EPFL

Thèse n° 9095

Methods for the generation of large combinatorial macrocycle libraries

Présentée le 29 octobre 2021

Faculté des sciences de base
Laboratoire de protéines et peptides thérapeutiques
Programme doctoral en chimie et génie chimique

pour l'obtention du grade de Docteur ès Sciences

par

Sevan HABESHIAN

Acceptée sur proposition du jury

Prof. Y. Aye, présidente du jury
Prof. C. Heinis, directeur de thèse
Prof. J. Kihlberg, rapporteur
Prof. H. Waldmann, rapporteur
Prof. J. Waser, rapporteur

École polytechnique fédérale de Lausanne

2021

Acknowledgements

First, I would like to thank Prof. Christian Heinis for allowing me to work with you on these exciting projects. You have given me a great opportunity, and I hope that my work continues to be an important contribution to the lab and to the field in the future. Your guidance and advice have made me into a better scientist!

I would also like to thank the members of my thesis committee, Professor Herbert Waldmann, Professor Jan Kihlberg, and Professor Jérôme Waser, for taking the time to evaluate my work from the past four years. Also thank you to Professor Yimon Aye for serving as jury president.

A huge thank you goes to Béatrice Bliesener-Tong for all of the help over the years, and for all of the work you do for our lab. You always make sure that we are comfortable and happy, and we all greatly appreciate it!

To all former (Yuteng, Ale, Sangram, Ganesh M, Ganesh S, Xudong, Vanessa, Patrick, Carl, Jun) and current (Mischa, Gontran, Manuel, Zsolt, Ed, Cristina, Xinjian, Alexander) members of LPPT during my time here - thank you for making me feel welcome, for your kindness and generosity, and for all the help and advice you have all given. You all made the lab a fun place to be every day. Also, all of the non-LPPT people - Tamara, Nora, Chloé, Eduard – you guys definitely made PhD life more fun and exciting!

I am grateful to everyone who helped contribute to my project. From LPPT: Manuel, Mischa, Ganesh S, Gontran, Ganesh M - from the physical work you did to the brainstorming sessions and advice. An extra thank you to Mischa for being the master of instruments, and for the creative equipment you made to make our lives easier. To my students, Adrian and Ilaria, thank you for all of your hard work, and for putting up with me on a daily basis (not always easy!). From the BSF, Jonathan and Julien, thank you for your contribution to my experiments, and for the advice about HTS. You were both always very kind and helpful! Finally, thank you to our external collaborators in Italy, Prof. Alessandro Angelini and Prof. Laura Cendron.

Thank you to my family for all of the support you have given me

Last but not least, thank you Hannah for your unconditional love and support. You have always believed in me, encouraged me, and made me feel confident. I could not have done this without you.

Abstract

Macrocycles are an attractive class of molecules due to their good binding properties and relatively small size that allows, in many cases, crossing membranes to reach intracellular targets. In comparison to classical small molecule drugs, macrocycles are generally better at binding to challenging targets such as proteins having flat, featureless surfaces, or protein-protein interactions. However, the development of macrocycle-based ligands to new targets has been challenging due to the lack of sufficiently large libraries of macrocyclic compounds that could be screened. The preparation of large numbers of macrocycles is limited due to the chromatographic purification that is typically required following low-yielding cyclization reactions. As a result, the full potential of macrocycles cannot be exploited in drug discovery efforts.

The goal of my thesis project was to develop new methods for the efficient generation of large libraries comprising tens of thousands of macrocycles. Importantly, the methods had to omit throughput-limiting purification steps, and they had to yield macrocycles of high purity. Towards this end, I sought to establish a new macrocycle library synthesis principle in which a large panel of *m* small cyclic peptides are combinatorially diversified with a large number of *n* chemical fragments, and the resulting $m \times n$ products screened in microwell plates without purification. For generating the large libraries efficiently and economically, I planned to perform the reactions and screens at a picomole scale and using contactless liquid transfer.

In a first project, I developed a method for efficiently producing hundreds of cyclic peptides needed for the described combinatorial approach. I aimed to develop the following method in which pure cyclic peptides are obtained directly from solid phase. Peptides are synthesized via a disulfide linker on solid phase and the side chains of the peptide deprotected while the peptide remains disulfide-linked to the resin. A thiol group, which was included at the N-terminus of the peptides during the synthesis, is then deprotonated by treatment with base which results in an intramolecular disulfide exchange with the disulfide linker, to release peptides cyclized by a disulfide bond. The method could be established, was tolerant of diverse peptide sequences, and released all tested compounds in high purity. With this method, I was able to synthesize cyclic peptides in 96-well plates, and thus hundreds of peptides in parallel with minimal effort. I screened a test library and identified a weak inhibitor of thrombin (K_i = 13 ± 1 μM), which validated the efficient synthesis method. Most importantly, the method allowed me to access large numbers of cyclic peptides that I required in the second project, as described in the following.

In the second project, I intended to establish the approach described above, in which large numbers of cyclic peptides are combinatorially diversified with chemical fragments, all at a picomole scale and by transferring reagents in nanoliter volumes. As the chemical ligation reaction, I chose to acylate amines built into the cyclic peptides, with carboxylic acid building blocks. A proof-of-concept experiment with macrocycles bearing peripheral amines showed high conversion to products. As a next test, I diversified a 384-member cyclic peptide library with 10 carboxylic acids in nanoliter volumes to give 3,840 macrocyclic compounds, which I screened against thrombin. A potent thrombin inhibitor was identified (K_i = 44 ± 1 nM), and a co-crystal structure was obtained which confirmed binding to the active site of the protease. In a final application, I generated and screened a library of 19,968 macrocycles (192 cyclic peptides × 104 carboxylic acids) against MDM2, an important oncology target forming a protein-protein interaction with p53. The screen identified a high nanomolar binder (K_D = 390 ± 40 nM) that I optimized to a high-affinity binder (K_D = 31 ± 6 nM) in two rounds of iterative library synthesis and screening.

The methods developed in the two projects are currently applied by several members of our laboratory, and for the development of macrocycle-based ligands to diverse disease targets. The future will show if the new methods can yield leads for drug development and for addressing currently unsolved medical challenges.

Key words: macrocycles, peptides, drug discovery, cyclative release, combinatorial synthesis, protein-protein interactions, disulfide bonds, therapeutics, high-throughput screening, hit optimization, nanoscale synthesis, acoustic droplet ejection, thrombin, MDM2

Zusammenfassung

Makrozyklen sind wegen ihren guten Bindungseigenschaften und der relative kleinen Grösse, welche oft den Zugang zu intrazellulären Protein erlaubt, eine interessante Molekülklasse zur Enwicklung von Medikamenten. Im Vergleich zu kleinen Molekülen können Makrozyklen generell besser an schwierige Zielproteine, wie zum Beispiel Proteine mit flachen Oberflächen oder an Protein-Protein Interaktionen, binden. Ein Problem in der Entwicklung von Makrozyklen für neue Zielproteine ist der Mangel an ausreichend grossen Bibliotheken von Markozyklen, welche getested werden könnten. Die Herstellung einer grossen Anzahl von Makrozyklen ist schwierig weil meist eine chromatographische Reinigung nach der Synthese benötigt wird, welche aufwändig ist und den Durchsatz einschränkt.

Das Ziel meiner Doktorarbeit war es, eine Methode zu entwickeln, welche schnell und einfach grössere Makrozyklenblibliotheken generieren kann. Dabei sollte die Methode auf durchsatz-limitierende Reinigungsschritte verzichten und Makrozyklen in hoher Reinheit herstellen können. Um dieses Ziel zu erreichen hatte ich die folgende Strategie ins Auge gefasst. Eine grosse Anzahl von *m* zyklischen Peptiden werden kombinatorisch mit vielen, *n* Fragmenten diversifiziert um eine Bibliothek von $m \times n$ Makrozyklen herzustellen. Die Reaktion sollte Produkte mit hoher Reinheit liefern damit keine Reinigung notwendig ist. Um solche Bibliotheken effiziet und kostenarm herzustellen hatte ich beabsitigt die Makrozyklen in äusserts kleinen, picomolaren Mengen, herzustellen und die Reagenzien mittles akustischen Wellen zu transferieren.

Im ersten Projekt hatte ich eine Methode für die effiziente Herstellung von zyklischen Peptiden entwickelt, welche ich für den gerade beschriebenen Ansatz benötigte. Ich hatte die folgende Strategie angestrebt welche reine zyklische Peptide direkt von der Festphase liefern sollte. Peptide werden via einem Disulfid-Linker auf Harz synthetisiert und die Aminosäure Seitenketten werden entfernt während die Peptide immobilisiert bleiben. Eine Thiolgruppe, welche am N-terminus von den Peptiden eingeführt wurde, wird dann mittels Zugabe einer Base deprotoniert was zu einer Disufidaustauschreaktion mit dem Disulfid-Linker führt, wobei das Peptid zyklisiert und vom Harz gelöst wird. Die Methode hatte sich als effizient herausgestellt, funktionierte mit verschiedenen Peptidsequenzen, und hatte reine Produkte geliefert. Die Methode war kompatibel mit der Synthese in 96-Mikrotiterplatten und hatte mir erlaubt, viele Peptide mit relativ kleinem Aufwand zu synthetisieren. Aus einer Testbibliothek konnte ich einen schwachen Inhibitor von Thrombin identifizieren (K_i = 13 ± 1 μM). Am wichtigsten war dass diese neue Methode es mir erlaubte, eine grosse Anzahl an

verschiedenen zyklichen Peptiden, welche ich für den anfangs beschriebenen kombinatorischen Ansatz gebrauchte, herzustellen.

Das Ziel vom zweiten Projekt war den kombinatorischen Ansatz zur Herstellung von grossen Makrozyklen Bibliotheken zu entwickeln. Als chemische Reaktion wählte ich die Azylierung von Aminogruppen in den zyklischen Peptiden mit Carbonsäure-Bausteinen. Ein erstes Experiment zeigte dass gute Ausbeuten mit wenig Nebenprodukten erzielt werden konnten. In einem weiteren Experiment wurde eine 384 Disulfidpeptidbibliothek mit zehn Carbonsäuren im Nanolitermassstab zu total 3'840 Makrozyklen diversifiziert und im biochemischen Thrombinassay getestet. Ein Thrombin Inhibitor wurde identifiziert (K_i = 44 ± 1 nM) und die Bindung des Inhibitors zum aktiven Zentrum der Protease wurde mittels einer co-Kristallstruktur bestätigt. In einem zweiten Durchtestverfahren wurden 19'968 (192 Sequenzen × 104 Carbonsäuren) Makrozyklen gegen MDM2, welches eine wichtige Protein-Protein Interaktion mit p53 bildet, getestet. Ein erster Binder wurde gefunden (K_D = 390 ± 40 nM) und durch zwei schnelle Optimierungsiterationen unter der Verwendung der gleichen Methode verbesser (K_D = 31 ± 6 nM).

Die Methoden, welche ich in den zwei Projekten entwickelt hatte, werden gegenwärtig von mehreren LaborkollegenInnen für die Erzeugung von Makrozyklenliganden gegen verschiedene Zielproteine angewendet. Die Zukunft wird zeigen ob die Methoden Moleküle zur Entwicklung von neuen Medikamenten liefern können.

Stichwörter: Makrozyklen, Peptide, Wirkstoffforschung, zyklische Freisetzung, kombinatorische Synthese, Protein-Protein Interaktionen, Disulfidbindungen, Therapeutika, Hoch-Durchsatz-Screening, Trefferoptimierung, Nanomassstabsynthese, akustischer Tröpfchenauswurf, Thrombin, MDM2

Table of contents

1 Introduction

1.1 Macrocycles

1.1.1 Overview

Macrocycles are cyclic compounds with a ring size greater than 12 atoms. Their molecular weights typically range between 500 and 2000 daltons, making them larger than most classical small molecule compounds, which usually have molecular weight < 500 daltons.[1] Macrocycles are found ubiquitously in nature, and are produced by a variety of plants, animals, bacteria, and fungi.[2] Synthetic macrocycles have also been developed, largely for therapeutic purposes (discussed further in section 1.1.2).

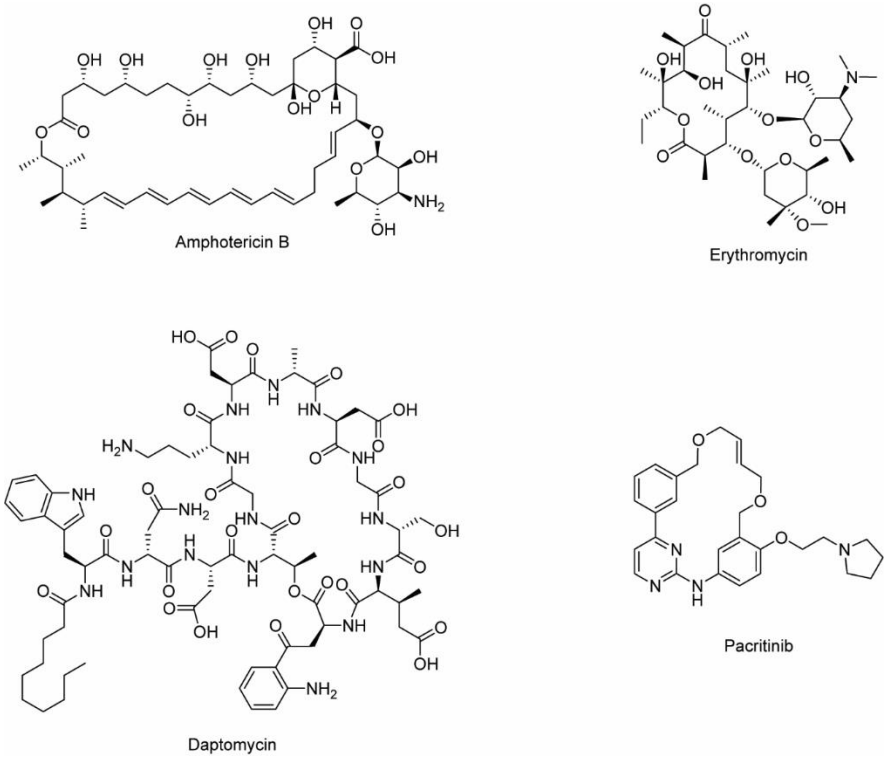


Figure 1. Examples of diverse macrocyclic structures. Amphotericin B and erythromycin are examples of polyketide natural products. Daptomycin is a natural product peptide macrocycle bearing an N-terminal lipid. Pacritinib is an example of a fully synthetic macrocycle.

Most macrocycles can be categorized into the following two main structural classifications: peptides and polyketides. Peptidic macrocycles are comprised of amino acids,

which are coupled via amide bond formation. Subclasses of peptides, such as depsipeptides, can include amino-alcohol monomers as well.[3] Notable examples of peptidic macrocycles are cyclosporine A (CsA) and daptomycin. Polyketides are synthesized in nature by polyketide synthase enzymes using acetyl-coenzyme A monomers. Cyclization is often achieved via macrolactonization, as in the well-known macrolide antibiotics. Glycosylation of peripheral alcohols on these structures is also common.[4] However, many macrocycles do not fit into these classifications. These include alkaloid natural products, fully synthetic therapeutics such as pacritinib, or hybrids of other classes. Example structures of all of these aforementioned macrocycles are shown in Figure 1. Cyclization of all classes of linear precursors can be achieved in a large variety of ways, and will be discussed in section 1.2.

1.1.2 Macrocycle therapeutics

As of 2014, 68 macrocycles are currently in use for therapeutic purposes, with more than 30 in clinical development. The most common indications treated with macrocycle drugs are infectious disease and oncology.[5] Perhaps the most well-known class of macrocycle drugs are the orally available macrolide antibiotics, which include erythromycin, azithromycin, and clarithromycin (Figure 2). These compounds are hydrocarbon lactones typically ranging from 12-16 atoms in size, with peripheral methyl, ethyl, and hydroxyl substitution. Glycosylation of hydroxyl groups with deoxy sugars such as cladinose and desosamine is common. The macrolides were first discovered in *Streptomyces* bacteria, and are effective antibiotics against many gram-positive and gram-negative bacteria. Their antibiotic activity is due to binding to the bacterial ribosome large subunits, thus interfering with protein synthesis.[6] Rifamycins are another category of antibiotic macrocycles. They are polyketide lactams, and contain a naphthoquinone moiety in the backbone ring (Figure 2). These compounds kill bacteria by inhibiting bacterial RNA polymerase, and are particularly effective against tuberculosis infection.[7] Peptidic macrocycle antibiotics are also known. Polymixins are an example of gram-negative antibiotics, usually administered as drugs of last resort due to their toxicity. They are polycationic peptides that contain lipophilic fatty acids at the N-terminus (Figure 2). These features enable polymixins to displace metal cations on bacterial lipopolysaccharides, destabilizing them and leading to a disruption of the outer membrane.[8]

Macrocycles are also important drugs for the treatment of cancer. Rapamycin is an inhibitor of the mTOR pathway, important in cell survival and proliferation. Like the previously mentioned polyketide antibiotics, rapamycin is a macrolide that was first discovered in *Streptomyces*. Several semisynthetic analogs of rapamycin have been developed. Modifications are made on a peripheral cyclohexyl substituent in order to improve

physicochemical properties.[9] One such analog, everolimus is used for the treatment of certain breast cancers, pancreatic cancer, renal cell carcinoma, gastrointestinal cancer, lung cancer, and subependymal giant cell astrocytoma.[10] Synthetic macrocyclic drugs have also been developed as cancer treatments. These macrocycles are neither polyketides nor peptides, but rather resemble traditional small molecule kinase inhibitors. Promiscuity is a problem with many kinase inhibitors, and can lead to undesirable side effects. Macrocycles offer an advantage over the typically linear kinase inhibitors by having increased selectivity.[11,12] Lorlatinib, an ALK and ROS1 inhibitor developed by Pfizer, is FDA approved for the treatment of ROS1 positive non-small cell lung cancer. Lorlatinib has a less promiscuous kinase inhibition profile than earlier ALK inhibitors, and is more active against certain ALK mutants.[13,14]. Another macrocycle, zotiraciclib, is a CDK9 inhibitor in clinical trials being investigated for the treatment of gliomas.[15]

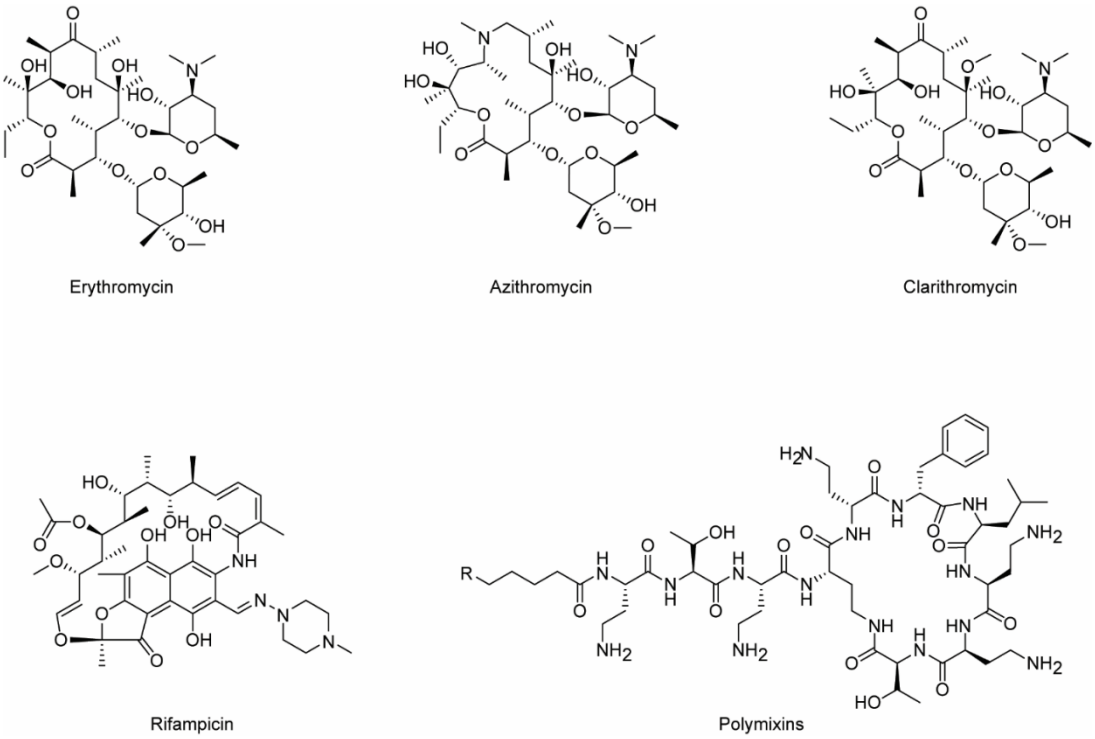


Figure 2. Examples of macrocyclic antibiotics. Erythromycin, azithromycin, and clarithromycin are common examples of the macrolide class of antibiotics. Rifampacin is a polyketide lactam, while the polymixins are lipidated macrolactam peptides.

In the pharmaceutical industry, macrocycles have raised much interest due to their ability to disrupt protein-protein interactions (PPIs). For many PPI disease targets, no therapeutic compounds yet exist. Small molecule compounds, while being orally available and cell permeable, are often not able to bind to surfaces of PPIs if they are flat, or can't disrupt large PPIs due to their diminutive nature. Biologic drugs, such as large peptides, antibodies, and fusion proteins, contain much larger surface area, and thus are able to bind to these

difficult targets. However, biologics are not cell permeable or orally available, which prevents their use for intracellular targets or oral administration. Macrocycles are typically larger than small molecules, but far smaller than biologics. This places them in a unique chemical space where they are able to bind to and modulate PPIs like biologics, but possess the cell permeability and oral availability properties of traditional small molecules.[1,16,17]

1.1.3 Properties of membrane permeable, orally available macrocycles

In small-molecule medicinal chemistry, Lipinski's rule-of-five guidelines are followed in order to ensure cell permeability and thus oral bioavailability. The rules state that a compound should have molecular weight less than 500 daltons, water/n-octanol partition coefficient (LogP) between 1 and 5, less than 5 hydrogen bond donors (HBDs), and less than 10 hydrogen bond acceptors (HBAs).[18] Subsequent studies by Veber have shown that rotatable bonds and polar surface area (PSA) are also important metrics, with fewer than 10 rotatable bonds and PSA less than 140 Å^2 being optimal for oral availability.[19] Macrocycles typically violate many of these guidelines due to their larger size. However, many of them are orally available compounds. The group of Jan Kihlberg has extensively studied the properties of "beyond rule-of-five" (BRo5) compounds. Their analyses show that molecular weight up to 1,000 daltons, and PSA up to 250 Å^2 are permissible so long as the two metrics scale linearly in a 4:1 ratio. However, LogP and HBDs must remain low (roughly within the Lipinski parameters) for the compounds to be orally available; too many HBDs will result in the desolvation energy being too high for membrane permeation, while a high LogP may favor the hydrophobic environment within the membrane (Figure 3). Many of the macrolide antibiotics adhere to these expanded rules and are indeed orally available compounds.[5,20,21]

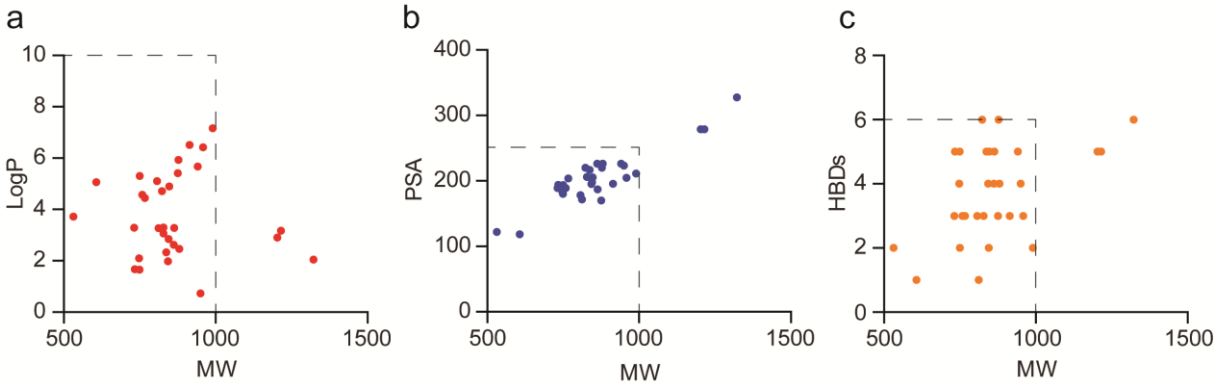


Figure 3. Physicochemical properties of approved drugs and clinical candidates with MW >500 daltons. The accepted beyond-Ro5 space is indicated with dashed lines. **A**-**C** show MW vs cLogP, PSA, and number of hydrogen bond donors (HBDs), respectively. The vast majority of orally available compounds fall in BRo5 permeability space. Approved drugs and clinical candidates selected from [5].

An additional parameter is crucial for the passive cell permeability of large compounds: the conformational dynamics of the molecules. This includes the ability to fold and shield polar functional groups while crossing lipophilic membranes, as well as the formation of intramolecular hydrogen bonds (IHBs). In BRo5 compounds, a larger number of rotatable bonds are actually favored due to dynamic flexibility that they impart. The standard PSA metric becomes less important for these compounds, and instead the three-dimensional exposed PSA (3D PSA) becomes more relevant. Computational modeling, NMR, and crystallographic analysis of known orally available drugs shows that 3D PSA will change depending on the environment the molecule is in: in aqueous solution, polar surface area will be exposed ensuring solubility, while in lipophilic environments, the molecules will fold and bury the polar groups. [22,23,24] CsA is a macrocycle that exhibits this behavior. In aqueous solution, the amide protons point towards the solvent, while the N-methylated amides point inward. However in organic solvent, the amide protons point inward, forming a network of IHBs with carbonyl oxygens. All of the four unmethylated amides undergo this bonding, leaving no “free” HBDs from the backbone (Figure 4).[25,26]

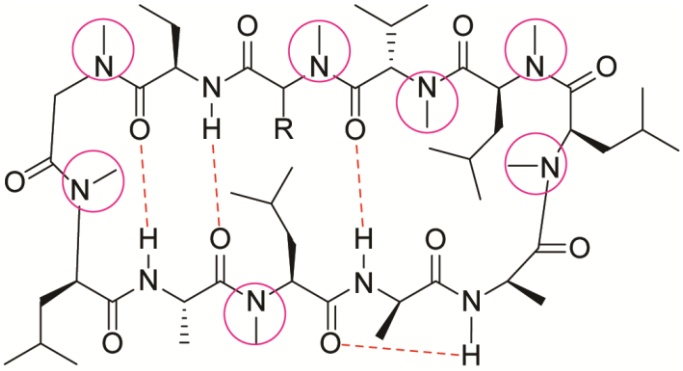


Figure 4. The structure of cyclosporin A allows for membrane permeability. Both the extensive intramolecular hydrogen bonding network formed in lipophilic media (shown as dashed red lines), and the amide bond N-methylation (circled in pink) contribute to the permeable nature of CsA.

It is important to note that the aforementioned physicochemical properties and dynamic properties are largely important for passive diffusion through cell membranes. However, transporter mediated and endosomal entry into cells are also relevant uptake methods for macrocycles. Overall, not much is known about active transport of macrocycles, or how to design compounds which are capable of utilizing this mechanism.[25,27] Thus, active transport and endosomal entry are not relevant topics for this thesis.

1.1.4 Strategies for making peptidic macrocycles cell permeable and orally available

In the previous section, the important properties for passive cell permeability of macrocycles were described. Here I will discuss how peptidic macrocycles can be designed so that they adhere to these guidelines. Peptides have inherent limitations when it comes to cell permeability: amide bonds contribute to PSA, HBDs, and HBAs. A simple Gly_6 macrocycle already exceeds Lipinski's guidelines, having six HBDs, and PSA of 174.6 $Å^2$. The addition of any polar side chains would increase PSA and HBD/HBAs further. Side chains with ionic charges would of course minimize permeability as well. Therefore, reducing the number of HBDs is a strategy that is widely used to improve oral availability of macrocycles. One way of achieving this is via N-methylation of backbone amide nitrogens. N-methylation has been shown to be an effective method for improving cell permeability. However, it can also greatly influence the structure of macrocycles, and could have a negative effect on binding affinity to a target. In addition, the impact on cell permeability is dependent on the specific amide in a molecule that is being methylated.[25,28,29] Perhaps the best known example of an N-methylated cyclic peptide is CsA, where seven of the eleven amide nitrogens have been methylated. Despite having a molecular weight of over 1,200 daltons, CsA is cell permeable. As previously mentioned, this is partly also due to the ability of CsA to adopt a folded confirmation in lipid bilayers.[25,26] Similar to N-methylation, peptoids are another solution to reduce the number of HBDs. Typically, cyclic peptoids contain peripheral groups bound to the macrocycle backbone via amine nitrogen atoms rather than on the alpha carbon. N-alkylated aromatic amino acids have also been used, termed benzylopeptoids.[30] Peptoids are more cell permeable than peptides due to the lack of an amide N-H bond, which is polar and acts as a HBD. One study found up to 20-fold improvement in permeability between peptides and the analogous peptoids.[31] Peptoids have been utilized in the pharmaceutical industry for the development of therapeutics. One example is a peptide-peptoid hybrid binder for CXCR7, which is a protein implicated in atherosclerosis and heart failure. The lipophilic lead compound contained three HBDs, bound the target with K_i = 9 nM, was permeable with $P_{app} = 6.2 \times 10^{-6}$ cm/s, and was 18% bioavailable in rats.[32] Depsipeptides are compounds where one or more amide bond has been replaced by an ester. They are found in nature, being produced by a variety of bacteria.[33,34] Depsipeptides can be more cell permeable due to the replacement of a hydrogen bond donating amide.

As mentioned above, peptide macrocycles such as CsA can be cell permeable if they can effectively bury HBDs and polar functionality while crossing membranes via the formation of IHBs. However, CsA is a natural product and was not rationally designed. For de novo

synthesized macrocycles, designing the ability to undergo intramolecular hydrogen bonding is more challenging. One way to do this is by including D-amino acids. Macrocycles made up of amino acids with L stereochemistry can only adopt a limited number of conformations for a given number of residues. Including D-amino acids at various positions will greatly increase the conformational diversity. The group of Lokey synthesized nine diastereomers of a cyclic hexapeptide that had D-amino acids at various positions. They found that the diastereomers have vastly different permeabilities, differing by up to ~100-fold. Conformational analysis of the best and the worst macrocycles showed that the IHB network of each was different, with the best macrocycle having more extensive bonding (Figure 5a).[35] The introduction of exocyclic peripheral groups can also affect the IHBs. The group of Yudin achieved this by synthesizing 18-membered macrocycles that contained an exocyclic amide bond. While the inclusion of an additional amide bond may sound counterintuitive for increasing cell permeability, it actually serves to establish a more extensive network of IHBs. The macrocycle with the exocyclic amide was 40 times more permeable than the analog without it (Figure 5b).[36]

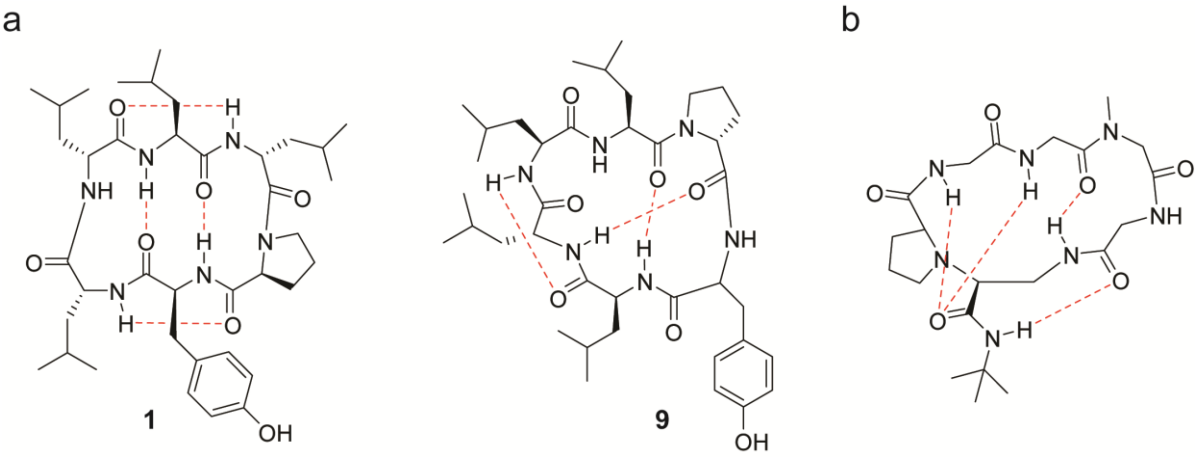


Figure 5. Intramolecular hydrogen bonding in macrocycles. a The difference in hydrogen bonding networks (determined by NMR, red dashed lines) for two peptides with D-amino acids at various positions. The extent of this network determines the cell permeability, with **1** being roughly 100x more permeable than **9**.[35] **b** An exocyclic amide bond allows for a conformation where many intramolecular hydrogen bonds can form.[36]

The aforementioned modifications of N-alkylation, replacement of amides by esters, and inclusion of D-amino acids have an additional benefit: increasing proteolytic stability of peptides. While small macrocycles already have an advantage over linear species and larger cyclic peptides, they can still be susceptible to degradation by various proteases. This represents an additional barrier for oral availability. Introducing amide bond modifications can help to minimize this undesired metabolism of peptidic macrocycles.[31,37,38] This can also be achieved by the use of other non-canonical amino acids, including β-amino acids.[39]

1.2 Synthesis of macrocycle libraries

1.2.1 Examples of existing macrocycle libraries

This chapter discusses macrocycle libraries amenable for screening in microwell plates, but does not include macrocycles libraries screened differently such as DNA-encoded macrocycle libraries screened by affinity selections, that are discussed in chapter 1.3. The chemical synthesis of macrocycles is known to be difficult, and is a major reason why macrocycles remain an underrepresented therapeutic modality.[40,17] Synthesis of linear precursors can be relatively simple, as in the case of peptides, which rely on iterative rounds of robust amide bond formation. However, synthesis of macrolide precursors is more complex, requiring many air and/or water sensitive reactions, stereoselective reactions, and purification steps.[41,42,43] For all approaches, cyclization remains the major limitation. Ideally, a library of macrocycles will have diverse backbone structures. However, this means that cyclization will often occur in low yields, and will necessitate chromatographic purification. Performing many such purifications is time, labor, and resource intensive. Thus, the size of macrocycle libraries is often limited.

Table 1. Examples of large macrocycle libraries from the literature and from commercial sources. While the library of Seiple et al. is significant smaller than the others, the complex polyketide nature of the compounds makes the library noteworthy.

Company/group	Library description	Library size	Citation
Seiple et al.	Macrolide antibiotics synthesized using a convergent synthesis of eight precursors.	300	[44]
Guo et al.	Rapamycin inspired peptides. Two FKBD binding motifs fused with 22,500 variable peptide sequences.	45,000	[45]
Marcaurelle et al.	Aldol reaction based. DOS strategy applied for generating many diastereomers, along with various cyclization chemistries.	14,500	[46]
ChemBridge	<800 daltons molecular weight. Many likely to be suitable for CNS application.	13,000	[47]
Asinex	<600 daltons molecular weight, small drug-like compounds.	35,000	[48]
Polyphor	MacroFinder and PEMfinder technologies. Includes large peptides also.	50,000	[49]

Despite the synthetic difficulties, several large macrocycle libraries have been reported in the literature (Table 1). The group of Meyers developed a convergent synthesis of 14 and 15-membered macrolide antibiotics, and synthesized a library of 300 compounds. Considering the complexity of the molecules, this represents a significant effort. The method utilizes only eight commercially available or easily synthesized precursors, which are assembled into diverse structures. They were able to identify many compounds with antibiotic activity against resistant gram-negative strains.[44] More recently, a large rapamycin-inspired library was synthesized. Two optimized FKBD binding motifs were combined with a total of 22,500 variable tetrapeptide regions to give 45,000 macrocycles (Figure 6). The library was purified by creating 3,000 pools of 15 macrocycles, and performing flash-column chromatography. In a cell-based screen, a dual FKBD12-ENT1 binder that prevented renal reperfusion injury in a mouse model was identified.[45] While quite a large number of macrocycles were synthesized, the thousands of column purifications in this work serve as a reminder of the strenuous work that is required to make such a large library.

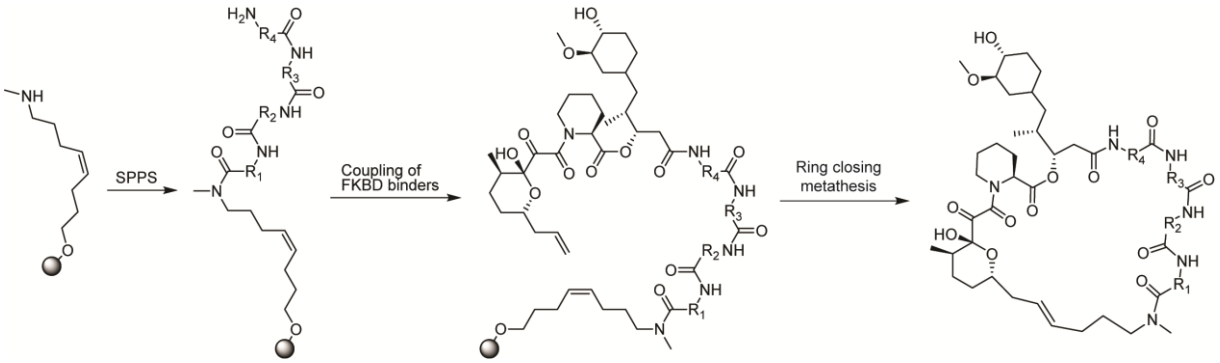


Figure 6. A large 45,000-member macrocycle library for the formation of ternary complexes. The "Rapafucin" library utilizes FKBD-binding regions in combination with a peptide sequence for binding to new targets. Standard SPPS conditions were used to synthesize the linear compounds, followed by ring-closing metathesis using the second generation Hoveyda-Grubbs catalyst.[45]

Diversity-oriented synthesis (DOS) is a technique that can be applied to generate macrocycle libraries. Traditional combinatorial libraries typically utilize one type of reaction at a time, limiting the overall chemical space that is represented, despite variations in building block structure. In contrast, starting from a single functional group, each step of DOS utilizes different types of reactions to produce multiple, diverse products. Repeating this divergent approach over several iterative steps results in a highly diverse library that covers large chemical space (Figure 7).[50,51] DOS has been applied extensively to macrocycles.[52] For example, the group of David Spring developed a strategy to make libraries of natural product-like macrocycles. In the first step, aromatic hydroxyl acids would be alkylated on the alcohol

by bromo-alkenes and bromo-azides. The next step was amidation with amino-alkenes or amino-alkynes. They then employed four different macrocyclization reactions: linear species bearing azides and alkynes were cyclized using copper/ruthenium catalyzed azide-alkyne cycloaddition (CuAAc/RuAAc) to give 1,4- or 1,5-triazoles, linear species bearing two alkenes were cyclized by ring closing metathesis (RCM), and species bearing an alkene and an alkyne were cyclized using ene-yne RCM. [53]

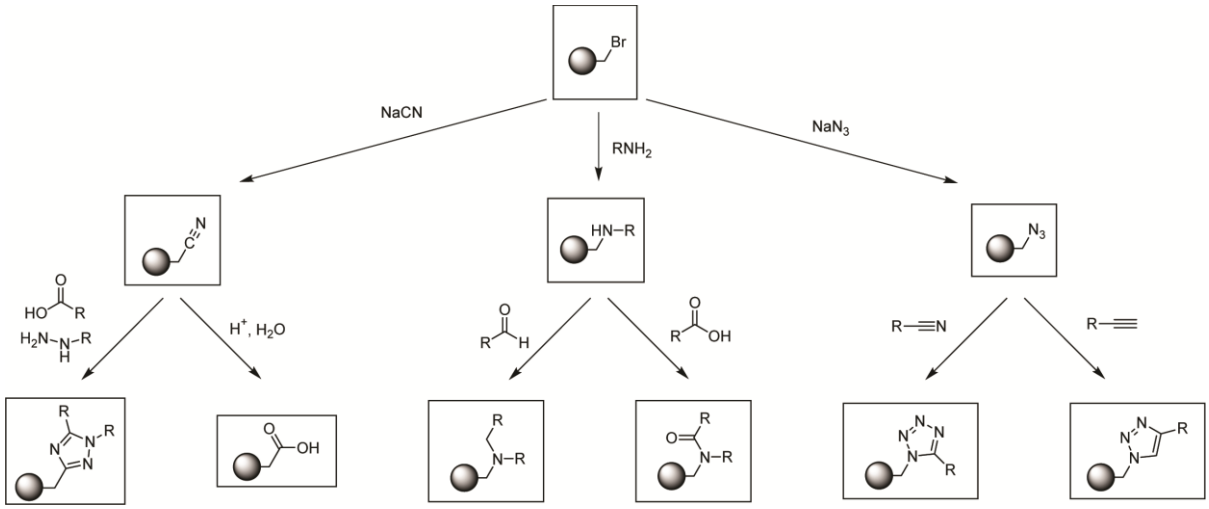


Figure 7. Diversity oriented synthesis. The general DOS principle illustrated by an example of a single alkyl halide to six different chemical motifs. Utilizing each starting material functional group for several reaction types rapidly expands the total diversity of a library.

