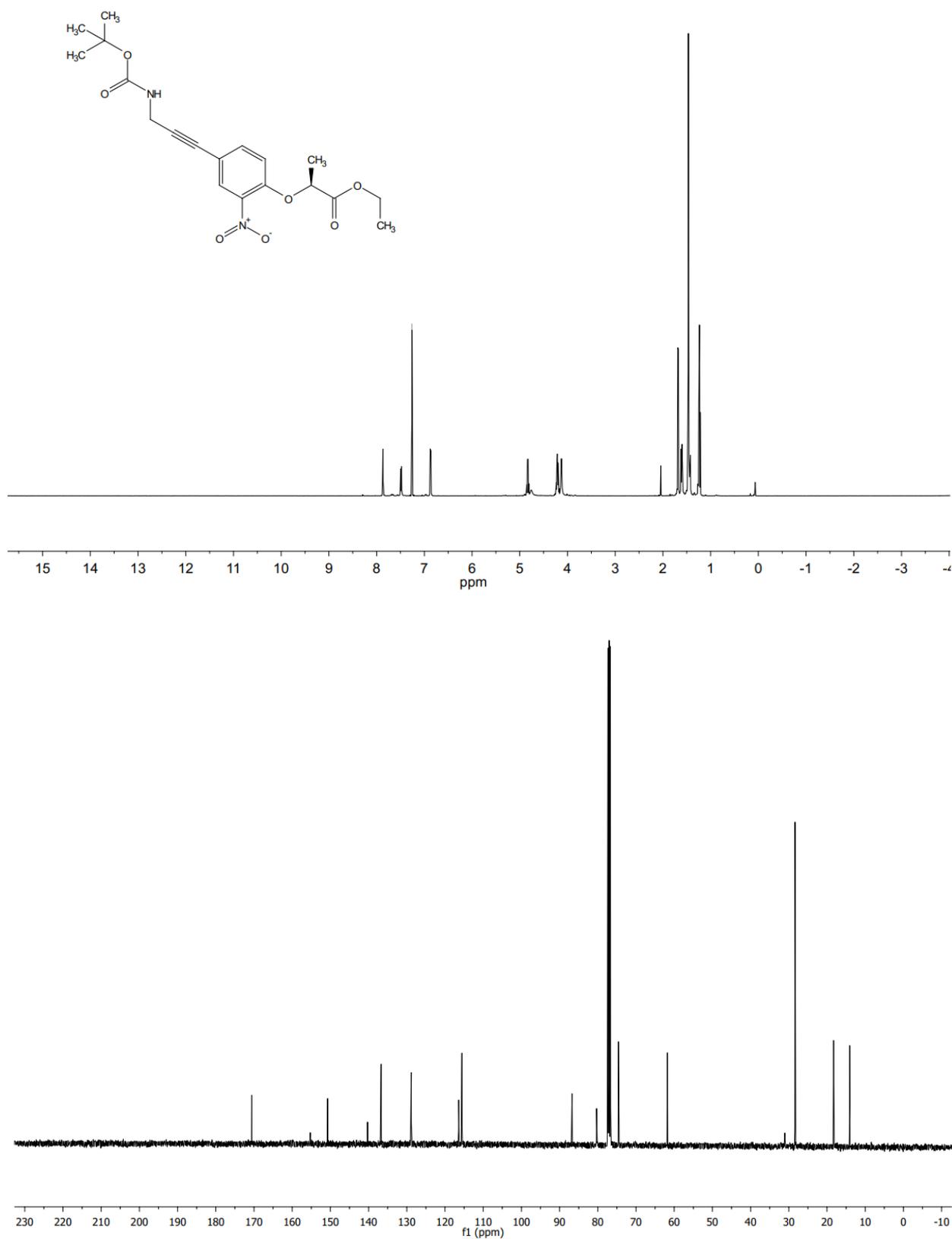


Figure S32. NMR spectra of **26**: ^1H NMR (400 MHz, CDCl_3) and ^{13}C NMR (101 MHz, CDCl_3)

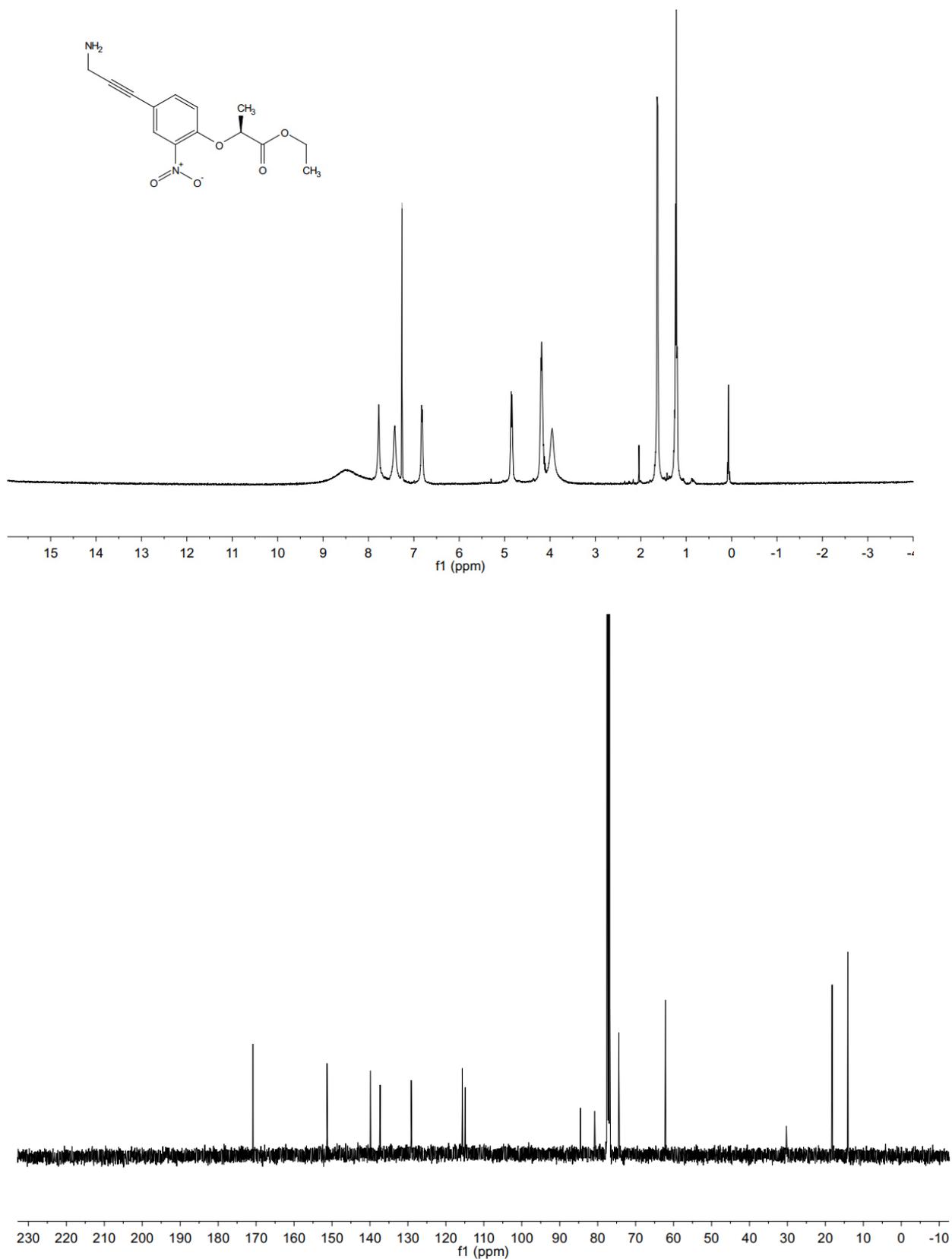


Figure S33. NMR spectra of **27**: ^1H NMR (400 MHz, CDCl_3) and ^{13}C NMR (101 MHz, CDCl_3).

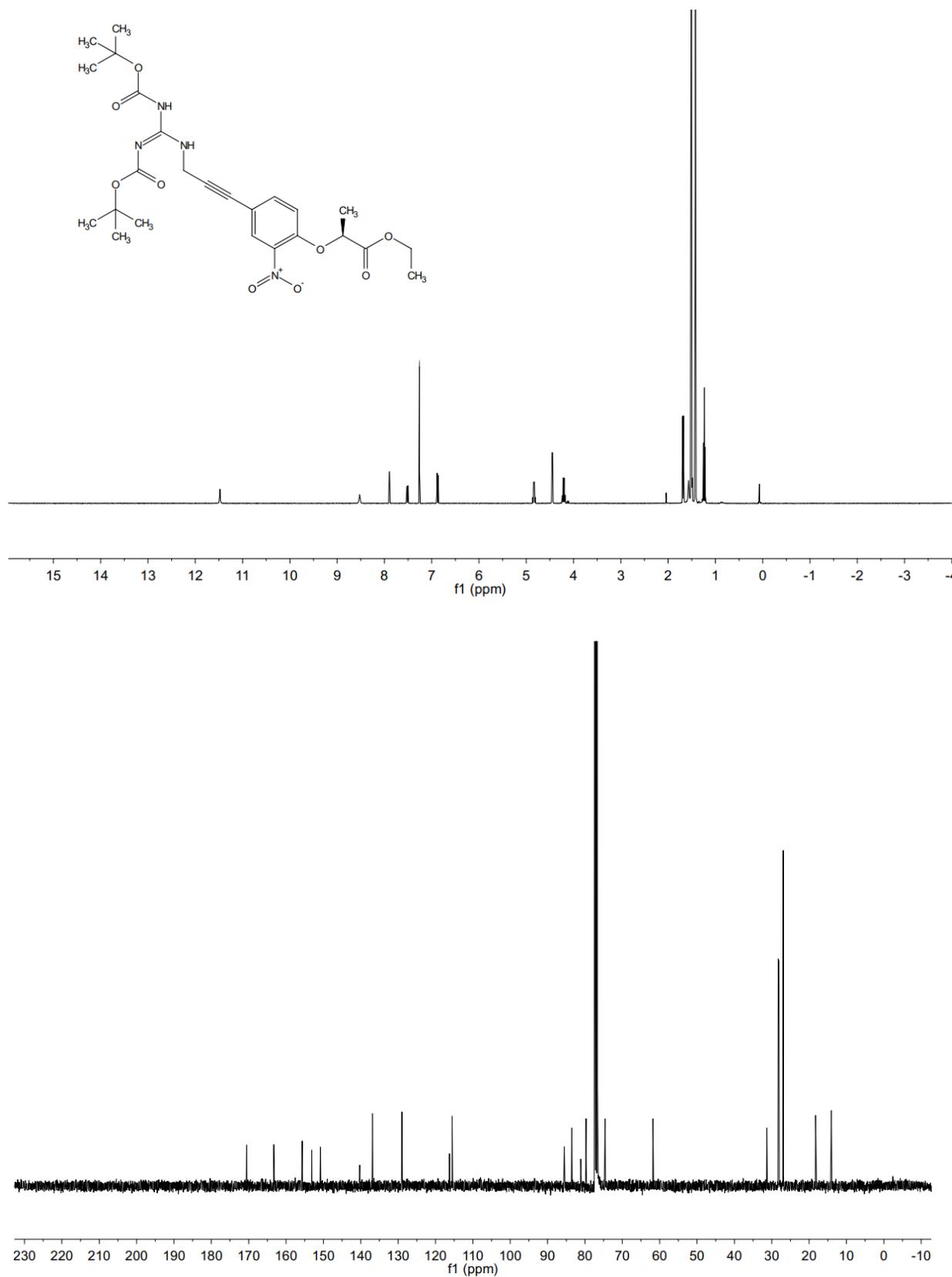


Figure S34. NMR spectra of **28**: ^1H NMR (400 MHz, CDCl_3) and ^{13}C NMR (101 MHz, CDCl_3).