In an example from the Broad Institute, the Aldol reaction was used to generate linear precursors with rich stereochemical diversity. These precursors were functionalized with various groups, and cyclized via S_NAr, RCM, or CuAAc/RuAAc to give 48 macrocyclic scaffolds bearing two protected amines. Orthogonal deprotection and capping with carboxylic acids, sulfonyl chlorides, isocyanates, and aldehydes gave 14,500 fully diversified macrocycles. The library was screened against HDAC2, and a low micromolar inhibitor was identified.[46] It is important to note that DOS approaches will also require many chromatographic purifications, and are thus difficult to carry out on larger scales.

While the previous examples of macrocycle libraries have been products of academic research institutions, several large commercial collections also exist. Chembridge, a contract research organization (CRO) advertises a 13,000 member library having molecular weight less than 800 daltons.[47] Asinex, another CRO, possesses a collection of roughly 35,000 macrocycles having molecular weight less than 600 daltons, with specialized subsets for certain therapeutic applications (e.g. neurological targets).[48] The pharmaceutical company Polyphor has generated a library of 50,000 macrocycles using its MacroFinder and PEMfinder

technologies.[49] However it is not clear how many of these macrocycles are small and drug-like, and how many are larger cyclic peptides.

1.2.2 Linear macrocycle precursors from solid-phase peptide synthesis (SPPS)

As previously mentioned, cyclic peptides are an important class of macrocycles. As with most other synthetically produced peptides, they can be made using SPPS. This technique, first developed by Merrifield,[54] allows for facile, robust, and fast synthesis of linear peptides tethered to an insoluble solid phase. Due to its ubiquitous use, all necessary reagents are commercially available and cheap. The amino acids that are utilized for this process have 9-fluorenylmethoxycarbonyl (Fmoc) protecting groups on the N-terminus, which can be removed by treatment with amine bases, most often piperidine. The amino acid side chain functionality is protected by acid labile groups, such as triphenylmethyl (trityl) or tert-butyloxycarbonyl (Boc). This allows for orthogonal N-terminal deprotection, and ensures that side chains do not undergo undesired reactions. Amino acids are activated in the presence of tertiary amine base with one of many activating reagents so that they may undergo acylation reactions. Upon completion of synthesis, the peptides on solid phase are cleaved from resin, and simultaneously globally deprotected to give the final desired peptide. This step is typically performed with acid, such as trifluoroacetic acid (TFA), but other methods can be utilized (discussed later).[55]

While the method is simple in theory, in fact many considerations must be taken into account for a successful synthesis. The first of which is the choice of resin and linker. Most resins that are used are polystyrene (PS) which has been cross-linked with divinylbenzene (DVB). The extent of crosslinking is important, with 1-2% DVB being optimal for efficient syntheses. These resins swell best in organic solvents such as dimethylformamide (DMF) or dichloromethane (DCM). The PS structure will have linkages that are used for attachment of the first amino acid. Commonly used are Rink-amide and 4-Methylbenzhydrylamine (MBHA), which provide an amine for acylation and upon cleavage produces a C-terminal carboxamide, or benzyl chloride (Merrifield) and Trityl-chloride, which result in an ester linkage and upon cleavage liberate a free carboxylic acid. In order to make the resin swell better in polar solvents such as water, methanol (MeOH) or dimethylsulfoxide (DMSO), polyethylene glycol (PEG) can be conjugated to the resin. It is often placed between the linkage and the polystyrene matrix, as in Tentagel resins. Other resins such as ChemMatrix and PEGA resins are not PS based, but are rather comprised of cross-linked PEG matrices. Importantly, these PEG containing resins also swell in more hydrophobic solvents, and not just in polar ones. Finding the right resin is important for high yields during synthesis.[55]

Carboxylic acids and amines will not form amides on their own under mild conditions. The acids must first be activated with another reagent to form a reactive species. Certain amino acids can be difficult to activate and couple, especially unnatural ones with two α-side chains such as α-aminoisobutyric acid (AIB). Amino acid racemization can also occur during coupling steps, leading to complex product mixtures. Therefore, the correct coupling reagent choice is a crucial part of SPPS. In the first step of activation, *O*-acyl isourea, *O*-acyl isouronium, *O*-acyl phosphonium, or other similar species are formed. These species are highly reactive and prone to racemization. Therefore, additives are frequently included as well. These are typically acidic alcohols that can displace the urea/uronium to form an active ester with a good leaving group. The active esters are reactive towards amines, but stable enough to be used without side product formation or degradation. Carbodiimides such as *N,N*-dicyclohexylcarbodiimide (DCC) were the earliest reagents utilized.[54] With these reagents, additives such as 1-hydroxybenzotriazole (HOBt) or *N*-hydroxysuccinimide (HOSu) must be added. Other reagents, such as *O*-(7-azabenzotriazol-1-yl)-*N,N,N′,N′*-tetramethyluronium-hexafluorphosphate (HATU), *N,N,N′,N′*-tetramethyl-*O*-(N-succinimidyl)-uronium-hexafluorophosphate (HSTU), and 1-[(1-(cyan-2-ethoxy-2-oxoethylidenaminooxy)-dimethylamino-morpholino)]-uronium-hexafluorophosphate (COMU) have a “built in” equivalent of additive. By changing the uronium portion and the additive portion, the reactivity of coupling reagents can be tuned. Carboxylic acids can also be converted to reactive acyl chlorides and fluorides using reagents such as cyanuric chloride and tetramethylfluoroformamidinium hexafluorophosphate (TFFH). All reagents are shown in Figure 8. Difficult couplings will require investigation of many reagents to find the best one.[56,57]

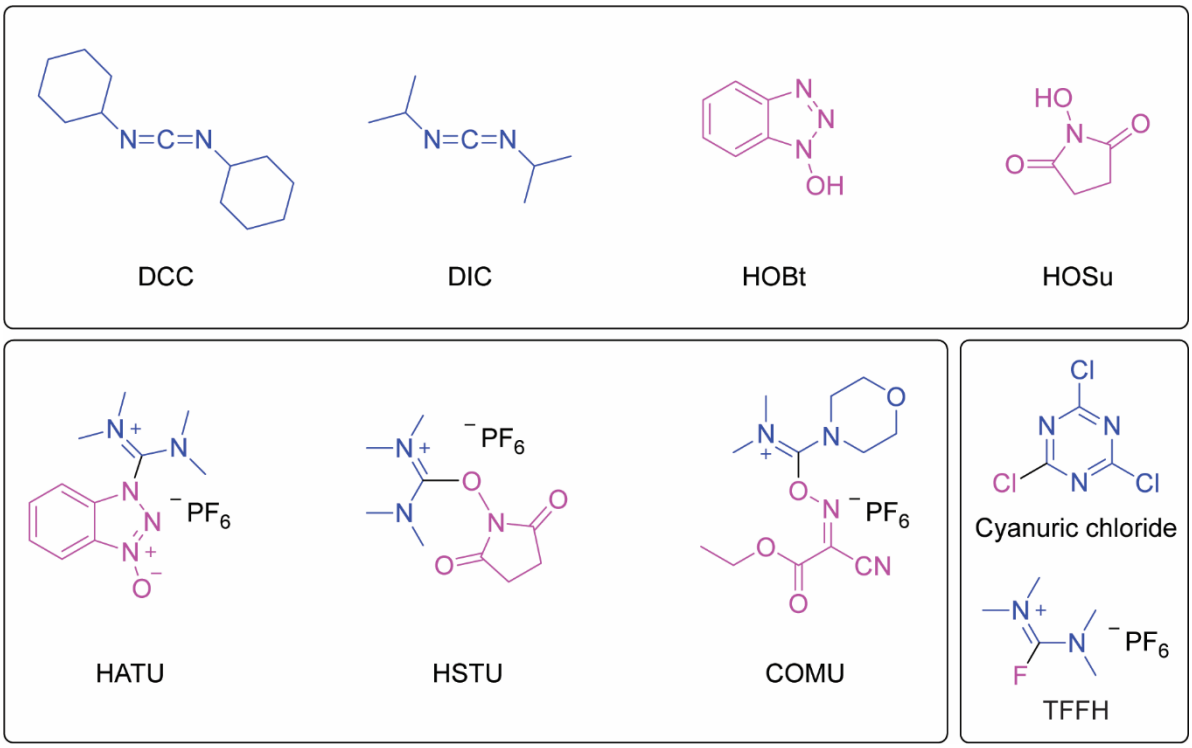


Figure 8. Examples of common coupling reagents used in SPPS. Primary activating reagents or components are shown in blue, while additives used to generate active species while minimizing racemization are shown in pink. For the bottom row, this additive is "built in" to the coupling reagent. In the case of cyanuric chloride and TFFH, acyl halides are generated instead of active esters.

Finally, cleavage must be performed with the appropriate cleavage cocktail. Side chain protecting groups can produce reactive electrophilic species upon treatment with acid, and must be scavenged in order to prevent the formation of side products. For example, Boc protecting groups will result in tert-butyl (tBu) cations, which form an intermediate reactive species with TFA. This tBu-OTFA species can alkylate cysteine and methionine side chains irreversibly. Therefore, cation scavengers must be included in the cleavage cocktail. These include triisoproylsilane, which acts as a hydride source, and phenol/anisole/thiophenol, which can quench cations via electrophilic aromatic substitution reactions. Thioanisole is often used as a scavenger, though it does not represent a permanent quenching of cations, and can itself act as an alkylating agent. Ethanedithiol (EDT) is frequently included to ensure that cysteine residues remain reduced during cleavage, but it can also act as a scavenger.[58,59] Depending on the resin linkage that is employed, TFA must make up to 95% v/v of such cleavage cocktails (for example with Rink-amide). 2-chlorotrityl resins, on the other hand, can be cleaved with only 1% TFA, though more may be required for simultaneous side-chain deprotection. Peptides with cationic amines/guanidines are typically recovered from cleavage mixture by precipitation from diethyl ether, which acts to remove protecting group and synthesis byproducts.[55]

So far, standard Fmoc-based SPPS conditions with acid-labile resin linkages have been described. However, there are linkages to resin which are not acid-labile. This orthogonality can be used to deprotect side chains while the peptide remains on resin thus eliminating the need for precipitation, and to introduce different C-terminal functional groups. The latter point is especially important if one wishes to utilize diverse chemical methods for macrocyclization. One such orthogonal method is basic cleavage. For this method, it crucial that secondary amine bases such as piperidine (used for Fmoc deprotection) are not able to cleave the peptides. Peptides linked to resin through linkages such as 4-(hydroxymethyl)benzoic acid (HMBA) esters can be cleaved with strong bases (hydroxide or alkoxide), or more mild bases (e.g. carbonates) in water or alcohols to give the corresponding acid or ester. If peptides are linked in the reverse direction (resins bearing carboxylic acids, peptide with C-terminal alcohol), these can also be liberated by treatment with base to generate free alcohols. When peptides are linked via formamidines, hydrazine can be used for cleavage. It can also be used to cleave peptides N-terminally linked though a 4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl (Dde) group.[60] While Dde linked through the 4 position is not completely stable to piperidine, similar analogs linked through the double bond have been developed that are better able to resist unwanted cleavage.[61]

A particularly fascinating example of orthogonal cleavage is with so called “safety catch” linkers. In these strategies, a linker will be inert until treated with a specific reagent, which then renders it susceptible to cleavage. These reagents are often oxidants (H_2O_2, *m*-chloroperbenzoic acid, *N*-bromosuccinamide) or alkylating/acylating reagents (CH_2N_2, Boc_2O). One example is the sulfide/sulfone safety catch linker. A peptide can be synthesized, via a C-terminal ester, on a phenol which is linked via a thioether in the *para* position to resin. Such esters are resistant to basic cleavage. However, when the thioether is oxidized to a sulfone with H_2O_2, the phenol becomes a good leaving group and the ester can be cleaved with a primary amine under mild conditions to generate an amide (Figure 9a).[62] The use of oxidants has limitations however: cysteine, methionine, and tryptophan will be oxidized and thus are not compatible with these strategies. An example of an alkylation-based safety catch linker is the one of Ellman’s group, itself an improvement of the work of Kenner et al.[63] Here, peptide synthesis occurs on a sulfonamide linker, generating an acyl sulfonamide. The linkage is stable to acidic and basic conditions until it is alkylated with iodoacetonitrile, after which cleavage can occur with amines and amino acids to give newly formed amide bonds (Figure 9b). The low pKa of the acylsulfonamide compared to those of amides allows for the selective alkylation.[64]

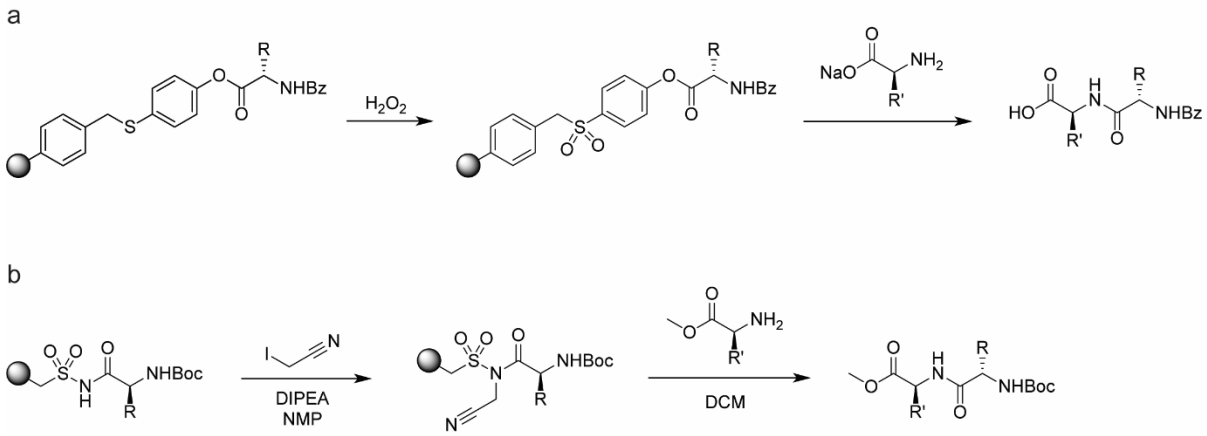


Figure 9. Safety catch linkers in SPPS. Linkers activated with hydrogen peroxide (**a**) and iodoacetonitrile (**b**) allow for C-terminal reaction with nucleophiles, resulting in cleavage and the formation of a new peptide bond. The un-activated linkers render the C-terminal carbonyls “inert”, which allows for selective cleavage under specific conditions.

There are few options for non-acidic cleavage of C-terminal thiols: thiols linked through groups such as trityl will be too acid labile for on-resin deprotection of side chains, and thioesters will not be stable to piperidine during synthesis. In contrast, orthogonal cleavage of C-terminal thiols is possible when disulfide linkages to solid phase are used. This linkage can be introduced by synthesizing amino acids containing an internal disulfide bond, such as aminoethyldithio-2-isobutyric acid (Figure 10a). Alternatively, a thiol-acid can first be acylated onto amine-bearing resin, followed by installation of a disulfide via disulfide exchange reagents (Figure 10b). The disulfide bond is stable to both TFA-based cleavage conditions, and piperidine-based Fmoc deprotection conditions. Instead, reducing agents such as dithiothreitol (DTT) and tris-(2-carboxyethyl)phosphine (TCEP) are used to cleave the disulfide bond. This cleavage is fast and nearly quantitative, and allows for the use of different solvents at various pH. When TCEP is used, resin-bound cationic scavengers can be employed to remove excess reagent after cleavage.[65,66]

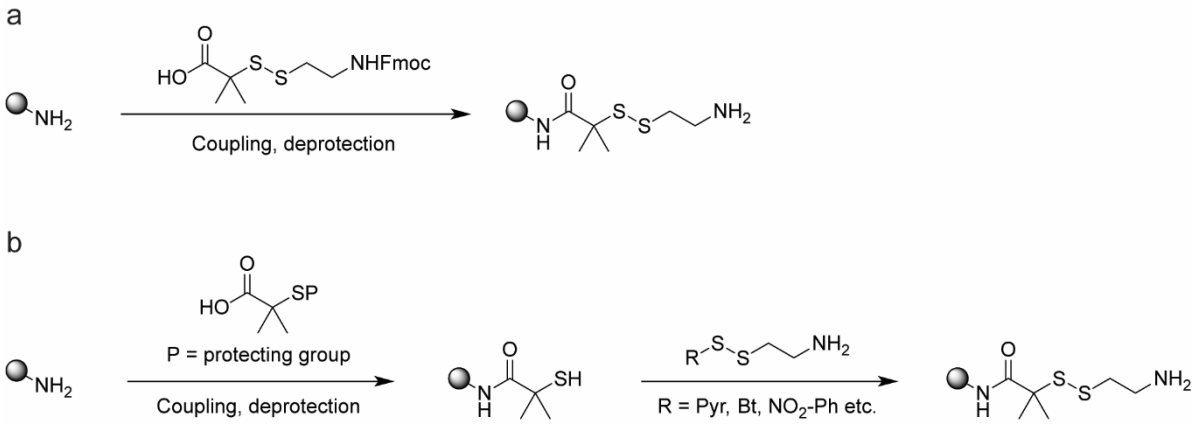


Figure 10. Disulfide linkers for SPPS of peptide containing C-terminal thiols. The disulfide can be introduced through acylation of a protected amino acid (**a**), or via coupling of a protected thiol-carboxylate, followed by disulfide exchange with an activated disulfide species (**b**). After either installation step, SPPS can proceed normally.

1.2.3 Cyclization methods

The previous section described how linear peptidic species can be synthesized using SPPS. Here, the cyclization of linear species will be explored. Macrocyclization is typically the most difficult step of synthesis. Therefore, reactions need to be very robust and broadly applicable. To further aid in enhancing yields, the utilization of certain types of amino acids (proline, D-proline, N-methylated) can result in secondary structures that favor cyclization. Other strategies have also been used, such as transition metals, which form chelates and template linear molecules into favorable conformations. Cyclization most often takes place in solution at low concentration to minimize intermolecular reactions, but many on-resin cyclization reactions have been reported. Cyclization of a linear peptide can occur via the following ways: head to tail (N-terminus to C-terminus), head to sidechain, tail to sidechain, or sidechain to sidechain (Figure 11).[67,68] With these considerations in mind, many different chemical reactions have been utilized for macrocyclization, and examples will be discussed further in this section.

The coupling of carboxylic acids to amines and alcohols gives rise to lactams and lactones, respectively. Lactamization and lactonization are commonly used macrocyclization reactions. In peptides, the C-terminus, aspartic acid, or glutamic acid can serve as the carboxylic acid component, while the N-terminus or lysine serves as the amine for macrolactamization, and serine or threonine serving as alcohols in macrolactonization. For these reactions, peptide coupling reagents, such as those previously described, are typically used (Figure 12a).[67,68] Both macrolactonization and macrolactamization have been used for the synthesis of the HDAC inhibitor FK228 (romidepsin). Using HATU as the coupling reagent

in DCM, a macrolactamization yield of 54% was obtained. This represented an improvement to the macrolactonization based synthesis, where much optimization was needed to overcome the steric hindrance of the carboxylic acid and stability of the alcohol.[69]

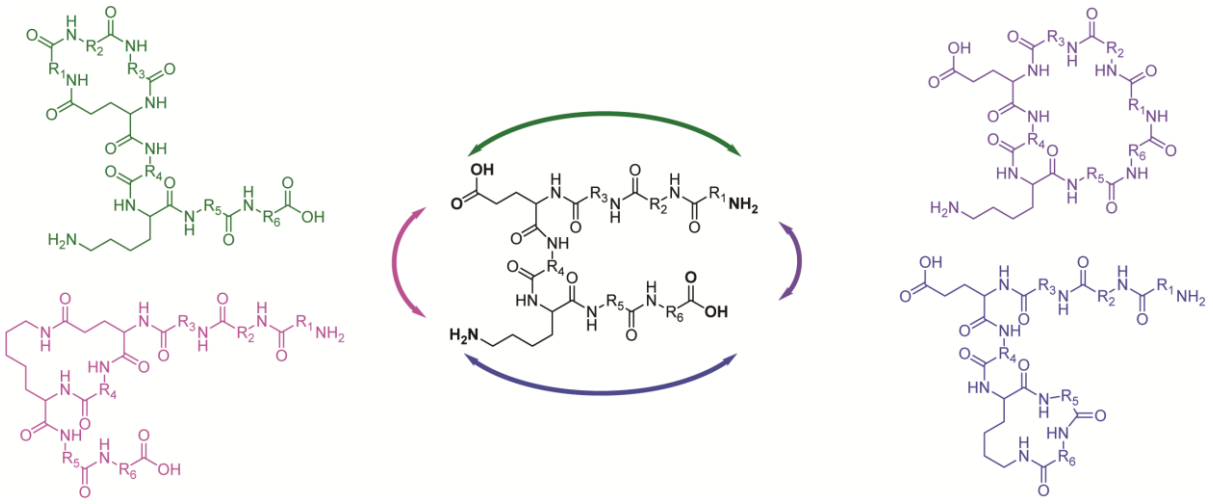


Figure 11. A generalized depiction of the possible ways a linear compound can by cyclized. A generic peptide can be cyclized head-to-tail (purple), head-to-sidechain (green), sidechain-to-sidechain (pink), or sidechain-to-tail (blue). Inspired by [67].

A variety of transition metal catalyzed macrocyclization reactions have been utilized (Figure 12b). One of the most well-known examples is RCM, which utilizes ruthenium catalysts to form carbon-carbon double bonds. The method is frequently utilized for the synthesis of stapled peptides, which have increased alpha-helical character and have been shown to be cell permeable.[70,71] Another metal catalyzed macrocyclization is the so called “click” reaction reaction (copper catalyzed azide-alkyne cycloaddition). In this reaction, an azide and an alkyne undergo cycloaddition with the aid of a Cu(I) catalyst to generate 1,2,3-triazoles. The reactions are robust, high yielding, and tolerant of many functional groups.[72,73] Finally, various metal catalyzed cross-coupling reactions have been utilized. These include Suzuki, Heck, and Buchwald-Hartwig reactions for generation of carbon-carbon/oxygen/nitrogen/sulfur bonds using palladium catalysts, and Ullmann type carbon-carbon/oxygen bond formation using copper species.[74,75]

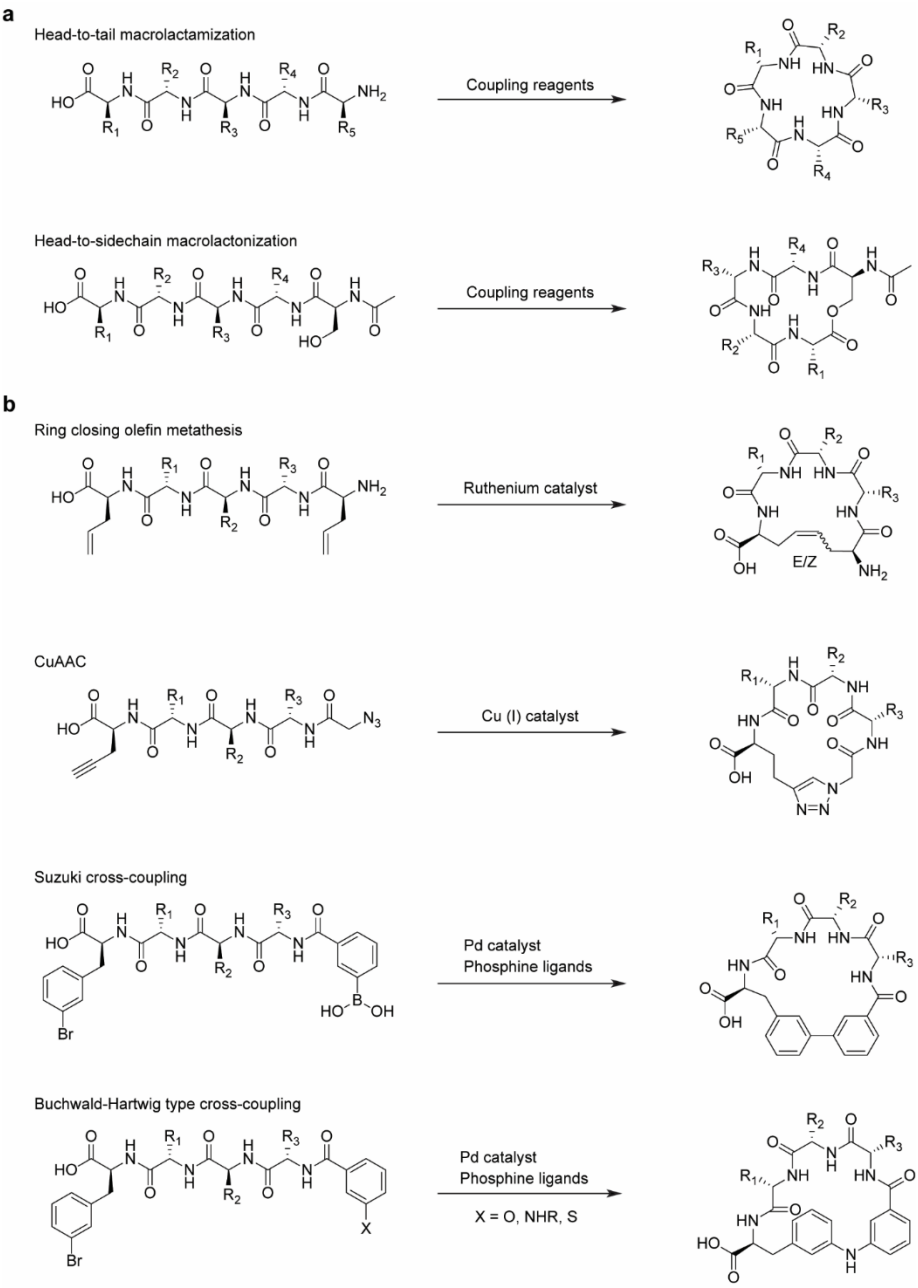
A very powerful macrocyclization reaction is nucleophilic substitution of alkyl halides with thiol or amine nucleophiles (Figure 12c). The group of Christian Heinis has utilized this chemistry extensively for the synthesis of monocyclic and bicyclic peptides derived from phage display. Bis- and tris-electrophilic linkers are utilized to react with cysteine side chains to generate the cyclic peptides.[76,77] The reaction is extremely robust, and can be performed on-

phage. Many potent cyclic peptides have been generated using this method, including coagulation factor XIa inhibitors, and interleukin-23 receptor binders.[78,79] More recently, the group has utilized bis-electrophilic linkers for the synthesis of smaller macrocycles with more favorable physicochemical properties. Both thiol to amine[80,81], and thiol to thiol[82] linkages have been used to generate macrocycle libraries that can be screened without prior purification (discussed further in the next section). Other groups have utilized S_N2 reactions for the generation of macrocycles. Suga and coworkers frequently utilize an N-terminal chloroacetic acid moiety for reaction with C-terminal cysteine residues to cyclize mRNA display derived peptides containing unnatural amino acids.[83,84,85] Ellman has used similar chemistry to cyclize smaller peptide/peptoids as β-turn mimetics.[65] Finally, nucleophilic aromatic substitution reactions have also been used to generate macrocycles. For example, the group of Hyun-Suk Lim utilized cyanuric chloride to react with thiols and amines, and generate cyclic peptoids.[86,87]

Disulfide bond formation is another way to generate macrocycles (Figure 12d). This motif is found in some important peptide therapeutics such as oxytocin and vasopressin. Disulfide bond generation often occurs with high conversion, making it an attractive cyclization option. Peptides containing two or more thiols can be treated with a variety of reagents to induce intramolecular disulfide bonding, including DMSO,[88] halogen based oxidants such as *N*-chlorosuccinamide[89] and I_2,[90] and the 3-nitro-2-pyridinesulfenyl group (Npys)[91]. Many of these reactions can be performed on either solid phase or in solution. Orthogonal cysteine protecting groups allow for selective disulfide bond formation in peptides with more than two cysteines.[92] It is important to note that disulfide bonds can be reduced intracellularly by glutathione, and therefore it may be necessary to replace the group with a non-reducible analog for compounds intended for therapeutic purposes. Analogs such as thioethers, 1,2,3-triazoles, and alkanes have been utilized with success.[93]

An important strategy in the synthesis of macrocycles is the use of cyclative cleavage from solid phase for the direct generation of macrocycles. This method offers the advantage of simultaneous cleavage and cyclization (fewer synthetic steps), along with higher macrocycle purity (only the desired cyclic product will be eluted from solid phase). Various chemical reactions can be used to perform cyclative release. One such case is macrolactamization from Kenner's "safety catch" linkers. In these systems, a stable C-terminal carboxylate is rendered reactive via alkylation with iodoacetonitrile. Mild deprotection of the N-terminal amine, followed by basification results in intramolecular amide bond formation. The method has been used for the synthesis of the natural product kahalalide A.[94,95] Macrolactamization can also be performed from N-acyl urea linkers, such as MeNBz. Activation of the linker with 4-nitrophenylchloroformate allows for nucleophilic attack on the C-terminal carbonyl by the thiol

of cysteine, which can then undergo acyl transfer to the neighboring amine.[96,97] The method has been used for efficient syntheses of sunflower trypsin inhibitor (SFTI-1).[98,99]



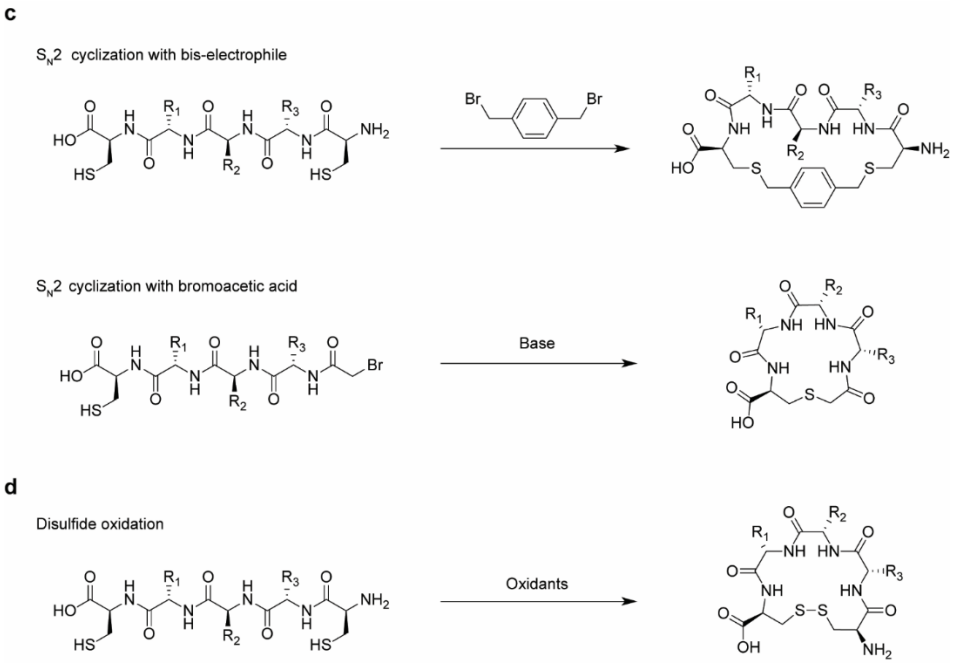


Figure 12. Examples of reactions that have been used for macrocyclization. A generic pentapeptide structure is used to demonstrate each reaction. **a** Macrolactamization and macrolactonization. **b** Metal-catalyzed reactions: ring-closing metathesis, CuAAC triazole formation, Suzuki coupling, and Buchwald-Hartwig type coupling. **c** S_N2 reaction between a dithiol peptide and a bis-electrophile, and a head-to-tail cyclization S_N2 reaction. **d** Oxidation of a dithiol peptide to a disulfide cyclized specie.

1.3 Screening technologies

1.3.1 High throughput screening (HTS) assays

The drug discovery process is often initiated by screening large numbers of molecules, including macrocycles, for a certain biochemical or biological activity relevant to the disease being studied. These high-throughput screens involve screening a large compound collection, often numbering in the millions at large pharma companies, for initial hits. This process is aided by automation, with robotic arms performing movement of tubes/plates, and pipetting compound solutions. Though tedious and expensive, HTS is still the primary method of hit discovery for pharmaceutical industry.[100,101] The assays utilized must be of very high quality if they are to be successful methods of hit identification. In general, the assays will result in the generation of a signal, which is detected by a plate reader. The detection is most often measured as fluorescence, fluorescence polarization, UV/Vis absorbance, or luminescence.[100,102] Some examples of assays using these detection methods will be discussed here.

Many enzymes are of therapeutic interest due to their involvement in key biological processes that lead to disease. One way to screen for enzyme inhibitors is to measure the activity of the proteins in the presence of compounds. Generally, a labeled substrate is present in the assay well, and the enzymatic activity of the protein on the substrate will result in the production of a signal. For example, in protease assays, often the substrate is an aminomethylcoumarin coupled peptide whose sequence is recognized by the enzyme. Upon hydrolysis of the amide bond by proteolytic action of the enzyme, the free aminocoumarin becomes fluorescent. In the presence of an inhibitor, the enzyme will hydrolyze less substrate. The increases in fluorescent signal over time can be measured, and normalized to that of a control to obtain the percent protease activity.[102,103]

In contrast to enzymes, many important therapeutic targets act not through catalytic activity, but via binding interactions. These include signaling PPIs important for controlling apoptosis and gene expression.[104,105] For these proteins, competitive fluorescence polarization (FP) and time-resolved Förster resonance energy transfer (TR-FRET) based assays are more relevant. In the case of FP, a fluorescent probe (small molecule or peptide) bound to target protein can be displaced by a competitive binder, which leads to a decrease in the observed fluorescence polarization due to the change in rotation of the small probe vs. large probe/protein complex.[106] In TR-FRET based systems, a fluorescent probe bound to protein is again used. In this case, an antibody-europium/terbium chelate is also used, frequently bound to an affinity tag on the protein. When excited by UV, the lanthanide donor complex and

acceptor fluorescent probe undergo a FRET transfer. Displacement of the probe will prevent this transfer from occurring, and a decrease in signal is observed. This method has the advantage of utilizing moieties with long fluorescence half-lives, preventing interference from fluorescent screening compounds.[100,107] As macrocycles can be particularly useful for the inhibition of PPIs, these competitive binding assays are of higher importance for macrocycle discovery.

As discussed in the chapter about macrocycle libraries, the number of macrocyclic compounds that are available for screening is limited. While microwell plate based screens can in principle be used to screen millions of compounds, this is not possible in the case of macrocycles due to the limited number of them in existing libraries.

1.3.2 One-bead-one-compound (OBOC) libraries

A method that allows for the construction of large libraries is OBOC. In this strategy, libraries are synthesized in a combinatorial fashion on solid phase resin using split and pool technique such that at the end, each bead contains many copies of a single compound.[108] After synthesis, the resin can be pooled and screened for binders by adding the desired protein of interest. Detection of hits is usually performed using colorimetric assays. For example, protein bound to solid phase via the inhibitor can be conjugated to alkaline phosphatase through its affinity tag, and adding the chromogenic substrate 5-bromo-4-chloro-3-indolyl phosphate will result in the production of visible color. The colored beads can then be identified and separated.[109]

As many times the OBOC library consists of amino acid derived peptides, Edman degradation and MS analysis can be used to identify the sequence of a hit compound from an isolated bead. However, this method only works for α-amino acids due to the mechanism of the Edman degradation process. For non-α-amino acids, building blocks other than amino acids, and for cyclic peptide macrocycles, other strategies must be used. One option is analyzing the isolated compound by MS/MS analysis. For peptide macrocycles, it may be more desirable to use an alternative method due to their complexity. It is possible to synthesize a peptide sequence simultaneously on two linkers on the same bead. Using orthogonal protecting groups, only one of these linear peptides can be cyclized, while the other remains linear for Edman degradation analysis (Figure 13).[109,110] Non peptidic macrocycle libraries can also be synthesized. A 161,000-member macrocyclic peptoid library was synthesized and screened against Skp2, an important regulatory protein, and a low micromolar binder was identified. Rather than generating a parallel amino acid sequence, in this case the cyclic peptoids were linearized by addition of oxidant, cleaved, then analyzed by MS/MS analysis.[111]

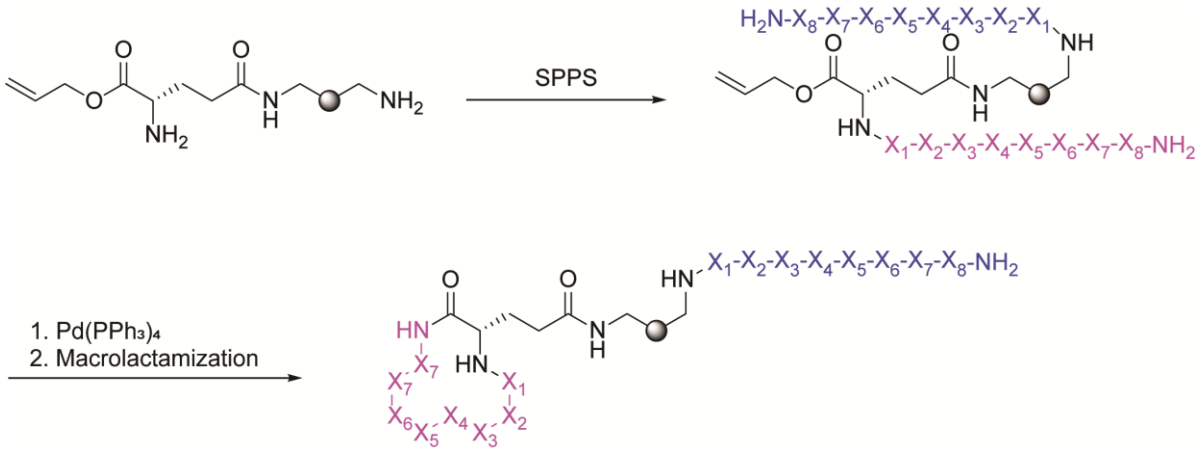


Figure 13. The OBOC macrocycle synthesis strategy. Synthesis on two linkers simultaneously allows for cyclization to obtain the macrocycle library on one end (pink), while utilizing the other for Edman degradation and identification of the chemical structure (blue). Inspired by [110].