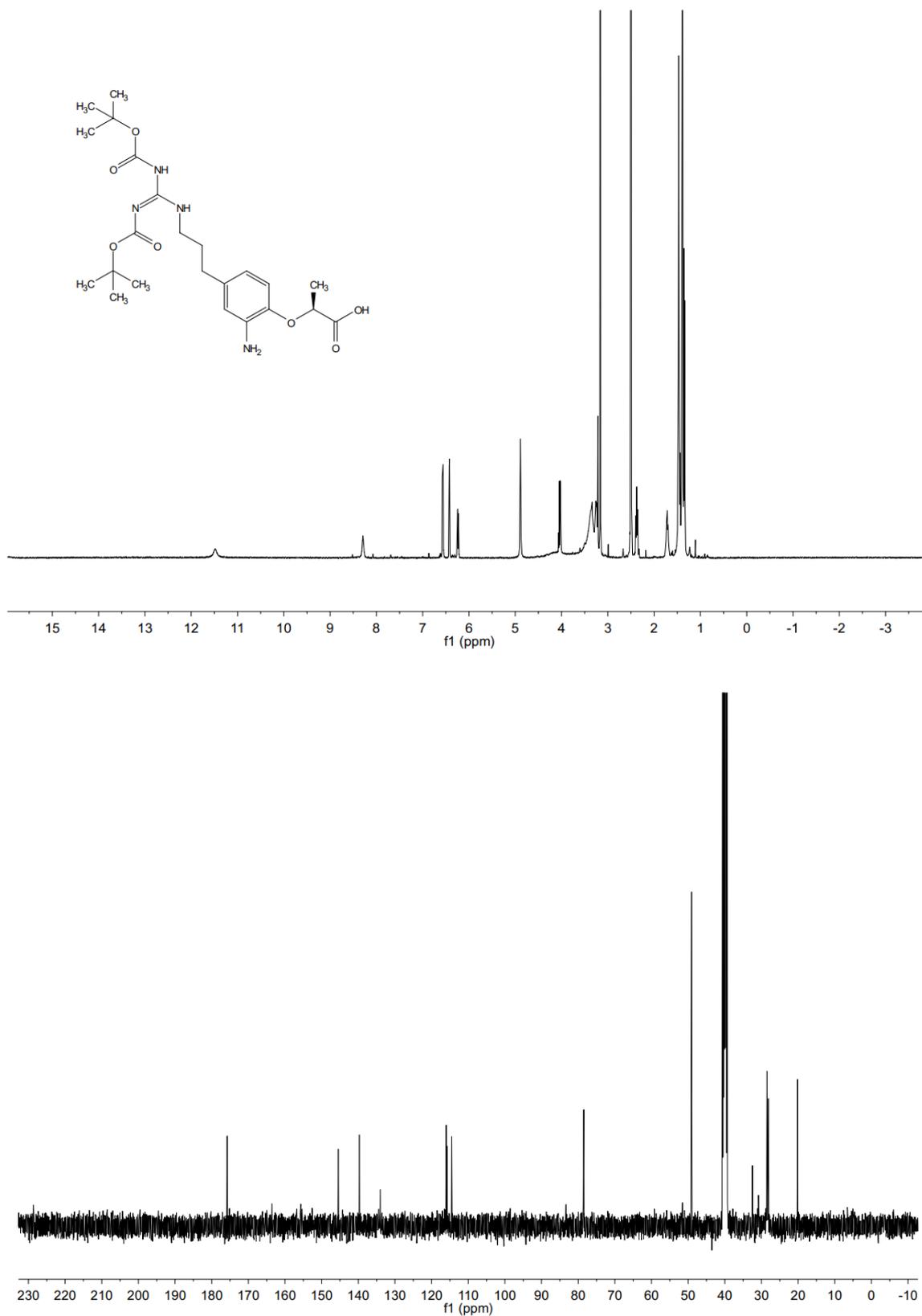


Figure S35. NMR spectra of **29**: ^1H NMR (400 MHz, CDCl_3) and ^{13}C NMR (101 MHz, CDCl_3).

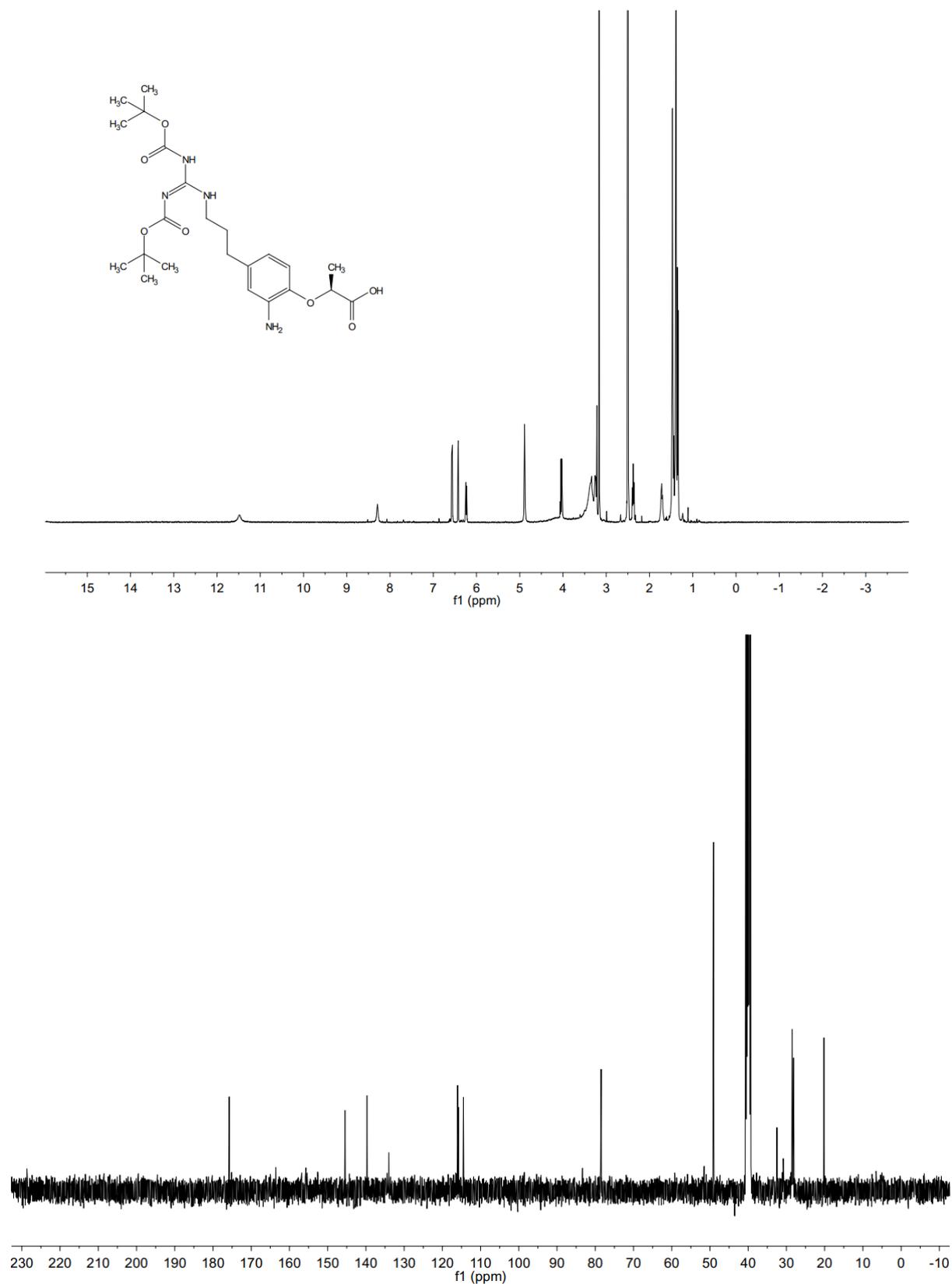


Figure S36. NMR spectra of **30**: ^1H NMR (400 MHz, $\text{DMSO}-d_6$) and ^{13}C NMR (101 MHz, $\text{DMSO}-d_6$)