1.3.3 DNA-encoded chemical libraries (DELs)

First described by Brenner and Lerner[112], DELs are a powerful tool for the generation of vast libraries, numbering into the billions of different compounds. The general principle of DELs is that a small molecule compound is built on a linkage to a unique sequence of DNA, which encodes for the structure of the small molecule. The libraries are screened using affinity selection to immobilized targets, and non-binders are washed away. The bound small molecule compounds can then be identified by sequencing the DNA tags. Generally, two main synthesis strategies are utilized: direct DNA recording, and DNA templated synthesis. In DNA recording, a strand of DNA with a chemical handle is split into multiple wells and reacted with chemical building blocks. This step is followed by the ligation of a unique DNA strand, which encodes for the chemical synthesis step. The library is then pooled, purified, and split again for the next cycle of synthesis (Figure 14a). DNA templated synthesis also starts with a reactive handle on DNA. However, in this case, each chemical building block used is also attached via a labile linkage to a strand of DNA that is complementary to a portion of the main encoding strand. Thus annealing the two strands brings the building block into proximity of the reactive handle, ensuring chemical reaction will only take place on the desired DNA molecule (Figure 14b). This method has the benefit of allowing all reactions to take place in a single container, but requires more work to prepare all building block-DNA conjugates.[113,114]

Examples of DNA-encoded libraries of macrocycles have been reported. A recent example utilized DNA recording synthesis to generate a 1.4 million member macrocycle DEL library, which utilized diverse amino acid building blocks to generate over 2000 unique

scaffolds, along with high side-chain diversity. The library was selected against α-1-acid glycoprotein, and low micromolar binders were identified.[115] The group of David Liu, a pioneer of DNA templated synthesis, recently synthesized a library of 256,000 macrocycles cyclized using the Wittig reaction to generate carbonyl conjugated alkene moieties. The library was selected against insulin-degrading enzyme, and a 40 nM inhibitor was found.[116]

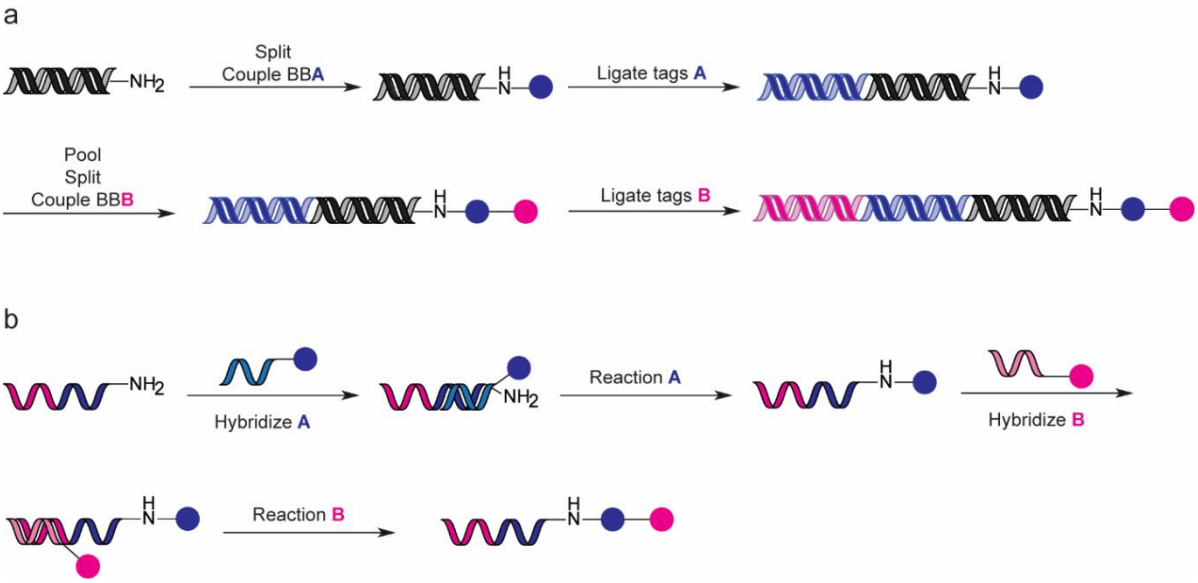


Figure 14. Most common DEL synthesis strategies. a DNA recoding method: building block A (BBA) is coupled to a DNA headpiece, followed by the ligation of a coding DNA unique to each BB. The library is pooled, then split again. Synthesis continues in this manner to construct the DEL. **b** Templated synthesis: building blocks linked to a DNA strand via a reactive linker are hybridized with an encoding DNA strand with an amine handle. The proximity of the building block to the amine ensures selective reaction, allowing all reactions to be performed in one pot.

1.3.4 Screening of crude reactions

A strategy that has been utilized rather recently by our laboratory for screening large libraries of macrocycles is based on assaying crude reaction mixtures in microwell plates, rather than individually purified compounds. This method has the benefit of greatly increasing synthetic throughput, as chromatographic purification is time and resource consuming. Of course, the reactions utilized during synthesis must result in high yields, and not produce byproducts that could interfere with biochemical assays. The strategy could be used as a primary screening method, or as a tool for hit optimization. Examples will be given below.

Our laboratory has synthesized libraries of macrocycles by combinatorially cyclizing *m* short linear peptides with *n* bis-electrophilic linkers to generate *m* × *n* macrocyclic products. For example a library of 1,284 linear peptides was cyclized with 7 linkers to give 8,988 macrocycles, which were screened without purification against thrombin and KLK5. A 42 nM

thrombin inhibitor, and high nanomolar KLK5 inhibitors were identified.[80] In another study, our group then further increased the complexity of the crude mixtures by performing two diversification steps in series. Fifteen linear peptides containing cysteine at the C-terminus, and bromoacetic acid at the N-terminus were each first reacted with 42 amines to give 630 peptide/peptoids, which were then cyclized with 6 bis-electrophilic linkers (Figure 15). Screening of the crude 3,780-member library identified a 4 nM thrombin inhibitor, thus demonstrating the power or the method.[81]

The group of Chi-Huey Wong has utilized screening of crude reaction mixtures extensively for the optimization of previously-identified inhibitor core structures (peptides or small molecules, but not macrocyclic compounds). A simple amide bond formation reaction was utilized for the improvement of α-fucosidase inhibitors based on a fuconojirimycin scaffold. The scaffold mimics the transition state of fucose during hydrolysis by the enzyme, and had a potency of 1.6 nM. Acylation with 60 carboxylic acids led to the discovery of a 0.6 pM inhibitor, representing a 35,000-fold improvement in potency.[117] Such a strategy using amide bond formation was also used for the optimization of HIV protease inhibitors, resulting in 1000-fold improvement in potency.[118] The group has also utilized more exotic chemical reactions, including CuAAC to make 1,2,3-triazoles, the Pictet-Spengler reaction, alkylation of nucleotides and sulfonamides using TBAF as base, and epoxide ring opening by amines. The reactions were shown to be compatible with enzymatic assays, FRET assays, and even screening in cells.[119] While Wong did not utilize these diversification reactions for the synthesis of macrocycles, the work demonstrates that such reactions can be used for synthesis and screening of crude macrocycles, as will be discussed in chapter 4.

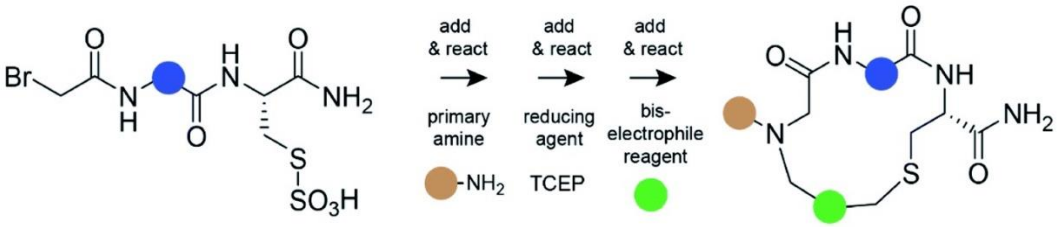


Figure 15. The "add & react" strategy used by Mothukuri et al. A peptide with N-terminal bromoacetamide is first reacted with 42 primary amines, then cyclized in an "S-to-N" fashion with 6 bis-electrophiles to generate a 3,780-member library. The library was then screened without prior purification. Reproduced from [81] with permission from the Royal Society of Chemistry.

1.3.5 Nanoscale synthesis and screening

So far the examples of crude reaction screens given have been performed on microliter-scale by pipetting, sometimes with the aid of automated liquid handlers. However, it is possible to perform these reactions on even smaller scale using more advanced instruments.

Combining nanoliter-scale synthesis with crude reaction screening represents an even further increase in throughput, while also minimizing consumption of precious reagents, and reducing costs since no pipette tips are typically used. These benefits are especially important for macrocycle synthesis and screening since the reactants are more complex than small molecules, and thus more precious.

A 2018 study from Merck & Co. utilized nanoscale synthesis (50-100 nmole scale) to screen metal catalyzed cross-coupling reaction efficiency, then applied the reactions to the crude screening of 345 compounds against kinases in an affinity selection mass spectrometry (ASMS) based assay.[120] ASMS can detect protein with bound ligand via mass spectrometry, and has the benefit of high sensitivity, and compatibility with various buffers, solvents, and reagents.[121] The group utilized a SPT Labtech mosquito instrument, which uses positive displacement via a piston in linked pipette tips to dispense down to 25 nl volumes.[122] While the overall screens were of limited size, the utility of nanoscale synthesis was demonstrated by the fact that over 3,000 reactions were performed and analyzed using only 123 mg of the common reactant.[120]

The group of Dömling has extensively utilized acoustic droplet ejection (ADE) technology for screening reaction conditions and substrate reactivity in multicomponent reactions.[123,124] In contrast to the previously described mosquito instrument, ADE does not require any pipetting tips. Instead, ultrasonic waves are used to dispense nanoliter droplets from a source plate into a destination plate.[125] Recently, the group has utilized nanoscale synthesis using ADE and crude screening for identifying inhibitors of the Menin-MLL PPI. The Groebcke-Blackburn-Bienaymé three-component heterocycle forming reaction was chosen, and 1536 random compounds were synthesized on 500 nmole scale from a total set of 71 isocyanides, 53 aldehydes, and 38 cyclic amidines. The crude mixtures were screened using differential scanning fluorimetry (DSF), and a low micromolar hit was identified.[126]

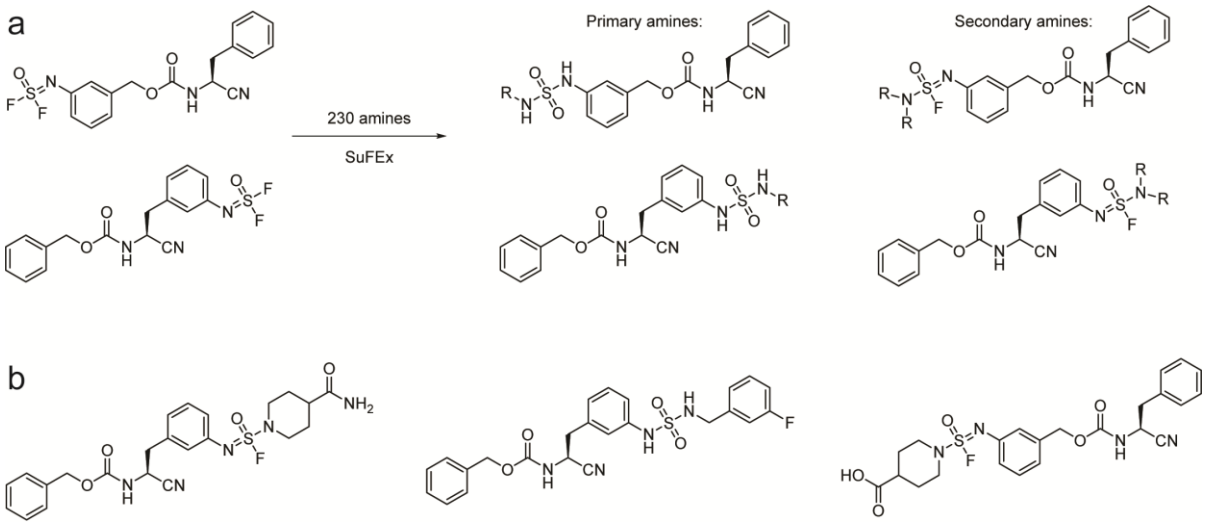


Figure 16. Optimization of an HTS hit to low-nanomolar inhibitors of SpeB using SuFEx chemistry on nanoscale and crude screening. a Two analogs of an HTS hit were fitted with reactive sulfur (VI) handles for SuFEx based chemistry, then reacted with 230 amines dispensed by ADE. In total, 460 crude compounds were screened. **b** The three double-digit nanomolar inhibitors identified using the method.[127].

In further examples, work from the Scripps Research Institute has been published which shows nanoscale synthesis and crude screening using the sulfur-fluorine exchange (SuFEx) click reaction. SuFEx utilizes hexavalent sulfur–fluorine species, such as sulfonyl fluorides, which have ideal stability and reactivity.[128] In a first example, the group diversified two weakly binding scaffolds to SpeB bacterial protease with 230 amines using nanoscale SuFEx chemistry with ADE, and screened them to identify double-digit nanomolar inhibitors (Figure 16).[127] In other example, a 7 μM inhibitor of the AF9/ENL YEATS domain proteins was diversified with 288 amines, and potency was improved to 120 nM in a FRET based screen. A next round of optimization and crude screening in living cells identified SR-0813, a 25 nM inhibitor with cell activity.[129]

Most recently, our laboratory has used ADE for the “add and react” synthesis of a 2,700-member macrocycle library, which was screened against the MDM2/P53 PPI.[130]

1.4 References

(1) Marsault, E.; Peterson, M. L. Macrocycles Are Great Cycles: Applications, Opportunities, and Challenges of Synthetic Macrocycles in Drug Discovery. *Journal of Medicinal Chemistry* **2011**, *54* (7), 1961–2004. https://doi.org/10.1021/jm1012374.

(2) Newman, D. J.; Cragg, G. M. Bioactive Macrocycles from Nature. *RSC Drug Discovery Series* **2015**, *Macrocycle* (40), 1–36.

(3) Vinogradov, A. A.; Yin, Y.; Suga, H. Macrocyclic Peptides as Drug Candidates: Recent Progress and Remaining Challenges. *Journal of the American Chemical Society* **2019**, *141* (10), 4167–4181. https://doi.org/10.1021/jacs.8b13178.

(4) Staunton, J.; Weissman, K. J. Polyketide Biosynthesis: A Millennium Review. *Natural Product Reports* **2001**, *18* (4), 380–416. https://doi.org/10.1039/a909079g.

(5) Giordanetto, F.; Kihlberg, J. Macrocyclic Drugs and Clinical Candidates: What Can Medicinal Chemists Learn from Their Properties? *Journal of Medicinal Chemistry* **2014**, *57* (2), 278–295. https://doi.org/10.1021/jm400887j.

(6) Dinos, G. P. The Macrolide Antibiotic Renaissance. *British Journal of Pharmacology* **2017**, *174* (18), 2967–2983. https://doi.org/10.1111/bph.13936.

(7) Campbell, E. A.; Korzheva, N.; Mustaev, A.; Murakami, K.; Nair, S.; Goldfarb, A.; Darst, S. A. Structural Mechanism for Rifampicin Inhibition of Bacterial RNA Polymerase. *Cell* **2001**, *104* (6), 901–912. https://doi.org/10.1016/S0092-8674(01)00286-0.

(8) Poirel, L.; Jayol, A.; Nordmann, P. Polymyxins: Antibacterial Activity, Susceptibility Testing, and Resistance Mechanisms Encoded by Plasmids or Chromosomes. *Clinical Microbiology reviews* **2017**, *30* (2), 557–596.

(9) Li, J.; Kim, S. G.; Blenis, J. Rapamycin: One Drug, Many Effects. *Cell Metabolism* **2014**, *19* (3), 373–379. https://doi.org/10.1016/j.cmet.2014.01.001.

(10) Everolimus - National Cancer Institute https://www.cancer.gov/about-cancer/treatment/drugs/everolimus (accessed Jun 27, 2021).

(11) Vendeville, S.; Cummings, M. D. *Synthetic Macrocycles in Small-Molecule Drug Discovery*, 1st ed.; Copyright © 2013 Elsevier Inc. All rights reserved., 2013; Vol. 48. https://doi.org/10.1016/B978-0-12-417150-3.00023-5.

(12) Mallinson, J.; Collins, I. Macrocycles in New Drug Discovery. *Future Medicinal Chemistry* **2012**, *4* (11), 1409–1438. https://doi.org/10.4155/fmc.12.93.

(13) Basit, S.; Ashraf, Z.; Lee, K.; Latif, M. First Macrocyclic 3rd-Generation ALK Inhibitor for Treatment of ALK/ROS1 Cancer: Clinical and Designing Strategy Update of Lorlatinib. *European Journal of Medicinal Chemistry* **2017**, *134*, 348–356. https://doi.org/10.1016/j.ejmech.2017.04.032.

(14) Pirker, R.; Filipits, M. From Crizotinib to Lorlatinib: Continuous Improvement in Precision Treatment of ALK-Positive Non-Small Cell Lung Cancer. *ESMO Open* **2019**, *4* (5), 1–4. https://doi.org/10.1136/esmoopen-2019-000548.

(15) 18. Zotiraciclib (TG02) Plus Dose-Dense or Metronomic Temozolomide Followed by Randomized Phase II Trial of Zotiraciclib (TG02) Plus Temozolomide Versus Temozolomide Alone in Adults With Recurrent Anaplastic Astrocytoma and Glioblastoma. . ClinicalTrials. https://www.clinicaltrials.gov/ct2/show/NCT02942264. (accessed Jun 27, 2021).

(16) Villar, E. A.; Beglov, D.; Chennamadhavuni, S.; Porco, J. A.; Kozakov, D.; Vajda, S.; Whitty, A. How Proteins Bind Macrocycles. *Nature Chemical Biology* **2014**, *10* (9), 723–731. https://doi.org/10.1038/nchembio.1584.

(17) Driggers, E. M.; Hale, S. P.; Lee, J.; Terrett, N. K. The Exploration of Macrocycles for Drug Discovery - An Underexploited Structural Class. *Nature Reviews Drug Discovery* **2008**, *7* (7), 608–624. https://doi.org/10.1038/nrd2590.

(18) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Advanced Drug Delivery Reviews* **2012**, *46* (1–3), 3–26. https://doi.org/10.1016/j.addr.2012.09.019.

(19) Veber, D. F.; Johnson, S. R.; Cheng, H. Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. Molecular Properties That Influence the Oral Bioavailability of Drug Candidates. *Journal of Medicinal Chemistry* **2002**, *45* (12), 2615–2623. https://doi.org/10.1021/jm020017n.

(20) Poongavanam, V.; Doak, B. C.; Kihlberg, J. Opportunities and Guidelines for Discovery of Orally Absorbed Drugs in beyond Rule of 5 Space. *Current Opinion in Chemical Biology* **2018**, *44*, 23–29. https://doi.org/10.1016/j.cbpa.2018.05.010.

(21) Doak, B. C.; Over, B.; Giordanetto, F.; Kihlberg, J. Oral Druggable Space beyond the Rule of 5: Insights from Drugs and Clinical Candidates. *Chemistry and Biology* **2014**, *21* (9), 1115–1142. https://doi.org/10.1016/j.chembiol.2014.08.013.

(22) Rossi Sebastiano, M.; Doak, B. C.; Backlund, M.; Poongavanam, V.; Over, B.; Ermondi, G.; Caron, G.; Matsson, P.; Kihlberg, J. Impact of Dynamically Exposed Polarity on Permeability and Solubility of Chameleonic Drugs beyond the Rule of 5. *Journal of Medicinal Chemistry* **2018**, *61* (9), 4189–4202. https://doi.org/10.1021/acs.jmedchem.8b00347.

(23) Danelius, E.; Poongavanam, V.; Peintner, S.; Wieske, L. H. E.; Erdélyi, M.; Kihlberg, J. Solution Conformations Explain the Chameleonic Behaviour of Macrocyclic Drugs. *Chemistry - A European Journal* **2020**, *26* (23), 5231–5244. https://doi.org/10.1002/chem.201905599.

(24) Over, B.; Matsson, P.; Tyrchan, C.; Artursson, P.; Doak, B. C.; Foley, M. A.; Hilgendorf, C.; Johnston, S. E.; Lee, M. D.; Lewis, R. J.; McCarren, P.; Muncipinto, G.; Norinder, U.; Perry, M. W. D.; Duvall, J. R.; Kihlberg, J. Structural and Conformational Determinants of Macrocycle Cell Permeability. *Nature Chemical Biology* **2016**, *12* (12), 1065–1074. https://doi.org/10.1038/nchembio.2203.

(25) Dougherty, P. G.; Sahni, A.; Pei, D. Understanding Cell Penetration of Cyclic Peptides. *Chemical Reviews* **2019**, *119* (17), 10241–10287. https://doi.org/10.1021/acs.chemrev.9b00008.

(26) Alex, A.; Millan, D. S.; Perez, M.; Wakenhut, F.; Whitlock, G. A. Intramolecular Hydrogen Bonding to Improve Membrane Permeability and Absorption in beyond

Rule of Five Chemical Space. *MedChemComm* **2011**, *2* (7), 669–674. https://doi.org/10.1039/c1md00093d.

(27) Zaretsky, S.; Yudin, A. K. Macrocyclic Peptides for Intracellular Drug Targets; Srivastava, V., Ed.; The Royal Society of Chemistry, 2017; pp 141–171.

(28) Wang, C. K.; Northfield, S. E.; Colless, B.; Chaousis, S.; Hamernig, I.; Lohman, R. J.; Nielsen, D. S.; Schroeder, C. I.; Liras, S.; Price, D. A.; Fairlie, D. P.; Craik, D. J. Rational Design and Synthesis of an Orally Bioavailable Peptide Guided by NMR Amide Temperature Coefficients. *Proceedings of the National Academy of Sciences of the United States of America* **2014**, *111* (49), 17504–17509. https://doi.org/10.1073/pnas.1417611111.

(29) White, T. R.; Renzelman, C. M.; Rand, A. C.; Rezai, T.; McEwen, C. M.; Gelev, V. M.; Turner, R. A.; Linington, R. G.; Leung, S. S. F.; Kalgutkar, A. S.; Bauman, J. N.; Zhang, Y.; Liras, S.; Price, D. A.; Mathiowetz, A. M.; Jacobson, M. P.; Lokey, R. S. On-Resin N-Methylation of Cyclic Peptides for Discovery of Orally Bioavailable Scaffolds. *Nature Chemical Biology* **2011**, *7* (11), 810–817. https://doi.org/10.1038/nchembio.664.

(30) Webster, A. M.; Cobb, S. L. Recent Advances in the Synthesis of Peptoid Macrocycles. *Chemistry - A European Journal* **2018**, *24* (30), 7560–7573. https://doi.org/10.1002/chem.201705340.

(31) Kwon, Y. U.; Kodadek, T. Quantitative Evaluation of the Relative Cell Permeability of Peptoids and Peptides. *Journal of the American Chemical Society* **2007**, *129* (6), 1508–1509. https://doi.org/10.1021/ja0668623.

(32) Boehm, M.; Beaumont, K.; Jones, R.; Kalgutkar, A. S.; Zhang, L.; Atkinson, K.; Bai, G.; Brown, J. A.; Eng, H.; Goetz, G. H.; Holder, B. R.; Khunte, B.; Lazzaro, S.; Limberakis, C.; Ryu, S.; Shapiro, M. J.; Tylaska, L.; Yan, J.; Turner, R.; Leung, S. S. F.; Ramaseshan, M.; Price, D. A.; Liras, S.; Jacobson, M. P.; Earp, D. J.; Lokey, R. S.; Mathiowetz, A. M.; Menhaji-Klotz, E. Discovery of Potent and Orally Bioavailable Macrocyclic Peptide-Peptoid Hybrid CXCR7 Modulators. *Journal of Medicinal Chemistry* **2017**, *60* (23), 9653–9663. https://doi.org/10.1021/acs.jmedchem.7b01028.

(33) Moore, R. Cyclic Peptides and Depsipeptides from Cyanobacteria: A Review. *Journal of Industrial Microbiology* **1996**, *16*, 134–143.

(34) Ueda, H.; Nakajima, H.; Hori, Y.; Fugita, T.; Nishimura, M.; Goto, T.; Okuhara, M. A Novel Antitumor Bicyclic Depsipeptide Produced By Chromobacterium Violaceum. *The Journal of Antibiotics* **1993**, *47* (3), 301–310.

(35) Rezai, T.; Yu, B.; Millhauser, G. L.; Jacobson, M. P.; Lokey, R. S. Testing the Conformational Hypothesis of Passive Membrane Permeability Using Synthetic Cyclic Peptide Diastereomers. *Journal of the American Chemical Society* **2006**, *128* (8), 2510–2511. https://doi.org/10.1021/ja0563455.

(36) Hickey, J. L.; Zaretsky, S.; St Denis, M. A.; Kumar Chakka, S.; Morshed, M. M.; Scully, C. C. G.; Roughton, A. L.; Yudin, A. K. Passive Membrane Permeability of Macrocycles Can Be Controlled by Exocyclic Amide Bonds. *Journal of Medicinal Chemistry* **2016**, *59* (11), 5368–5376. https://doi.org/10.1021/acs.jmedchem.6b00222.

(37) Räder, A. F. B.; Reichart, F.; Weinmüller, M.; Kessler, H. Improving Oral Bioavailability of Cyclic Peptides by N-Methylation. *Bioorganic and Medicinal Chemistry* **2018**, *26* (10), 2766–2773. https://doi.org/10.1016/j.bmc.2017.08.031.

(38) Sivanathan, S.; Scherkenbeck, J. Cyclodepsipeptides: A Rich Source of Biologically Active Compounds for Drug Research. *Molecules* **2014**, *19* (8), 12368–12420. https://doi.org/10.3390/molecules190812368.

(39) Gentilucci, L.; De Marco, R.; Cerisoli, L. Chemical Modifications Designed to Improve Peptide Stability: Incorporation of Non-Natural Amino Acids, Pseudo-Peptide Bonds, and Cyclization. *Current Pharmaceutical Design* **2010**, *16* (28), 3185–3203. https://doi.org/10.2174/138161210793292555.

(40) Terrett, N. K. Methods for the Synthesis of Macrocycle Libraries for Drug Discovery. *Drug Discovery Today: Technologies* **2010**, *7* (2), e97–e104. https://doi.org/10.1016/j.ddtec.2010.06.002.

(41) Ferrié, L.; Fenneteau, J.; Figadère, B. Total Synthesis of the Marine Macrolide Amphidinolide F. *Organic Letters* **2018**, *20* (11), 3192–3196. https://doi.org/10.1021/acs.orglett.8b01020.

(42) DeGruyter, J. N.; Maio, W. A. The Taumycin A Macrocycle: Asymmetric Total Synthesis and Revision of Relative Stereochemistry. *Organic Letters* **2014**, *16* (19), 5196–5199. https://doi.org/10.1021/ol5025585.

(43) Maddess, M. L.; Tackett, M. N.; Watanabe, H.; Brennan, P. E.; Spilling, C. D.; Scott, J. S.; Osborn, D. P.; Ley, S. V. Total Synthesis of Rapamycin. *Angewandte Chemie - International Edition* **2007**, *46* (4), 591–597. https://doi.org/10.1002/anie.200604053.

(44) Seiple, I. B.; Zhang, Z.; Jakubec, P.; Langlois-Mercier, A.; Wright, P. M.; Hog, D. T.; Yabu, K.; Allu, S. R.; Fukuzaki, T.; Carlsen, P. N.; Kitamura, Y.; Zhou, X.; Condakes, M. L.; Szczypiński, F. T.; Green, W. D.; Myers, A. G. A Platform for the Discovery of New Macrolide Antibiotics. *Nature* **2016**, *533* (7603), 338–345. https://doi.org/10.1038/nature17967.

(45) Guo, Z.; Hong, S. Y.; Wang, J.; Rehan, S.; Liu, W.; Peng, H.; Das, M.; Li, W.; Bhat, S.; Peiffer, B.; Ullman, B. R.; Tse, C. M.; Tarmakova, Z.; Schiene-Fischer, C.; Fischer, G.; Coe, I.; Paavilainen, V. O.; Sun, Z.; Liu, J. O. Rapamycin-Inspired Macrocycles with New Target Specificity. *Nature Chemistry* **2019**, *11* (3), 254–263. https://doi.org/10.1038/s41557-018-0187-4.

(46) Marcaurelle, L. A.; Comer, E.; Dandapani, S.; Duvall, J. R.; Gerard, B.; Kesavan, S.; Lee, M. D.; Liu, H.; Lowe, J. T.; Marie, J. C.; Mulrooney, C. A.; Pandya, B. A.; Rowley, A.; Ryba, T. D.; Suh, B. C.; Wei, J.; Young, D. W.; Akella, L. B.; Ross, N. T.; Zhang, Y. L.; Fass, D. M.; Reis, S. A.; Zhao, W. N.; Haggarty, S. J.; Palmer, M.; Foley, M. A. An Aldol-Based Build/Couple/Pair Strategy for the Synthesis of Medium- and Large-Sized Rings: Discovery of Macrocyclic Histone Deacetylase Inhibitors. *Journal of the American Chemical Society* **2010**, *132* (47), 16962–16976. https://doi.org/10.1021/ja105119r.

(47) Macrocycle Screening Library Compounds https://www.chembridge.com/screening_libraries/targeted_libraries/macrocycle-screening-library/index.php (accessed Jun 27, 2021).

(48) Macrocycles http://www.asinex.com/libraries-macrocyclic-5-html/ (accessed Jun 27, 2021).

(49) Macrocycle Discovery Platform https://www.polyphor.com/macrocycle-discovery-platform/ (accessed Jun 27, 2021).

(50) Galloway, W. R. J. D.; Isidro-Llobet, A.; Spring, D. R. Diversity-Oriented Synthesis as a Tool for the Discovery of Novel Biologically Active Small Molecules. *Nature Communications* **2010**, *1* (6), 1–13. https://doi.org/10.1038/ncomms1081.

(51) Burke, M. D.; Schreiber, S. L. A Planning Strategy for Diversity-Oriented Synthesis. *Angewandte Chemie - International Edition* **2004**, *43* (1), 46–58. https://doi.org/10.1002/anie.200300626.

(52) Mortensen, K. T.; Osberger, T. J.; King, T. A.; Sore, H. F.; Spring, D. R. Strategies for the Diversity-Oriented Synthesis of Macrocycles. *Chemical Reviews* **2019**, *119* (17), 10288–10317. https://doi.org/10.1021/acs.chemrev.9b00084.

(53) Ciardiello, J. J.; Galloway, W. R. J. D.; O'Connor, C. J.; Sore, H. F.; Stokes, J. E.; Wu, Y.; Spring, D. R. An Expedient Strategy for the Diversity-Oriented Synthesis of Macrocyclic Compounds with Natural Product-like Characteristics. *Tetrahedron* **2016**, *72* (25), 3567–3578. https://doi.org/10.1016/j.tet.2015.10.061.

(54) Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *Journal of the American Chemical Society* **1963**, *85* (14), 2149–2154. https://doi.org/10.1021/ja00897a025.

(55) Jensen, K. J.; Shelton, P. T.; Pedersen, S. L. *Peptide Synthesis and Applications*, second edi.; Jensen, K. J., Pedersen, S. L., Eds.; Humana Press, 2013.

(56) El-Faham, A.; Albericio, F. Peptide Coupling Reagents, More than a Letter Soup. *Chemical Reviews* **2011**, *111* (11), 6557–6602. https://doi.org/10.1021/cr100048w.

(57) Valeur, E.; Bradley, M. Amide Bond Formation: Beyond the Myth of Coupling Reagents. *Chemical Society Reviews* **2009**, *38* (2), 606–631. https://doi.org/10.1039/b701677h.

(58) Lundt, B. F.; Johansen, Nils, L.; Volund, A.; Markussen, J. Removal of T-Butyl and t-Butoxycarbonyl Protecting Groups with Trifluoroacetic Acid. *Int. J. Peptide Protein Res* **1978**, *12*, 258–268.

(59) Yang, Y. *Peptide Global Deprotection/Scavenger-Induced Side Reactions*; Elsevier Inc., 2016. https://doi.org/10.1016/b978-0-12-801009-9.00003-3.

(60) Guillier, F.; Orain, D.; Bradley, M. Linkers and Cleavage Strategies in Solid-Phase Organic Synthesis and Combinatorial Chemistry. *Chemical Reviews* **2000**, *100* (6), 2091–2157. https://doi.org/10.1021/cr980040+.

(61) Chhabra, S. R.; Khan, A. N.; Bycroft, B. W. Versatile Dde-Based Primary Amine Linkers for Solid Phase Synthesis. *Tetrahedron Letters* **1998**, *39* (21), 3585–3588. https://doi.org/10.1016/S0040-4039(98)00555-3.

(62) Marshall, D. L.; Liener, I. E. A Modified Support for Solid-Phase Peptide Synthesis Which Permits the Synthesis of Protected Peptide Fragments. *Journal of Organic Chemistry* **1970**, *35* (3), 867–868. https://doi.org/10.1021/jo00828a089.

(63) Kenner, G. W.; McDermott, J. R.; Sheppard, R. C. The Safety Catch Principle in Solid Phase Peptide Synthesis. *Journal of the Chemical Society D: Chemical Communications* **1971**, No. 12, 636–637. https://doi.org/10.1039/C29710000636.

(64) Backes, B. J.; Ellman, J. A. An Alkanesulfonamide “safety-Catch” Linker for Solid-Phase Synthesis. *Journal of Organic Chemistry* **1999**, *64* (7), 2322–2330. https://doi.org/10.1021/jo981990y.

(65) Virgilio, A. A.; Schfirer, S. C.; Ellman, J. A. Expedient Solid-Phase Synthesis of Putative B-Turn Mimetics Incorporating the I+1, I+2, and I+3 Sidechains. **1996**, *37* (39), 6961–6964.

(66) Tegge, W.; Bautsch, W.; Frank, R. Synthesis of Cyclic Peptides and Peptide Libraries on a New Disulfide Linker. *Journal of Peptide Science* **2007**, *13* (10), 693–699. https://doi.org/10.1002/psc.879.

(67) White, C. J.; Yudin, A. K. Contemporary Strategies for Peptide Macrocyclization. *Nature Chemistry* **2011**, *3* (7), 509–524. https://doi.org/10.1038/nchem.1062.

(68) Chow, H. Y.; Zhang, Y.; Matheson, E.; Li, X. Ligation Technologies for the Synthesis of Cyclic Peptides. *Chemical Reviews* **2019**, *119* (17), 9971–10001. https://doi.org/10.1021/acs.chemrev.8b00657.

(69) Wen, S.; Packham, G.; Ganesan, A. Macrolactamization versus Macrolactonization: Total Synthesis of FK228, the Depsipeptide Histone Deacetylase Inhibitor. *Journal of Organic Chemistry* **2008**, *73* (23), 9353–9361. https://doi.org/10.1021/jo801866z.

(70) Schafmeister, C. E.; Po, J.; Verdine, G. L. An All-Hydrocarbon Cross-Linking System for Enhancing the Helicity and Metabolic Stability of Peptides. *Journal of the American Chemical Society* **2000**, *122* (24), 5891–5892. https://doi.org/10.1021/ja000563a.

(71) Kim, Y. W.; Grossmann, T. N.; Verdine, G. L. Synthesis of All-Hydrocarbon Stapled ±-Helical Peptides by Ring-Closing Olefin Metathesis. *Nature Protocols* **2011**, *6* (6), 761–771. https://doi.org/10.1038/nprot.2011.324.

(72) Pasini, D. The Click Reaction as an Efficient Tool for the Construction of Macrocyclic Structures. *Molecules* **2013**, *18* (8), 9512–9530. https://doi.org/10.3390/molecules18089512.

(73) Kutovaya, I. V.; Nenajdenko, V. G. Synthesis of Macrocyclic Depsipeptides via the Passerini–Click Reaction Strategy. *Russian Journal of Organic Chemistry* **2020**, *56* (4), 559–568. https://doi.org/10.1134/S1070428020040016.

(74) Rivera, D. G.; Ojeda-Carralero, G. M.; Reguera, L.; Van Der Eycken, E. V. Peptide Macrocyclization by Transition Metal Catalysis. *Chemical Society Reviews* **2020**, *49* (7), 2039–2059. https://doi.org/10.1039/c9cs00366e.

(75) Heravi, M. M.; Hashemi, E. Recent Advances in Application of Intramolecular Suzuki Cross-Coupling in Cyclization and Heterocyclization. *Monatshefte fur Chemie* **2012**, *143* (6), 861–880. https://doi.org/10.1007/s00706-012-0746-0.

(76) Heinis, C.; Rutherford, T.; Freund, S.; Winter, G. Phage-Encoded Combinatorial Chemical Libraries Based on Bicyclic Peptides. *Nature Chemical Biology* **2009**, *5* (7), 502–507. https://doi.org/10.1038/nchembio.184.

(77) Kale, S. S.; Villequey, C.; Kong, X. D.; Zorzi, A.; Deyle, K.; Heinis, C. Cyclization of Peptides with Two Chemical Bridges Affords Large Scaffold Diversities. *Nature Chemistry* **2018**, *10* (7), 715–723. https://doi.org/10.1038/s41557-018-0042-7.

(78) Carle, V.; Wu, Y.; Mukherjee, R.; Kong, X.-D.; Rogg, C.; Laurent, Q.; Cecere, E.; Villequey, C.; Konakalla, M. S.; Maric, T.; Lamers, C.; Díaz-Perlas, C.; Butler, K.; Goto, J.; Stegmayr, B.; Heinis, C. Development of Selective FXIa Inhibitors Based on Cyclic Peptides and Their Application for Safe Anticoagulation. *Journal of Medicinal Chemistry* **2021**, *64* (10), 6802–6813. https://doi.org/10.1021/acs.jmedchem.1c00056.

(79) Kong, X. D.; Moriya, J.; Carle, V.; Pojer, F.; Abriata, L. A.; Deyle, K.; Heinis, C. De Novo Development of Proteolytically Resistant Therapeutic Peptides for Oral Administration. *Nature Biomedical Engineering* **2020**, *4* (5), 560–571. https://doi.org/10.1038/s41551-020-0556-3.

(80) Kale, S. S.; Bergeron-Brlek, M.; Wu, Y.; Kumar, M. G.; Pham, M. V.; Bortoli, J.; Vesin, J.; Kong, X. D.; Franco Machado, J.; Deyle, K.; Gonschorek, P.; Turcatti, G.; Cendron, L.; Angelini, A.; Heinis, C. Thiol-to-Amine Cyclization Reaction Enables Screening of Large Libraries of Macrocyclic Compounds and the Generation of Sub-Kilodalton Ligands. *Science Advances* **2019**, *5* (8), eaaw2851. https://doi.org/10.1126/sciadv.aaw2851.

(81) Mothukuri, G. K.; Kale, S. S.; Stenbratt, C. L.; Zorzi, A.; Vesin, J.; Bortoli Chapalay, J.; Deyle, K.; Turcatti, G.; Cendron, L.; Angelini, A.; Heinis, C. Macrocycle Synthesis Strategy Based on Step-Wise “Adding and Reacting” Three Components Enables Screening of Large Combinatorial Libraries. *Chemical Science* **2020**, *11* (30), 7858–7863. https://doi.org/10.1039/d0sc01944e.

(82) Wu, Y.; Zorzi, A.; Williams, J.; Heinis, C. A Releasable Disulfide-Linked Peptide Tag Facilitates the Synthesis and Purification of Short Peptides. *Chemical Communications* **2020**, *56* (19), 2917–2920. https://doi.org/10.1039/c9cc09247a.

(83) Otero-Ramirez, M. E.; Passioura, T.; Suga, H. Structural Features and Binding Modes of Thioether-Cyclized Peptide Ligands. *Biomedicines* **2018**, *6* (4). https://doi.org/10.3390/biomedicines6040116.

(84) Passioura, T.; Suga, H. A RaPID Way to Discover Nonstandard Macrocyclic Peptide Modulators of Drug Targets. *Chemical Communications* **2017**, *53* (12), 1931–1940. https://doi.org/10.1039/c6cc06951g.

(85) Iwasaki, K.; Goto, Y.; Katoh, T.; Suga, H. Selective Thioether Macrocyclization of Peptides Having the N-Terminal 2-Chloroacetyl Group and Competing Two or Three Cysteine Residues in Translation. *Organic and Biomolecular Chemistry* **2012**, *10* (30), 5783–5786. https://doi.org/10.1039/c2ob25306b.

(86) Lee, J. H.; Kim, H. S.; Lim, H. S. Design and Facile Solid-Phase Synthesis of Conformationally Constrained Bicyclic Peptoids. *Organic Letters* **2011**, *13* (19), 5012–5015. https://doi.org/10.1021/ol201773f.

(87) Lee, J. H.; Meyer, A. M.; Lim, H. S. A Simple Strategy for the Construction of Combinatorial Cyclic Peptoid Libraries. *Chemical Communications* **2010**, *46* (45), 8615–8617. https://doi.org/10.1039/c0cc03272g.

(88) Tam, J. P.; Wu, C. R.; Liu, W.; Zhang, J. W. Disulfide Bond Formation in Peptides by Dimethyl Sulfoxide. Scope and Applications. *Journal of the American Chemical Society* **1991**, *113* (17), 6657–6662. https://doi.org/10.1021/ja00017a044.

(89) Postma, T. M.; Albericio, F. N-Chlorosuccinimide, an Efficient Peptide Disulfide Bond-Forming Reagent in Aqueous Solution. *RSC Advances* **2013**, *3* (34), 14277–14280. https://doi.org/10.1039/c3ra43149e.

(90) Rietman, B. H.; Smulders, R. H. P. H.; Eggen, I. F.; van Vliet, A.; van de Werken, G.; Tesser, G. I. Protected Peptide Disulfides by Oxidative Detachment from a Support. *International Journal of Peptide and Protein Research* **1994**, *44* (3), 199–206. https://doi.org/10.1111/j.1399-3011.1994.tb00161.x.

(91) Rentier, C.; Fukumoto, K.; Taguchi, A.; Hayashi, Y. The 3-Nitro-2-Pyridinesulfenyl Group: Synthesis and Applications to Peptide Chemistry. *Journal of Peptide Science* **2017**, *23* (7–8), 496–504. https://doi.org/10.1002/psc.2964.

(92) Postma, T. M.; Albericio, F. Disulfide Formation Strategies in Peptide Synthesis. *European Journal of Organic Chemistry* **2014**, *2014* (17), 3519–3530. https://doi.org/10.1002/ejoc.201402149.

(93) Gori, A.; Gagni, P.; Rinaldi, S. Disulfide Bond Mimetics: Strategies and Challenges. *Chemistry - A European Journal* **2017**, *23* (60), 14987–14995. https://doi.org/10.1002/chem.201703199.

(94) Yang, L.; Morriello, G. Solid Phase Synthesis of “head-to-Tail” Cyclic Peptides Using a Sulfonamide “Safety-Catch” Linker: The Cleavage by Cyclization Approach. *Tetrahedron Letters* **1999**, *40* (47), 8197–8200. https://doi.org/10.1016/S0040-4039(99)01701-3.

(95) Bourel-Bonnet, L.; Rao, K. V.; Hamann, M. T.; Ganesan, A. Solid-Phase Total Synthesis of Kahalalide a and Related Analogues. *Journal of Medicinal Chemistry* **2005**, *48* (5), 1330–1335. https://doi.org/10.1021/jm049841x.

(96) Blanco-Canosa, J. B.; Dawson, P. E. An Efficient Fmoc-SPPS Approach for the Generation of Thioester Peptide Precursors for Use in Native Chemical Ligation. *Angewandte Chemie - International Edition* **2008**, *47* (36), 6957–6961. https://doi.org/10.1002/anie.200705471.

(97) Blanco-Canosa, J. B.; Nardone, B.; Albericio, F.; Dawson, P. E. Chemical Protein Synthesis Using a Second-Generation N-Acylurea Linker for the Preparation of Peptide-Thioester Precursors. *Journal of the American Chemical Society* **2015**, *137* (22), 7197–7209. https://doi.org/10.1021/jacs.5b03504.

(98) Arbour, C. A.; Belavek, K. J.; Tariq, R.; Mukherjee, S.; Tom, J. K.; Isidro-Llobet, A.; Kopach, M. E.; Stockdill, J. L. Bringing Macrolactamization Full Circle: Self-Cleaving Head-to-Tail Macrocyclization of Unprotected Peptides via Mild N-Acyl Urea Activation. *Journal of Organic Chemistry* **2019**, *84* (2), 1035–1041. https://doi.org/10.1021/acs.joc.8b02418.

(99) Gless, B. H.; Olsen, C. A. Direct Peptide Cyclization and One-Pot Modification Using the MeDbz Linker. *Journal of Organic Chemistry* **2018**, *83* (17), 10525–10534. https://doi.org/10.1021/acs.joc.8b01237.

(100) Busby, S. A.; Carbonneau, S.; Concannon, J.; Dumelin, C. E.; Lee, Y.; Numao, S.; Renaud, N.; Smith, T. M.; Auld, D. S. Advancements in Assay Technologies and

Strategies to Enable Drug Discovery. *ACS Chemical Biology* **2020**, *15* (10), 2636–2648. https://doi.org/10.1021/acschembio.0c00495.

(101) MacArron, R.; Banks, M. N.; Bojanic, D.; Burns, D. J.; Cirovic, D. A.; Garyantes, T.; Green, D. V. S.; Hertzberg, R. P.; Janzen, W. P.; Paslay, J. W.; Schopfer, U.; Sittampalam, G. S. Impact of High-Throughput Screening in Biomedical Research. *Nature Reviews Drug Discovery* **2011**, *10* (3), 188–195. https://doi.org/10.1038/nrd3368.

(102) Inglese, J.; Johnson, R. L.; Simeonov, A.; Xia, M.; Zheng, W.; Austin, C. P.; Auld, D. S. High-Throughput Screening Assays for the Identification of Chemical Probes. *Nature Chemical Biology* **2007**, *3* (8), 466–479. https://doi.org/10.1038/nchembio.2007.17.

(103) Acker, M. G.; Auld, D. S. Considerations for the Design and Reporting of Enzyme Assays in High-Throughput Screening Applications. *Perspectives in Science* **2014**, *1* (1–6), 56–73. https://doi.org/10.1016/j.pisc.2013.12.001.

(104) Bakail, M.; Ochsenbein, F. Targeting Protein-Protein Interactions, a Wide Open Field for Drug Design. *Comptes Rendus Chimie* **2016**, *19* (1–2), 19–27. https://doi.org/10.1016/j.crci.2015.12.004.

(105) Arkin, M. R.; Tang, Y.; Wells, J. A. Small-Molecule Inhibitors of Protein-Protein Interactions: Progressing toward the Reality. *Chemistry and Biology* **2014**, *21* (9), 1102–1114. https://doi.org/10.1016/j.chembiol.2014.09.001.

(106) Hall, M. D.; Yasgar, A.; Peryea, T.; Braisted, J. C.; Jadhav, A.; Simeonov, A.; Coussens, N. P. Fluorescence Polarization Assays in High-Throughput Screening and Drug Discovery: A Review. *Methods and Applications in Fluorescence* **2016**, *4* (2), 022001. https://doi.org/10.1088/2050-6120/4/2/022001.

(107) Ergin, E.; Dogan, A.; Parmaksiz, M.; Elçin, A.; Elçin, Y. Time-Resolved Fluorescence Resonance Energy Transfer [TR-FRET] Assays for Biochemical Processes. *Current Pharmaceutical Biotechnology* **2016**, *17* (14), 1222–1230. https://doi.org/10.2174/1389201017666160809164527.

(108) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierskit, W. M.; Knappt, R. J. A New Type of Synthetic Peptide Library for Identifying Ligand-Binding Activity. *Nature* **1991**, *354* (6348), 82–84. https://doi.org/10.1038/354082a0.

(109) Lam, K. S.; Lebl, M.; Krchňák, V. The “One-Bead-One-Compound” Combinatorial Library Method. *Chemical Reviews* **1997**, *97* (2), 411–448. https://doi.org/10.1021/cr9600114.

(110) Qian, Z.; Upadhyaya, P.; Pei, D. Synthesis and Screening of One-Bead-One-Compound Cyclic Peptide Libraries. In *Peptide Libraries: Methods and Protocols*; Derda, R., Ed.; Springer Science+Business Media: New York, 2015; Vol. 1248, pp 39–53. https://doi.org/10.1007/978-1-4939-2020-4_3.

(111) Oh, M.; Lee, J. H.; Moon, H.; Hyun, Y. J.; Lim, H. S. A Chemical Inhibitor of the Skp2/P300 Interaction That Promotes P53-Mediated Apoptosis. *Angewandte Chemie - International Edition* **2016**, *55* (2), 602 –606. https://doi.org/10.1002/anie.201508716.

(112) Brenner, S.; Lerner, R. A. Encoded Combinatorial Chemistry. *Proceedings of the National Academy of Sciences of the United States of America* **1992**, *89* (12), 5381–5383. https://doi.org/10.1073/pnas.89.12.5381.

(113) Goodnow, R. A.; Dumelin, C. E.; Keefe, A. D. DNA-Encoded Chemistry: Enabling the Deeper Sampling of Chemical Space. *Nature Reviews Drug Discovery* **2017**, *16* (2), 131–147. https://doi.org/10.1038/nrd.2016.213.

(114) Favalli, N.; Bassi, G.; Scheuermann, J.; Neri, D. DNA-Encoded Chemical Libraries – Achievements and Remaining Challenges. *FEBS Letters*. 2018, pp 2168–2180. https://doi.org/10.1002/1873-3468.13068.

(115) Stress, C. J.; Sauter, B.; Schneider, L. A.; Sharpe, T.; Gillingham, D. A DNA-Encoded Chemical Library Incorporating Elements of Natural Macrocycles. *Angewandte Chemie - International Edition* **2019**, *58* (28), 9570–9574. https://doi.org/10.1002/anie.201902513.

(116) Usanov, D. L.; Chan, A. I.; Maianti, J. P.; Liu, D. R. Second-Generation DNA-Templated Macrocycle Libraries for the Discovery of Bioactive Small Molecules. *Nature Chemistry* **2018**, *10* (7), 704–714. https://doi.org/10.1038/s41557-018-0033-8.

(117) Chang, C. F.; Ho, C. W.; Wu, C. Y.; Chao, T. A.; Wong, C. H.; Lin, C. H. Discovery of Picomolar Slow Tight-Binding Inhibitors of α-Fucosidase. *Chemistry and Biology* **2004**, *11* (9), 1301–1306. https://doi.org/10.1016/j.chembiol.2004.07.009.

(118) Brik, A.; Lin, Y. C.; Elder, J.; Wong, C. H. A Quick Diversity-Oriented Amide-Forming Reaction to Optimize p-Subsite Residues of HIV Protease Inhibitors. *Chemistry and Biology* **2002**, *9* (8), 891–896. https://doi.org/10.1016/S1074-5521(02)00184-9.

(119) Brik, A.; Wu, C. Y.; Wong, C. H. Microtiter Plate Based Chemistry and in Situ Screening: A Useful Approach for Rapid Inhibitor Discovery. *Organic and Biomolecular Chemistry* **2006**, *4* (8), 1446–1457. https://doi.org/10.1039/b600055j.

(120) Gesmundo, N. J.; Sauvagnat, B.; Curran, P. J.; Richards, M. P.; Andrews, C. L.; Dandliker, P. J.; Cernak, T. Nanoscale Synthesis and Affinity Ranking. *Nature* **2018**, *557* (7704), 228–232. https://doi.org/10.1038/s41586-018-0056-8.

(121) Prudent, R.; Annis, D. A.; Dandliker, P. J.; Ortholand, J. Y.; Roche, D. Exploring New Targets and Chemical Space with Affinity Selection-Mass Spectrometry. *Nature Reviews Chemistry* **2021**, *5* (1), 62–71. https://doi.org/10.1038/s41570-020-00229-2.

(122) Positive displacement pipetting https://www.sptlabtech.com/innovation/positive-displacement-pipetting/ (accessed Jul 4, 2021).

(123) Shaabani, S.; Xu, R.; Ahmadianmoghaddam, M.; Gao, L.; Stahorsky, M.; Olechno, J.; Ellson, R.; Kossenjans, M.; Helan, V.; Dömling, A. Automated and Accelerated Synthesis of Indole Derivatives on a Nano-Scale. *Green Chemistry* **2019**, *21* (2), 225–232. https://doi.org/10.1039/c8gc03039a.

(124) Wang, Y.; Shaabani, S.; Ahmadianmoghaddam, M.; Gao, L.; Xu, R.; Kurpiewska, K.; Kalinowska-Tluscik, J.; Olechno, J.; Ellson, R.; Kossenjans, M.; Helan, V.; Groves, M.; Dömling, A. Acoustic Droplet Ejection Enabled Automated Reaction Scouting. *ACS Central Science* **2019**, *5* (3), 451–457. https://doi.org/10.1021/acscentsci.8b00782.

(125) Echo Acoustic Liquid Handling https://www.labcyte.com/echo-liquid-handling (accessed Jul 4, 2021).

(126) Gao, K.; Shaabani, S.; Xu, R.; Zarganes-Tzitzikas, T.; Gao, L.; Ahmadianmoghaddam, M.; Groves, M. R.; Dömling, A. Nanoscale, Automated, High Throughput Synthesis and Screening for the Accelerated Discovery of Protein Modifiers. *RSC Medicinal Chemistry* **2021**, *12* (5), 809–818. https://doi.org/10.1039/d1md00087j.

(127) Kitamura, S.; Zheng, Q.; Woehl, J. L.; Solania, A.; Chen, E.; Dillon, N.; Hull, M. V.; Kotaniguchi, M.; Cappiello, J. R.; Kitamura, S.; Nizet, V.; Sharpless, K. B.; Wolan, D. W. Sulfur(VI) Fluoride Exchange (SuFEx)-Enabled High-Throughput Medicinal Chemistry. *Journal of the American Chemical Society* **2020**, *142* (25), 10899–10904. https://doi.org/10.1021/jacs.9b13652.

(128) Barrow, A. S.; Smedley, C. J.; Zheng, Q.; Li, S.; Dong, J.; Moses, J. E. The Growing Applications of SuFEx Click Chemistry. *Chemical Society Reviews* **2019**, *48* (17), 4731–4758. https://doi.org/10.1039/c8cs00960k.

(129) Garnar-Wortzel, L.; Bishop, T. R.; Kitamura, S.; Milosevich, N.; Asiaban, J. N.; Zhang, X.; Zheng, Q.; Chen, E.; Ramos, A. R.; Ackerman, C. J.; Hampton, E. N.; Chatterjee, A. K.; Young, T. S.; Hull, M. V.; Sharpless, K. B.; Cravatt, B. F.; Wolan, D. W.; Erb, M. A. Chemical Inhibition of ENL/AF9 YEATS Domains in Acute Leukemia. *ACS Central Science* **2021**, *7* (5), 815–830. https://doi.org/10.1021/acscentsci.0c01550.

(130) Sangouard, G.; Zorzi, A.; Wu, Y.; Ehret, E.; Schüttel, M.; Kale, S.; Diaz Perlas, C.; Vesin, J.; Chapalay, J. B.; Turcatti, G.; Heinis, C. *Angewandte Chemie - International Edition* **2021**. https://doi.org/10.1002/anie.202107815.

2 Aim of the thesis

Macrocycles are a promising modality for the development of cell permeable or orally available drugs for challenging targets. However, due to their synthetic complexity, existing macrocycle libraries are rather small, typically containing less than 10,000 compounds. Thus the chances of finding ligands to targets of interest in such libraries is rather small.

To address this gap, the goal of my thesis was to establish a new macrocycle library synthesis principle in which a large panel of *m* small cyclic peptides are combinatorially ligated in microwell plates with a large number of *n* chemical fragments, and the resulting $m \times n$ products are screened without purification. For making this approach particularly efficient and economical, I planned to perform the reactions and screens at a nanomole scale and using contactless liquid transfer.

The first step towards this end was to develop a synthetic method for the rapid generation of libraries of small cyclic peptides needed in the combinatorial diversification approach. As the largest bottleneck in macrocycle synthesis is the need to purify following cyclization, the primary criteria was that no purification be needed. Other criteria for the method included the following: capable of being performed in 96-well format, have high cyclization yield, utilize robust and broadly applicable reactions, produce no or minimal impurities, be fast and cost efficient, and be compatible with the subsequent diversification with the fragments.

In a second step, my goal was to massively expand the diversity of the above-mentioned cyclic peptide libraries. While the first goal would produce libraries in the hundreds/thousands, I wished to increase the size to tens/hundreds of thousands through the use of combinatorial modification of the macrocycle backbones, as described above. The criteria for this goal were the following: the fragments used must be commercially available, the reaction must be nearly quantitative, no or only small quantities of side products should be formed, and all components must be compatible with direct screening in biochemical assays. I also wished that the reactions could be performed using contactless acoustic dispensing, since this would be even faster, cheaper, and consume less material than traditional pipetting.

A final goal of my thesis was to apply the new method for generating libraries comprising thousands of macrocyclic compounds and to screen them against targets for identifying binders.

3 A cyclative release strategy to obtain pure cyclic peptides directly from solid phase

3.1 Summary

A cyclative release strategy to obtain pure cyclic peptides directly from solid phase

Sevan Habeshian,[1] Ganesh Sable,[1] Mischa Schüttel,[1] Manuel Merz,[1] and Christian Heinis[1]

[1]Institute of Chemical Sciences and Engineering, School of Basic Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland.

Author contributions: I, along with Christian Heinis conceptualized the method and developed the strategy. I performed proof-of-concept experiments, synthesis and screening of the peptide library, and hit compound synthesis and characterization. I also helped to make all figures, wrote the first draft of the manuscript, and contributed to editing. Ganesh Sable synthesized and analyzed the peptoid library. Mischa Schüttel prepared resins for proof-of-concept reactions, and performed reductive release experiments. Manuel Merz developed the Ellman's reagent based strategy for cyclic peptide quantification.

This chapter is based on a manuscript for publication.

3.2 Abstract

Cyclic peptides are an attractive molecular format for drug development, with more than a hundred currently in clinical studies. The development process for cyclic peptide drugs typically requires the synthesis and testing of large numbers of peptide variants, which is facilitated by automated parallel solid-phase synthesis. However, there is currently no approach that yields already pure cyclic peptides from solid phase, and the process thus remains inefficient due to the required subsequent chromatographic purification. Herein, we have developed a strategy in which only pure cyclic pep-tides are detached from the solid phase. Peptides with an N-terminal thiol group are synthesized on solid phase via a C-terminal disulfide linker, their sidechain protecting groups are removed while the peptides remain on solid phase, and the peptides are finally released via a cyclative mechanism by addition of a based that deprotonates the N-terminal thiol group and triggers an intra-molecular disulfide exchange reaction. The method yields disulfide-cyclized peptides, which is a format on which many important peptide drugs are based, such as oxytocin, vasopressin, and octreotide. We demonstrate that the method is applicable for facile synthesis in 96-well plates, and allows for synthesis and screening of hundreds of cyclic peptides.

3.3 Introduction

Cyclic peptides have received much attention in recent years for their ability to target challenging proteins to which small molecule ligands are difficult or impossible to obtain. Currently, more than 40 cyclic peptides are approved as drugs, and more than 100 are being evaluated at different stages in clinical trials.[1] The development of peptide drugs typically involves multiple iterative cycles of synthesizing dozens to hundreds of variants to improve key properties such as binding affinity, specificity, stability, pharmacokinetic properties, etc., and thus requires the preparation of large numbers of peptides. Due to the typically low-yielding macrocyclization reactions,[2] and the presence of reagents and scavengers added for peptide release and side-chain deprotection, the chromatographic purification of each individual peptide is generally required. Even in a semi-automated fashion, this purification is expensive and time consuming, and limits the number of cyclic peptide variants that can be synthesized.

Herein, we aimed at developing an approach in which essentially pure cyclic peptides are released from solid phase, and that is compatible with synthesis in 96-well plates, in order to access large libraries of pure cyclic peptides that can readily be screened. We proposed the approach shown in Figure 1a, in which peptides bearing two thiol groups at their two ends, and being disulfide-linked to the solid phase via the C-terminal thiol, are synthesized on solid phase. We envisioned that all protecting groups would be removed while the peptides are on solid phase, allowing efficient elimination of protecting groups and reagents by bead washing, and that the peptides would be released in a disulfide exchange reaction during which the peptides are cyclized. We proposed that this cyclative release could be initiated by deprotonating the N-terminally located thiol group, which would attack the disulfide bridge in an intramolecular fashion. The release mechanism would ensure that only cyclic peptide is released, omitting contamination of the product with linear peptide.

To our knowledge, the proposed strategy of cyclative disulfide release is completely new and has not been reported. Strategies that come closest to our plans are the oxidative release of thioether-immobilized peptides,[3,4] but we consider them as not suited for library generation due to the low yields, dimeric side products, and the presence of oxidants in the eluted product that would need removal by purification. Cyclative peptide release strategies were developed before using elegant strategies and yielding peptide cyclized via different chemical bonds.[5] Strategies based on cyclization from thioester linkers,[6] olefin linkers for ring-closing metathesis,[7] and Dawson's MeDbz linker for native chemical cyclization[8,9] were used to synthesize and screen peptide libraries, but the released peptides still required chromatographic purification prior to screening.

3.4 Results and discussion

In order to test the strategy proposed in Figure 1a for cyclative release of peptides from solid phase by intramolecular disulfide exchange, we synthesized the test peptide Mpa-Gly-Gln-Trp-Mea, with Mpa representing 3-mercaptopropionic acid (cysteine without the amino group) and Mea representing 2-mercaptoethylamine (cysteamine; cysteine without carboxylic acid), that was linked via a disulfide bridge to resin (Figure 1b). For this synthesis, we first conjugated the Mea building block to thiol-functionalized resin as described in the Supplementary Results, and then coupled the amino acids and Mpa using standard Fmoc chemistry. In order to test the synthesis, on-resin side chain deprotection, and cyclative release in different microenvironments, we tested five resins: two polar polyethylene glycol (PEG) resins (resins 1 and 2), one polar PEG-modified polystyrene (PS) resin (resin 3), and two apolar PS resins (resins 4 and 5) (5 μmol scale; Figure 1b and Table S1). Weighting of the resins before and after synthesis indicated that peptides were efficiently synthesized on all resins (Table S2). Side-chain deprotection followed by disulfide bond reduction released peptide with the correct mass but overall low yield (Supplementary results, Figure S1, and Table S3 and S4).

Next, we tested whether the peptide can be released from the resin via the anticipated cyclization mechanism by deprotonating the sulfhydryl group at the N-terminal end of the peptide. Incubating the resins in DMSO with the base DIPEA (150 mM) led to highly efficient release for the apolar resins 4 and 5 (Figure 1d and Table S5). The yields of peptide in the eluate were 1.05 ± 0.25 μmol (resin 4) and 2.2 ± 0.2 μmol (resin 5), and thus not too far from the maximal amount that could be expected based on the resin loading (5 μmol), and far more sufficient than required for screening. More important than the high yields were the high concentrations of 5.1 mM (resin 4) and 11.6 mM (resin 5), that allowed convenient use of the peptide for high-throughput screening without concentration or solvent or base removal. For example, stocks in DMSO may be diluted 1000-fold for testing the cyclic peptides at 10 μM, which is a typical screening concentration, leaving negligible amounts of DMSO (0.1%) and DIPEA (150 μM) in the assay that are unlikely to affect biological assays. Liquid chromatography-mass spectrometry (LC-MS) analysis of the products showed high purities of 95 ± 4% (resin 4) and 93 ± 1% (resin 5) for the disulfide-cyclized peptide (Figure 1d and S2). The only side product observed was found in small quantities (6% on average) and corresponded to a cyclic peptide dimer, most likely formed by disulfide exchange-mediated transfer of one peptide to a neighboring one on the resin and the subsequent cyclative release of a cyclic dimer.

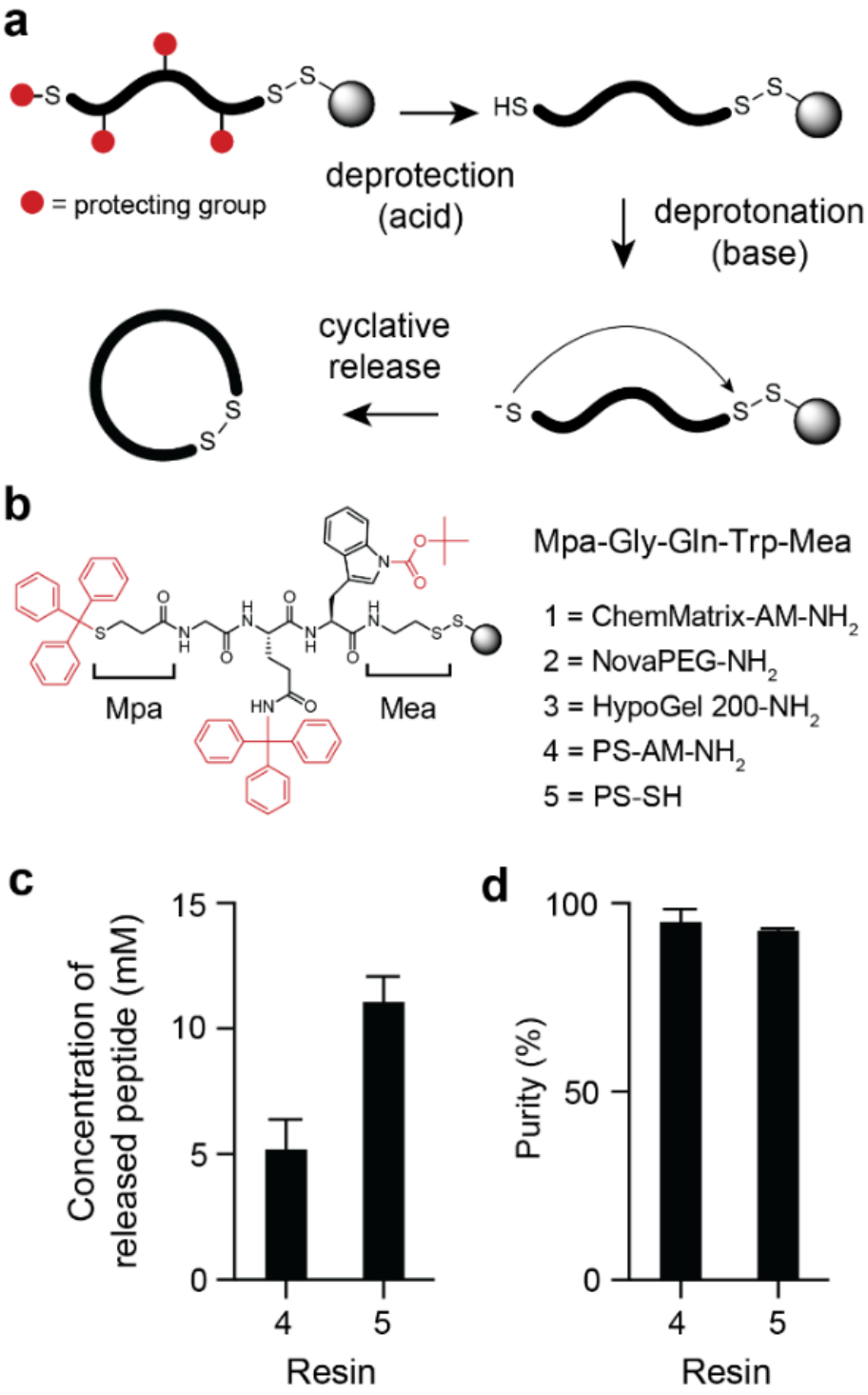


Figure 1. Cyclative disulfide release strategy. a Schematic representation of short peptides synthesized via a disulfide bridge on solid phase. Protecting groups (red) are removed on solid phase for an efficient separation of the cleavage reagents. Treatment with a base deprotonates the N-terminal thiol, which induces an intramolecular disulfide exchange to generate the cyclic product. **b** Chemical structure of test peptide Mpa-Gly-Gln-Trp-Mea disulfide-linked to a solid support and the commercial resins used. **c** Recovery of disulfide-cyclized peptide Mpa-Gly-Gln-Trp-Mea synthesized on resins 4 and 5 and released with 150 mM DIPEA in DMSO. Concentrations were determined by measuring absorbance at 280 nm. Reactions were performed in triplicate. **d** Purity of disulfide-cyclized peptide of panel c. Purity was determined by LC-MS, measuring the AUC of all species at 220 nm UV absorbance.

To assess the substrate scope of the cyclative disulfide release strategy, we synthesized four peptides with the sequence Mpa-Gly-Ala-Xaa-Mea, where “Xaa” are amino acids that imposed various levels of rigidity into the peptide backbone (Trp, Amb, Pro, Nip; Figure 2a). HPLC analysis of the products obtained by base-induced cyclative release showed a dominant peak for each one of the four peptides, and MS analysis confirmed that these main products were the desired cyclic peptides. Yields ranged from 54-100% (13.5 to 26 μmol), as calculated via absorbance and weight after lyophilization (Table S6). For all four peptides, only small quantities of a limited number of side products were observed (Figure 2b and Table S6). The main side products were cyclic dimers, eluting as two close peaks, most likely corresponding to the head-to-head/tail-to-tail dimer and the head-to-tail version. Linearization of the products with TCEP and subsequent HPLC analysis showed linear peptide products with purities of 96, 96, 100, and 92% for peptides **1**–**4**, respectively (Figure 2b, right chromatograms). The TCEP-linearization experiment suggests that dimeric peptides can be reduced and then re-cyclized at low concentrations that favor the intramolecularly cyclized monomer.

To see if shorter peptides could be generated following the cyclative release strategy, which is important to test due to the strain of a smaller ring, we repeated the experiment with four new peptides that were one amino acid shorter than in the previous experiment (Figure S3a). These four peptides with the format Mpa-Ala-Xaa-Mea (Xaa = Trp, Amb, Pro, Nip) were efficiently released, and the main peak in the HPLC profile was the desired cyclic peptide (Figure S3b and Table S7). The fraction of cyclic dimer was slightly higher overall, probably due to a less efficient circularization resulting from conformational constraints of the short backbones.

To determine whether the cyclative disulfide release strategy can be applied for library synthesis and screening, we designed and synthesized cyclic peptides in a 96-well plate (5 μmol scale; Figure S4). We prepared 96 random disulfide-cyclized peptides with three different formats, which contained three random amino acids (“Xaa”) flanked by Mpa and Mea (Figure 3a). Because variation of the peptide backbone is critical for generating structurally diverse macrocycle libraries, two of these random amino acids in each peptide were selected from four structurally diverse α-, β-, and γ-amino acids to generate diverse cyclic peptide backbones (amino acids I to IV; Figure 3a). The third random amino acid was Trp or Tyr, which allowed for quantification of peptide yields by absorption measurement at 280 nm. Cyclative release of the peptides and subsequent absorption measurement showed a high average peptide yield (2.7 μmol), a high average concentration (13.3 mM) and a narrow concentration distribution for 90 of the peptides between 8.9 mM (1.5-fold below average) and 20 mM (1.5-fold above average). Three of the peptides were

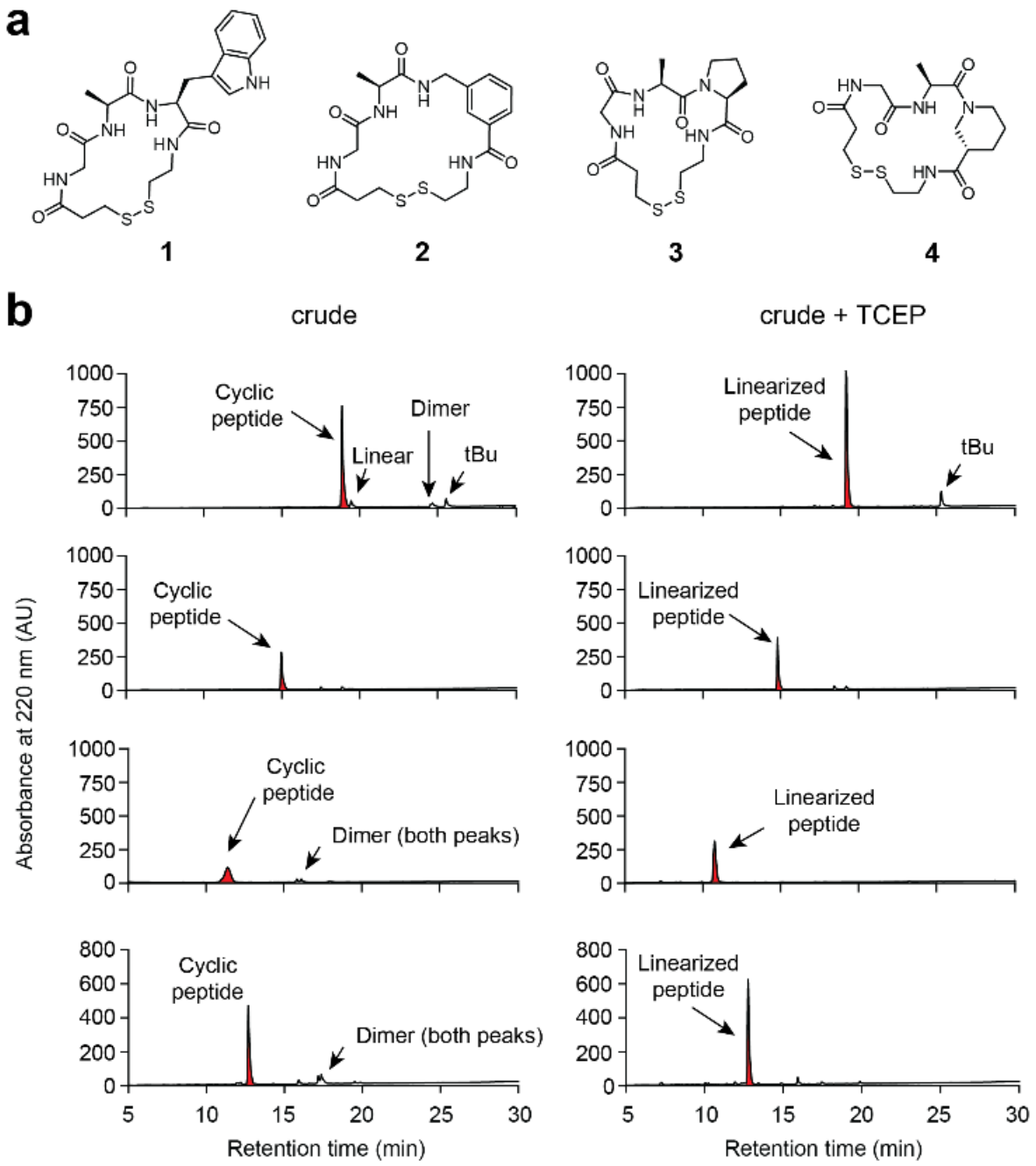


Figure 2. Cyclative release of peptides with variable sequences. **a** Chemical structures of the desired four cyclic peptides. The peptides are based on the linear sequence Mpa-Gly-Ala-Xaa-Mea, with "Xaa" being an amino acid with variable conformational flexibility in the backbone (Trp, Amb, Pro, Nip). **b** Analytical HPLC chromatograms of the crude peptide after cyclative release. The chromatograms on the right show the peptides after disulfide-bond reduction with TCEP. For impurities that were not identified, the mass is indicated (assuming that the species are singly charged). tBu = peptide with tBu group.

not synthesized or released at all. Analysis of 12 randomly selected cyclic peptides by LC-MS showed excellent purities (Figure S5 and S6).

We performed a screen against thrombin, a target for developing anticoagulation therapeutics, using a compound concentration of approximately 10 μM. The most active peptide reduced thrombin activity by around 50%, which was remarkable considering that the peptide did not contain amino acids Arg or Lys that form key interactions with the thrombin substrate recognition pocket S1 (Figure 3c). Repetition of the screen identified the same cyclic peptide as the most active hit (green dots in Figure 3c). The HPLC-purified disulfide-cyclized peptide Mpa-Tyr-II-Pro-Mea inhibited thrombin with a K_i of 13 ± 1 μM (Figure S7). Additionally, it is important to note that the DMSO and DIPEA, present after release, did not affect the biological screen.

Regarding the development of cell permeable peptides, we tested whether smaller, *N*-substituted cyclic peptides could be synthesized by the new cyclative disulfide release strategy. *N*-substitution removes one H-bond donor from the peptide, thus increasing the chances of membrane permeability. In these peptides, one of four building blocks was *N*-substituted, which we reasoned should still cyclize efficiently due to the higher rotational flexibility around the peptide bond between Mpa and the "peptoid" amino acid, even with the shorter length of only four building blocks. To obtain the two cyclic peptide formats shown in Figure 4a, the *N*-substituted amino acids, either glycine (sub-library 1) or 3-(aminomethyl)benzoic acid (sub-library 2), were introduced using the sub-monomer approach, by first appending bromoacetic acid or 3-(bromomethyl)benzoic acid and then adding a primary amine. We combinatorially synthesized a total of 342 peptides by using nine different amino acids, two sub-monomer building blocks, and 19 different primary amines (Figure S8a). The majority of chosen amino acids and amines were not charged, and had a limited polarity to achieve cyclic peptides with physicochemical properties that resembled those of approved cell permeable or orally available macrocyclic drugs.[10] Indeed, there was a good overlap between the 342 cyclic peptides and 34 orally available macrocyclic drugs or drug candidates when comparing the molecular weight, the calculated water/n-octanol partition coefficient (clogP), the number of H-bond donor groups, and the polar surface area (Figure S8b).