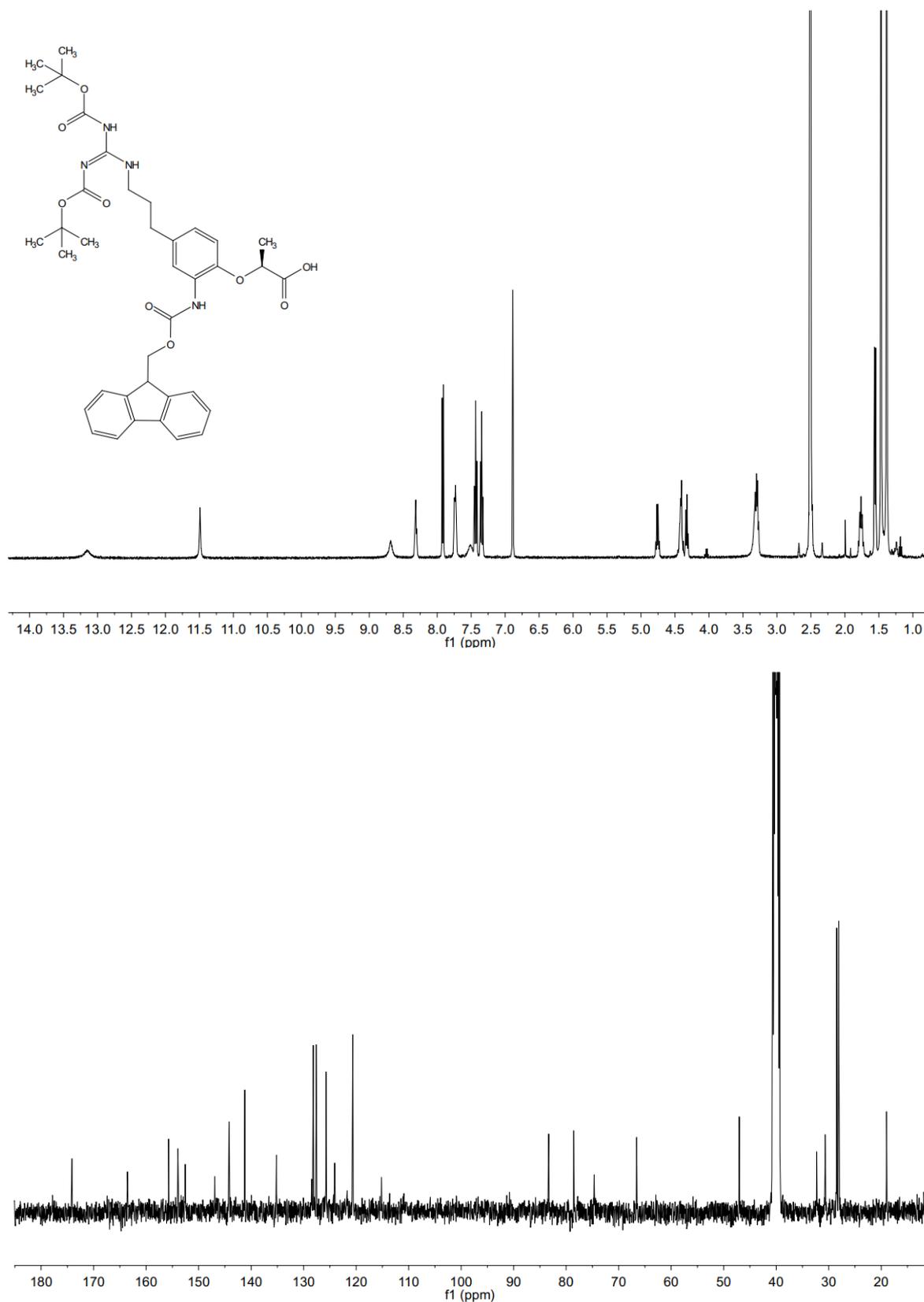


Figure S37. NMR spectra of **31**: ^1H NMR (400 MHz, $\text{DMSO}-d_6$) and ^{13}C NMR (101 MHz, $\text{DMSO}-d_6$)

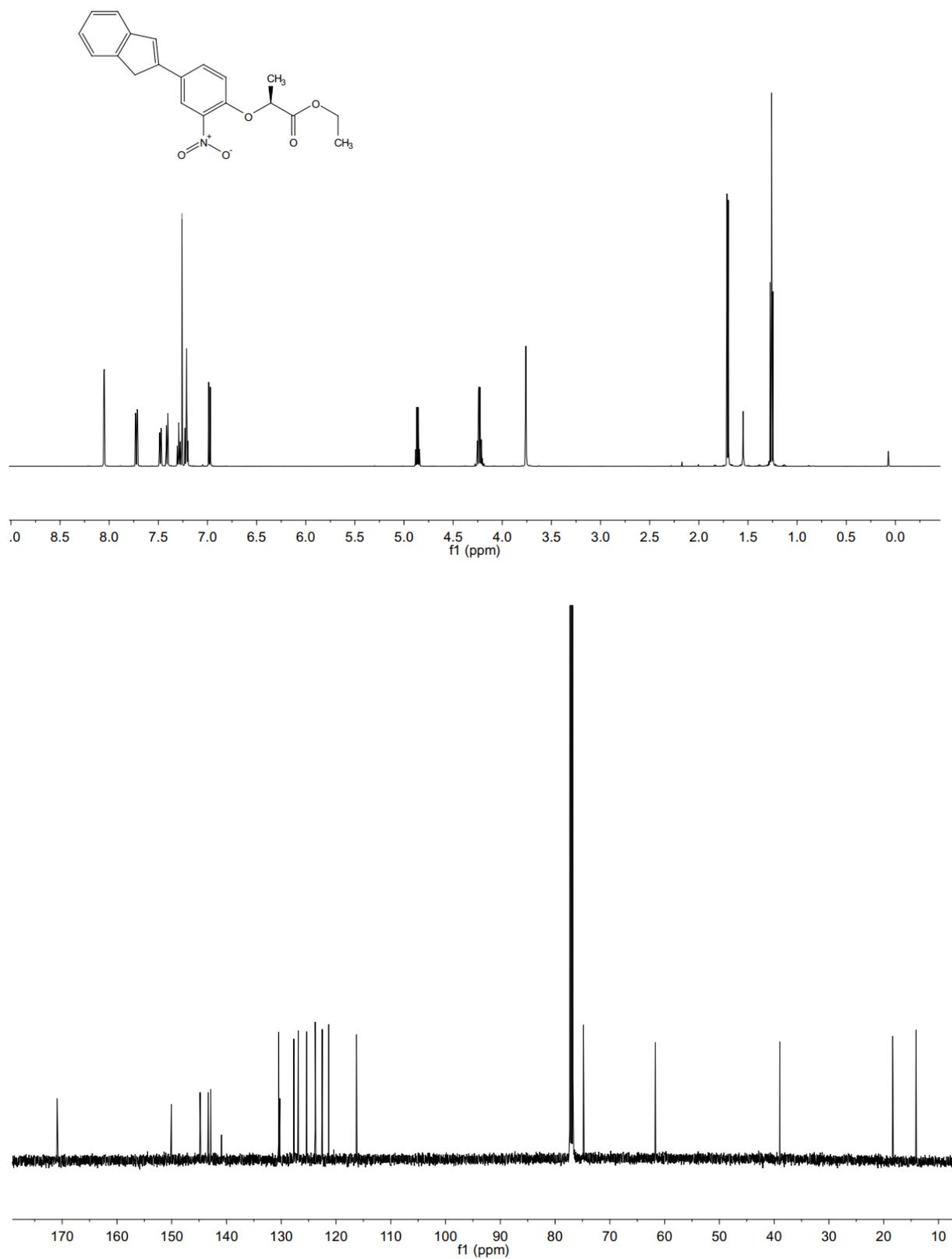


Figure S38. NMR spectra of **32**: ^1H NMR (400 MHz, CDCl_3) and ^{13}C NMR (101 MHz, CDCl_3).

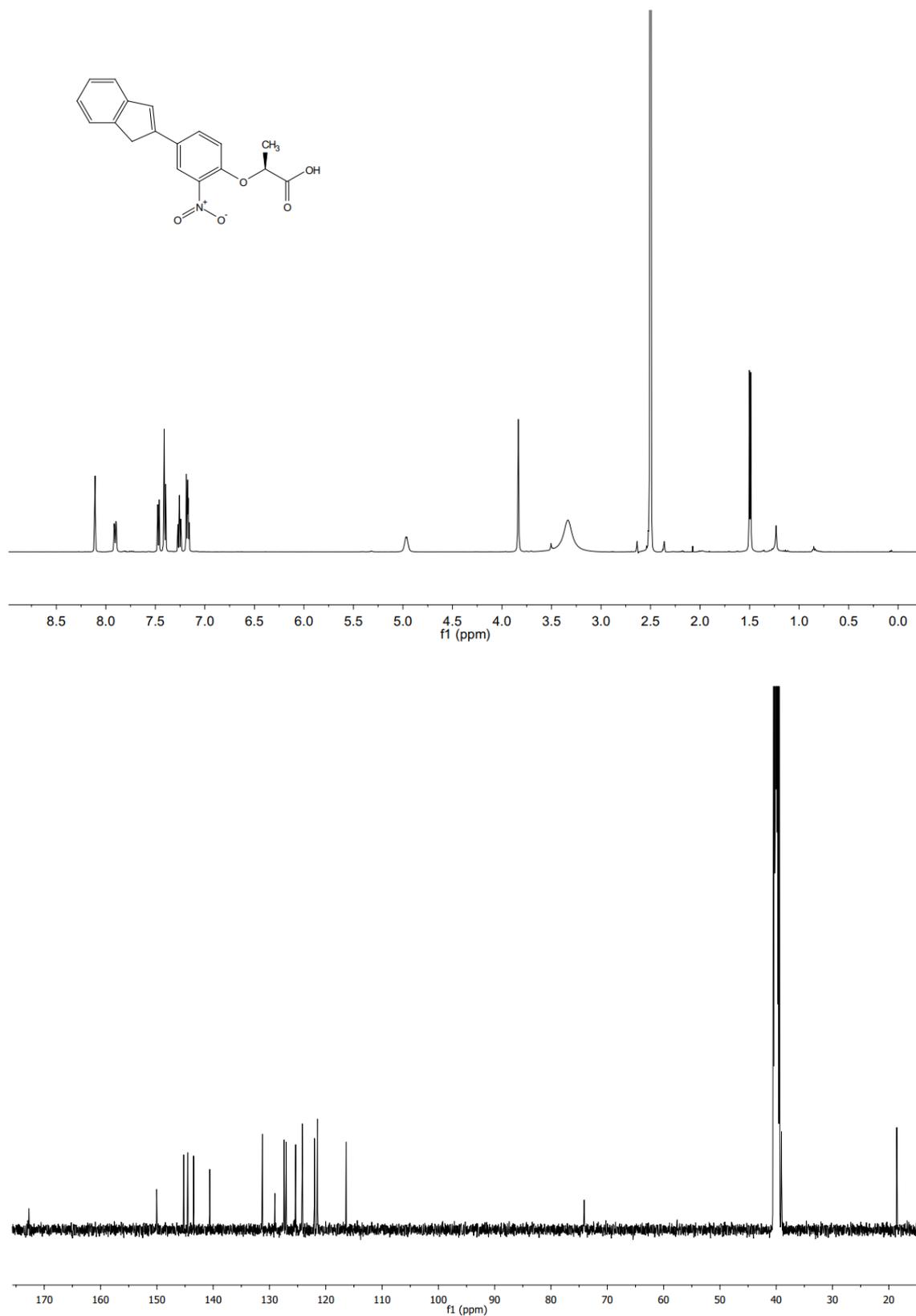


Figure S39. NMR spectra of **33**: ^1H NMR (400 MHz, $\text{DMSO}-d_6$) and ^{13}C NMR (101 MHz, $\text{DMSO}-d_6$).