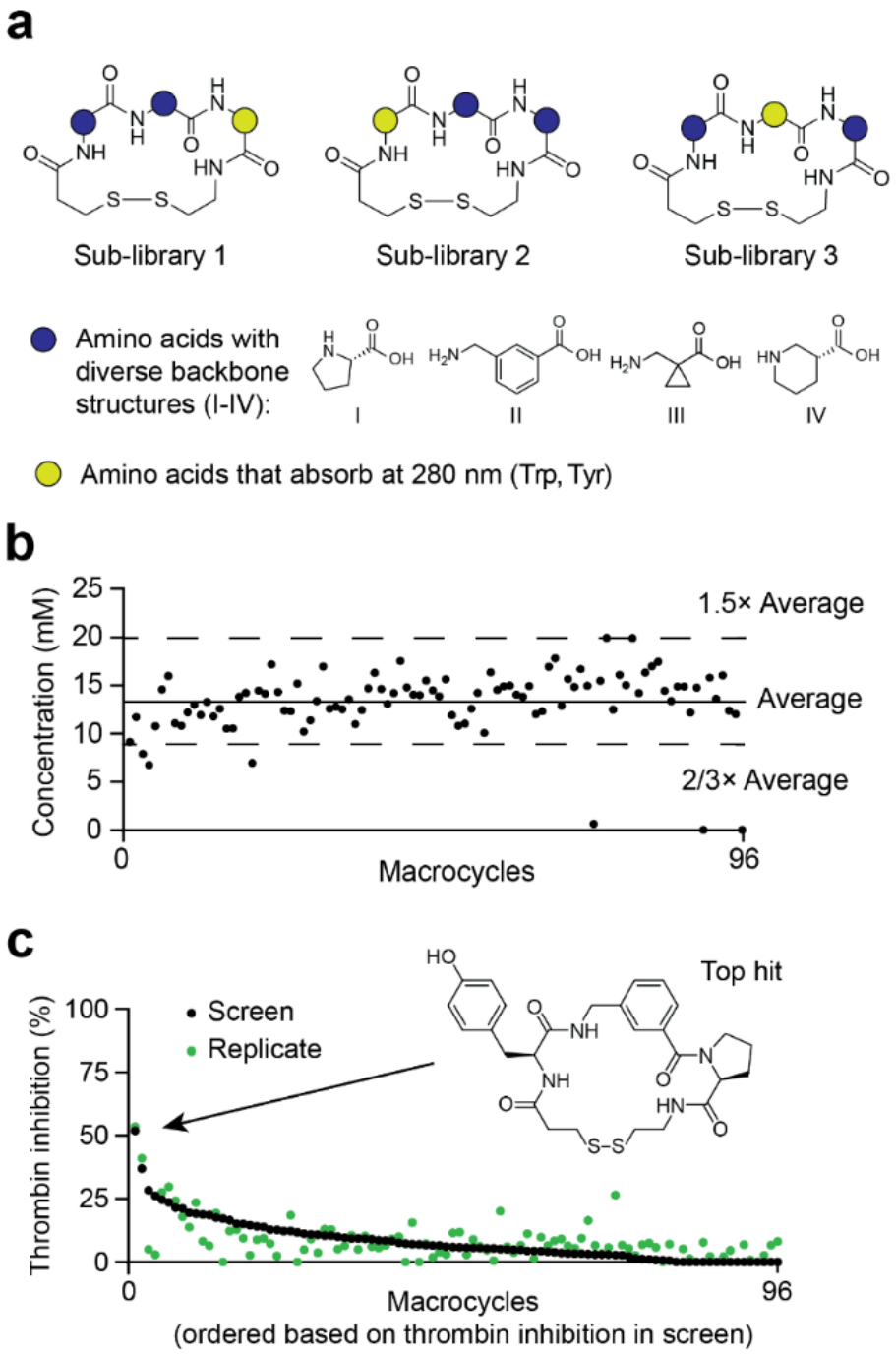


Figure 3. Library design, peptide recovery quantification by absorption, and thrombin inhibition of the cyclic peptide library. **a** Design of library comprising 96 different peptides and structures of unnatural amino acids used in the library. **b** Scatter plot of cyclic peptide concentrations (mM in DMSO) for the 96 synthesized cyclic peptides quantified by absorption at 280 nm. The average recovery, along with 1.5× and 0.67× of this value, is indicated on the chart. **c** Thrombin inhibition measured at an average cyclic peptide concentration of 11 μM. The peptides are ordered according to their thrombin inhibition activity in the first screen (black dots; highest to lowest activity). Green dots indicate thrombin inhibition for the same cyclic peptides measured in a second screen using the same conditions. The chemical structure of the most active inhibitor is shown (K_i = 13 ± 1 μM).

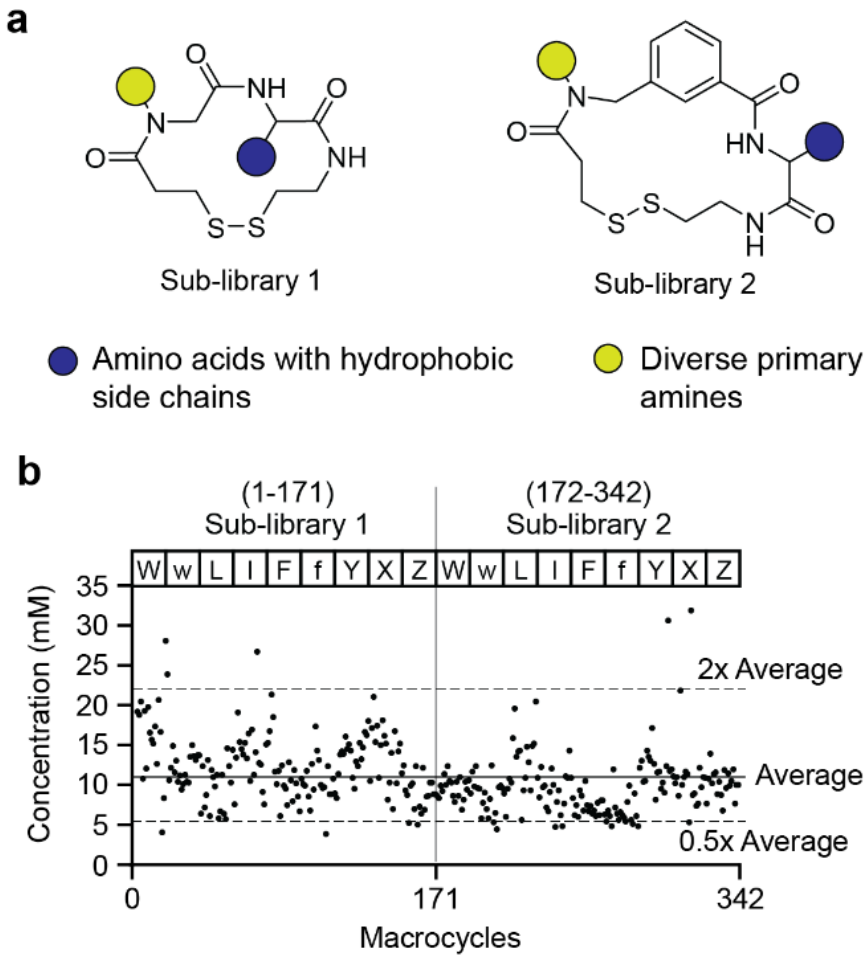


Figure 4. Library composed of 342 small cyclic N-alkylated peptides. **a** Scaffolds of the two sub-libraries. Building blocks used for the synthesis of the combinatorial library are shown in Figure S8. **b** Concentrations of 342 cyclic peptides determined by an Ellman's reagent-based assay. The peptides are sorted first according to the amino acids, and then by amine building blocks (N1–N19 shown in Figure S8).

Because most entries in the library did not contain Trp or Tyr, whose absorption at 280 nm form the established method for measuring peptide concentrations, we next established a procedure to determine concentration based on the presence of thiols. In brief, we incubated the released peptides with immobilized TCEP to reduce the disulfide bridge, transferred the reduced peptides to wells of a microtiter plate containing Ellman's reagent, and quantified the number of sulfhydryl groups by measuring the absorption of released 2-nitro-5-thiobenzoic acid (TNB) species at 412 nm. In parallel, we determined the concentrations of the 26 peptides that contained Trp or Tyr by measuring their absorbance at 280 nm and found that the concentrations determined based on the thiol groups correlated well with the values measured by UV absorption (Figure S9). We found that 320 of the 342 library entries were obtained in high yields, with an average yield of 2.2 μmol (11 mM average concentration) and a narrow yield and concentration distribution (Figure 4b), indicating an efficient release by the cyclative

disulfide method despite the short peptide sequences and potential conformational constraints. Overall, the cyclic peptides containing *N*-substituted amino acid and having shorter macrocyclic rings were obtained in good yields showing that molecules with properties of oral drugs can be released well. LC-MS analysis of 26 sample peptides, half of which pertain to each sub-library and with all 19 different amine building blocks represented, showed a high purity of the desired peptides (85% average for 26 peptides; Figure S10).

3.5 Conclusions

To address the lack of strategies for synthesizing cyclic peptide libraries that do not need purification before screening, we have developed a cyclative peptide release strategy based on a disulfide exchange reaction that releases highly pure disulfide-cyclized peptides directly from the solid support. Because of the high purity and a volatile base as cleavage reagent, which can be removed by evaporation, these peptides can be readily screened using bioassays without prior purification. Importantly, the yields of the library components containing different building blocks and sequences were narrowly distributed, permitting screening without the need to determine or adjust the concentrations of hundreds of compounds. We have shown that the approach is applicable for the generation of libraries comprising hundreds of peptides, which can be synthesized and screened in only a few days. Our rapid and facile method allows for the access of chemical space that was previously difficult to access. Importantly, the method can be applied to any molecule linked via a disulfide bond to solid phase and containing another thiol, allowing for even larger increases in chemical diversity. As a next step, we intend to further increase our synthetic throughput and screen targets relevant for the treatment of disease.

3.6 Materials and methods

General considerations

Unless otherwise noted, all reagents were purchased from commercial sources and used with no further purification. Solvents were not anhydrous, nor were they dried prior to use. All water was purified using a MilliQ Integral 5 water purification system (Merck).

Synthesis of model peptide on different resins

ChemMatrix-AM-NH_2 (100 mg, 100 μmol, 1.0 mmol/g; Sigma-Aldrich), NovaPEG-NH_2 (189 mg, 100 μmol, 0.53 mmol/g; Novabiochem), HypoGel-200-NH_2 (111 mg, 100 μmol, 0.9 mmol/g; Sigma Aldrich), PS-AM-NH_2 (71.9 mg, 100 μmol, 1.39 mmol/g; Aapptec) and PS-SH (105 mg, 100 μmol, 0.95 mmol/g; Rapp Polymer) were introduced into five different 5 ml polypropylene synthesis columns (MultiSyntech GmbH, V051PE076) and SPPS was performed manually.

To all resins except to PS-SH, the thiol source *S*-trityl-3-mercaptopropanoic acid (Trt-MPA) was coupled by reacting the carboxylic acid of Trt-MPA with amino groups on the resins. For this, the resins were pre-washed as follows: 2 × 3 ml MeOH, 2 × 3 ml DCM, 2 × 3 ml 1:99 TFA/DCM (v/v), 2 × 3 ml 1:19 DIPEA/DCM (v/v), 2 × 3 ml DCM. Immediately before coupling the Trt-MPA, the amino resins were washed with 2 × 3 ml DMF. Trt-MPA was coupled twice using Trt-MPA (150 mM, 3 equiv.), HATU (150 mM, 3 equiv.) and DIPEA (300 mM, 6 equiv.) in 2 ml DMF. The two consecutive coupling reactions were performed for one hour at room temperature under rotation (20 rpm). After the coupling, the resins were washed with 3 × 3 ml DMF. Subsequently, a capping step was applied using 2 ml capping solution containing 5:6:89 Ac_2O/2,6-lutidine/DMF (v/v/v) for 5 min. After the capping, the resins were washed with 3 × 3 ml DMF. The Trt group was deprotected using 2 × 3 ml of 10:1:89 TFA/TIS/DCM (v/v/v) for 2 × 1 hour at room temperature under rotation (20 rpm) followed by washing with 3 × 3 ml DCM.

Before the disulfide exchange reaction, all the resins were washed with 3 × 3 ml 3:7 MeOH/DCM (v/v). *S*-(2-pyridylthio)cysteamine hydrochloride (98.0 mg, 0.44 mmol, 4.4 equiv.) in 3.5 ml of 3:7 MeOH/DCM (v/v) was added to the resins followed by the addition of DIPEA (76.6 μl, 0.44 mmol, 4.4 equiv.). The disulfide exchange reactions were performed by incubation for 3 hours at room temperature. The resins were washed with 3 × 3 ml 3:7 MeOH/DCM v/v, and 3 × 3 ml DMF.

Afterwards, the different peptide building blocks were installed in the following order using SPPS: Gln, Trp, Gly, and Trt-MPA. All couplings were carried out twice for each building block using protected amino acid/carboxylic acid derivative (150 mM, 3 equiv.), HATU (150 mM, 3 equiv.), and DIPEA, (300 mM, 6 equiv.) in 2 ml DMF. Coupling reactions were performed for one hour at room temperature under rotation (20 rpm). After each double-coupling, 2 ml capping solution (5% Ac_2O and 6% 2,6-lutidine in DMF) was applied for 5 min at room temperature under rotation (20 rpm). After each coupling and capping step, the resins were washed with 3 × 3 ml DMF. Fmoc protecting groups were removed with 2 × 3 ml 1:4 piperidine/DMF (v/v) for 5 min. The resins were washed with 3 × 3 ml DMF and 3 x 3 ml DCM. The extent of coupling was qualitatively assessed after each step using Kaiser and/or Ellman's tests. At the end of the synthesis, the resins were washed with DCM. The excess of DCM was allowed to evaporate on air for 4 hours, and the samples were placed under vacuum for drying until the measured mass stopped decreasing (24 hours). The dry weights of the resins were measured after the synthesis.

The side chain protecting groups were removed after transfer of the resins to wells of 96-well plates (Orochem). Resin carrying 5 μmoles of peptide were transferred to each well. The weight transferred for each resin/peptide is shown in the last row of the Table S4. For side chain removal, the bottom of the 96-well synthesis plate was sealed by pressing the plate onto a soft 6 mm thick ethylene-vinyl acetate pad, and the resin in each well was incubated with around 500 μl of 38:1:1 TFA/TIS/ddH_2O (v/v/v) for one hour. The plates were covered with a polypropylene adhesive seal, then weighed down by placing a weight (1 kg) on top to ensure that no leakage occurred. After 1.5 hours, the synthesis plates were placed onto 2 ml deep-well plates, and the TFA mixture was allowed to drain. The wells were washed three times with approximately 500 μl of DCM (added with syringe), then allowed to air dry for 3 hours.

Kaiser and Ellman's tests

For the Kaiser test, three droplets of each of the following solutions A, B and C were given into a vial containing a few test beads. The test mixture was gently warmed with a heat gun until a blue/purple coloration was visible. Solution A: 200 μM KCN in H_2O:pyridine (1:49), Solution B: 280 mM ninhydrin in 1-butanol, Solution C: 21 M phenol in 1-butanol. As a negative control, PS-SH resin was used (no coloration). As a positive control, AM PS-NH_2 resin was used (dark blue).

For the Ellman's test, a spatula tip of test beads were pre-swelled by adding THF (100 μl, 10 min) followed by the addition of 100 μl Ellman's reagent (5 mM in MeOH) and DIPEA (1 μl). The sample was further diluted with methanol (299 μl). A yellow coloration appears in presence

of free thiols. As a negative control, AM PS-NH_2 resin was used (colorless). As a positive control, PS-SH resin was used (yellow coloration).

Reductive release

The bottom of the 96-well synthesis plate was sealed by pressing the plate onto a soft 6 mm thick ethylene-vinyl acetate pad (Rayher Bastelkunst). A volume of 200 μl of DMSO:H_2O (9:1, v/v) containing 25 μmol TCEP·HCl (5 equiv.) were added to the wells. The plate was covered with an adhesive aluminium lid (Thermo Scientific), a weight (1 kg) was placed on top of it to prevent detachment of the plate from the ethylene-vinyl acetate pad and leakage of the solutions, and incubated overnight at room temperature. The plate was placed into a 96-deep well plate (ThermoScientific, 278752) and the solutions transferred by applying vacuum. The concentration of peptide in the filtrates were determined by measuring absorption at 280 nm (nanodrop) and calculated assuming that mainly Trp absorbs (A_{280}, $\varepsilon_{Trp,\ 280\ nm}$ = 5,500 M^{-1}*cm^{-1}, d = 0.1 cm).

Cyclative release

Plates were pressed into foam pads as described above to plug the openings, and 200 μl of 150 mM DIPEA in DMSO (6 equiv.) were added to each well. The plates were sealed with an adhesive foil and weighed down (1 kg), and left overnight. The next day, the synthesis plates were placed onto 2 ml deep-well plates and centrifuged at around 200 g (1000 rpm with a Thermo Heraeus Multifuge 3L-R centrifuge) for one minute to collect the released cyclic peptides in DMSO.

LC-MS analysis

Peptides were analyzed by LC-MS analysis with a UHPLC and single quadrupole MS system (Shimadzu LCMS-2020) using a C18 reversed phase column (Phenomenex Kinetex 2.1 mm × 50 mm C18 column, 100 Å pore, 2.6 μM particle) and a linear gradient of solvent B (acetonitrile, 0.05% formic acid) over solvent A (H_2O, 0.05% formic acid) at a flow rate of 1 ml/min. Mass analysis was performed in positive ion mode.

Model peptides eluted in DMSO/DIPEA were diluted around 100-fold with water to reach a concentration of 100 μM, and 5 μl of the samples were injected and analyzed using a 0 to 60% gradient of acetonitrile over 5 min.

For peptides of sub-libraries 1 to 3, 1 μl of the DMSO/DIPEA eluates were diluted into 110 μl of water to give a cyclic peptide concentration of around 120 μM. 5 μl of the samples were injected and analyzed using a 0 to 60% gradient of acetonitrile over 5 min.

For peptides of the peptoid library, 1 or 2.5 μl of the DMSO/DIPEA eluates were diluted with 9 or 7.5 μl acetonitrile to reach a cyclic peptide concentration of around 2 mM. To these dilutions was added 4 μl of TCEP·HCl solution (24 mM in 100 mM MES buffer pH 6.6). 4 μl of the samples were injected and analyzed using a 5 to 80% gradient of acetonitrile over 7 min.

Analytical HPLC

The purity of peptides was analyzed by reversed-phase analytical HPLC (Agilent Technologies, 1260 Infinity). Peptide released from solid phase were run over a reversed-phase C18 column (Agilent Zorbax 300SB, 300 Å pore, 5 μm particle, 4.6 mm × 250 mm) using a linear gradient of solvent B (acetonitrile, 0.1% TFA) over solvent A (H_2O, 0.1% TFA) from 0−50% in 30 min at a flow rate of 1 ml/min. Typically, 100 nmol of cyclic peptide (around 5 μg) were injected. The mass of the purified cyclic peptides was confirmed by ESI-MS.

Preparation of polystyrene-S-S-cysteamine resin for library synthesis

The following procedure was applied to prepare polystyrene-S-S-cysteamine resin needed for the synthesis of 4 × 96 peptides at a 5 μmol scale in four 96-well plates. Into each of four 20 ml plastic syringes was added 589 mg resin (Rapp Polymere Polystyrene AM SH resin, 200-400 mesh, 0.85 mmol/gram loading), corresponding to a 0.5 mmol scale per syringe, and a 2 mmol scale in total. The resin of each syringe was washed with 15 ml of DCM, then swelled in 15 ml of 3:7 MeOH/DCM v/v for 15 minutes. 2-(2-pyridinyldithio)-ethanamine hydrochloride (1.96 grams, 8.8 mmoles, 4.4 equiv.) was dissolved in 21.12 ml of MeOH, then 49.28 ml of DCM and 1.53 ml of DIPEA (8.8 mmoles, 4.4 equiv.) were added. 17.7 ml of this solution was pulled into each syringe, which were then shaken at room temperature for 3 hours. After this time, the 2-(2-pyridinyldithio)-ethanamine solutions were discarded, and the resins were washed with 2 × 20 ml 3:7 MeOH/DCM v/v, then 2 × 20 ml DMF. The resins were combined into a single syringe as a suspension in DMF and washed with 11.8 ml of 1.2 M DIPEA solution in DMF for 5 minutes to ensure that all amines were neutral. This solution was discarded, and the resin was washed with 2 × 20 ml DMF, 4 × 20 ml DCM, then kept under vacuum overnight to yield a free-flowing powder.

Peptide library synthesis in 96-well plates

Automated solid-phase peptide synthesis was performed on an Intavis Multipep RSi synthesizer. To a 50 ml tube was added 565 mg of polystyrene-S-S-cysteamine resin (0.48 mmol S-S-cysteamine assuming that thiol groups were quantitatively modified with cysteamine) and 20 ml of DMF. The tube was shaken to ensure the resin was uniformly suspended, and 200 μl (5.88 mg resin, 5 μmoles) were transferred to each well of a 96-well solid phase synthesis plate (Orochem). The resin was washed with 6 × 150 μl DMF. Coupling was performed with 53 μl of amino acids (500 mM, 5.3 equiv.), 50 μl HATU (500 mM, 5 equiv.), 12.5 μl of *N*-methylmorpholine (4 M, 10 equiv.), and 5 μl *N*-methylpyrrolidone. All components were premixed for one minute, then added to the resin (one hour reaction, no shaking). The final volume of the coupling reaction was 120.5 μl and the final concentrations of reagents were 220 mM amino acid, 208 mM HATU and 415 *N*-methylmorpholine. Coupling was performed twice, then the resin was washed with 6 × 225 μl of DMF. Fmoc deprotection was performed using 120 μl of 1:5 piperidine/ DMF v/v for 5 minutes, and was performed twice. The resin was washed with 8 × 225 μl DMF. At the end of the peptide synthesis, the resin was washed with 2 × 200 μl of DCM.

Peptoid library synthesis

The peptoid library was synthesized in 96-well plates as the peptide library. For the incorporation of peptoid building blocks, 62 μl bromoacetic acid in DMF (1 M, 12 equiv.) or 3-(bromomethyl)benzoic acid in DMF (1 M, 12 equiv.) and 63 μl of *N,N*-diisopropylcarbodiimide in DMF (1 M, 12 equiv.) were pre-incubated for one minute, added to the wells, and incubated for one hour without shaking. The reactions were repeated once and the resin washed with 5 × 225 μl DMF. Next, 125 μl of amine in DMF (2 M in DMF, 40 equiv.) was added to the wells and the reactions incubated for two hours. The wells were washed with 5 × 225 μl DMF.

Amino acid side chain deprotection in 96-well plates

The procedure is described above in the last paragraph of the chapter "Synthesis of model peptide on different resins".

Cyclative release of library peptides in 96-well plates

Plates were pressed into foam pads as described above to plug the openings, and 200 μl of 150 mM DIPEA in DMSO (6 equiv.) were added to each well. The plates were sealed with an

adhesive foil and weighed down with a pipette tip box full of sand, and left overnight. The next day, the synthesis plates were placed onto 2 ml deep-well plates and centrifuged at around 200 g (1000 rpm with a Thermo Heraeus Multifuge 3L-R centrifuge) for one minute to collect the cleaved cyclic peptides in DMSO.

Library peptide quantification by absorption

Absorbance measurements were performed with a Nanodrop 8000 spectrophotometer (Thermo Scientific) at a wavelength of 280 nm using a 1 mm path length. Peptides containing Trp were diluted 10-fold and those containing Tyr 5-fold using DMSO so that absorbance measurement values would be in a range from 0.5 to 1.25 units. DMSO with 150 mM DIPEA containing no peptide, diluted 10- and 5-fold with DMSO, showed absorption of 0.023 and 0.039, respectively, showing that the solvent contributed less than 10% to the absorption. The Beer-Lambert law was used to calculate the concentration of the peptides. Extinctions coefficients Trp ε_{280} = 5500 $M^{-1}cm^{-1}$ and Tyr ε_{280} = 1490 $M^{-1}cm^{-1}$ were used. The exact extinction coefficient value for Cl-Trp was not determined and assumed to be ε_{280} = 5500 $M^{-1}cm^{-1}$.

Library peptide quantification by Ellman's reagent

The concentration of peptides was determined by fully reducing the disulfide bonds and quantifying the concentration of free thiol group using the Ellman's assay. The peptides were reduced using TCEP immobilized on beads so that the reducing agent could be removed by filtration before adding Ellman's reagent. A volume of 40 μl of peptide in DMSO, 150 mM DIPEA at an average concentration of around 14 mM (0.56 μmol peptide) was added to 35 mg TCEP beads (2.21 μmol TCEP, around 4 equiv. relative to peptide) in wells of 96-well U-bottom plates and incubated for 2 hours under shaking (140 rpm) at room temperature. The solutions were transferred into a 96-well fritted filter plate which was placed on top of a 2 ml deep-well plate and centrifuged at 2000 rpm for 2 minutes to collect the reduced peptides.

In a clear 96-well flat-bottom plate was added 78 μl per well of 3:7 DMSO/60 mM NH_4HCO_3 buffer pH 8 (v/v). 2 μl of the reduced peptide stocks were added to the plate, followed by 20 μl of Ellman's reagent 20 mM 5,5'-dithiobis-(2-nitrobenzoic acid) in the same DMSO/buffer mixture. The final concentrations of reduced peptide in this assay was in average around 0.28 mM. The final concentration of Ellman's reagent was 4 nM.

Absorbance at 412 nm was measured with a Tecan Infinite M200 Pro plate reader. The absorbance value for Ellman's reagent (4 mM) alone was 0.21 units. The absorbance of peptides, after subtracting 0.21 units ranged from 0.55 to 3.20 units. A standard calibration

curve was established by following the above procedure using a peptide with known concentration.

Library screening

Thrombin inhibition by cyclic peptides was assessed by measuring residual activity of thrombin in presence of the cyclic peptides at 11 μM average final concentration. The assays were performed in 384-well plates using Tris buffer at pH 7.4 (100 mM Tris-Cl, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM $CaCl_2$, 0.1% w/v BSA, 0.01% v/v Triton-X100, and 1% v/v DMSO) using thrombin at a final concentration of 2 nM and the fluorogenic substrate Z-Gly-Gly-Arg-AMC at a final concentration of 50 μM.

To 384 low-volume assay plates, 12.5 nl cyclic peptide (in DMSO containing 150 mM DIPEA; average cyclic peptide concentration of 13.5 mM) were transferred using acoustic liquid transfer (ECHO 550, Labcyte). Thrombin (7.5 μl, 4 nM) in the Tris buffer described above was added to each peptide using a Gyger Certus Flex liquid dispenser, and incubated for 10 minutes at room temperature. The fluorogenic substrate (7.5 μl, 100 μM) in the same butter was added using the Gyger Certus Flex liquid dispenser, and the florescence intensity measured with a Tecan Infinite M200 Pro fluorescence plate reader (excitation at 360 nm, emission at 465 nm) at 25°C for a period of 30 min with a read every three min.

Cyclic peptide purification

A quantity of 50 nmoles of the most active inhibitor was purified on a Thermo Dionex Ultimate 3000 HPLC (Waters Nova-Pak C18 Column, 60 Å pore, 6 μm particle, 7.8 mm × 300 mm) using a 10 to 60% acetonitrile gradient over 20 minutes. The fraction containing the cyclic peptide was dissolved first in 4 μl of DMSO, followed by addition of 196 μl of water.

Determination of the inhibitory constants of thrombin inhibitors

The IC_{50}s of the best cyclic peptide was determined by measuring the residual activity of thrombin at different peptide concentration with the same assay as in the screen. The reactions were performed in volumes of 150 μl in 96-well plates. 50 μl of 2-fold dilutions of cyclic peptide in assay buffer at pH 7.4 (100 mM Tris-Cl, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM $CaCl_2$, 0.1% w/v BSA, 0.01% v/v Triton-X100) containing 3% DMSO v/v were pipetted to the wells. A volume of 50 μl of thrombin (6 nM) in assay buffer was added to each well to reach a final concentration of 2 nM and was incubated for 10 min. A volume of 50 μl of the fluorogenic substrate Z-Gly-

Gly-Arg-AMC (150 μM) in assay buffer containing 1% DMSO v/v was added to reach a final substrate concentration of 50 μM, and 1.33% DMSO. Fluorescence intensity was measured for 30 min using a Tecan Infinite M200 Pro plate reader (excitation at 360 nm, emission at 465 nm), with a read every three minutes. Sigmoidal curves were fitted to the data using GraphPad Prism 6 software, and the following dose-response equation:

$$y = 100/1+10(\log IC_{50} - x)p$$

where x is cyclic peptide concentration, y is % protease activity, and p is the Hill slope. IC_{50} values were derived from the fitted curve.

The inhibitory constants (K_i) were calculated using the following equation of Cheng and Prusoff:

$$K_i = IC_{50}/(1 + [S]_0/Km)$$

where IC_{50} is the functional strength of the inhibitor, $[S]_0$ is the total substrate concentration, and K_m is the Michaelis-Menten constant. The K_m for thrombin and the substrate Z-Gly-Gly-Arg-AMC was determined to be 168 μM.

3.7 Supplementary information

Supplementary Tables

Supplementary Table 1. Solid phase resins. Properties of five commercial resins used in this work.

	Solid support				
	ChemMatrix NH_2 (1)	NovaPEG-NH_2 (2)	Hypogel 200-NH_2 (3)	Polystyrene AM PS- NH_2 (4)	Polystyrene A SH (5)
Provider	Sigma	Novabiochem	Sigma	Aapptec	Rapp Polymere
Polymer	PEG	PEG	PEGylated polystyrene	Polystyrene	Polystyrene
Particle size	35-100 mesh	35-100 mesh	100-140 mesh	100-200 mesh	200-400 mesh
Capacity	1 mmol/g	0.53 mmol/g	0.90 mmol/g	1.39 mmol/g	0.95 mmol/g
Property	polar	polar	polar	apolar	apolar
Functional group	NH_2	NH_2	NH_2	NH_2	SH

Supplementary Table 2. Synthesis of peptides immobilized via a disulfide bridge. Concentrations in μM and % recovery are indicated. Values are means of two independent syntheses and cleavage reactions.

	Solid support				
	ChemMatrix NH_2 (1)	NovaPEG-NH_2 (2)	Hypogel 200-NH_2 (3)	Polystyrene AM PS- NH_2 (4)	Polystyrene A SH (5)
Scale (μmol)	100	100	200	100	100
Syringe tare (mg)	2234.6	2244.5	2232.5	2251.1	2236.5
Initial resin weight (mg)	101.6	189.0	225.5	72.4	105.9
Gross weight after S-S exchange (mg)	107.6	205	236.3	79.2	109.5
Gross weight after SPPS (mg)	131.1	233.6	270.2	119.6	195.4
Amount of peptide (protected) on resin (mg)	23.5	28.6	33.9	40.4	85.9
Amount of resin transferred to 96-well plate for 5 μmoles (mg)	6.6	11.7	6.8	6.0	9.8

Supplementary Table 3. Reductive cleavage in 9:1 DMSO/water. Concentrations in μM and % recovery are indicated. Values are means of two independent syntheses and cleavage reactions.

Reducing agent	Performance	Solid support				
		ChemMatrix NH_2 (1)	NovaPEG-NH_2 (2)	Hypogel 200- NH_2 (3)	Polystyrene AM PS- NH_2 (4)	Polystyrene A SH (5)
TCEP	Concentration	3.0	3.2	21	60	73
	% recovery	0.039	0.042	0.16	0.15	0.22
Solvent only	Concentration	1.2	1.1	0.1	3.5	1.0
	% recovery	0.005	0.004	0.001	0.034	0.013

Supplementary Table 4. Reductive cleavage in 9:1 MeCN/water. Concentrations in μM and % recovery are indicated. Values are means of two independent syntheses and cleavage reactions. * 85% ACN and 15% water were used as a solvent to fully dissolve TCEP.

Reducing agent	Performance	Solid support				
		ChemMatrix NH_2 (1)	NovaPEG-NH_2 (2)	Hypogel 200- NH_2 (3)	Polystyrene AM PS- NH_2 (4)	Polystyrene A SH (5)
TCEP*	Concentration	8.3	9.7	37	35	51
	% recovery	0.013	0.013	0.084	0.25	0.31
solvent only	Concentration	2.6	4.7	0.2	7.6	3.3
	% recovery	0.012	0.018	0.001	0.087	0.042

Supplementary Table 5. Yields and purity of cyclic peptide Mpa-Gly-Gln-Trp-Mea obtained with the cyclative release strategy. The peptide Mpa-Gly-Gln-Trp-Mea was synthesized on five resins at a 5 μmol scale, followed by on-resin removal of side-chain protecting groups. Cleavage was performed with either 200 μl of 150 mM DIPEA in DMSO, or 200 μl of 1:9 buffer/DMSO (v/v) overnight (buffer = 150 mM NH_4HCO_3 pH 8.0). All steps were performed in triplicate. Percent recovery was determined by measuring absorbance at 280 nm. Purity was determined by measuring the AUC of UHPLC chromatograms recorded at 220 nm. The detection limit of the UHPLC was estimated to be around 0.5 % recovery. N.D. = not determined.

Cyclative release condition	Performance	Solid support				
		ChemMatrix NH_2 (1)	NovaPEG-NH_2 (2)	Hypogel 200- NH_2 (3)	Polystyrene AM PS- NH_2 (4)	Polystyrene A SH (5)
150 mM DIPEA in DMSO	% Recovery	< 0.5	< 0.5	< 0.5	21 ± 5	44 ± 4
	% Purity of cleaved cyclic peptides	N. D.	N. D.	N. D.	95 ± 4	93 ± 1
150 mM DIPEA in 1:9 buffer:DMSO	% Recovery	< 0.5	< 0.5	< 0.5	N. D.	N. D.
	% Purity of cleaved cyclic peptides	N. D.	N. D.	N. D.	N. D.	N. D.

Supplementary Table 6. Quantification of cyclic peptides and side products. The relative abundance of all species was determined based on the area under the curve (AUC) of the analytical HPLC spectra shown in Figure 2b.

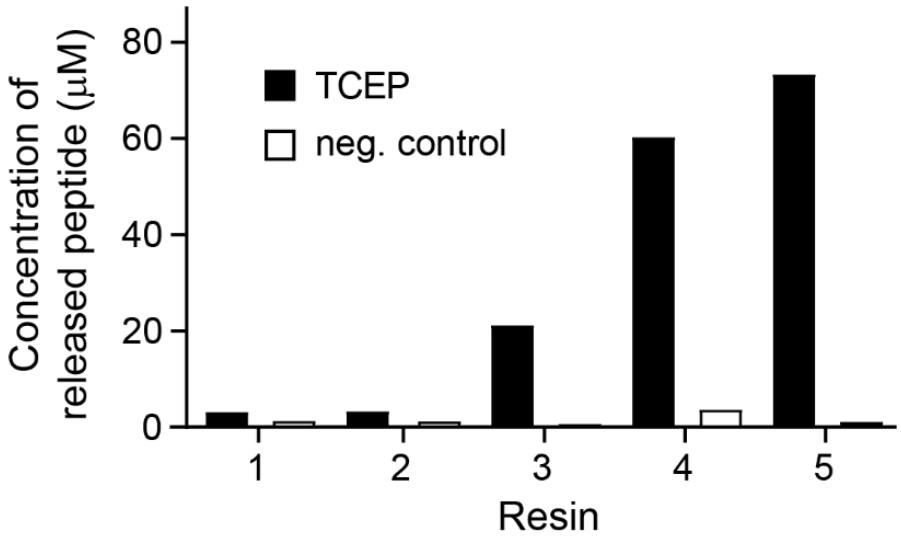
Cyclic peptide	Recovery[a]	Product	Side products			
		Cyclic peptide	Linear	Dimer	tBu	Other
1	67%	86%	6%	4%	4%	0%
2	94%	96%	0%	0%	0%	4%[b]
3	108%	90%	0%	10%	0%	0%
4	54%	75%	0%	17%	0%	8%[c]

[a]Relative to yield expected based on resin loading. [b]Side products M/Z = Product +53 or +133 Da. [c]Side product M/Z = product + 100 Da.

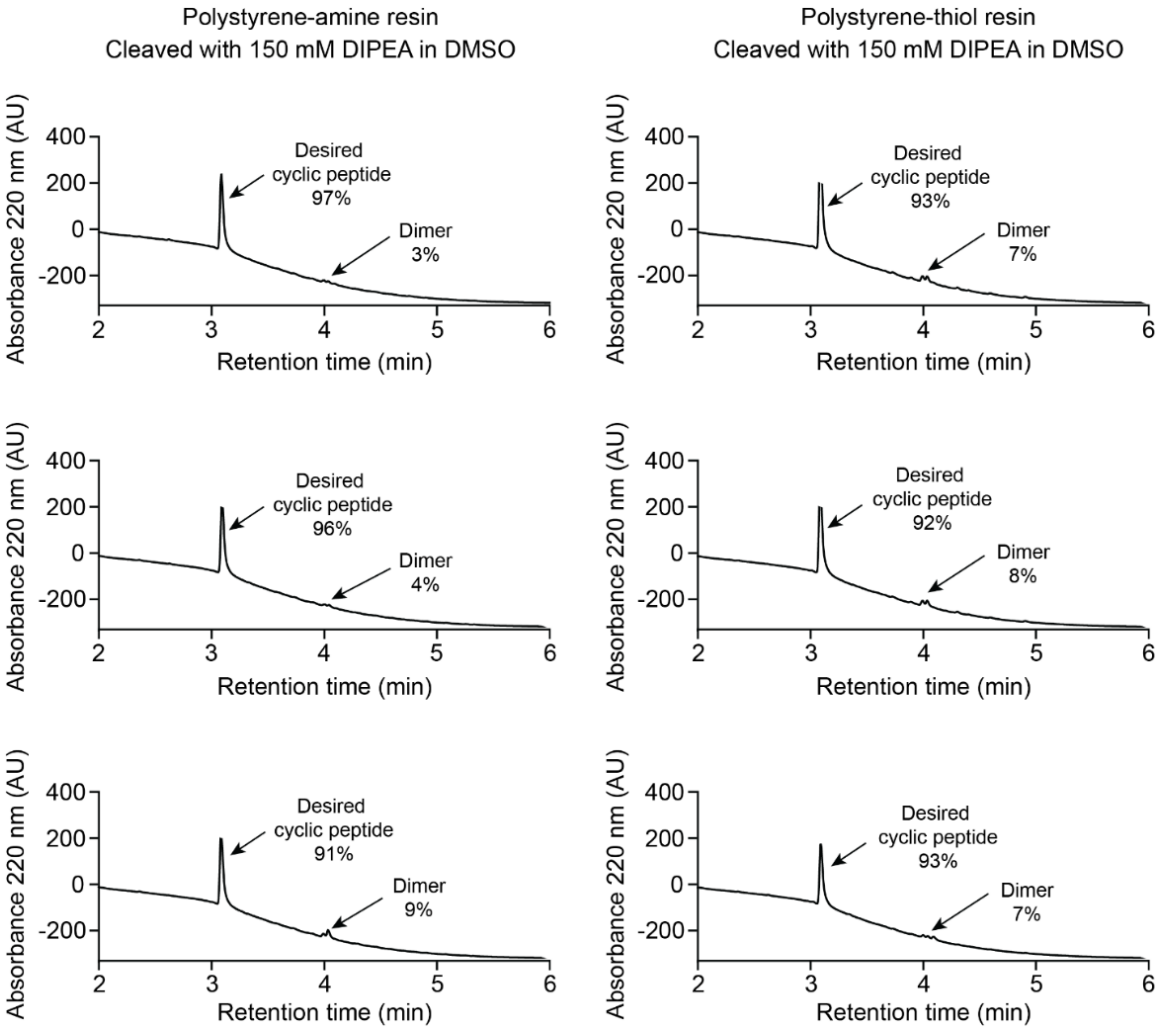
Supplementary Table 7. Quantification of cyclic peptides and side products. The relative abundance of all species was determined based on the AUC of the analytical HPLC spectra shown in Figure S3b. [a]Relative to yield expected based on resin loading. [b]Side product M/Z = Product +53 Da. [c]Side product M/Z = Product + 133 Da.

		Products				
Amino acid (cyclic peptide)	Recovery[a]	Cyclic peptide	Linear	Dimer	tBu	Other
Trp (5)	67%	68%	6%	18%	8%	0%
3-AMBA (6)	71%	68%	0%	15%	0%	17%[b]
Pro (7)	79%	49%	7%	37%	0%	7%[c]
Nipecotic (8)	31%	82%	0%	18%	0%	0%

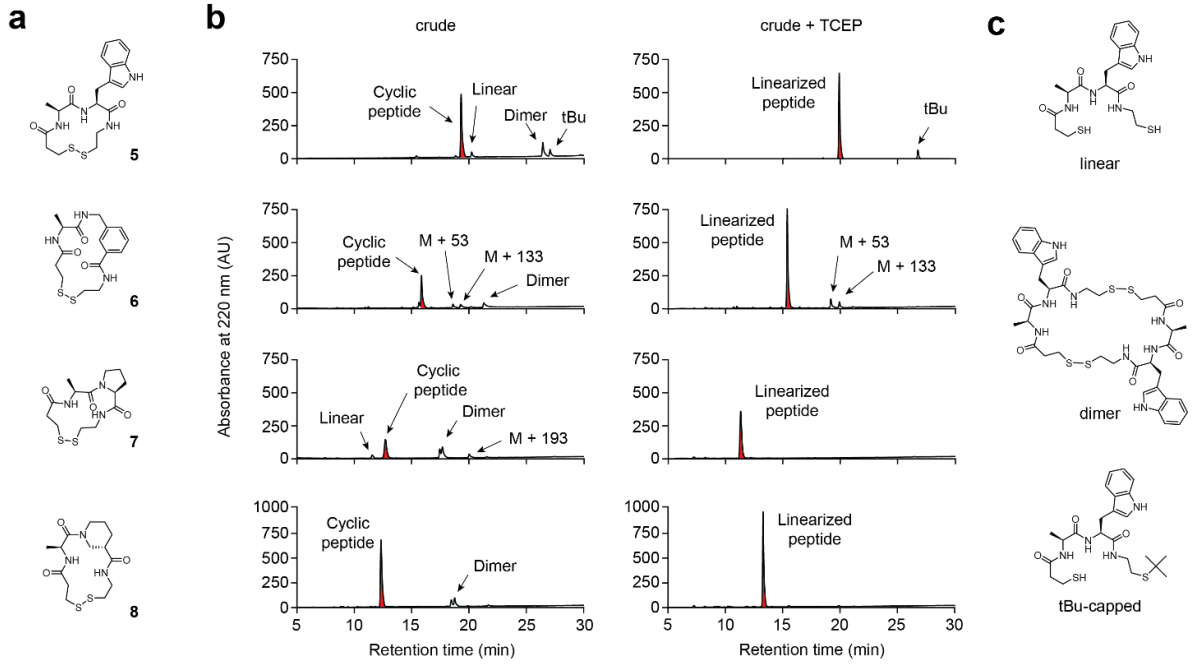
Supplementary Figures



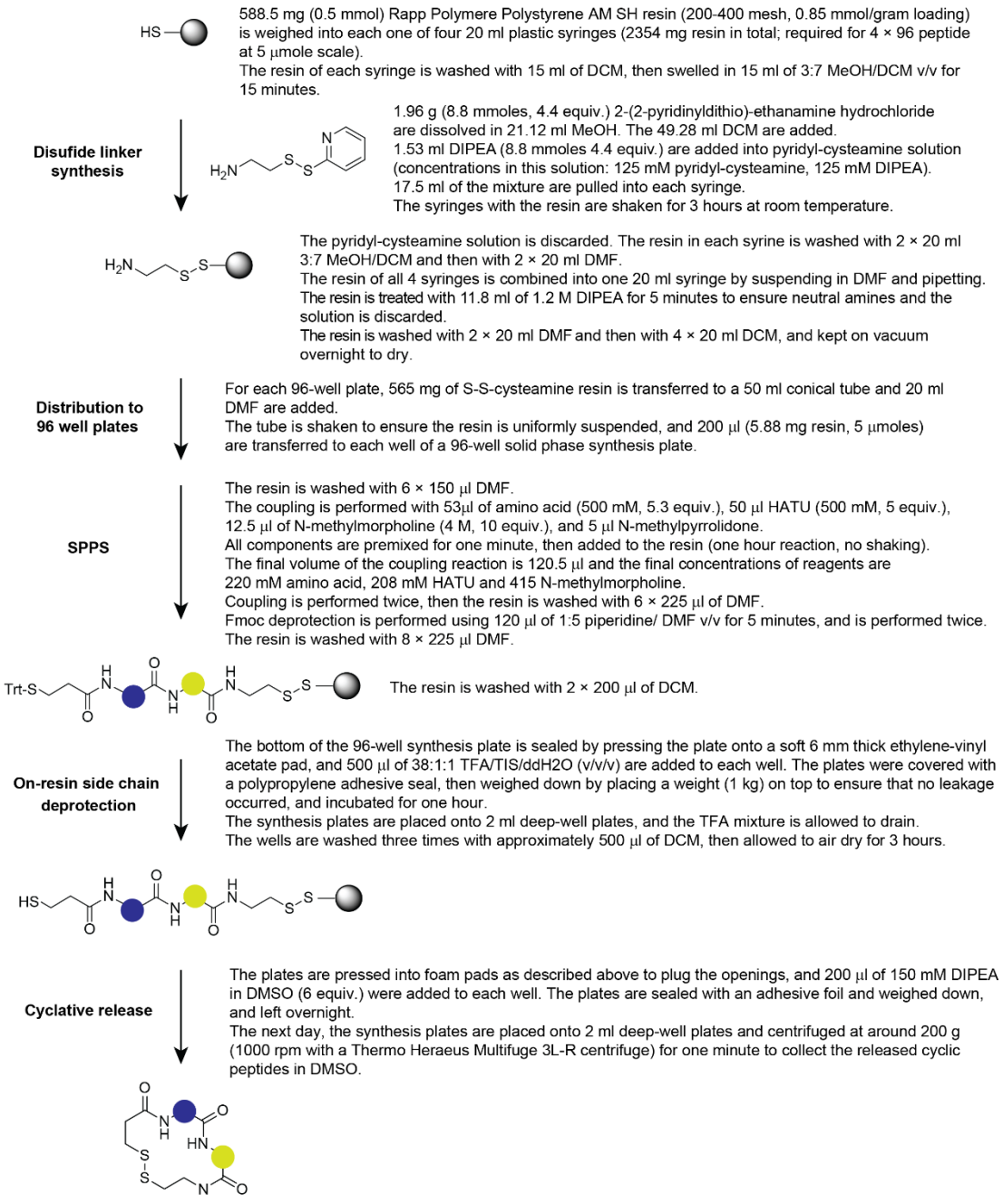
Supplementary Figure 1. Release of peptide by disulfide bridge reduction with TCEP. Yields of Mpa-Gly-Gln-Trp-Mea upon incubation of the five resins (5 μmol scale) with the reducing agent TCEP in 9:1 DMSO/water (v/v), wherein the negative control represents solvent only. Yields are indicated according to the concentration in the 200 μl eluates, determined by measuring absorbance at 280 nm. Reactions were perfomed in duplicate, and the average recovery values are shown.



Supplementary Figure 2. LC-MS chromatograms of model peptide Mpa-Gly-Gln-Trp-Mea obtained by cyclative release. The peptide was synthesized on resin 4 (polystyrene-amine resin) or resin 5 (polystyrene-thiol resin). Following SPPS and on-resin removal of side-chain protecting groups, the peptides were released by treatment with 150 mM DIPEA in DMSO. Quantities of product and side-products were determined by measuring the AUC of peaks in the UHPLC chromatograms at 220 nm absorbance. Synthesis and cleavage were performed in triplicate for each resin.



Supplementary Figure 3. Cyclative release of peptides containing four building blocks. (a) Chemical structures of the desired four cyclic peptides. The peptides are based on the linear sequence Mpa-Ala-Xaa-Mea with “Xaa” being an amino acid with variable conformational flexibility in the backbone. (b) Analytical HPLC chromatograms of the crude peptides after cyclative release. The chromatograms on the right show the peptides after disulfide-bond reduction with TCEP. For impurities that were not identified, the mass is indicated. (c) Examples of chemical structures that fit with the molecular masses of the identified impurities.



Supplementary Figure 4. One-page recipe for the synthesis and cyclative release of peptides in four 96-well plates (384 peptides in total). Note: the protocol is slightly different than procedures described in the materials and methods section for model peptides, as the resin was prepared in batch at a larger scale, and as some procedures were slightly optimized.

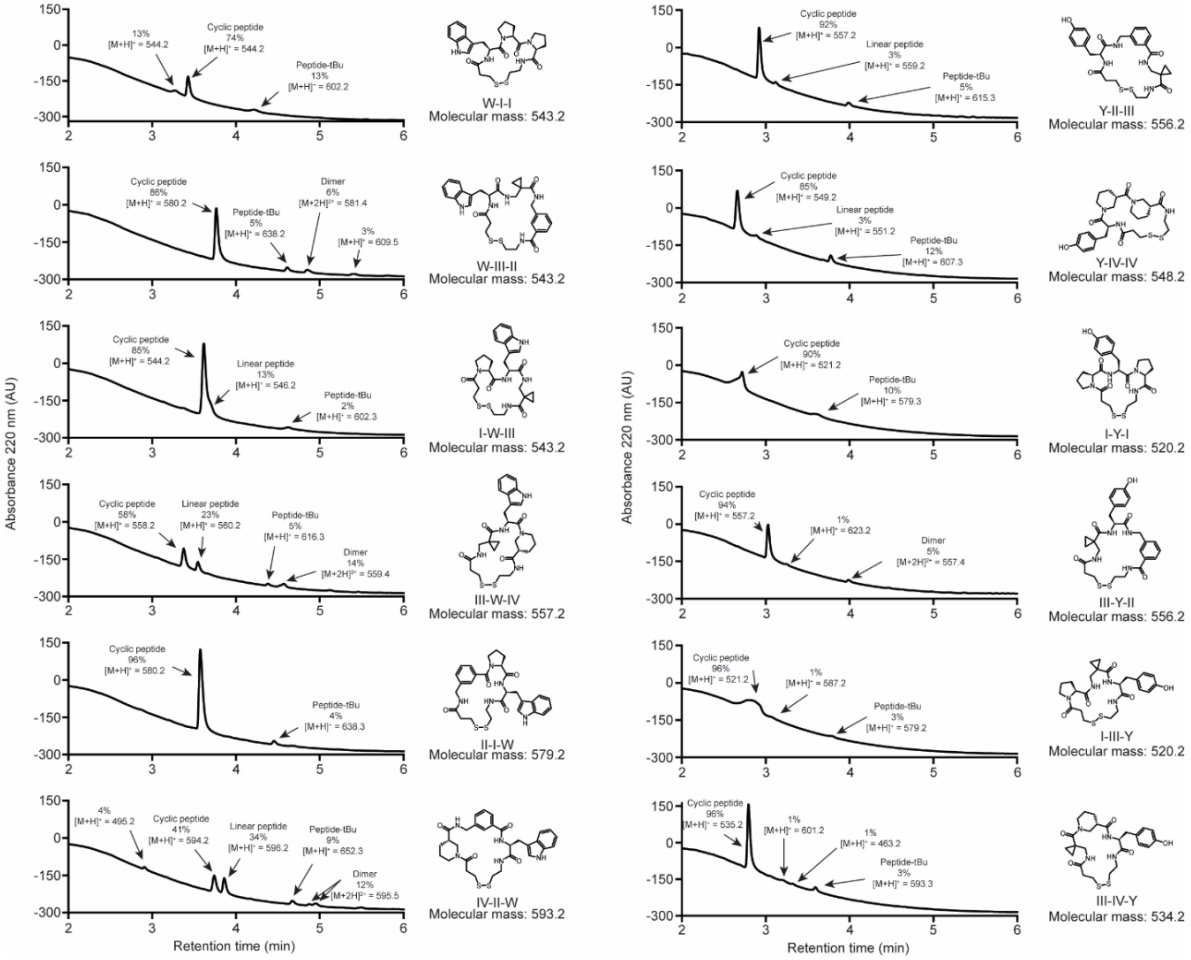
a

	1	2	3	4	5	6	7	8	9	10	11	12
A	W-I-I	W-III-I	I-W-I	III-W-I	I-I-W	III-I-W	Y-I-I	Y-III-I	I-Y-I	III-Y-I	I-I-Y	III-I-Y
B	W-I-II	W-III-II	I-W-II	III-W-II	I-II-W	III-II-W	Y-I-II	Y-III-II	I-Y-II	III-Y-II	I-II-Y	III-II-Y
C	W-I-III	W-III-III	I-W-III	III-W-III	I-III-W	III-III-W	Y-I-III	Y-III-III	I-Y-III	III-Y-III	I-III-Y	III-III-Y
D	W-I-IV	W-III-IV	I-W-IV	III-W-IV	I-IV-W	III-IV-W	Y-I-IV	Y-III-IV	I-Y-IV	III-Y-IV	I-IV-Y	III-IV-Y
E	W-II-I	W-IV-I	II-W-I	IV-W-I	II-I-W	IV-I-W	Y-II-I	Y-IV-I	II-Y-I	IV-Y-I	II-I-Y	IV-I-Y
F	W-II-II	W-IV-II	II-W-II	IV-W-II	II-II-W	IV-II-W	Y-II-II	Y-IV-II	II-Y-II	IV-Y-II	II-II-Y	IV-II-Y
G	W-II-III	W-IV-III	II-W-III	IV-W-III	II-III-W	IV-III-W	Y-II-III	Y-IV-III	II-Y-III	IV-Y-III	II-III-Y	IV-III-Y
H	W-II-IV	W-IV-IV	II-W-IV	IV-W-IV	II-IV-W	IV-IV-W	Y-II-IV	Y-IV-IV	II-Y-IV	IV-Y-IV	II-IV-Y	IV-IV-Y

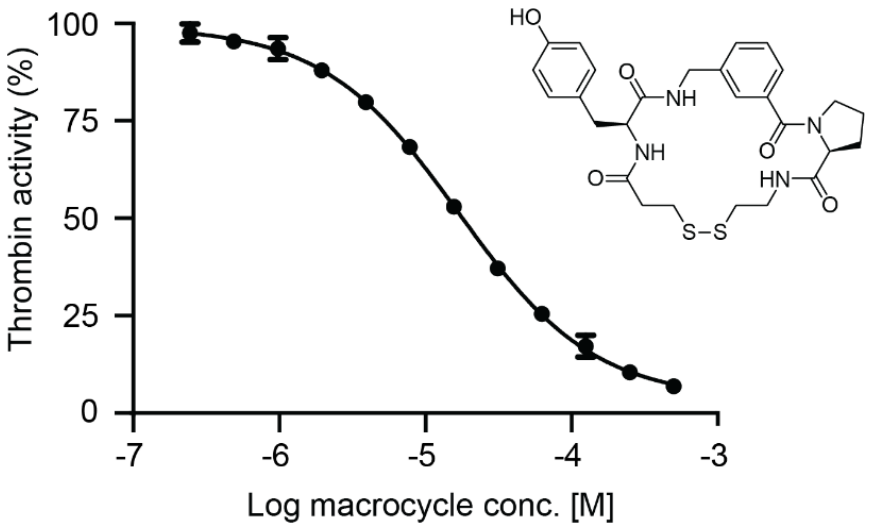
b

Well	Sequence	% Product	% Linear	% tBu	% Dimer	% other
A1	Mpa-W-I-I-Mea	87	0	13	0	0
B2	Mpa-W-III-II-Mea	86	0	5	6	3
C3	Mpa-I-W-III-Mea	85	13	2	0	0
D4	Mpa-III-W-IV-Mea	58	23	5	14	0
E5	Mpa-II-I-W-Mea	96	0	4	0	0
F6	Mpa-IV-II-W-Mea	41	34	9	12	4
G7	Mpa-Y-II-III-Mea	92	3	5	0	0
H8	Mpa-Y-IV-IV-Mea	85	3	12	0	0
A9	Mpa-I-Y-I-Mea	90	0	10	0	0
B10	Mpa-III-Y-II-Mea	94	0	0	5	1
C11	Mpa-I-III-Y-Mea	96	0	3	0	1
D12	Mpa-III-IV-Y-Mea	92	0	6	0	2

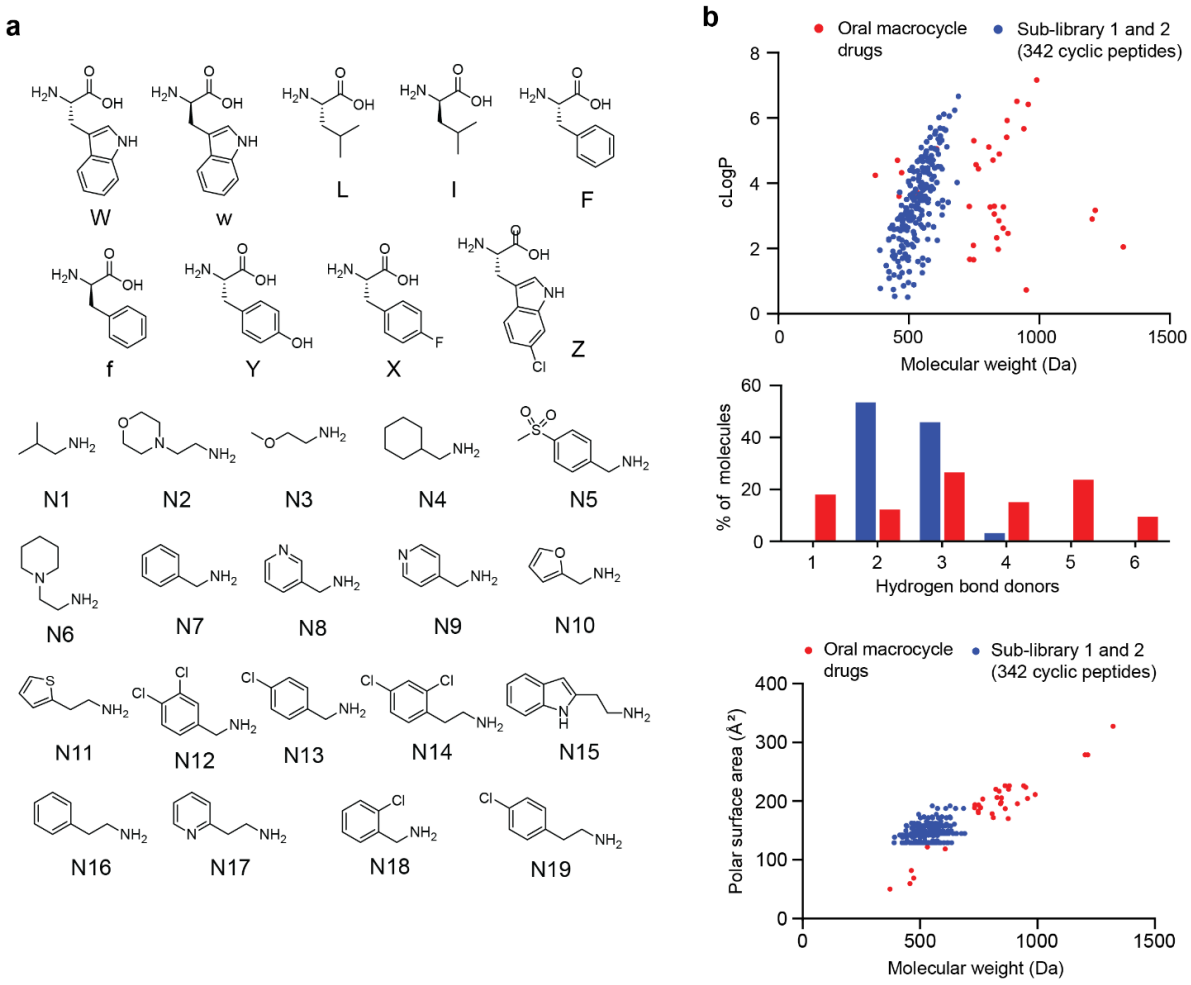
Supplementary Figure 5. Sample peptides analyzed from the 96 cyclic peptide library. (a) Layout of the synthesis plate showing the sample peptides that were randomly selected (yellow). (b) Purity of peptides analyzed by LC-MS analysis. The AUC of peaks recorded at 220 nm was determined and the % of each species calculated. The codes for the amino acid building blocks shown in Figure 3 are indicated.



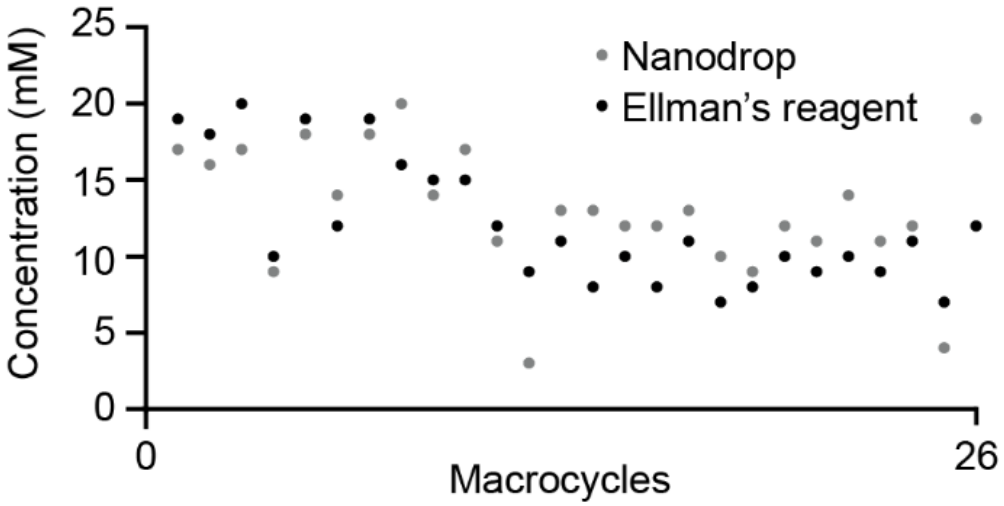
Supplementary Figure 6. UHPLC chromatograms of the 12 sample peptides analyzed from the 96 cyclic peptide sub-libraries 1 to 3. Samples were run on a 0 to 60% MeCN gradient over 6 minutes. For impurities that were not identified, only the observed mass is indicated. The chemical structures and molecular masses (monoisotopic masses) of the desired cyclic peptides are indicated.



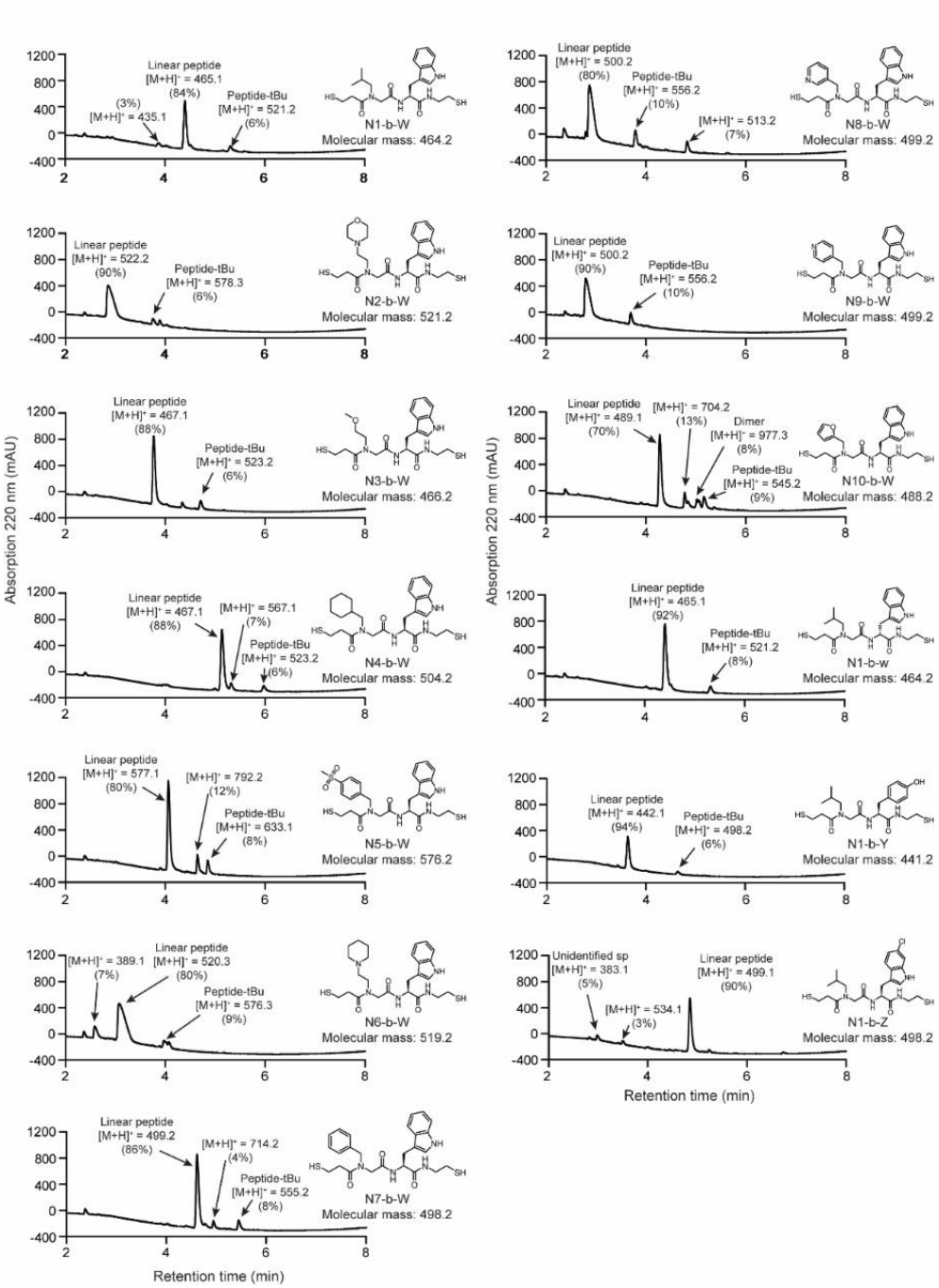
Supplementary Figure 7. Activity of the most active hit found in the thrombin screen. Residual activity of thrombin is shown for increasing concentrations of the HPLC-purified cyclic peptide Mpa-Tyr-II-Mea. Mean values and standard deviations of three independent measurements are shown.



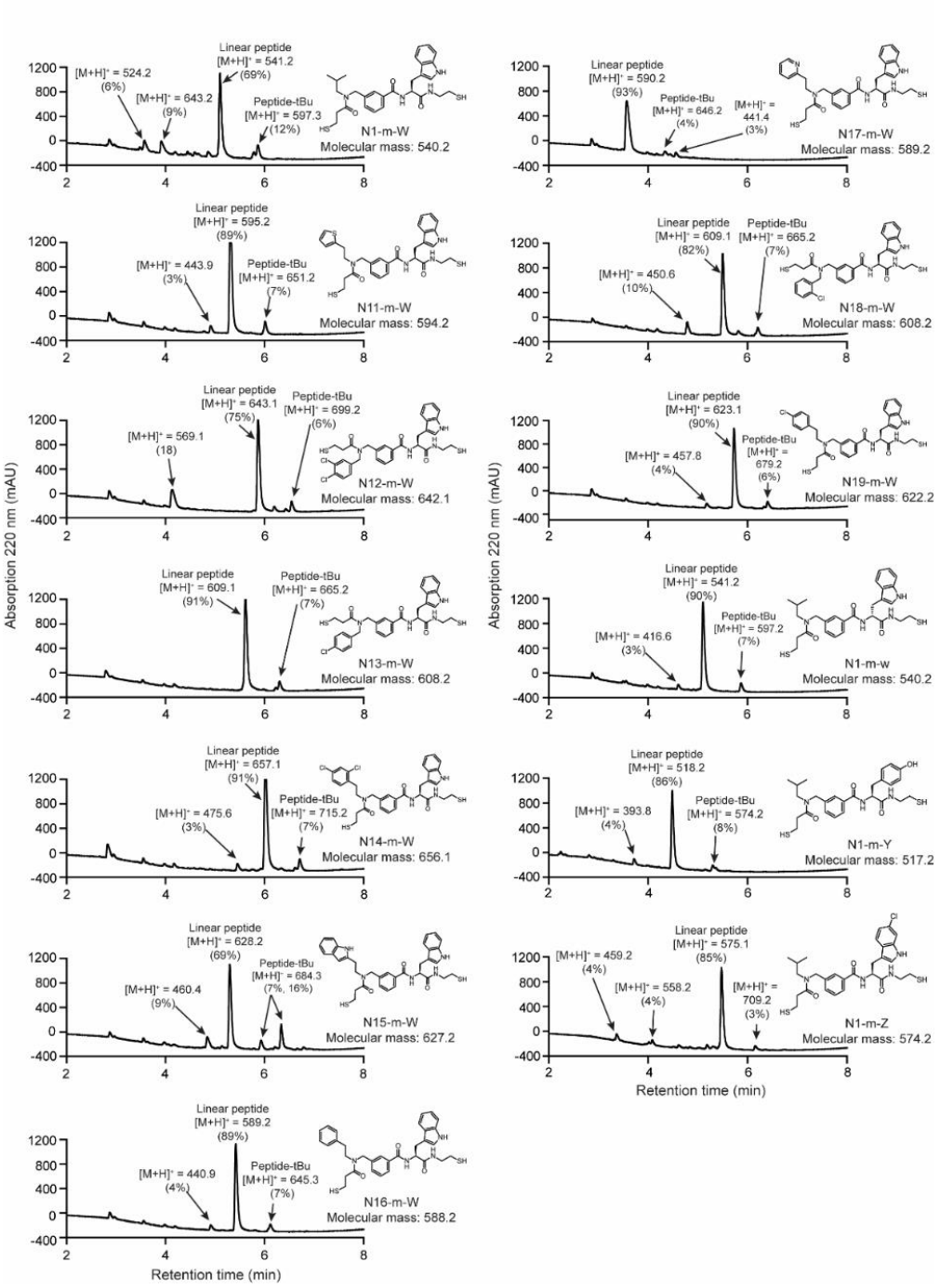
Supplementary Figure 8. Library composed of 342 small cyclic N-alkylated peptides. (a) Building blocks used for the synthesis of the combinatorial library. (b) Comparison of molecular weight, calculated logP, number of H-bond donors, and polar surface area in cyclic peptides of the library and 34 oral macrocycle drugs or drug candidates. Values were calculated using DataWarrior. Oral macrocycle drugs were taken from Kihlberg and colleagues.



Supplementary Figure 9. Quantification of cyclic peptides based on absorbance and number of thiol groups.



Supplementary Figure 10. LC-MS analysis of the peptide/peptoid library. 26 samples of the 342-member library, reduced for concentration determination using Ellman's reagent, were analyzed by LC-MS. Sample peptides were chosen as follows: half of the peptides are from sub-library 1 and 2, respectively. From sub-library 1, peptides containing L-Trp and the amine building blocks N1-N9 were analyzed, as well as three peptides containing D-Trp, L-Tyr, and L-Cl-Trp. From sub-library 2, peptides containing L-Trp and the amine building blocks N1 and N10-N19 were analyzed, as well as three peptides containing D-Trp, L-Tyr, and L-Cl-Trp. The percentage of desired peptide was determined by quantifying the AUC of product and side-product peaks recorded at 220 nm. Samples were run on a 5-80% MeCN gradient over 6 minutes.



Supplementary Figure 10. Continued.

3.8 References

(1) Zorzi, A.; Deyle, K.; Heinis, C. Cyclic Peptide Therapeutics: Past, Present and Future. *Current Opinion in Chemical Biology*. 2017, pp 24–29. https://doi.org/10.1016/j.cbpa.2017.02.006.

(2) White, C. J.; Yudin, A. K. Contemporary Strategies for Peptide Macrocyclization. *Nat. Chem.* **2011**, *3* (7), 509–524. https://doi.org/10.1038/nchem.1062.

(3) Rietman, B. H.; Smulders, R. H. P. H.; Eggen, I. F.; Van Vliet, A.; Van De Werken, G.; Tesser, G. I. Protected Peptide Disulfides by Oxidative Detachment from a Support. *Int. J. Pept. Protein Res.* **1994**, *44* (3), 199–206. https://doi.org/10.1111/j.1399-3011.1994.tb00161.x.

(4) Zoller, T.; Ducep, J. B.; Tahtaoui, C.; Hibert, M. Cyclo-Release Synthesis of Cyclic Disulfides on Solid Phase. *Tetrahedron Lett.* **2000**, *41* (51), 9989–9992. https://doi.org/10.1016/S0040-4039(00)01852-9.

(5) Ganesan, A. Cyclative Cleavage as a Solid-Phase Strategy. In *Linker Strategies in Solid-Phase Organic Synthesis*; Scott, P. J. H., Ed.; John Wiley & Sons, Ltd, 2009. https://doi.org/10.1002/9780470749043.ch4.

(6) Marsault, E.; Hoveyda, H. R.; Gagnon, R.; Peterson, M. L.; Vézina, M.; Saint-Louis, C.; Landry, A.; Pinault, J. F.; Ouellet, L.; Beauchemin, S.; Beaubien, S.; Mathieu, A.; Benakli, K.; Wang, Z.; Brassard, M.; Lonergan, D.; Bilodeau, F.; Ramaseshan, M.; Fortin, N.; Lan, R.; Li, S.; Galaud, F.; Plourde, V.; Champagne, M.; Doucet, A.; Bhérer, P.; Gauthier, M.; Olsen, G.; Villeneuve, G.; Bhat, S.; Foucher, L.; Fortin, D.; Peng, X.; Bernard, S.; Drouin, A.; Déziel, R.; Berthiaume, G.; Dory, Y. L.; Fraser, G. L.; Deslongchamps, P. Efficient Parallel Synthesis of Macrocyclic Peptidomimetics. *Bioorganic Med. Chem. Lett.* **2008**, *18* (16), 4731–4735. https://doi.org/10.1016/j.bmcl.2008.06.085.

(7) Guo, Z.; Hong, S. Y.; Wang, J.; Rehan, S.; Liu, W.; Peng, H.; Das, M.; Li, W.; Bhat, S.; Peiffer, B.; Ullman, B. R.; Tse, C. M.; Tarmakova, Z.; Schiene-Fischer, C.; Fischer, G.; Coe, I.; Paavilainen, V. O.; Sun, Z.; Liu, J. O. Rapamycin-Inspired Macrocycles with New Target Specificity. *Nat. Chem.* **2019**, *11* (3), 254–263. https://doi.org/10.1038/s41557-018-0187-4.

(8) Gless, B. H.; Olsen, C. A. Direct Peptide Cyclization and One-Pot Modification Using the MeDbz Linker. *J. Org. Chem.* **2018**, *83* (17), 10525–10534. https://doi.org/10.1021/acs.joc.8b01237.

(9) Arbour, C. A.; Belavek, K. J.; Tariq, R.; Mukherjee, S.; Tom, J. K.; Isidro-Llobet, A.; Kopach, M. E.; Stockdill, J. L. Bringing Macrolactamization Full Circle: Self-Cleaving Head-to-Tail Macrocyclization of Unprotected Peptides via Mild N-Acyl Urea Activation. *J. Org. Chem.* **2019**, *84* (2), 1035–1041. https://doi.org/10.1021/acs.joc.8b02418.

(10) Giordanetto, F.; Kihlberg, J. Macrocyclic Drugs and Clinical Candidates: What Can Medicinal Chemists Learn from Their Properties? *J. Med. Chem.* **2014**, *57* (2), 278–295. https://doi.org/10.1021/jm400887j.

4 Synthesis and screening of large macrocycle libraries by late stage modification at picomole scale

4.1 Summary

Synthesis and screening of large macrocycle libraries by late stage modification at picomole scale

Sevan Habeshian[1], Manuel Merz[1], Gontran Sangouard[1], Mischa Schüttel[1], Jonathan Vesin[2], Julien Bortoli Chapalay[2], Gerardo Turcatti[2], Laura Cendron[3], Alessandro Angelini[4,5], and Christian Heinis[1]*

[1]Institute of Chemical Sciences and Engineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland.

[2]Biomolecular Screening Facility, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.

[3]Department of Biology, University of Padova, 35131 Padova, Italy.

[4]Department of Molecular Sciences and Nanosystems, Ca' Foscari University of Venice, Via Torino 155, Venezia Mestre, Venice 30172, Italy.

[5]European Centre for Living Technologies (ECLT), Ca' Bottacin, Dorsoduro 3911, Calle Crosera, Venice 30124, Italy.

Author contribution: I, along with Christian Heinis conceptualized the method and developed the strategy. I performed proof-of-concept experiments, synthesis and screening of both libraries, hit compound synthesis and characterization, and hit optimization steps. I also helped to make all figures, wrote the first draft of the manuscript, and contributed to editing. Manuel Merz contributed to acylation reaction optimization. Gontran Sangouard established the MDM2 FP assay that was utilized. Mischa Schüttel implemented the screening instrumentation into our laboratory. Jonathan Vesin, Julien Chapalay, and Gerardo Turcatti provided access to and operation of the instrumentation at the EPFL biomolecular screening facility. Laura Cendron and Alessandro Angelini obtained and solved the co-crystal structure of thrombin bound to inhibitor macrocycle.

This chapter is based on a manuscript for publication.