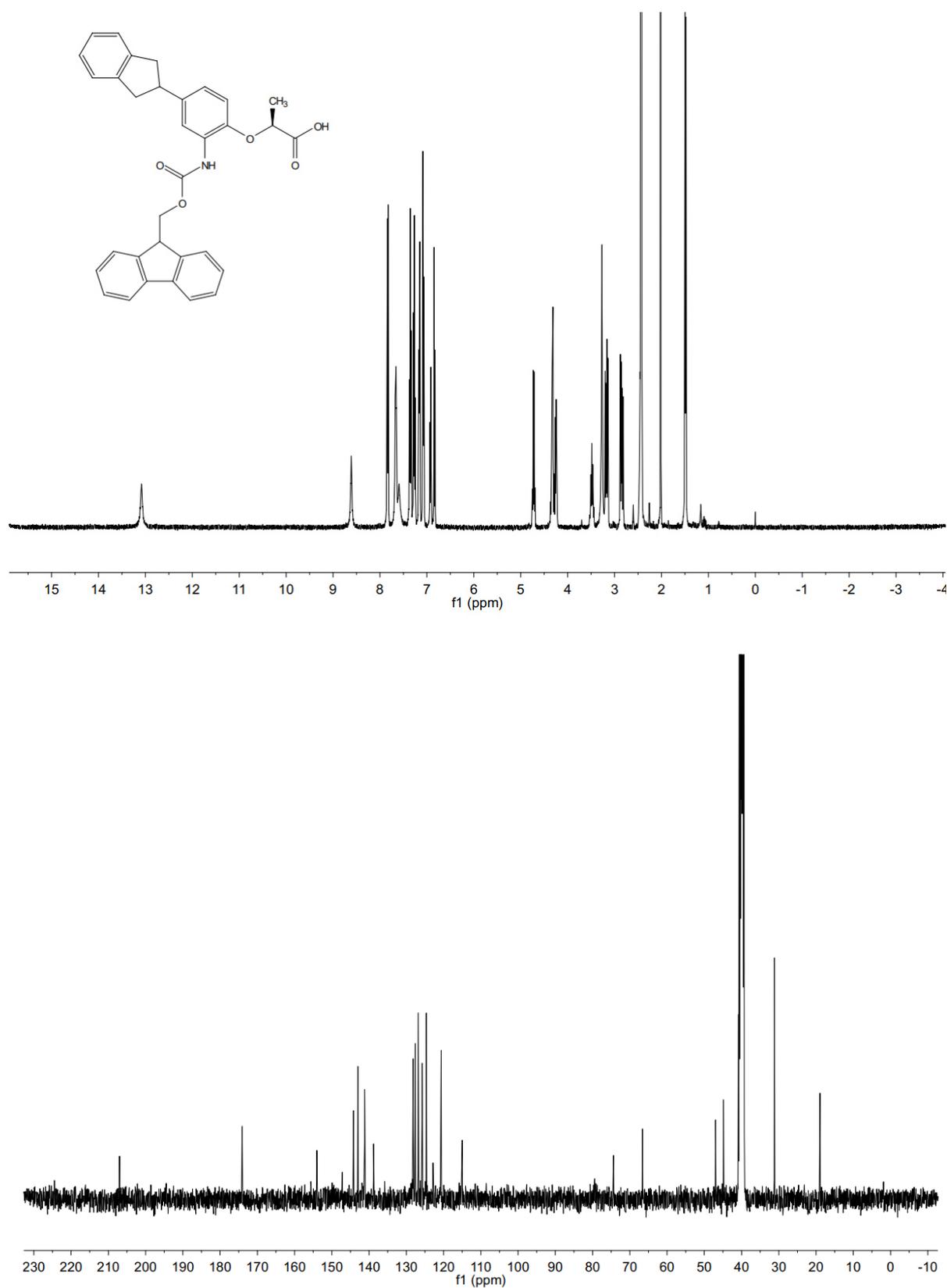


Figure S40. NMR spectra of **35**: ^1H NMR (400 MHz, $\text{DMSO-}d_6$) and ^{13}C NMR (101 MHz, $\text{DMSO-}d_6$).

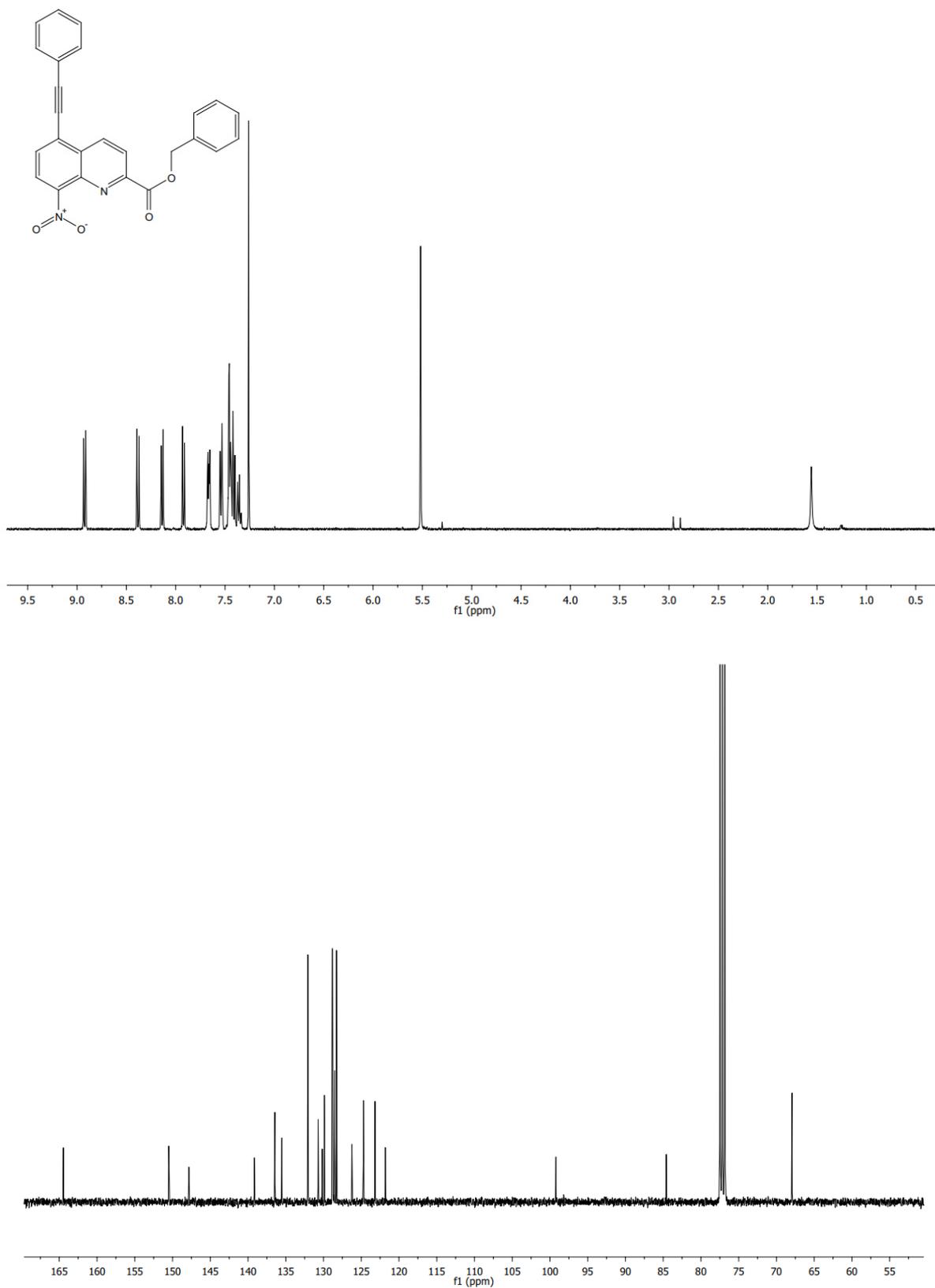


Figure S41. NMR spectra of **37**: ^1H NMR (400 MHz, CDCl_3) and ^{13}C NMR (101 MHz, CDCl_3).

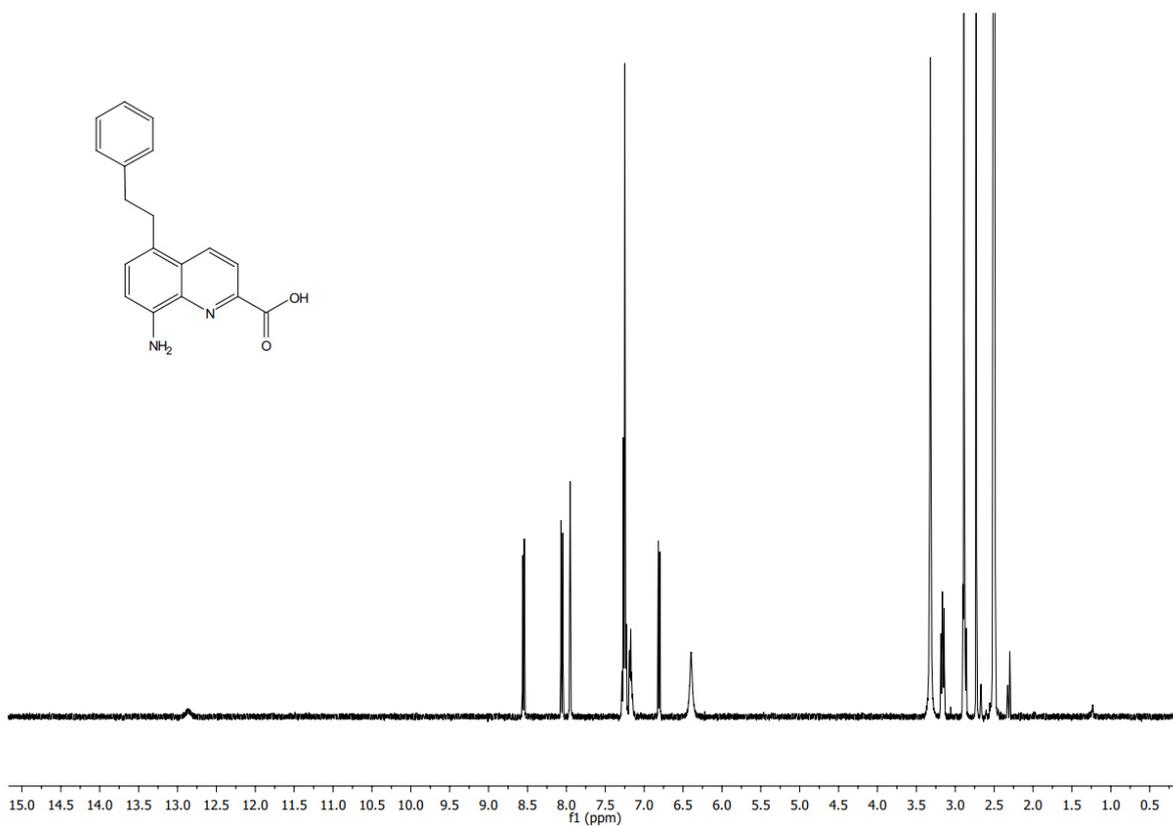


Figure S42. NMR spectra of **38**: ^1H NMR (400 MHz, $\text{DMSO}-d_6$).