4.2 Abstract

Macrocycles hold great promise as a therapeutic modality due to their ability to bind challenging targets while also being able to cross membranes. However, the generation of macrocycle-based ligands to new targets is currently hindered by a lack of large macrocycle libraries needed in HTS campaigns. We have overcome this limitation by establishing an approach based on tethering large numbers of chemical fragments to peripheral groups of structurally diverse macrocyclic scaffolds in a combinatorial fashion and at a picomole scale in nL volumes using acoustic droplet ejection technology. In a proof-of-concept, we generated a library of 19,968 macrocycles by conjugating 104 carboxylic acid fragments to 192 macrocyclic scaffolds. The high reaction efficiency and small number of side products of the acylation reactions allowed for HTS without purification and thus a large throughput. In screens, we identified nanomolar inhibitors of thrombin (K_i = 44 ± 1 nM) and the MDM2:p53 protein-protein interaction (K_i = 31 ± 6 nM). The approach offers a dramatic increase in the number and rate at which macrocycles can be synthesized and screened and is generally applicable to any protein target.

4.3 Introduction

Omics technologies have greatly expanded the number of potential drug targets over the last years, offering enormous opportunities for addressing unmet medical needs. However, for many proteins such as those with flat, featureless surfaces, or for protein-protein interactions, it has been difficult or impossible to generate ligands based on classical small molecules[1]. Monoclonal antibodies (mAbs), a modality that can bind to highly challenging targets, cannot offer a solution for intracellular targets due to their membrane impermeability. Prominent examples of currently undruggable targets are MYC transcription factors, RAS family proteins, and β-catenin[2]. A molecule class that could potentially fill the important gap left by small molecules and mAbs is that of macrocycles. These ring-shaped structures can bind to challenging targets, and yet can cross cell membranes and be orally available if they have a reasonably small size (well below 1 kDa) and polar surface area (below 200 $Å^2$).[3,4] A number of macrocyclic compounds, most of them derived from natural sources and a few developed *de novo*, are used as successful drugs and showcase the enormous potential of the drug class of macrocycles. Some examples include tacrolymus, rifampicin, erythromycin, lorlatinib, and glecaprevir[5].

For generating macrocyclic ligands to targets of interest, various principles and approaches have been developed and applied, including diversity oriented synthesis strategies[6], one-bead-one-compound methods[6], biological display methods such as phage and mRNA display[7,8], and DNA-encoding of chemical libraries[9–11]. A common challenge in methods involving combinatorial synthesis principles are product mixtures from incomplete reactions, such as the usually difficult macrocyclization step, which can complicate hit identification, in particular for methods where libraries are pooled such as in split-and-mix approaches. Biological display methods give cleaner products but have other limitations, the major one being that ligands developed so far have been rather large (> 1 kDa) and thus showed limited membrane permeability. Classical high-throughput screening in microwell plates could in principle offer a robust alternative to the newer methods, but the macrocycle libraries currently offered by major providers such as Asinex (12,076 compounds), ChemBridge (11,000 compounds) or Polyphor[12] are rather small, limiting the chance of finding ligands to targets of interest. For developing macrocycle-based ligand to challenging targets, it would be desirable to have access to large libraries comprising ten- to hundred-thousands of highly structurally and chemically diverse macrocyclic compounds that can be screened in microwell plates, which offers maximal flexibility in assay format and readout.

Herein, we proposed and established a general principle to rapidly generate and screen thousands of macrocyclic compounds that is based on late-stage diversification of *m* macrocyclic scaffolds with *n* fragments (Fig. 1a). We intended to chemically link the fragments to peripheral groups of the scaffolds in nanoliter droplets and at picomole scale, to access large libraries at high speed and with minimal reagent consumption. For example, 200 scaffolds are combinatorially reacted with 100 fragments to generate a library of 20,000 macrocyclic compounds. Key in this approach would be the use of efficient coupling reactions so that library size-limiting purification steps could be omitted. We envisioned that the *m* macrocyclic scaffolds and the *n* fragments are simply "mixed and reacted" in wells of 384 or 1,536 plates, and directly assayed in the same wells by adding target protein and assay reagents. For transferring reagents, we intended to employ contactless acoustic dispensing technology that we and others recently found highly suited for the combinatorial synthesis of compound libraries[13,14].

We had previously generated macrocyclic libraries by combinatorially cyclizing short peptides with diverse linkers, but due to the challenging nature of macrocyclization reactions, the reaction yields varied substantially. In addition, the limited number of commercially available linker reagents (< 20) had limited the diversities of these libraries[14–16]. In contrast, the herein proposed approach is based on a late-stage modification of peripheral groups (e.g. amines) in macrocyclic scaffolds, which allows more efficient chemical diversification with thousands of commercially available fragments (e.g. carboxylic acids), and offers access to chemically and structurally enormously diverse macrocycle libraries.

4.4 Results and discussion

Acylation of macrocyclic scaffolds in nanoliter volumes

For combinatorially diversifying macrocyclic scaffolds as shown in Fig. 1a, we chose to modify peripheral amines with carboxylic acid-based fragments. Such *N*-acylation reactions are efficient and selective, as demonstrated in the synthesis of DNA-encoded chemical libraries[17] or the affinity maturation of small molecule ligands[18,19], and many chemically and structurally diverse carboxylic acids are commercially provided. We tested the reaction with the model scaffold 1 shown in Fig. 1b, carrying a primary amine as a peripheral group (lysine side chain, amine in blue), reacting it with eight structurally diverse carboxylic acids **1**—**8** (Fig. 1c), initially in large, 4 μL reaction volumes and thus at a rather large scale (40 nmol scaffold). In order to efficiently convert the scaffold into the desired products, we chose to react it with a 4-fold molar excess of carboxylic acid (final conc. of 10 mM scaffold, 40 mM carboxylic acid) and HBTU as activating agent and DIPEA as base. Excess of carboxylic acid over scaffold was chosen as the acids were expected to contribute less in target-binding than the scaffold moiety due to the smaller size, and non-reacted excess of the acids was likely not going to hide the activity of the synthesized macrocyclic compounds in subsequent screens of the crude products. The only expected byproducts of the acylation reaction were excess carboxylic acid, HOBt, and tetramethylurea, none of which should be incompatible with biochemical assays. LC-MS analysis after three hours of reaction showed full conversion of the scaffold with all but one of the eight acids (top numbers in Fig. 1c and Supplementary Fig. 1).

We subsequently tested the combinatorial diversification of the same cyclic peptide scaffold in 80 nL reaction volume, and thus a 50-fold smaller scale (800 pmol) that we planned for the high-throughput synthesis and screening. The sub-nanomole scale was important for us so that single-digit μmol quantities of scaffold, that could be purchased in milligram quantities or synthesized in wells of 96-well plates (5 μmol scale)[20], was sufficient to synthesize several thousands of macrocycles based on each scaffold. In addition, we aimed at applying acoustic dispensing technology for transferring reagents, which is suitable for transfering small quantities of reagent in nanoliter volumes but not large ones. Acoustic dispensing has the additional great advantage that reagents can be transferred contactless, which does not require pipetting tips, accelerating the speed of dispensing, and reducing waste and cost[13,14]. Application of the reaction conditions described above for 4 μL volumes to picomole scale reactions in 80 nL volumes and reagent transfer by acoustic waves led to incomplete acylation of the model scaffold 1 (first of the two lower numbers in Fig. 1c; Supplementary Fig. 1), and called for optimization of the acylation reaction. We hypothesized that the low yields were

related to the use of DIPEA, as the base is relatively volatile and may be partially evaporated during the transfer in 2.5 nL droplets by acoustic waves. We conducted tests with the non-volatile bases DABCO and the sodium salt of HEPES, and the volatile NMM that could be applied at higher concentration due to a high solubility in DMSO. With all new bases, the macrocycles were quantitatively acylated by test acids (**1**, **4**, **7**), despite the transfer in nanoliter droplets (Supplementary Fig. 2). Application of the preferred base DABCO to further acids (**1**—**8**) showed that the conditions were suitable for quantitative modification of a peripheral primary amine at nanomole scale (second one of the two lower numbers in Fig. 1c; Supplementary Fig. 3).

While the lysine ε-amine in model scaffold 1 is exposed and thus well accessible for acylation, primary or secondary amino groups closer to or within the macrocyclic backbones might be harder to modify. To assess the reactivity of such amines, we ordered from a commercial library provider four random macrocyclic scaffolds containing primary or secondary amines (model scaffolds 2 to 5; Fig. 1d) and applied the same reaction conditions and carboxylic acids **1**—**8** as before and at the same scale (80 nL volume, 800 pmol, DABCO as a base) (Fig. 1c). The model scaffolds 2, 3, and 5 were quantitatively acylated with all acids, and scaffold 4 showed high product yields (acids **1**, **4**—**8**: 100%, acid **2**: 93%, acid **3**: 69%; Supplementary Fig. 4), showing that the picomole-scale acylation in nanoliter volumes worked efficiently also for less accessible amines.

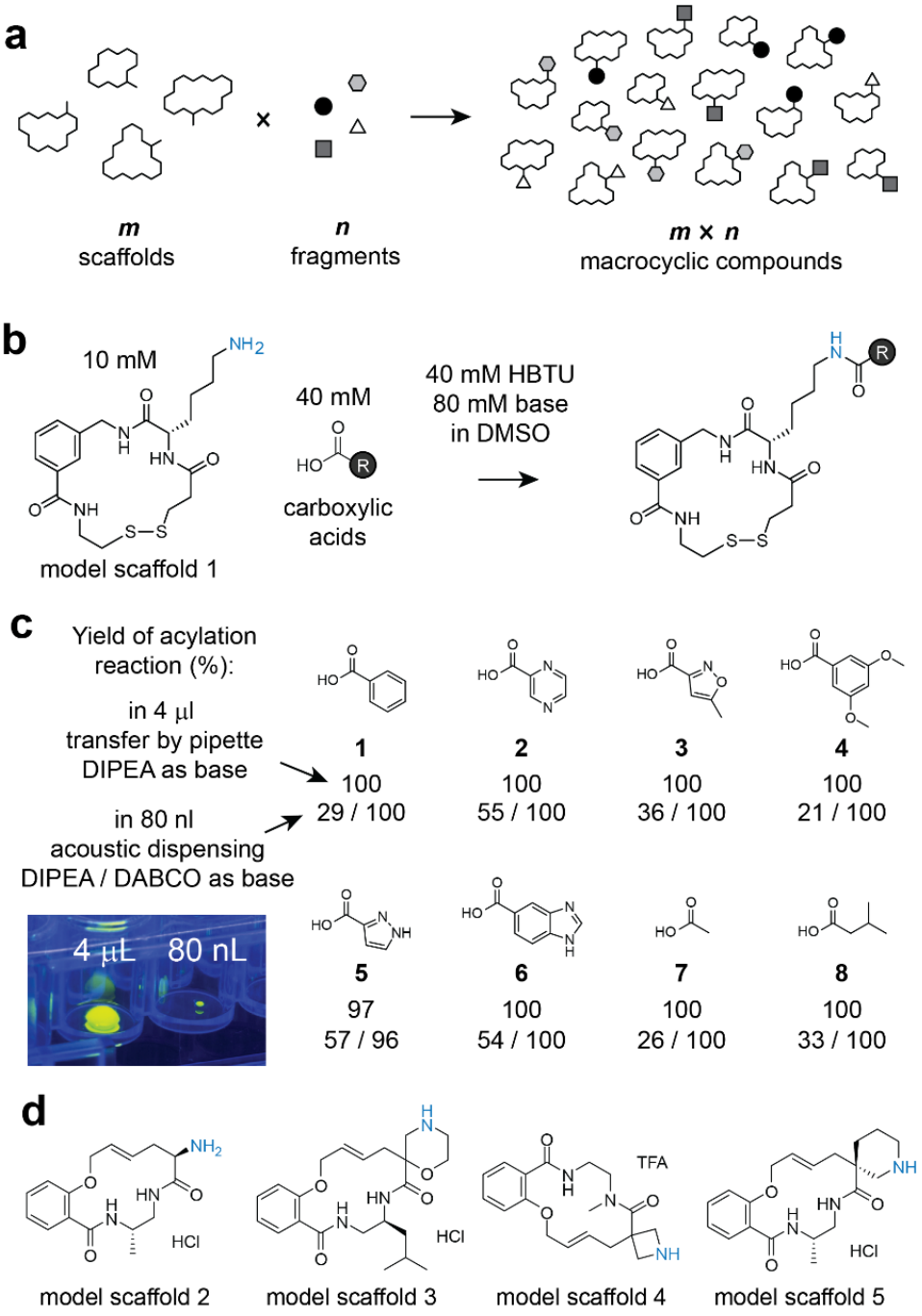


Figure 1 Diversification of macrocyclic scaffolds by combinatorially appending fragments to peripheral groups. a General principle of approach. **b** Model macrocycle scaffold containing a peripheral primary amine is modified by acylation. **c** Reaction of model macrocycle with indicated acids. The upper number shows conversion in 4 µL volume by pipetting and the lower number in 80 nL and acoustic liquid transfer, the first number with DIPEA and the second one with DABCO. Images of two droplets in a 96-well plate are shown to demonstrate the difference in scale. The droplets contain fluorescein and were exposed to UV light for visualization. **d** Randomly chosen, non-peptide scaffolds containing less accessible amino groups (in blue).

Scaffold and macrocycle library synthesis

We next assessed if the picomole scale synthesis procedure is suitable for generating a large macrocycle library. We synthesized 384 cyclic peptide scaffolds having random structures and one amino group, using a recently developed approach for producing large numbers of small cyclic peptides in 96-well plates[20]. In brief, short peptides were synthesized on solid phase via a disulfide linker, the protecting group removed, and the peptide released in a cyclative reaction, yielding disulfide-cyclized peptides with 90% or higher purity (Fig. 2a). We synthesized scaffolds that contained three amino acids, one having a primary amine in the side chain (chosen from six aa), one an α-amino acid with a random side chain (chosen from 15 aa), and one a random backbone structure (chosen from six aa) (Fig. 2b and 2c; scaffold formats 1a-f). In a smaller sub-set of scaffolds, the primary amino groups were introduced through N-terminal cysteine residues (Supplementary Fig. 5; Scaffolds 1g-l). Of the 3,780 scaffolds that could theoretically be assembled based on the indicated formats and amino acid building blocks, we synthesized 384 randomly chosen ones (10%). Quantification of cyclic peptides containing a Trp residue (45 scaffold) by absorption showed that most molecules were obtained in high yields (average conc. = 8.1 mM; Fig. 2d). Given a relatively narrow distribution of the yields, we did not normalize the concentrations for further use.

We next combinatorially reacted the 384 scaffolds with 12 carboxylic acids by acoustic dispensing, yielding 4,608 different macrocycles (Fig. 2e). Compared to the test acylation reactions described above, we reduced the reagent transfer volumes slightly (20 nL of 8.1 mM scaffold and 20 nL of 80 mM pre-activated acid) in order to increase the speed of dispensing and to reduce scaffold consumption (160 pmol scale). We also applied a higher excess of acid (10-fold; final concentrations of 4 mM scaffold and 40 mM carboxylic acid) to ensure quantitative modification of scaffolds even for less accessible amines (Fig. 2f). The contactless liquid transfer for one 384-well plate took 4 minutes and the synthesis of the 4,606 compound library approximately one hour. After incubation for five hours, we quenched the reaction by the addition of 5 μL of 100 mM Tris buffer and incubation overnight. Computational prediction of the physicochemical properties of the 4,608-member library showed that the vast majority of the macrocycles fall into a space that is predicted to be cell permeable[21,22] (Fig. 2g and Supplementary Fig. 6), suggesting that hits identified in the library could be developed into membrane permeable or orally available lead compounds.

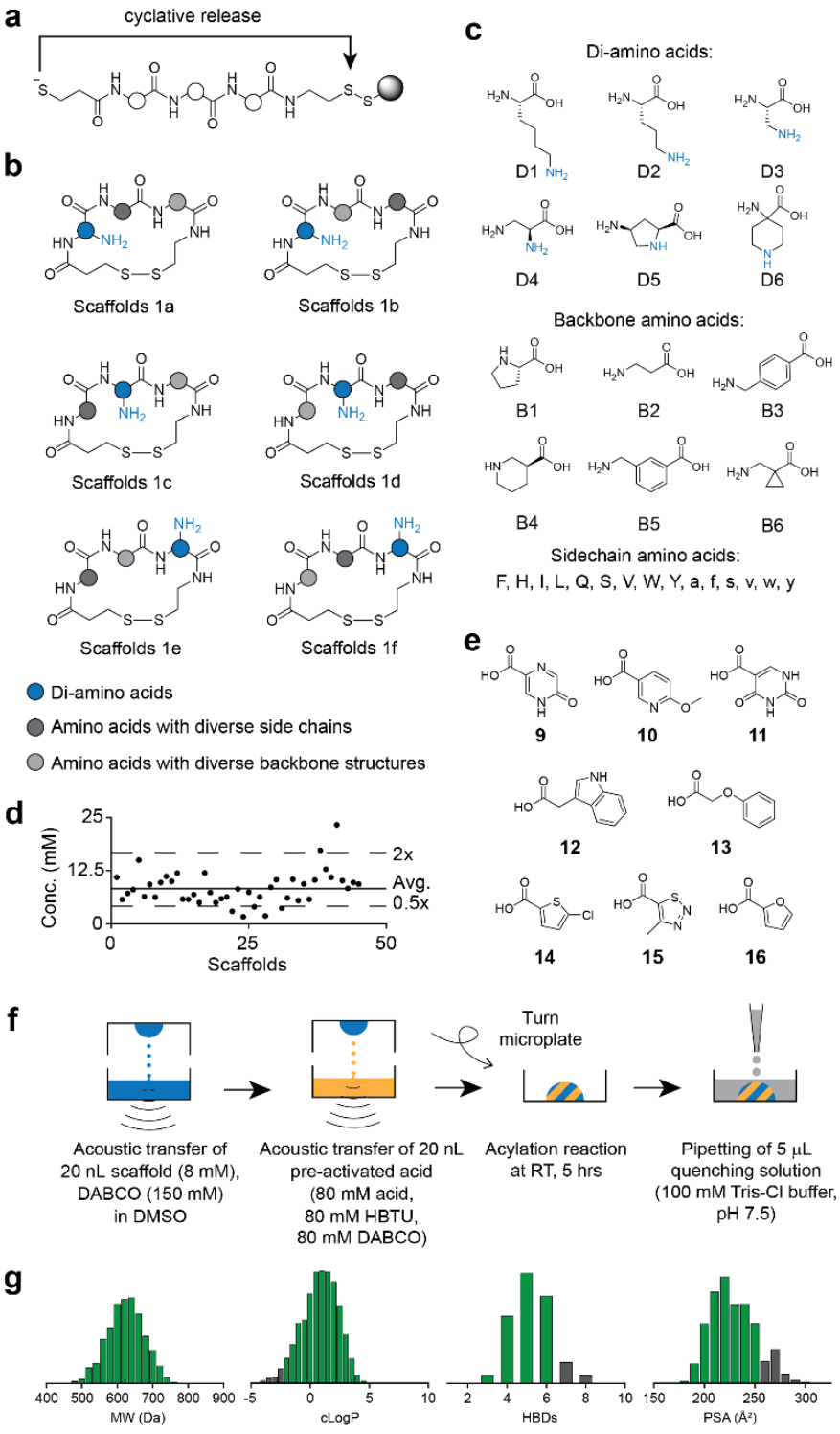


Figure 2 Acylation of amines in macrocyclic scaffolds and library generation. a Cyclative disulfide release of side chain-deprotected peptides. **b** Formats of cyclic peptide scaffolds. **c** Amino acids used for scaffold synthesis. **d** Yields of 45 tryptophan-containing scaffolds determined by absorption measurement. **e** Carboxylic acids **9** to **16** that were used along with acids **1** to **3** to diversify the scaffolds shown in panel b. **f** Schematic procedure for macrocycle library synthesis by acoustic liquid transfer. Reaction conditions are indicated. **g** Computationally predicted physicochemical properties of macrocycles in Library 1. Molecules in green regions have a value for the indicated property that is within a range predicated to be suitable for developing membrane permeable macrocycles.

Screening 4,608 macrocycles for thrombin inhibition

We next screened the macrocycle library against the coagulation protease and therapeutic target thrombin. A thrombin inhibitor, dabigatran etexilate, is already used as an anti-thrombosis drug but suffers from a low oral availability (6.5%) despite its prodrug nature, having two removable charge-shielding groups[23]. All macrocycles of the 4,608-member library are non-charged and could potentially offer leads for developing drugs with higher oral availability. We screened the macrocycle library by dispensing to the 384-well reaction plates, containing in each well a different macrocycle, 5 μL of thrombin (2 nM final conc.) and subsequently 5 μL of a fluorogenic substrate to measure the residual thrombin activity and thus the extent of protease inhibition. The concentration of macrocyclic compounds in the screen was around 10 μM. A fraction of 0.2% of the reactions (9 out of 4,608) inhibited thrombin > 50% (Fig. 3a), all of them being macrocycles containing the acid **14**, chlorothiophene carboxylic acid. Repetition of all macrocycle synthesis reactions that involved chlorothiophene carboxylic acid (384 scaffolds × acid **14**) and the thrombin screen identified essentially the same top hits and thus showed a high reproducibility for both, the picomole-scale diversification reactions and the activity screen (Fig. 3b).

Given that a control well with **14** only (no scaffold) did not inhibitor thrombin, and also that most scaffolds modified with acid **14** did not inhibit thrombin either, we expected that the acid fragment alone has a rather weak affinity for thrombin. Indeed, chlorothiopheneamide, a derivative of **14** that mimics the peripheral structure of macrocycles containing **14**, inhibited thrombin with only a weak affinity (K_i of 380 μM). This finding indicated that the cyclic peptide scaffolds contributed much of the binding energy. The best three hits, M1, M2, and M3, were highly related in structure, all being based on scaffolds of the format cyclo(D3-B5-Xaa) wherein the amino acids Xaa were all α-amino acids with hydrophobic side chains (D-Val, L-Phe, L-Val; Fig. 3c).

We next assessed if the activity observed in the screens derived from the anticipated macrocyclic compounds or side products. Towards this end, we repeated the reaction of scaffold and carboxylic acid **14** for two hits (M1 and M2) at a 250-fold larger scale (40 nmol), ran the reactions over a RP-HPLC column to separate 20 fractions each combining products that eluted in one minute, lyophilized the fractions and measured the thrombin inhibition activity (Fig. 3d). For both reactions, the fractions containing the desired macrocyclic products showed the highest activity, indicating that the hits were indeed identified based on activities of the anticipated products. The purified macrocycles M1, M2, and M3 inhibited thrombin with K_is of 44 ± 1 nM, 165 ± 38 nM, and 125 ± 8 nM, respectively (Fig. 3c, Supplementary Fig. 7). Depending on the therapeutic application, the reducible disulfide bonds present in all scaffolds screened herein are not desired, and we thus tested if they could be replaced by non-reducible

linkages. We synthesized M4 and M5 that contained dithioacetal or thioether linkers (Supplementary Fig. 7). M4 and M5 inhibited thrombin with K_is of 83 ± 8 nM and 135 ± 16 nM, showing that the disulfide bond could easily be replaced without losing much activity.

X-ray crystallographic analysis of M1 bound to human thrombin at 2.35 Å resolution (PDB 6Z48) revealed that M1 was an active-site inhibitor, with the chlorothiophene group pointing into the S1 specificity pocket and the macrocyclic ring filling the S2 (disulfide region) and S4 (valine) spaces (Fig. 3e). The interaction of a chlorothiophene group with the S1 site of a trypsin-like serine protease has previously been reported for the FX inhibitor rivaroxaban[23]. As found for the FX inhibitor, the chlorothiophene forms an aryl chloride-π interaction with a tyrosine residue at the bottom of the pocked (Fig. 3f). The structural information of M1 bound to thrombin visually illustrated how both the macrocyclic scaffold and the carboxylic acid fragment contribute to the overall binding and to the inhibition of thrombin.

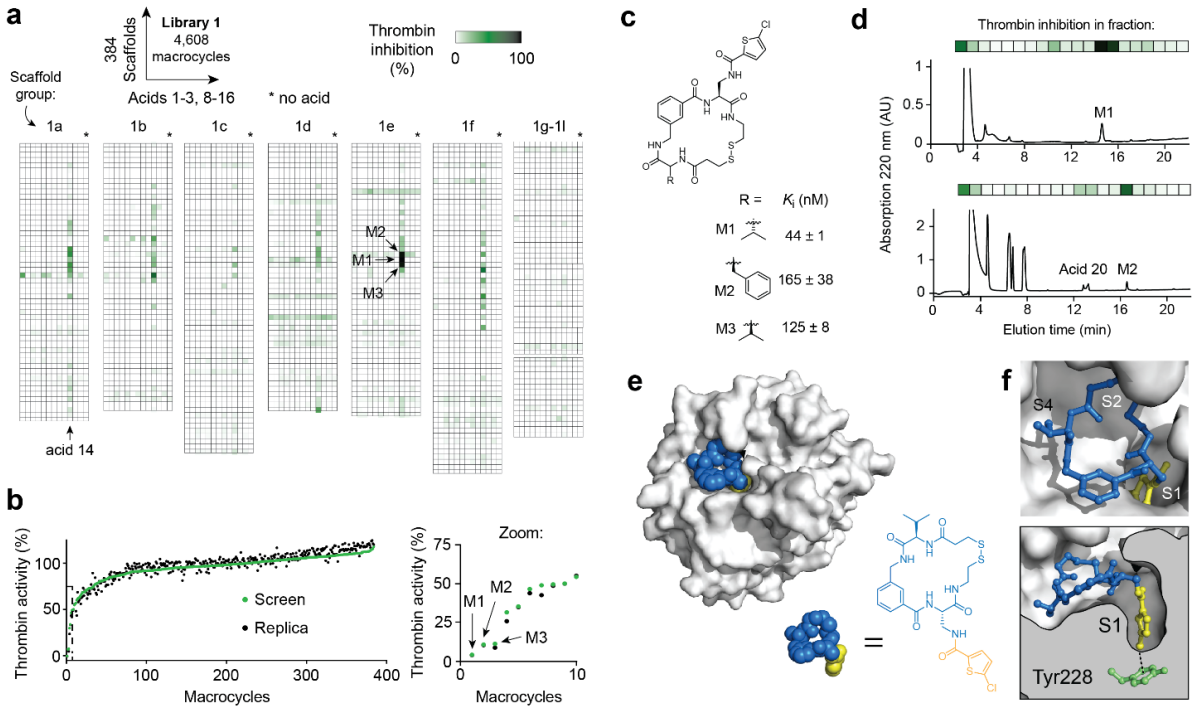


Figure 3 Screening of macrocyclic compound library against thrombin, hit identification and structure. a Heat map showing thrombin inhibition for each macrocycle. The amino acid sequences of the scaffolds are provided in Supplementary Data 1. **b** Replica reaction and screen of all macrocycles containing acid **14**. **c** Chemical structures and activities of top three hits M1 to M3. Mean values and SDs of three independent measurements are shown. **d** Chromatographic separation acylation reaction yielding M1 and M2 and analysis of fractions for thrombin inhibiting species. **e** X-ray structure of human thrombin with bound M1 solved at 2.35 Å resolution (PDB 6Z48). The inhibitor is shown in space fill model with the scaffold in blue and the carboxylic acid in yellow. **f** Zoomed structure of thrombin with the sub-sites indicated. The chlorothiophene group fills the S1 sub-site and forms an interaction with Tyr228.

Protein-protein interaction inhibitors and MDM2 screen

Macrocycles have received much interest due to their ability to inhibit protein-protein interactions (PPIs). A prototypical PPI disease target is MDM2:p53 for which much effort has been invested to develop inhibitors. Overexpression of MDM2 inhibits the activity of the tumor repressor p53, and MDM2 binders blocking the MDM2-p53 interaction are of interest for developing new anti-cancer therapies[23]. To test the new approach with this PPI target, we synthesized 192 structurally diverse cyclic peptide scaffolds, all based on three random amino acids of which one contained an amino group for lateral diversification (Supplementary Fig. 9; Scaffolds 2a-f). In order to increase the chances of identifying binders, we included in all scaffolds either tryptophan or phenylalanine, two amino acids that form key interactions in stapled peptides that bind MDM2 and inhibit the MDM2:p53 interaction[23]. The cyclic peptide scaffolds were synthesized in 96-well plates as described for Library 1 above and were obtained in an average concentration of 12.9 mM and a purity of around 90%. As before, we modified the scaffolds by acylation in a combinatorial fashion at picomole scale, this time using 104 carboxylic acids (Supplementary Fig. 10). Starting from 192 scaffolds, we generated 19,968 macrocyclic compounds, and thus a library that is more than 100-fold larger than the number of initial macrocyclic peptides (scaffolds).

We screened Library 2 by dispensing the target protein MDM2 and reporter peptide to the 384-microwell plates containing the macrocycle reactions. The fluorescent reporter peptide binds to the p53-binding interface of MDM2 (K_d = 0.5 μM) and its displacement by macrocycles can be followed by measuring the change in fluorescence polarization. We displayed the screening result again in an array of scaffold (vertical) and fragment (horizontal) combinations with the color indicating the extent of reporter peptide displacement from MDM2 (Fig. 4a). The heat map showed a rather speckled pattern due to high noise in the FP competition assay. However, a group of hits found on a horizontal line, and thus macrocycles sharing the same scaffold, cyclo[Trp-D3-B4], stood out, with the best hits being M6 (acid **14**), M7 (acid **25**), and M8 (acid **91**), as highlighted in Fig. 4a. Repetition of the reactions confirmed that the same scaffold and acid combinations yielded the most active products (Supplementary Fig. 11a). Activity measurements of HPLC-separated products and side-products from hits M6, M7, and M8 revealed that the active species were the anticipated macrocyclic structures (Supplementary Fig. 11b).

The purified macrocycles M6, M7, and M8 displaced the fluorescent peptide probe from MDM2 efficiently (Fig. 5b). As the competition assay was not suited for measuring binding constants below 1 μM due to the requirement of using MDM2 at a minimal concentration of 1.2

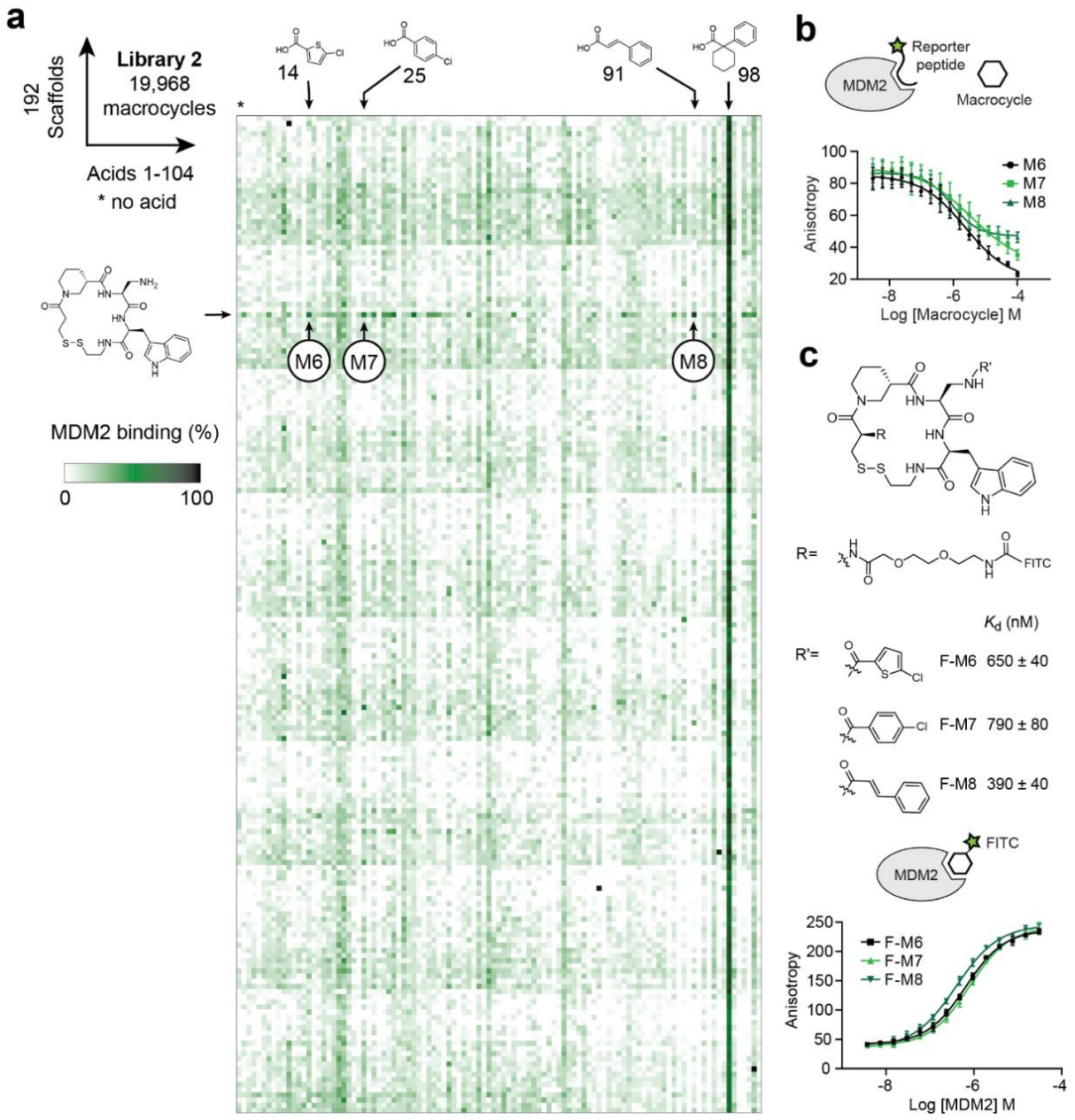


Figure 4 Screen against MDM2 and hit characterization. a 192 macrocycle scaffolds (Supplementary Fig. 8) were combinatorially acylated with 104 carboxylic acids, and the products screened for displacing a p53-based fluorescent peptide derived from human MDM2. Final macrocycle concentrations were 10 μM. The amino acid sequences of the scaffolds are provided in Supplementary Data 1. **b** Displacement of fluorescent reporter peptide from MDM2 by HPLC-purified hit macrocycles M6 to M8 measured by fluorescence polarization (FP). **c** Chemical structures of M6 to M8 labeled with 5(6)-FAM for measuring the affinity in a direct binding assay by FP. K_d values are means and SDs of three independent measurements.

μM, we synthesized the three macrocycles as conjugates with fluorescein and measured the binding affinities in a direct fluorescence polarization assay (Fig. 5c). The conjugates showed K_d values of 650 ± 50 nM (F-M6), 790 ± 80 nM (F-M7), and 340 ± 40 nM (F-M8).

Iterative picomole scale synthesis and screening

The method for facile generation of macrocyclic compounds allows for iterative synthesis of macrocycle sub-libraries based on hit compounds, and to screen for variants with improved binding affinity. For enhancing the potency of the macrocycle M8 identified in the initial screen, we synthesized the 63 scaffolds shown in Fig. 5a, using amino acids similar to those in M8, being three analogs of nipecotic acid (Nip), three analogs of diamino propionic acid (Dap), and seven analogs of tryptophan (Trp) (3×3×7 = 63). The 63 scaffolds were then diversified with 14 carboxylic acids, which included repeats from the initial library that yielded hits (carboxylic acids **14**, **25**, **91**), as well as related structures mostly being analogs of cinnamic acid **91**, the carboxylic acid of the macrocycle hit M8 (**105**-**115**; Supplementary Fig. 10). In order to identify binders with nanomolar affinity, we screened the 882 macrocycles (63×14) at a 13-fold lower concentration (750 nM) than in the first screen, which corresponded to a 30 pmol scale (Fig. 5b). While the most active macrocycles remained based on the original macrocycle scaffold, the screen identified carboxylic acids that yielded more potent macrocycles, namely **109** that led to 68% reporter peptide displacement at 750 nM (macrocycle M9), compared to 21% for M8 (acid **91**; Fig. 5b). Although the scaffold was not improved, the screen with 63 scaffold variants provided structure-activity relationship data for the macrocyclic ring and showed that all building blocks were essential. We subsequently performed a third cycle of library synthesis and screening, this time acylating the nine most promising scaffolds of the previous screens with 15 additional carboxylic acids, mostly cinnamic acid derivatives with larger substituents. The screen identified M10, a macrocycle based on the original scaffold and acylated with acid **120** that displaced the reporter peptide to 84% from MDM2 at 750 nM, and thus more efficiently than the parental compounds M8 and M9 (Fig. 5c). The macrocycle M10, conjugated to fluorescein, bound MDM2 with a K_d of 31 ± 6 nM, as measured by FP (Fig. 5d).

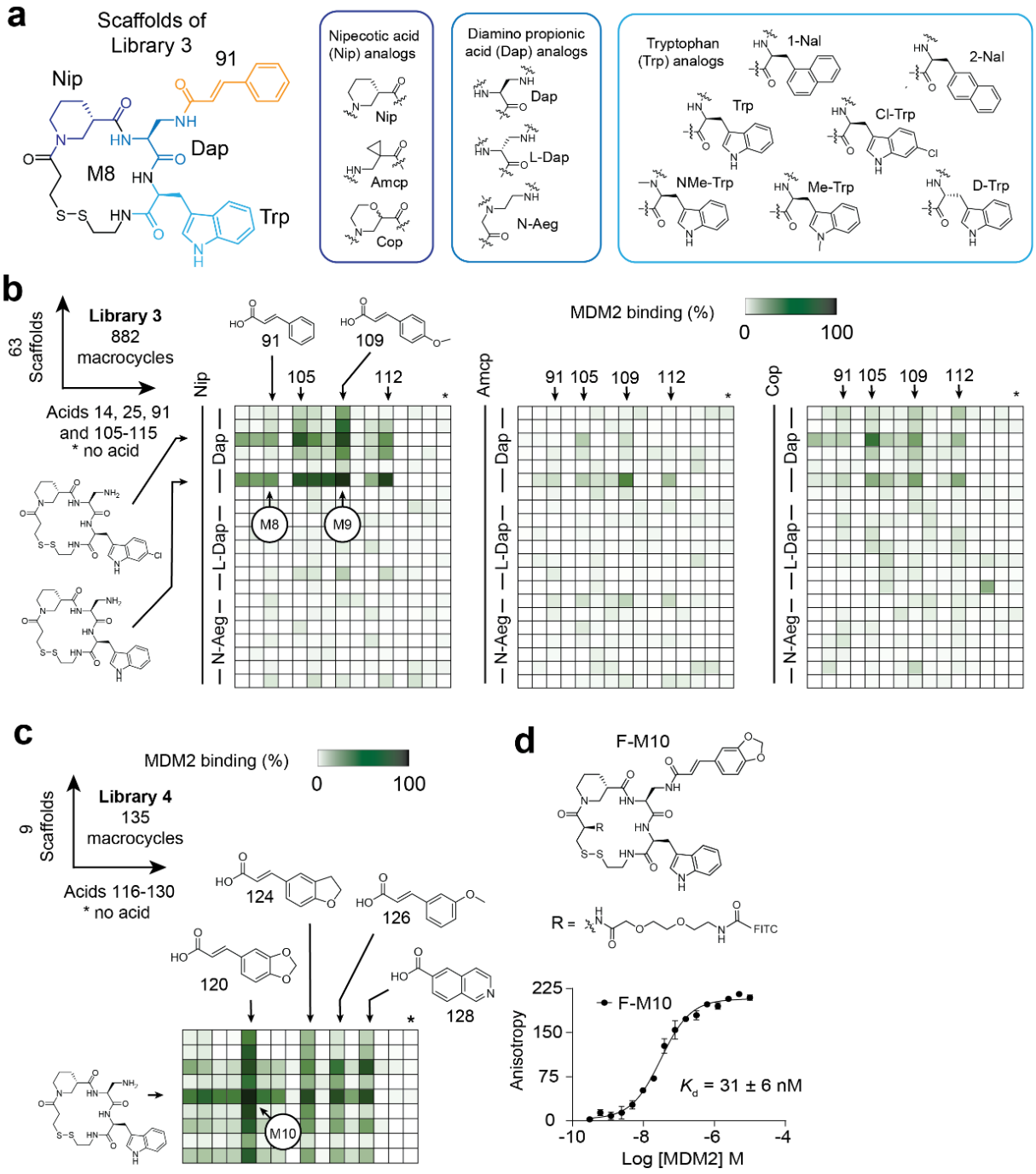


Figure 5 Affinity maturation of MDM2:p53 inhibitor. a Scaffolds of Library 3 are based on M8 wherein the amino acids shown in blue colors are diversified. Amino acid building blocks are shown in the three frames. **b** The 63 scaffolds were combinatorially acylated with the indicated 15 carboxylic acids at a 30 picomole scale. Binding to MDM2 was measured by displacement of fluorescent peptide probe by macrocycles at a concentration of 750 nM. **c** Screening of Library 4 based on the best nine scaffolds from the previous screens and 15 additional carboxylic acids. **d** Binding of fluorescein-labeled and HPLC-purified macrocycle M10 (F-M10) to MDM2 measured by FP.