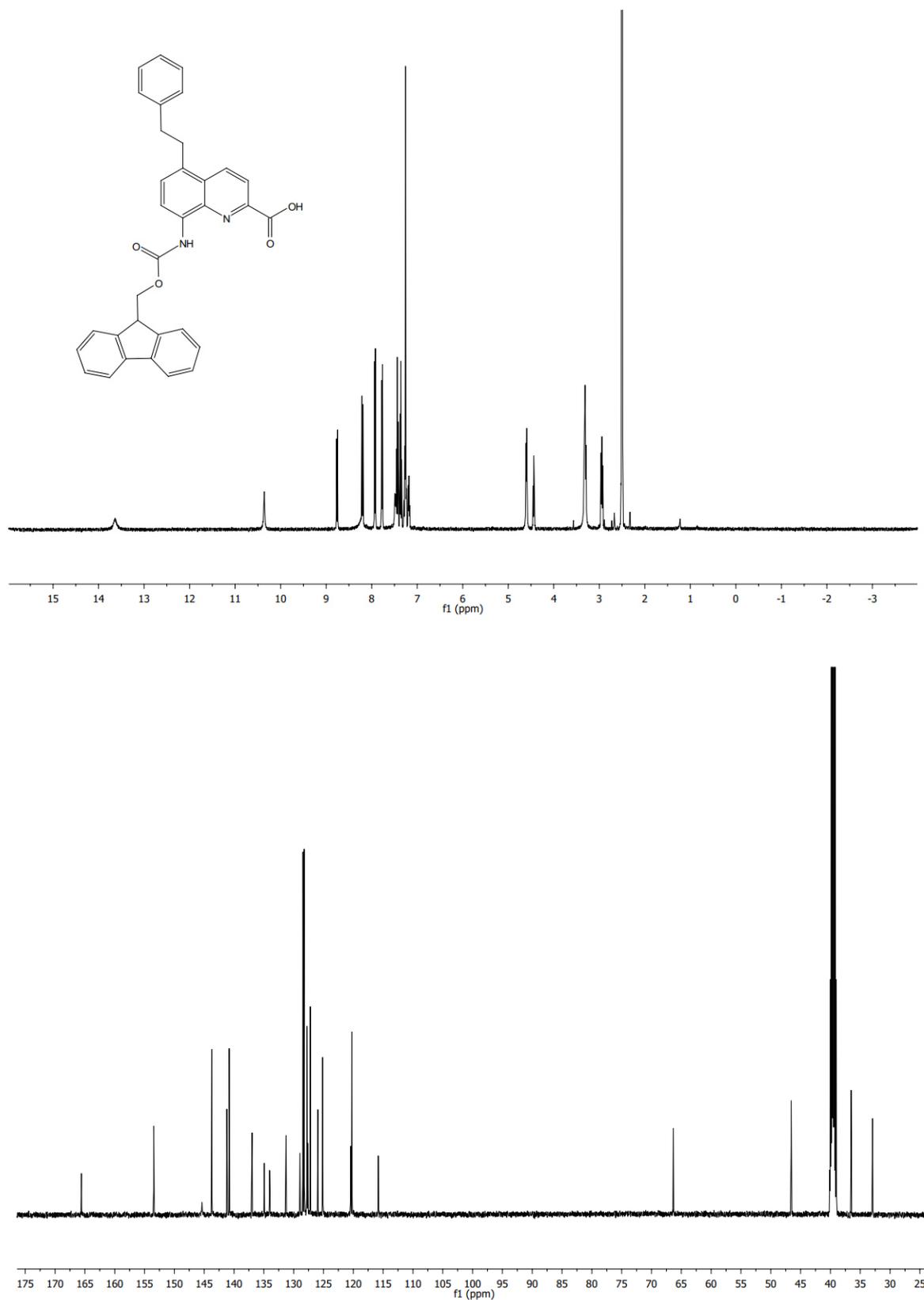


Figure S43. NMR spectra of **39**: ^1H NMR (400 MHz, $\text{DMSO}-d_6$) and ^{13}C NMR (101 MHz, $\text{DMSO}-d_6$).

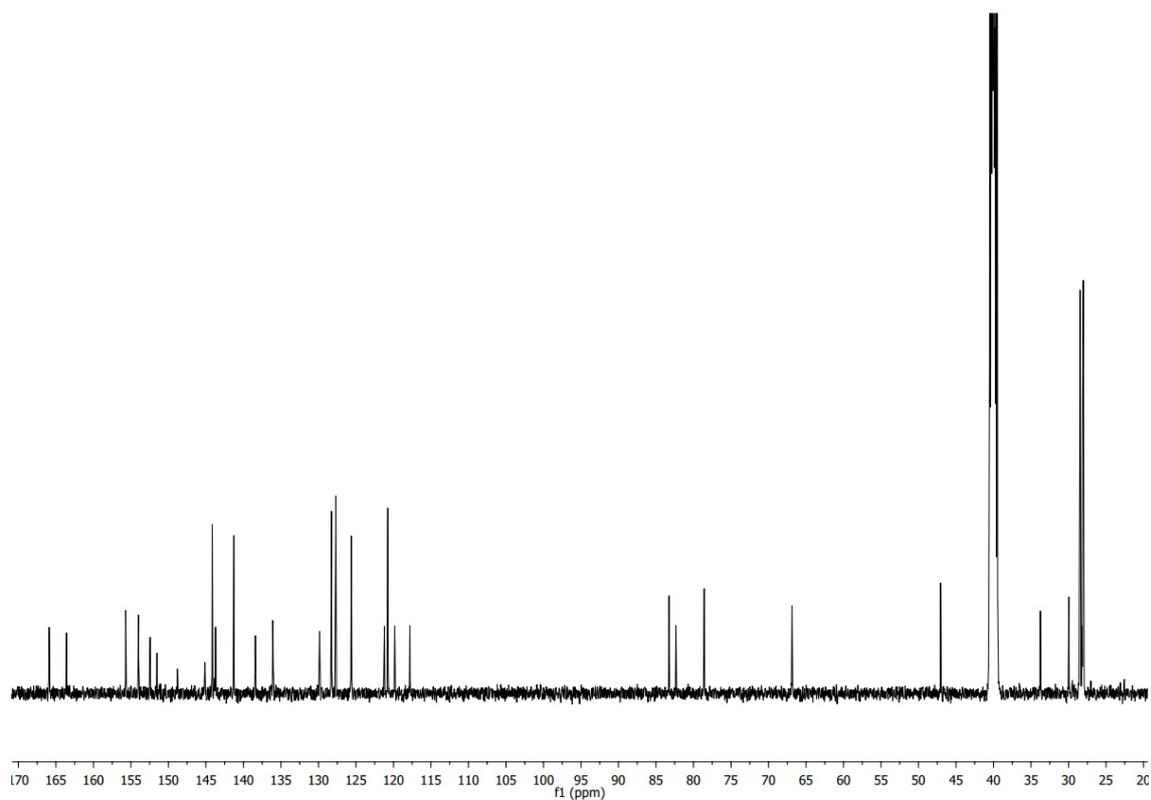
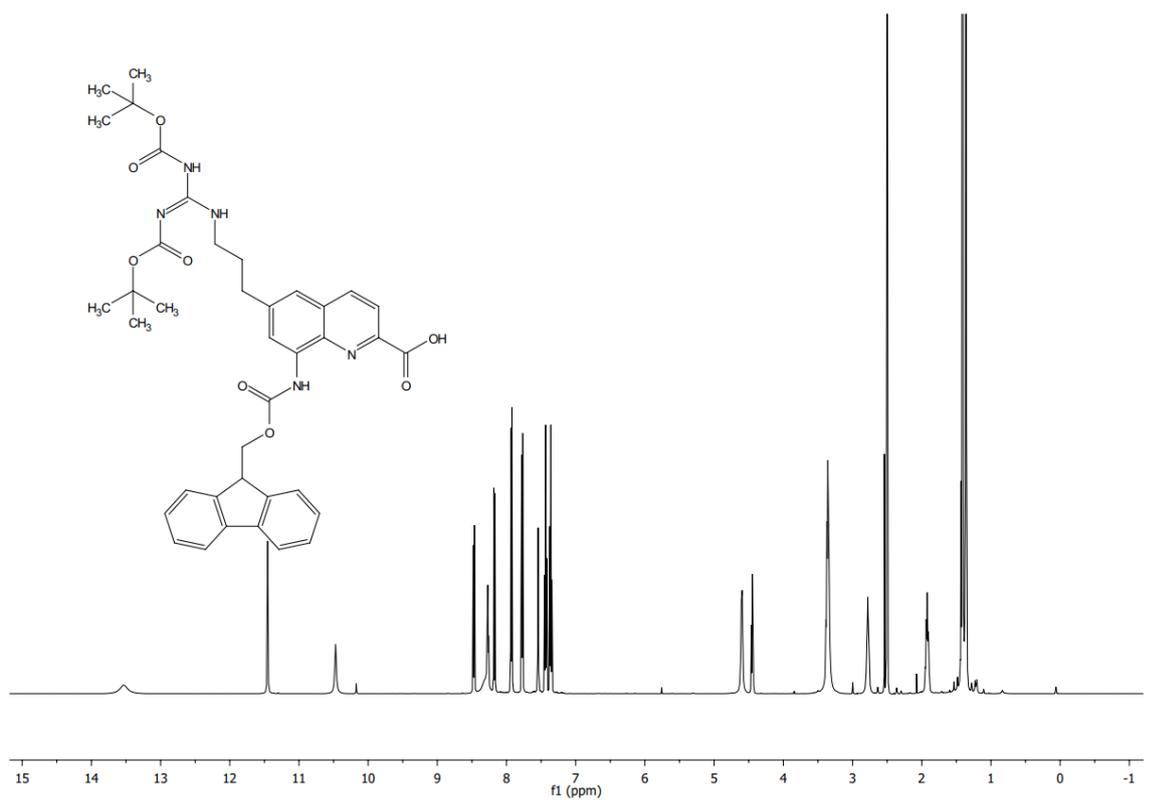


Figure S44. NMR spectra of **41**: ^1H NMR (400 MHz, $\text{DMSO-}d_6$) and ^{13}C NMR (101 MHz, $\text{DMSO-}d_6$).

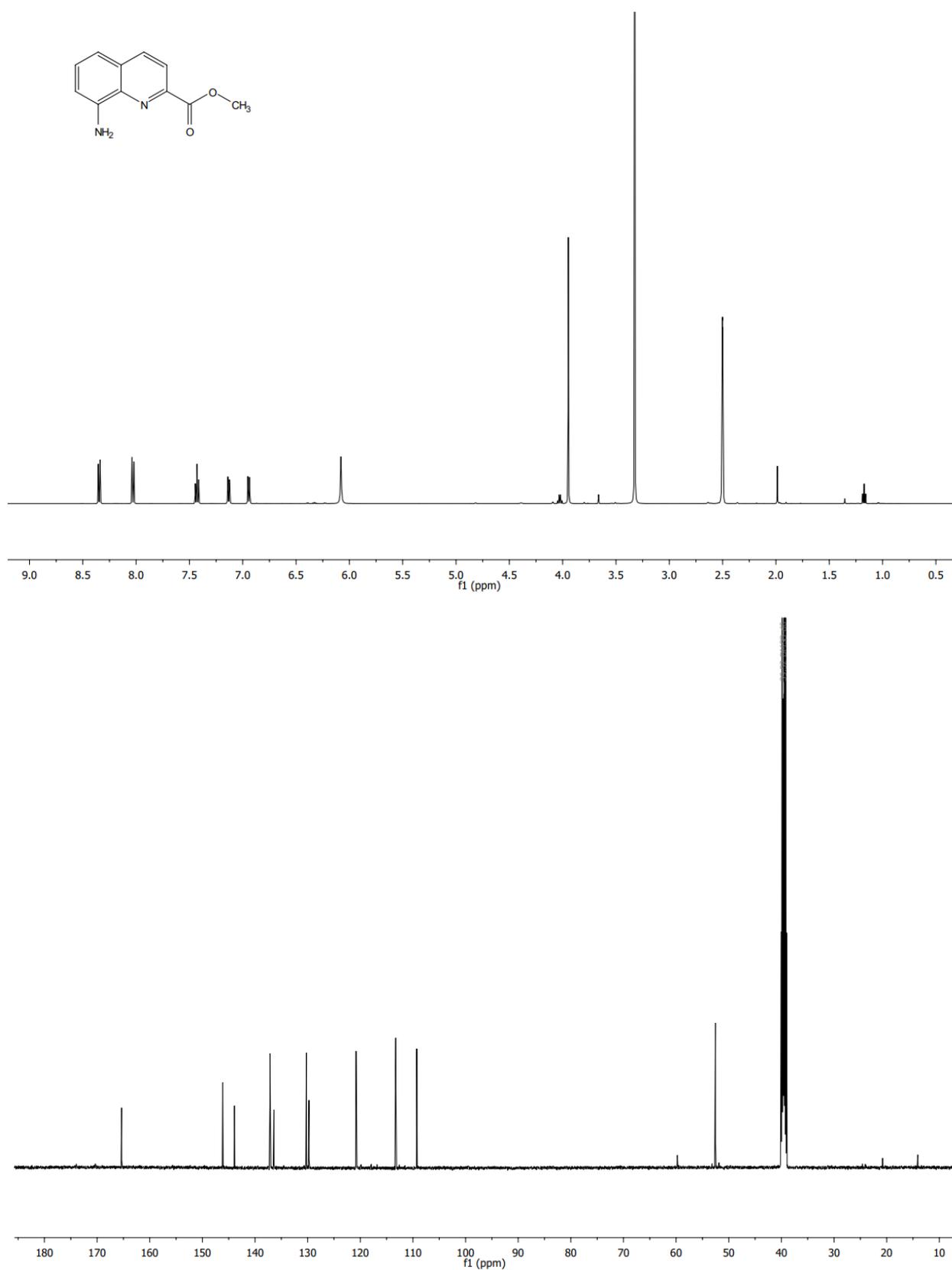


Figure S45. NMR spectra of **43**: ^1H NMR (500 MHz, $\text{DMSO-}d_6$) and ^{13}C NMR (126 MHz, $\text{DMSO-}d_6$).

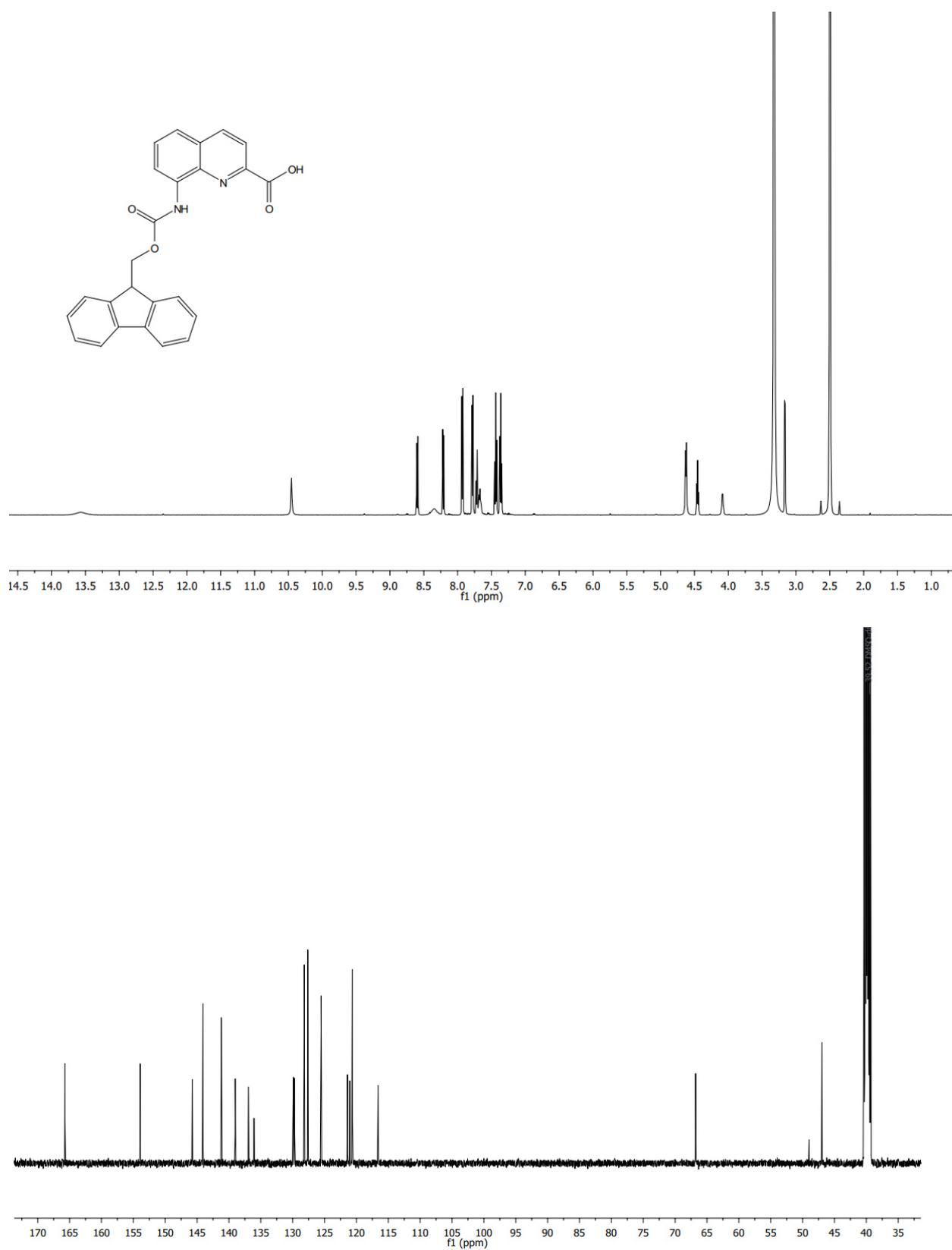


Figure S46. NMR spectra of **44**: ^1H NMR (500 MHz, $\text{DMSO}-d_6$) and ^{13}C NMR (101 MHz, $\text{DMSO}-d_6$)

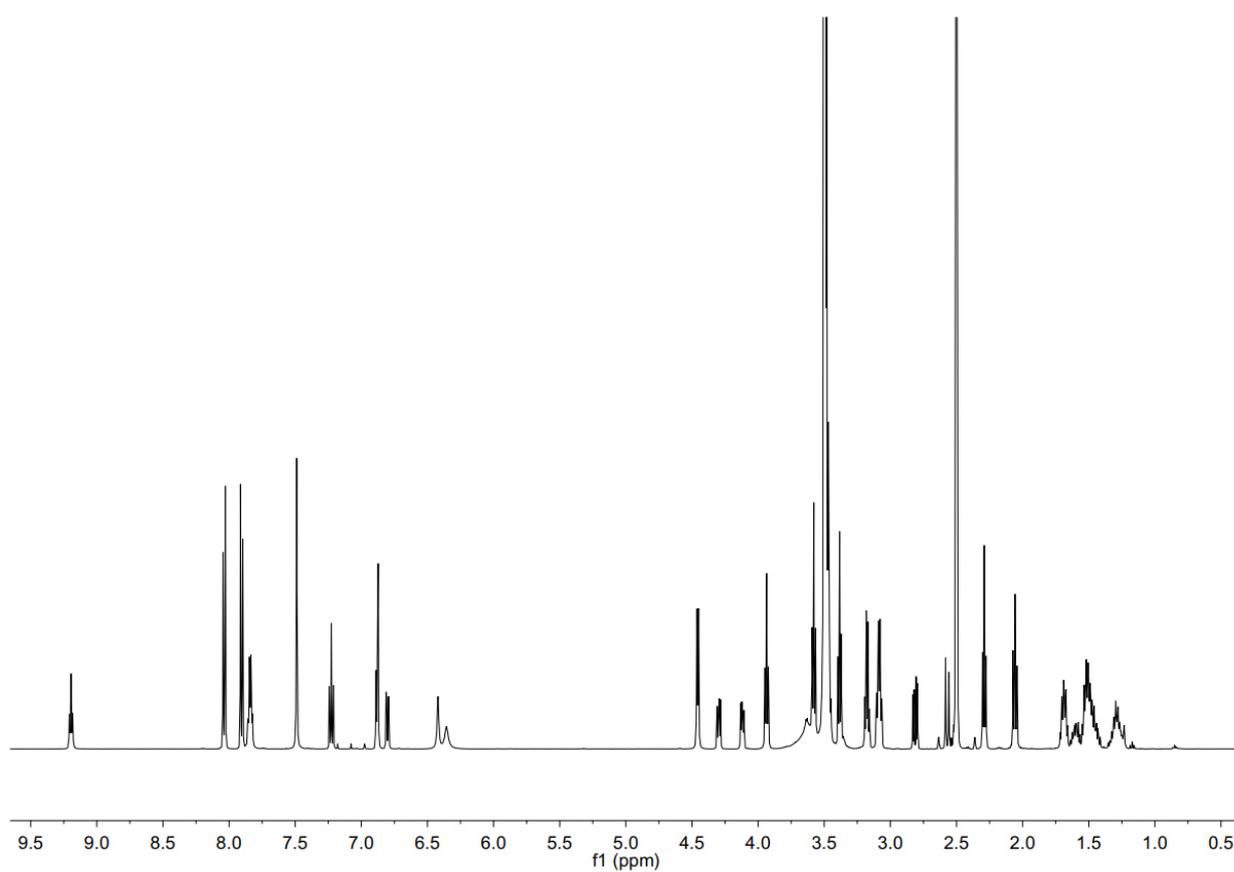
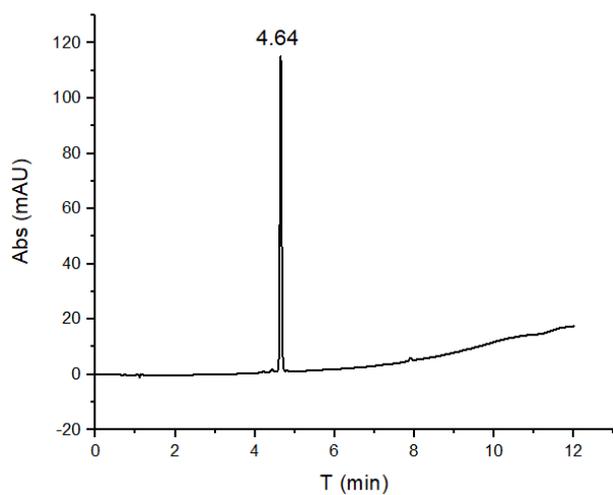


Figure S47. HPLC profile of purified compound **24** (10 to 100 B% over 10 min, 50 °C, $\lambda = 254$ nm) and ^1H NMR spectrum (500 MHz, $\text{DMSO-}d_6$, 25 °C)

10.6 References

1. J. Katigbak, H. Li, D. Rooklin and Y. Zhang, *J. Chem. Inf. Model.*, 2020, **60**, 1494.
2. O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 455.
3. D. A. A. Case, H.M.; Belfon, Kellon A.A.; Ben-Shalom, Ido Y.; Berryman, Joshua T.; Brozell, Scott R.; Cerutti, David S.; Cheatham, III, Thomas E.; Cisneros, G. Andrés; Cruzeiro, Vinícius Wiliam D.; Darden, Thomas A.; Duke, Robert E.; Giambasu, George Madalin; Gilson, Michael K.; Gohlke, Holger; Götz, Andreas W.; Harris, Robert C.; Izadi, Saeed; Измайлов, Сергей Александрович; Kasavajhala, Koushik; Kaymak, Mehmet Cagri ; King, Edward; Kovalenko, Andriy F.; Kurtzman, Tom; Lee, Tai-Sung; Le Grand, Scott; Li, Pengfei; Lin, Charles; Liu, Jian; Luchko, Tyler; Luo, Ray; Machado, Matias; Man, Viet; Manathunga, Madushanka; Merz, Kenneth; Miao, Yinglong; Михайловский, Олег Владимирович; Monard, Gérald; Nguyen, Hai; O'Hearn, Kurt A.; Onufriev, Alexey; Pan, Feng; Pantano, Sergio; Qi, Ruxi; Rahnamoun, Ali; Roe, Daniel R.; Roitberg, Adrian; Sagui, Celeste; Schott-Verdugo, Stephan; Shajan, Akhil; Shen, Jana; Simmerling, Carlos; Скрынников, Николай Русланович; Smith, Jamie; Swails, Jason; Walker, Ross C.; Wang, Junmei; Wang, Jinan; Wei, Haixin; Wolf, Romain M.; Wu, Xiongwu; Xiong, Yeyue; Xue, Yi; York, Darrin; Zhao, Shiji; Kollman, Peter A., *Amber 2022*, University of California, Sanfrancisco, 2022.
4. Z. Liu, A. M. Abramyan and V. Pophristic, *J. Chem.*, 2015, **39**, 3229.
5. F. Y. Dupradeau, A. Pigache, T. Zaffran, C. Savineau, R. Lelong, N. Grivel, D. Lelong, W. Rosanski and P. Cieplak, *Phys. Chem. Chem. Phys.*, 2010, **12**, 7821.
6. J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696.
7. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157.
8. M. Jewginski, L. Fischer, C. Colombo, I. Huc and C. D. Mackereth, *Chembiochem.*, 2016, **17**, 727.
9. W. Kabsch, *Acta Crystallogr D Biol Crystallogr*, 2010, **66**, 125; W. Kabsch, *Acta Crystallogr. D Biol. Crystallogr.*, 2010, **66**, 133.
10. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, *J. Appl. Crystallogr.*, 2007, **40**, 658.
11. P. S. Reddy, B. Langlois d'Estaintot, T. Granier, C. D. Mackereth, L. Fischer and I. Huc, *Chem. – Eur. J.*, 2019, **25**, 11042
12. L. A. Woods, O. Dolezal, B. Ren, J. H. Ryan, T. S. Peat and S.-A. Poulsen, *J. Med. Chem.*, 2016, **59**, 2192.
13. G. N. Murshudov, A.A. Vagin and E.J. Dodson, *Acta Crystallogr. D Biol. Crystallogr.*, 1997, **D53**, 240.
14. P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart, *Acta Crystallogr. D Biol. Crystallogr.*, 2010, **66**, 213.
15. P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr.*, 2010, **66**, 486.
16. A. W. Schuttelkopf and D. M. van Aalten, *Acta Crystallogr D Biol Crystallogr*, 2004, **60**, 1355; N. W. Moriarty, R. W. Grosse-Kunstleve and P. D. Adams, *Acta Crystallogr. D Biol. Crystallogr.*, 2009, **65**, 1074.
17. I. W. Davis, A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. Arendall, 3rd, J. Snoeyink, J. S. Richardson and D. C. Richardson, *Nucleic. Acids. Res.*, 2007, **35**, W375.

18. P. Koutnik, E. G. Shcherbakova, S. Gozem, M. G. Caglayan, T. Minami and P. Anzenbacher, *Chem.*, 2017, **2**, 271.
19. P. Gans, A. Sabatini and A. Vacca, *Talanta*, 1996, 43, 1739; P. Gans, A. Sabatini and A. Vacca, *Anal. Chim.*, 1999, **89**, 45.
20. S. Dengler, P. K. Mandal, L. Allmendinger, C. Douat and I. Huc, *Chem. Sci.*, 2021, **12**, 11004.
21. X. Hu, S. J. Dawson, P. K. Mandal, X. de Hatten, B. Baptiste and I. Huc, *Chem. Sci.*, 2017, **8**, 3741.
22. M. Vallade, P. Sai Reddy, L. Fischer and I. Huc, *European Journal of Organic Chemistry*, 2018, 2018, 5489.
23. V. Corvaglia, F. Sanchez, F. S. Menke, C. Douat and I. Huc, *Chemistry*, 2023, 29, e202300898.
24. J. Buratto, C. Colombo, M. Stupfel, S. J. Dawson, C. Dolain, B. Langlois d'Estaintot, L. Fischer, T. Granier, M. Laguerre, B. Gallois and I. Huc, *Angew. Chem., Int. Ed.*, 2014, **53**, 883.
25. N. G. Léonard, W. N. Palmer, M. R. Friedfeld, M. J. Bezdek and P. J. Chirik, *ACS Catal.*, 2019, **9**, 9034.
26. M. Zwillinger, P. S. Reddy, B. Wicher, P. K. Mandal, M. Csékei, L. Fischer, A. Kotschy and I. Huc, *Chem. – Eur. J.*, 2020, **26**, 17366.

11. Summary and Perspective

11.1 Summary for published/submitted work

This work showed the immense potential of aromatic oligoamide foldamers to interact with biological molecules. The crystal structure of foldamer first illustrated that the side chain arrangement formed several faces which can possibly recognize the surface of protein. By pull-down assay, a twelve consecutive quinoline units foldamer was tested against all the proteins in cell lysate and found to possess nanomolar affinity with 3 proteins. The BLI test showed that both *P*- and *M*- helix had high binding affinity of Rad 52 protein, which is responsible for DNA double-strand break repair and homologous recombination. The similar binding affinity of two enantiomers might indicate that the interaction was not selective, but these new results suggested that with a confined structure as foldamer showed the potential as competitors in affecting PPIs and other pharmacological applications.

Based on the positive results from the first work, we endeavour to explore the potential interaction between peptide binders and our foldamer sequences. We designed the sequences to form two faces by arranging the position of side chains: one face was supposed to point to the solvent and the other face to point to the peptide binders. To increase the water solubility and further facilitate the binding test, seven quinoline monomers with tetraethylglycol side chains were used in 12mer-foldamer (position 2,4,5,7,9,10,12, from N terminal to C terminal), forming one interface and the other interface was decorated with various proteinogenic side chains (position 1,3,6,8,11). The energy-minimized structure showed the two faces as we predicted. Two foldamer sequences of 12 quinoline units coupled by biotin moiety were first synthesized. One sequence was served as counter-control sequence, in which the position arrangement of proteinogenic side chains was different from the other. After the selection process, several linear and peptide macrocycles were chosen and synthesized to facilitate binding measurements, obtaining micromolar range affinity by SPR. One monomer was observed not to be stable, but not dramatically influencing the selection process. Two monomers carrying similar residues were synthesized as replacement with higher stability. We also found that the interaction between foldamer and peptide was diastereoselective which means that peptide selected from the RaPID system were able to discriminate one helix handedness over the other.

In HCAII foldamer-HCAII interface design project, several monomers carrying proteinogenic side chains have been designed and synthesized based on the previous foldamer-protein complex structure. Since our main focus was to confirm whether the new side chains could bring a cumulative effect upon decorating helix with more side chains, so we planned to preserve overall the helical shape of foldamer compared to the original sequence as much as possible. First generation of foldamer-protein complex was regarded as a starting point and backbone constitution should be kept similar during the structure iteration. Moreover, the previous crystal structure showed foldamer-foldamer contact, which might

facilitate the crystal packing. The first proposed phenylalanine side chain was inserted to the pocket on protein surface as designed. Based on this success, many other monomers with proteinogenic side chains were synthesized, *i.e.* guanidinium side chains, indane side chains. We found that decorating side chains on the surface of helical foldamer does not change the overall shape of helix. The location where foldamer binds to the protein seems to be fixed and the proposed side chains positioned themselves as computer modelling predicted. The binding affinity of foldamer sequences did not show dramatic change upon stepwise adding of proteinogenic side chains, neither increase nor decrease. Upon binding test by BLI, we found that the dissociation of foldamer-protein complexes was generally slow as well as the binding procedure, which blocked the possibility of binding affinity measurement by SPR. This might be also a hint that the side chains of foldamer or the foldamer backbone slowed down the process where ligand dissociated from the protein. The insertion of chiral **B** monomer gave quantitative handedness bias and save the time for incubation of protein-foldamer complex before we went crystal growth.

11.2 Challenges for protein surface recognition and future designs

Although many successful cases have been made to target protein surface, designing protein-binders are still challenging specifically when there is no pre-existing binding ligand. Most of studies are stemming from the improvement of existed ligand-protein complex and undergo rounds of structural modifications. The recognition between foldamer and protein is mostly determined by side chains, since the side chains are residues pointed out of the helix, the backbone functionalized as support where side chains located. The quinoline-based monomers have been developed in our group for many years, yet the diversity of side chains is still limited because of the synthetical difficulty and problem of scaling up. The stability of monomers could not precisely be predicted since the monomer would be brought to an environment different in pure organic or inorganic solvent. Unpredictable foldamer-foldamer interaction in the solvent and the stacking of backbones might also have effects to the stability of side chains of monomers. The computer simulation might propose hundreds or thousands of side chain proposals in one week, but the success of one single monomer synthesis might take one to two months. The introduction of monomers also needs to take nucleophilicity of amine function into consideration since the yield of amide condensation should not be too low and preventing the sequence elongation by solid phase synthesis. X-ray is the main method to we use to illustrate the structural information of protein-foldamer complex. Therefore, we also need to consider the effect of crystal packing, solubility issues when we design the sequences in the first place.

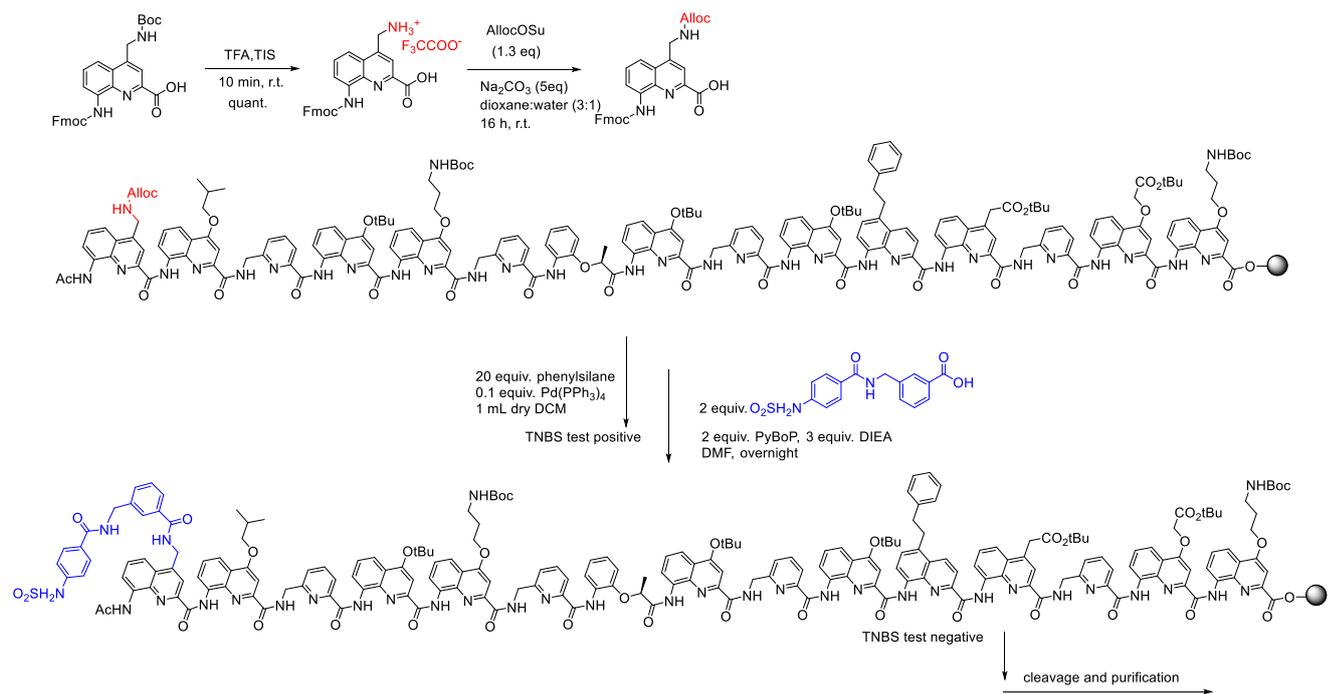
Quantify the binding affinity between foldamer and target plays also an important role in the projects. Indeed, many new technologies have been developed in the past few decades like BLI and SPR, understanding the binding mode and mathematical process provided by the machine is also important. During the structural modification of a candidate substance, the binding mode between the candidate

substance and the protein may change. Developing a reliable method for binding affinity measurement is essential to prove that the change of structure really brings the improvement in binding affinity. We replace SPR binding measurement by fluorescence competition test because new generation of foldamers possess a low disassociation process. The competition test by fluorescence spectroscopy is an indirect way to measure the binding constant. Although the K_D we obtained for one reference compounds in the fluorescence experiments was similar to previously obtained with the SPR method, we still cannot conclude the consistency of two methods. We also observed the difference between fluorescence assay and BLI, since the mechanism of these two techniques are not quite same, we could also not give a strong conclusion, which makes it hard to evaluate the effect of introduction of new side chains.

The role of HCA-II ligand, a derivate of a benzene sulfonamide, is important at the start of the project. The high binding affinity of sulfonamide moiety enables that the conjugated foldamers could reach the surface of protein. However, the binding affinity of ligand is already in nanomolar range, the binding affinity change brought by side chains alteration would not be easy to observe. Possibilities exist that the new side chains bring micromolar binding affinity improvement, but this level of increase or decrease is already below the error range of the techniques which measured the K_D . The ranking score of proposal of side chains focus mostly on cavity complementary and although molecular dynamics proves the overall stability of protein-foldamer complex, it does not guarantee the tight binding between side chain residues and protein surface.

Future direction of removing the ligand from the foldamer as well as keeping tight binding might depend on further elongation of foldamer sequence. The ligand is located on the N terminal of foldamer, which means that the elongation of foldamers can only follow the direction along the C terminal, blocking the possibility of exploring the protein surface on other direction. Relocating the ligand to the side chain of foldamer is one way to expand the length of foldamer on both C- and N terminals. New monomers with alloc protecting group could be synthesized based on previous reported Q monomer and the coupling of the ligand could be performed on the solid support. The general synthetical scheme is showed in **Figure 12a**. In the modelling, we could see that the overall shape of helix is still conserved. When we further expand the N terminals of a 15mer, the helix could cover a full length of protein when nine extra quinoline units were added. The energy minimized model of 24mer helical foldamer with HCA was showed in **Figure 12**. The ligand was shifted to the side chain of 9th quinoline unit (starting from N terminal). The foldamer possesses larger interface compared to the previous 15mer, which opens more chance for future design.

a)



b)

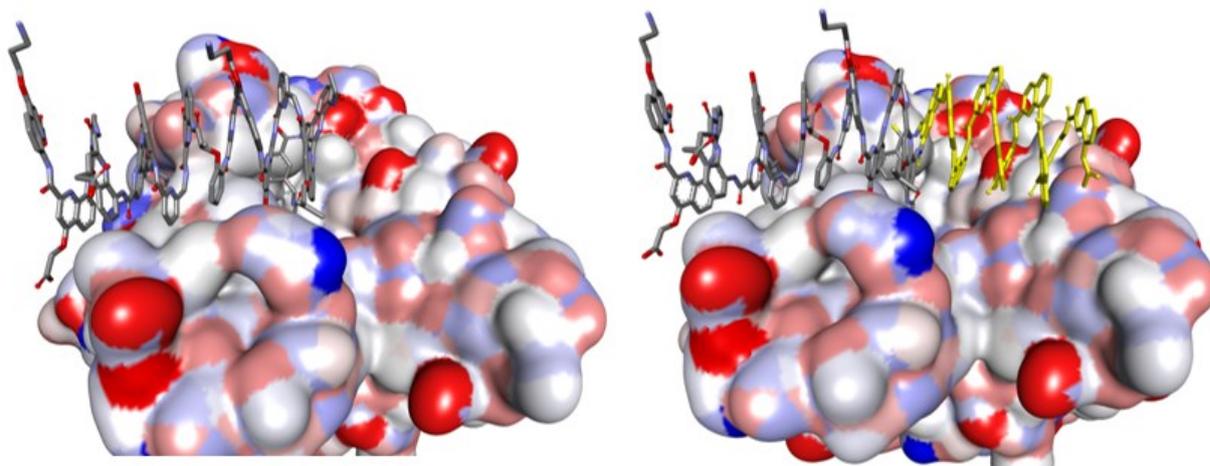


Figure 12: a) synthetical route of new Q monomer and general synthetical scheme of ligand coupling on side chain. b) energy minimized models of the 15mer (structure showed in figure 12a)-HCAII complex and model of 24mer-HCAII. The 24mer is elongated based on 15mer with 9 more quinoline units without side chains, labelled in yellow colour.

12. Acknowledgement

First of all, I am deeply grateful to Prof. Dr. Ivan Huc for giving me the opportunity to embark on the journey of foldamer research, starting with my master's thesis and continuing with my PhD work in his group.

Secondly, my deepest thanks go to my co-supervisor, Dr. Céline Douat-Casassus, for her patient guidance—from bench work to thesis writing—and for her constant encouragement.

I would also like to thank Dr. Pedro Mateus for his help in analyzing the data from fluorescence competition tests and for guiding me in using Hypspec software.

My gratitude extends to Johannes Sigl, Dr. Tulika Chakraborty, and Dr. Vasily Morozov for their guidance in protein expression, crystal structure elucidation, and the use of SEC.

Thanks to Dr. Wei Tuo, Jiaojiao Wu, and Dr. Shuhe Wang for the enjoyable Chinese dinners we shared.

Many thanks to Florian Sanchez, Cyrille Riviere, Dr. Mathieu Denis, Dr. Valentina Corvaglia, Dr. Siyuan Wang, and Dr. Alexandru Grozavu for their valuable discussions on my research project.

I am grateful to Tim Arczynski for joining me in the HCA project and for our valuable discussions about Chinese restaurants.

Special thanks to Dr. Lucile Fischer and Dr. Yann Ferrand for hosting me during my time in Bordeaux.

I am also thankful to Dr. David Konrad, Dr. Ryan Howard, and Dr. Mahshid Alizadeh for their support in my job search.

I appreciate Melis Cabbar, Manuel Loos, and Niklas Böcher for teaching me German slang.

Furthermore, I must express my gratitude to all the other group members for providing such a pleasant and supportive working environment.

I am deeply thankful to my family and friends for always backing me up during times of difficulty and frustration, especially to Yuqing Zheng, who continuously motivated and supported me through challenging times.

It has already been more than eight years since the day I first arrived in Düsseldorf Germany, and ten years since I began preparing to go to Germany for my master's studies. Looking back on the road I have travelled, it was challenging and filled with struggles, but every step was worth it. These eight years have been among the most valuable experiences of my life.

Finally, I also want to thank my younger self—ten years ago, a naive bachelor student, dreaming of being admitted to a top university in Germany. At that time, deciding to start learning a new language

in parallel of bachelor study was not an easy one. If I had the chance to speak to him, I would thank him for his courage in chasing those dreams. I would tell him: "This journey will be full of darkness, uncertainties, and troubles. You will have a hard time to find a room to live. You might not pass your German language course very fast. Receiving decline from university is frustrating. You might be the last one in your class to pass an exam where other classmates can pass with one time in good score. You are not talented in research so you need to work harder than others, Saturday and Sunday might not be always your holidays. But you have the power to get over it. Do not give it up. It will still be a wonderful journey, leading you to a bright and promising future."