4.5 Conclusions

We have addressed a major limitation in current macrocycle drug development by developing a method that can access larger libraries of macrocyclic compounds for HTS in microwell plates. While the largest macrocycle compound libraries currently offered by commercial providers for HTS contain only around 10,000 compounds, we generated a 20,000 macrocycle library with ease while establishing the new method of this study. From ongoing work in our laboratory, we know that even larger libraries comprising 100,000 compounds can be produced with the same method, and in the future, libraries exceeding a million structures could be generated. Steps that could additionally facilitate the synthesis of such large libraries are the change from 384- to 1,536-microwell plates (which have also been shown to work in our laboratory), and the production of the cyclic peptide scaffolds in smaller reactors. With the synthesizer used for this work, we were limited to 384 cyclic peptides per day and instrument (4×96-well plate reactors), but larger numbers may be produced if the reactor size can be reduced further.

Essential for accessing the large numbers of macrocyclic compounds were i) the strategy of combinatorially diversifying macrocyclic scaffolds by fragments via peripheral groups using an efficient chemistry, ii) the downscaling of the reactions to a picomole scale using contactless liquid transfer in nL volumes, and iii) the screening of the macrocyclic products as crude reactions, omitting a throughput-limiting purification step. While we had previously applied related "mix, react, and screen" approaches for generating and screening macrocyclic compounds libraries, they were all based on cyclization reactions which brought major limitations: we had combinatorially cyclized *m* short linear peptides by *n* different cyclization linkers to generate *m* × *n* macrocyclic compounds, but due to the challenging nature of cyclization reactions, the products typically had a limited purity and hits were occasionally based on side products. Another major limitation of cyclization reactions was that the number of *n* cyclization reagents that were commercially offered was limited to around twenty. By turning to a more efficient diversification reaction in this work—the acylation of a peripheral amino group—we reduced the reaction complexity and increased yield and purity of the macrocyclic compound products. Most importantly, the number of commercially available carboxylic acids exceeds 10,000, which opens the door for much larger macrocyclic diversities. Another important advantage of the diversification via a peripheral group vs. cyclization was that the reaction could be performed at high mM reagent concentrations, allowing downsizing of the reaction volume to 40 nL. In previous work with cyclization reactions, the highest concentrations allowed were 1 mM due to intermolecular reactions between peptide and

cyclization linkers. The small reagent volumes of 20 nL transferred herein accelerated also much the speed of liquid transfer: on the latest generation acoustic dispenser (ECHO 550) used in this work, liquids are transferred in pulses of 2.5 nL, and a transfer of 20 nL thus required only 8 pulses. For assembling 384 reactions in one plate, we required less than 5 minutes. While we have developed the picomole scale, late stage diversification strategy for macrocyclic compounds, it is tempting to speculate that it could be applied for the generation and screening of other molecule formats, as for example DNA aptamers, or even chemically diversified proteins.

The power of the new macrocycle synthesis approach and the good quality of the libraries are reflected by the high affinity of the identified binders, the 31 ± 6 nM thrombin inhibitor and the 44 ± 1 nM MDM2 binder inhibiting the p53:MDM2 protein-protein interaction. Importantly, for both of these two ligands, a clear structure-activity relationship was observed, indicating that they bind to their targets though defined molecular contacts. For the thrombin inhibitor, the crystal structure of the protein-ligand complex nicely confirmed the defined molecular contacts. For the binders of both targets, we found that the carboxylic acids had substantially increased the binding affinity of the cyclic peptide scaffolds, supporting the strategy of developing ligands that bind via both a macrocyclic scaffold and a peripheral substituent group. In fact, many natural product macrocycles bind in such a way, as for example the 4-hydroxy-3-methoxy-cyclohexyl substituent of rapamycin, or the cladinose and desosamine substituents of erythromycin[4]. A computational analysis of the physicochemical properties of the macrocycles in the libraries showed that a large portion of them are expected to be membrane permeable. Next steps will be the application of the technology to important and difficult targets, to demonstrate the ability of the macrocycles to enter cells, and to demonstrate *in vivo* the therapeutic effects of macrocycles identified with the new method.

4.6 Materials and methods

General considerations

Unless otherwise noted, all reagents were purchased from commercial sources and used with no further purification. Solvents were not anhydrous, nor were they dried prior to use. The following abbreviations are used: DIPEA (*N,N*-diisopropylethylamine), DABCO (1,4-diazabicyclo[2.2.2]octane), NMM (4-methylmorpholine), HBTU (*N,N,N′,N′*-tetramethyl-*O*-(1H-benzotriazol-1-yl)-uronium-hexafluorophosphate), HATU (*N,N,N′,N′*-tetramethyl-*O*-(7-azabenzotriazol-1-yl)uronium-hexafluorphosphate), HEPES (4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid)

Synthesis of model scaffold 1

The cyclic peptide model scaffold 1 was synthesized using the cyclative disuflide release strategy (CDR) described in the paper Habeshian, S. *et al.* 2021 (ref). The linear peptide precursor was synthesized on a 25 μmol scale in a 5 mL polypropylene synthesis column (MultiSyntech GmbH, V051PE076) using Rapp Polymere Polystyrene AM SH resin (200-400 mesh), 0.95 mmol/gram loading resin and following the procedure described in Habeshian, S. *et al.* 2021. The peptide was released as follows. For deprotection of the side chains, the resin was incubated with 2 mL of 38:1:1 TFA/TIS/ddH_2O (v/v/v) for one hour and then washed five times with approximately 4 mL of DCM. For cyclative peptide release, the resin was treated with 1 mL of DMSO containing 150 mM DIPEA (6 equiv.) overnight. The resin was removed by filtration. The crude mixture was purified by RP-HPLC using a Waters HPLC system (2489 UV detector, 2535 pump, Fraction Collector III), a 19 mm×250 mm Waters XTerra Prep MS C18 OBD column (125 Å pore, 10 μm particle), solvent systems A (H_2O, 0.1% v/v TFA) and B (MeCN, 0.1% v/v TFA), and a gradient of 0-25% solvent B over 30 min. The fraction containing the model scaffold was lyophilized and dissolved in DMSO to reach a concentration of 40 mM.

Acylation of model scaffold 1 by pipetting of reagents

The model scaffold was acylated at a 40 nmol scale in volumes of 4 μL as follows. The scaffold (20 μL of a 40 mM stock in DMSO) was supplemented with base (20 μL of 160 mM DIPEA dissolved in DMSO), and 2 μL of the mixture were transferred to wells of a PCR plate. The carboxylic acids were prepared as 160 mM stocks in DMSO containing 160 mM DIPEA. Equal volumes of HBTU (160 mM in DMSO) were added to each acid stock, and 2 μL of the resulting active esters (80 mM) were added to the same PCR plate. The reactions were allowed to

proceed for three hours at room temperature. After this time, 1 μL of the reaction was transferred into 99 μL of 100 mM Tris-HCl in water pH 7.5, incubated for six hours to allow quenching of activated acids with Tris, and the reactions analyzed by LC-MS.

Acylation of model scaffold 1 by acoustic reagent transfer

The model scaffold was acylated at an 800 pmol scale in volumes of 80 nL as follows. Scaffold 1 (20 μL of a 40 mM stock in DMSO) was supplemented with base (20 μL of 160 mM DIPEA, 20 μL of 160 mM DABCO, 20 μL of 160 mM HEPES sodium salt or 20 μL of 1 M NMM dissolved in DMSO) and 10 μL of the mixtures were transferred to an ECHO source plate (Labcyte Echo Qualified 384-well Low dead volume microplate). The concentrations in the source plate were 20 mM model scaffold and 80 mM DIPEA (4 equiv.), or 80 mM DABCO (4 equiv.), or 800 mM NMM (40 equiv.). The carboxylic acids were prepared as 160 mM stocks in DMSO containing either 160 mM DIPEA, 160 mM DABCO, or 1 M NMM. An equal volume of HBTU (160 mM in DMSO) was added to each acid stock and the active esters (80 mM) were added to the same source plate. The source plate was centrifuged at 950 g (2,000 rpm with a Thermo Heraeus Multifuge 3L-R centrifuge) for three minutes to remove potential bubbles. Using a Labcyte Echo 650 acoustic dispenser, 40 nL of the model scaffold 1 (800 pmol) were transferred to a Nunc 384 well low volume polystyrene plate, followed by 40 nL of the active esters (3.2 nmol, 4 equiv.). The transfers were performed in duplicate in order to have enough material for LC-MS analysis. The plates were sealed, and the reactions were allowed to proceed for six hours at room temperature. After this time, 8 μL of 100 mM Tris-HCl in water pH 7.5 were added to each one of the duplicate reactions, the duplicates pooled, incubated for three hours to allow quenching of activated acids with Tris, and the reactions analyzed by LC-MS.

Acylation of model scaffolds 2-5 by acoustic reagent transfer

The model scaffolds 2-5 purchased from Enamine were obtained as 1.1 to 1.2 mg powders. The scaffolds were dissolved in 62 to 88 μL DMSO to obtain 40 mM stocks. The scaffolds were acylated using DABCO as a base, as described for the model scaffold 1 above, with the following differences: 6 hour reaction time. Before LC-MS analysis, 720 nL of DMSO was dispensed to each well, then 7.2 μL of 100 mM Tris-HCl in water pH 7.5 was dispensed. Quenching took place overnight.

Design of scaffolds and amino acid sequences

The cyclic peptide scaffolds used for Library 1 were prepared by randomly choosing amino acid sequences. The number of different sequences that could theoretically be generated based on the chosen scaffold formats and amino acid building blocks was much larger than the number of scaffolds that were synthesized for Library 1 (384), as described in the following:

Theoretical number of scaffolds for Library 1:

- Scaffolds containing di-amino acids: 3,240
  # scaffold formats (6) × # di-aa (6) × # backbone aa (6) × # side chain aa (15)
- Scaffolds containing cysteine: 540
  # scaffold formats (6) × # backbone aa (6) × # side chain aa (15)

For randomly choosing 384 amino acid sequences, all building blocks were assigned an alphanumeric identifier, and every possible permutation was enumerated manually. The peptides were assigned numbers from 1 to 3,780. A random sequence generator (https://www.random.org/ sequences/) was then used to re-order the numbers, and the first 384 were chosen for synthesis.

Preparation of polystyrene-S-S-cysteamine resin for library synthesis

The following procedure was applied to prepare polystyrene-S-S-cysteamine resin for the synthesis of 4 × 96 peptides at a 5 μmol scale in four 96-well plates, as needed for the synthesis of the scaffolds for Library 1 (thrombin screen). Into each of four 20 mL plastic syringes was added 589 mg resin (Rapp Polymere Polystyrene AM SH resin, 200-400 mesh), 0.85 mmol/gram loading, corresponding to a 0.5 mmol scale. The resin was washed with 15 mL of DCM, then swelled in 15 mL of 3:7 MeOH/DCM v/v for 20 minutes. 2-(2-pyridinyldithio)-ethanamine hydrochloride (1.96 grams, 8.8 mmoles, 4.4 equiv.) was dissolved in 21.12 mL of MeOH, then 49.28 ml of DCM and 1.53 ml of DIPEA were added. 17.7 mL of this solution was pulled into each syringe, which were then shaken at room temperature for 3 hours. After this time, the 2-(2-pyridinyldithio)-ethanamine solutions were discarded, and the resins were washed with 2 × 20 mL 3:7 MeOH/DCM v/v, then 2 × 20 mL DMF. The resins were combined into a single syringe as a suspension in DMF, then was washed with 11.8 mL of 1.2 M DIPEA solution in DMF for 5 minutes to ensure that all amines were neutral. This solution was discarded, and the resin was washed with 2 × 20 mL DMF, 4 × 20 mL DCM, then kept under vacuum overnight to yield a free-flowing powder. For the synthesis of scaffolds for Library (MDM2 screen), the resin loading was 0.95 mmoles/gram, and thus 526 mg of resin (Rapp Polymere Polystyrene AM SH resin, 200-400 mesh) was added to each of two syringes.

Peptide library synthesis in 96-well plates

Automated solid-phase peptide synthesis was performed on an Intavis Multipep RSi synthesizer. For the thrombin library, to a 50 mL tube was added 565 mg of polystyrene-S-S-cysteamine resin (0.48 mmol cysteamine assuming that thiol groups were quantitatively modified with cysteamine) and 20 ml of DMF. For the MDM2 library, 505 mg of functionalized resin was added instead. The tube was shaken to ensure the resin was uniformly suspended, and 200 μL (5.88 mg resin, 5 μmoles) were transferred to each well of a 96-well solid phase synthesis plate (Orochem). The resin was washed with 6 × 150 μL DMF. Coupling was performed with 53 μL of amino acids (500 mM, 5.3 equiv.), 50 μL HATU (500 mM, 5 equiv.), 12.5 μL of *N*-methylmorpholine (4 M, 10 equiv.), and 5 μL *N*-methylpyrrolidone. All components were premixed for one minute, then added to the resin (one hour reaction, no shaking). The final volume of the coupling reaction was 120.5 μL and the final concentrations of reagents were 220 mM amino acid, 208 mM HATU and 415 *N*-methylmorpholine. Coupling was performed twice, then the resin was washed with 6 × 225 μL of DMF. Fmoc deprotection was performed using 120 μL of 1:5 piperidine/ DMF v/v for 5 minutes, and was performed twice. The resin was washed with 8 × 225 μL DMF. At the end of the peptide synthesis, the resin was washed with 2 × 200 μL of DCM.

Library side chain protecting group removal

For side chain protecting group removal, the bottom of the 96-well synthesis plate was sealed by pressing the plate onto a soft 6 mm thick ethylene-vinyl acetate pad, and the resin in each well was incubated with around 500 μl of 38:1:1 TFA/TIS/ddH2O (v/v/v) for one hour. The plates were covered with a polypropylene adhesive seal, then weighed down by placing a weight (1 kg) on top to ensure that no leakage occurred. After 1.5 hours, the synthesis plates were placed onto 2 mL deep-well plates, and the TFA mixture was allowed to drain. The wells were washed three times with approximately 500 μl of DCM (added with syringe), then allowed to air dry for 3 hours.

Library cyclative release of library peptides in 96-well plates

Plates were pressed into foam pads as described above to plug the openings, and 200 μL of 150 mM DABCO in DMSO (6 equiv.) were added to each well. The plates were sealed with an adhesive foil and weighed down (1 kg), and left overnight. The next day, the synthesis plates were placed onto 2 mL deep-well plates and centrifuged at around 200 g (1000 rpm with a

Thermo Heraeus Multifuge 3L-R centrifuge) for one minute to collect the cleaved macrocycles in DMSO.

Library peptide quantification by absorption

Absorbance measurements were performed with a Nanodrop 8000 spectrophotometer (Thermo Scientific) at a wavelength of 280 nm using a 10 mm path length. Cleaved peptides containing Trp and D-Trp were diluted 250 fold into water for the thrombin library, and 125 fold into water for the MDM2 library. The Beer-Lambert law was used to calculate the concentration of the peptides. Extinctions coefficient Trp ε_{280} = 5500 $M^{-1}cm^{-1}$ was used.

LC-MS analysis

Peptides were analyzed by LC-MS analysis with a UHPLC and single quadrupole MS system (Shimadzu LCMS-2020) using a C18 reversed phase column (Phenomenex Kinetex 2.1 mm × 50 mm C18 column, 100 Å pore, 2.6 μM particle) and a linear gradient of solvent B (acetonitrile, 0.05% formic acid) over solvent A (H_2O, 0.05% formic acid) at a flow rate of 1 mL/min. Mass analysis was performed in positive ion mode.

For the LC-MS analysis, the samples of the various experiments were prepared as follows. For analyzing the acylation proof of concept reactions, 160 nL of reaction mixtures were diluted into 16 μl of Tris-HCl buffer pH 7.5 to give a peptide concentration of 100 μM. For analyzing the scaffolds synthesized for Library 1, 1 μL of the DMSO/DABCO eluates were diluted into 80 μl of water to give cyclic peptide concentration of around 120 μM. For analyzing the scaffolds synthesized for Library 2, 1 μl of the DMSO/DABCO eluates were diluted into 128 μl of water to give cyclic peptide concentration of around 120 μM. For all analysis, 5 μl of the samples were injected, and typically using a 0 to 60% gradient of acetonitrile over 5 min.

Calculation of physicochemical properties of macrocycles

The physicochemical properties molecular weight, calculated water/n-octanol partition coefficient (cLogP), number of hydrogen bond donors (HBDs), number of hydrogen bond acceptors (HBAs), polar surface area (PSA), and number of rotatable bonds (NRotB) were calculated using DataWarrior (openmolecules.org) software. The structures of the scaffolds and the carboxylic acids were drawn in ChemDraw and copied as SMILES strings into two CSV files, one for the scaffolds and one for the acids. Both CSV files were opened in DataWarrior. The “enumerate combinatorial library” functionality was used to define the

desired amide bond forming reaction between the macrocycle scaffolds and the carboxylic acids. The macrocycle amines were defined such that existing amide bonds and amines in heterocyclic rings would not react. Following combinatorial enumeration, the desired properties were calculated from the structures.

Acylation of scaffolds to generate Library 1

Cyclic peptides scaffolds released, in the solvent used to release the peptides from resin (DMSO containing 150 mM DABCO) were transferred to a Labcyte Echo Qualified 384-well Low dead volume microplate (10 µl per well). The concentrations of the cyclic peptide scaffolds were around 8.1 mM in average. Carboxylic acids were dissolved to 160 mM in DMSO containing 160 mM DABCO. An equal volume of HBTU (160 mM in DMSO) was added to each acid stock. The active esters (80 mM) were added to another low dead volume source plate. The source plates were centrifuged at 850 g (2,000 rpm with a Thermo Heraeus Multifuge 3L-R centrifuge) for three minutes to remove potential bubbles. Using a Labcyte Echo 550 acoustic dispenser, 20 nL of the scaffolds (160 pmol) were transferred to Nunc 384 well low volume polystyrene plates, followed by 20 nL of the active esters (1.6 nmol, 10 equiv.). The plates were sealed, and the reaction was allowed to proceed for six hours at room temperature. After this time, 5 µL of Tris buffer (100 mM Tris-Cl, pH 7.5, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM $CaCl_2$, 0.1% w/v BSA, 0.01% v/v Triton-X100) was dispensed into each well using a BioTek MultiFlo microplate dispenser. The reaction were quenched overnight at room temperature.

Thrombin inhibition screen

Thrombin inhibition by the macrocycles of the Library 1 was assessed by measuring residual activity of thrombin in presence of the cyclic peptides at 11 µM average final concentration. The assays were performed in 384-well plates using Tris buffer at pH 7.4 (100 mM Tris-Cl, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM $CaCl_2$, 0.1% w/v BSA, 0.01% v/v Triton-X100, and 0.6% v/v DMSO) using thrombin at a final concentration of 2 nM and the fluorogenic substrate Z-Gly-Gly-Arg-AMC at a final concentration of 50 µM. Thrombin (5 µL, 6 nM) in the Tris buffer described above was added to each peptide using a BioTek MultiFlo microplate dispenser, and incubated for 10 minutes at room temperature. The fluorogenic substrate (5µL, 150 µM) in the same butter was added using the BioTek MultiFlo microplate dispenser, and the florescence intensity measured with a Tecan Infinite M200 Pro fluorescence plate reader (excitation at 360 nm, emission at 465 nm) at 25°C for a period of 30 minutes with a read every three minutes. The slope of each activity measurement curve was calculated by Excel. For the negative controls (20 wells containing no macrocycle but an equivalent volume of DMSO), an

average slope was calculated. The percent of thrombin inhibition was calculated by dividing the slopes and multiplying the results by 100.

Acylation of scaffolds to generate Library 2

Cyclic peptides scaffolds released, in the solvent used to release the peptides from resin (DMSO containing 150 mM DABCO) were transferred to a Labcyte Echo Qualified 384-well polypropylene microplate (40 μl per well). The concentrations of the cyclic peptide scaffolds were around 12.9 in average. Carboxylic acids were dissolved to 184 mM in a 184 mM DMSO solution of DABCO. An equal volume of HBTU (184 mM in DMSO) was added to each acid stock. The active esters (92 mM) were added to the same polypropylene source plate. The source plates were centrifuged at 950 g (2,000 rpm with a Thermo Heraeus Multifuge 3L-R centrifuge) for three minutes to remove potential bubbles. Using a Labcyte Echo 650 acoustic dispenser, 12.5 nL of macrocycles (161 pmol) were transferred to Nunc 384 well low volume polystyrene plates, followed by 17.5 nL of the active esters (1.61 nmol, 10 equiv.). The plates were sealed, and the reaction was allowed to proceed for six hours at room temperature. After this time, 5 μL of Tris buffer was dispensed into each well using a Gyger Certus Flex liquid dispenser, and the reactions were quenched overnight at room temperature.

MDM2 binding screen

MDM2 binding by cyclic peptides was assessed by measuring displacement of a fluorescent p53 peptide probe in presence of the cyclic peptides at 11 μM average final concentration. The assays were performed in 384-well plates using PBS buffer at pH 7.4 (100 mM Na_2HPO_4, 18 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl, 0.01% v/v Tween-20, and 3% v/v DMSO), MDM2 at a final concentration of 1.2 μM, and the fluorescent p53 peptide probe (FP53, sequence = 5(6)-FAM-SQETFSDLWKLLPEN) at a final concentration of 25 nM. Premixed MDM2 and FP53 (10 μL, 1.8 mM MDM2, 37.5 nM FP53) in the PBS buffer described above was added to each peptide using a Gyger Certus Flex liquid dispenser, and incubated for 30 minutes in the dark at room temperature. A single fluorescence anisotropy reading was taken with a Tecan Infinite F200 Pro fluorescence plate reader (excitation at 485 nm, emission at 535 nm) at 25°C.The percentage of probe displacement was calculated using to the following formula,

$$\%\ probe\ displacement = \frac{N - X}{N - P}\ x\ 100$$

where N is the average anisotropy of the negative controls (no inhibition), X is the value obtained for each well, and P is the average anisotropy of only the probe.

Identification of active species in reactions from hits

The macrocycles identified as hits in the thrombin screen were resynthesized at a 40 nmol scale by reacting 5 μL of 8 mM cyclic peptide scaffolds in DMSO containing 150 mM DABCO with 5 μL of 80 mM carboxylic acid, 80 mM HBTU and 80 mM DABCO for 5 hours at room temperature. Remaining activated ester was quenched by addition of 1.25 ml of Tris buffer (100 mM Tris-Cl, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM $CaCl_2$) and incubation over night. The next day, 240 μL of MeCN and 1 mL of water were added, and the reactions were run over a C18 column (7.8 mm × 300 mm Waters NovaPak C-18 column, 60 Å pore, 6 μm particle) on a Thermo Dionex HPLC using solvent A (H_2O, 0.1% v/v TFA) and a 10-80% gradient of solvent B (MeCN, 0.1% v/v TFA) over 20 minutes, and fractions were collected every minute. Fractions were lyophilized, dissolved in 120 μL of 2% DMSO in water. The activities of products in the fractions were measured using the same assays as described above, but in 96-well plates. 50 μL of each fraction was transferred to a Greiner flat bottom transparent 96-well assay plate, followed by 50 μL of thrombin (6 nM in buffer). After 10 minutes of incubation, 50 μL of fluorogenic thrombin substrate (Z-Gly-Gly-Arg-AMC, 150 μM in buffer, 1% DMSO) was added and the plates were read and the data processed as described above. Compounds in active fractions were identified by mass spectrometry.

For hits from the MDM2 screen, reactions were performed in the same way but at a 50 nmol scale, and purified with the same method but a 10-80% gradient of solvent B and over 22 minutes. Fractions were lyophilized and dissolved in 40 μL of DMSO, 160 μL of water was added, and 5 μL of each fraction was transferred to a 384-well plate, and 15 μL of premixed MDM2/FP53 peptide were added (final concentrations: 1.2 μM MDM2, 50 nM FP53, 5% DMSO).

Crystallization of thrombin with M1

Human α-thrombin was purchased from Haematologic Technologies (Catalogue number: HCT-0020). Protein-stabilizing agent was removed using a PD-10 desalting column (GE Healthcare) equilibrated with 20 mM Tris-HCl, 200 mM NaCl, pH 8.0 and the same buffer as solvent. Buffer exchanged human α-thrombin was incubated with the macrocycle M1 at a molar ratio of 1:3 and subsequently concentrated to 7.5 mg/mL by using a 3,000 MWCO Vivaspin ultrafiltration device (Sartorius-Stedim Biotech GmbH). Further M1 macrocycle was added during the concentration to ensure that a 3-fold molar excess is preserved. Crystallization trials of the complex were carried out at 293 K in a 96-well 2-drop MRC plate (Hampton Research, CA, USA) using the sitting-drop vapor-diffusion method and the

Morpheus and LMB crystallization screens (Molecular Dimensions Ltd, Suffolk, UK). Droplets of 600 nL volume (with a 1:1 protein:precipitant ratio) were set up using an Oryx 8 crystallization robot (Douglas Instruments Ltd, Berkshire, UK) and equilibrated against 80 μL reservoir solution. Best crystals were obtained by applying micro-seeding to fresh drops that had been allowed to equilibrate for 2–3 days using the following mixture as precipitant agent: 20 mM sodium formate, 20 mM ammonium acetate, 20 mM sodium citrate tribasic dihydrate, 20 mM potassium sodium tartrate tetrahydrate, 20 mM sodium oxamate, 100 mM MOPS/sodium HEPES pH 7.5, 12.5% w/v PEG 1000, 12.5% w/v PEG 3350, 12.5% v/v MPD.

Crystallization, data collection and structure determination

For X-ray data collection, crystals were mounted on LithoLoops (Molecular Dimensions Ltd, Suffolk, UK) and flash-cooled in liquid nitrogen. X-ray diffraction data of human α-thrombin in complex with M1 were collected at the i04 beamline at Diamond Light Source Ltd (DLS, Oxfordshire, UK). The best crystals diffracted to 2.27 Å maximum resolution. Crystals belong to the $P2_1$ space group, with unit cell dimensions a = 56.25 Å, b = 100.57 Å, c = 108.90 Å and α = 90°, β = 90.11°, γ = 90°. The asymmetric unit contains four molecules, corresponding to a Matthews coefficient of 2.78 $Å^3$/Da and a solvent content of about 56% of the crystal volume. Frames were indexed and integrated with software XIA2, merged and scaled with AIMLESS (CCP4i2 crystallographic package). The structure was solved by molecular replacement with software PHASER using as a template the model 6GWE. Refinement was carried on using REFMAC and PHENIX. Since the first cycles of refinement, a wide electron density corresponding to the bound ligand was clearly visible in the electron density map. Building of the macrocycle was performed by Molview, restraint file generated and optimized by Phenix eLBOW. The macrocycle was fitted manually by graphic software COOT. The final model contains 9275 protein atoms, 164 ligand atoms and 430 water molecules. The final crystallographic R factor is 0.184 (R_{free} 0.244). Geometrical parameters of the model are as expected or better for this resolution. The solvent excluded volume and the corresponding buried surface were calculated using PISA software and a spherical probe of 1.5 Å radius. Intra-molecular and inter-molecular hydrogen bond interactions were analyzed by PROFUNC, LIGPLOT, and PYMOL software.

Acylation of scaffolds to generate Libraries 3 and 4

The cyclic peptide scaffolds required for Libraries 3 and 4 were synthesized as described for those used in Library 2. Due to the presence of many N-methylated amino acids which are more difficult to couple, 200 mM HOAt was applied together with HATU. The scaffolds were

diluted to 2 mM in DMSO, and 15 nL were transferred using acoustic dispensing, followed by 15 nL DMSO containing carboxylic acids (40 mM, 20 equiv.), HBTU (40 mM) and DABCO (40 mM). After six hour reaction at room temperature, 370 nL of DMSO was added to each well, followed by 5 μL of 100 mM Tris-Cl pH 7.4 containing 0.01% v/v Tween 20, for quenching overnight. For the MDM2 binding screen, F-M6 was used as a fluorescent probe because of its higher affinity for MDM2, allowing to use the target protein at a lower concentration (750 nM, leading to around 55% bound probe). A volume 35 μL of PBS buffer at pH 7.4 (100 mM Na_2HPO_4, 18 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl, 0.01% v/v Tween-20 containing 28.6 nM F-M6 and 823 nM MDM2 were added to each well and the displacement of reporter probe determined as described above. The concentrations of the macrocycles were 720 nM.

Synthesis of macrocycles at mg scale

Automated solid-phase peptide synthesis was performed on an Intavis Multipep RSi synthesizer. To a 5 ml syringe was added 25 μmoles of polystyrene-S-S-cysteamine resin. The resin was washed with 6 × 150 μL DMF. Coupling was performed with 210 μL of amino acids (500 mM, 4.2 equiv.), 200 μL HATU (500 mM, 4 equiv.), 50 μL of *N*-methylmorpholine (4 M, 8 equiv.), and 5 μL *N*-methylpyrrolidone. All components were premixed for one minute, then added to the resin (one hour reaction, no shaking). The final volume of the coupling reaction was 465 μL and the final concentrations of reagents were 226 mM amino acid, 215 mM HATU and 430 *N*-methylmorpholine. Coupling was performed twice, then the resin was washed with 2 × 600 μL of DMF. Fmoc deprotection was performed using 450 μL of 1:5 piperidine/ DMF v/v for 5 minutes, and was performed twice. The resin was washed with 7 × 600 μL DMF. At the end of the peptide synthesis, the resin was washed with 2 × 600 μL of DCM.

After SPPS, the resin was incubated with 2 mL of 38:1:1 TFA/TIS/ddH_2O (v/v/v) for one hour. The TFA solution was discarded, and the resin was washed five times with 4 mL DCM. After air drying for three hours, 1 mL of 150 mM DIPEA in DMSO was pulled in, and the syringes were shaken overnight at room temperature. The following day, the DMSO solutions were pushed into 50 mL conical tubes.

Carboxylic acids were typically coupled according to the previously described procedure. After three hours at room temperature, 8 mL of water was added and the tubes were frozen, then lyophilized for two days to remove DMSO. The contents of the tubes were dissolved in 3 mL of MeCN, followed by addition of 7 mL of water.

The crude mixtures were purified by RP-HPLC using a Waters HPLC system (2489 UV detector, 2535 pump, Fraction Collector III), a 19 mm×250 mm Waters XTerra Prep MS C18

OBD column (125 Å pore, 10 μm particle), solvent systems A (H_2O, 0.1% v/v TFA) and B (MeCN, 0.1% v/v TFA), and typically a gradient of 30-70% solvent B over 30 min.

Determination K_is of thrombin inhibitors

Purified thrombin inhibitors (10 mM in DMSO) were diluted to 80 μM in 125 μL of Tris buffer (100 mM Tris-Cl, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM $CaCl_2$) containing 0.1% w/v BSA, 0.01% v/v Triton-X100 and 0.2% DMSO. The macrocycles were diluted two-fold in Tris buffer containing 0.1% w/v BSA, 0.01% v/v Triton-X100 and 1% DMSO in buffer. The thrombin activity was measured in 96-well plates and the residual activity calculated as described above in the assay used to measure activities of HPLC-separated fractions of screening hits. The residual activity was plotted against the Log of the corresponding macrocycle concentrations, and sigmoidal curves were fitted using the following four-parameter equation in GraphPad Prism 6:

$$Y = Bottom + \frac{Top - Bottom}{(1 + 10^{(LogIC50 - X) * HillSlope})}$$

K_i values were determined from the IC_{50} values using the Cheng-Prusoff equation (K_m = 168 μM for thrombin and the applied substrate):

$$K_i = \frac{IC_{50}}{1 + \frac{[S]}{K_m}}$$

Determination of IC_{50} of MDM2-binding macrocycles

The concentrations at which MDM2 macrocycles displaced the reporter peptide for 50% of the protein (IC_{50}) were determined with the above described fluorescence polarization competition assay. Volumes of 5 μL of purified macrocycles (20 mM in DMSO) were serial diluted two-fold in 100% DMSO in a low dead-volume ECHO source plate. Using acoustic droplet transfer, 150 nL of each dilution was transferred to a Nunc 384 well low volume polystyrene plate. A volume of 15 μL of MDM2/FP53 probe premix (1.2 μM MDM2, 25 nM FP53 probe) in PBS buffer pH 7.4 (100 mM Na_2HPO_4, 18 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl, 0.01% v/v Tween-20) containing 1% v/v DMSO was added to each well, and incubated for 30 minutes in the dark. Fluorescence anisotropy was measured as described above. The percentage of bound inhibitor was calculated using the following equation

$$\%\ bound\ inhibitor = \frac{N - X}{N - P}\ x\ 100$$

where N is the average anisotropy of the DMSO controls, X is the anisotropy value obtained for each well, and P is the average anisotropy of the unbound probe.

The IC_{50}s were determined by plotting the percent of bound inhibitor against the logarithm of the corresponding macrocycle concentration, and the curves were fitted in GraphPad Prism 6 as described above.

Synthesis of fluorescein-labeled macrocycles

Fluorescein-labeled macrocycles were synthesized essentially as described in the mg-scale macrocycle synthesis procedure. For 5(6)-FAM, manual coupling was performed using 4eq of acid (180 mM, 556 μL), 4eq HATU (500 mM, 200 μL), 10 eq NMM (4M, 62.5 μL), all in DMF. The coupling was performed 1 x 2 hours, then washed as previously described.

Determining K_ds of fluorescein-labeled MDM2 binders

Fluorescein labeled macrocycle stocks (20 mM in DMSO) were diluted to a concentration of 10 μM by adding 0.5 μL into 999.5 μL of PBS. These dilutions were further diluted to a concentration of 100 nM by transferring 10 μL to 990 μL PBS, and 7.5 μL were transferred to wells of a Nunc 384 well low volume polystyrene plate. Volumes of 7.5 μL of 2-fold dilutions of MDM2 in PBS were pipetted to the wells. The final concentrations of fluorescent macrocycles were 50 nM. After incubation of the plate for 30 minutes in the dark at room temperature, the fluorescence anisotropy was measured with a Tecan Infinite F200 Pro fluorescence plate reader (excitation at 485 nm, emission at 535 nm) at 25°C. Anisotropy was plotted against the logarithm of the corresponding MDM2 concentrations and sigmoidal curves were fitted as described above.

Synthesis of thrombin inhibitors containing thiother bonds

Linear peptides containing the three amino acids and the C-terminal cysteamine were synthesized by automated SPPS as described above for the synthesis of mg scale cyclic peptide, but on 50 μmol scale and Novabiochem Cysteamine 4-methoxytrityl resin (1% DVB, 200 - 400 mesh, 0.92 mmol/gram). To this peptide still on resin, 4-bromobutyric acid (500 μL, 500 mM, 10 equiv.) was coupled manually using *N,N'*-diisopropylcarbodiiimide (DIC, 500 μL, 500 mM, 10 equiv.) as an activating reagent in DMF. The acid and coupling reagent were premixed for one minute, then added to the resin (1 hour reaction with shaking). The final volume of the coupling reaction was 1 mL and the final concentrations of reagents were 250

mM amino acid, 250 mM DIC. Coupling was performed twice, then the resin was washed with 4 × 4 mL of DMF, then 2 × 4 mL DCM.

Side chain protecting group removal and cleavage was performed by incubating the resin with 2 mL of 38:1:1 TFA/TIS/ddH_2O (v/v/v) for one hour with shaking. After this time, 50 mL of cold diethyl ether was added to the solution to precipitate the peptide. The mixture was stored at -20°C for 30 minutes, then centrifuged for 30 minutes at 3,800 g (4,000 rpm on a Thermo Heraeus Multifuge 3L-R centrifuge) at 4°C. The ether was decanted, and the peptide pellet allowed to air dry for 15 minutes.

The peptide was dissolved in 50 mL of freshly de-gassed 1:4 water/acetonitrile and 200 μL (1.15 mmol, 23 equiv.) of neat DIPEA was added. The cyclization reaction was allowed to proceed at room temperature for 90 minutes, then frozen and lyophilized.

Carboxylic acid **14** was coupled as before. After three hours at room temperature, 8 mL of water was added and the tubes were frozen and lyophilized for two days to remove DMSO. The contents of the tubes were dissolved in 3 mL of MeCN followed by addition of 7 mL of water. The crude mixtures were purified by RP-HPLC as described above.

4.7 Supplementary information

Supplementary Results

Overall structure of human α-thrombin in complex with M1

Human α-thrombin consists of two polypeptide chains of 36 (L-chain) and 259 amino acid residues (H-chain) covalently linked via a disulfide bridge (Cys122 of H-chain with Cys1 of L-chain). The L-chain of human α-thrombin can be traced unambiguously from Glu1C to Ile14K. The first five amino terminal residues (Thr1H to Gly1D) and the carboxyl-terminal residues Asp14L, Gly14M and Arg15 are undefined and not visible in the Fourier map. The electron density of the H-chain is clearly visible for all residues with the exception of few amino acids part of the surface flexible autolysis loop (Trp148 to Val149C). The carboxyl-terminal residue Glu247 lacks adequate electron density. Four identical copies of H- and L-chains of human α-thrombin are present in the asymmetric unit. Only minor differences occur at the level of flexible and less defined loops or in the orientation of exposed peripheral side chains. The overall structure of human α-thrombin in bound to the macrocycle does not show any striking rearrangements of the main backbone if compared to other human α-thrombin structures, either in the apo form or in complex with inhibitors.

Overall structure of macrocycle M1

The electron density of the macrocycle M1 is well defined allowing an unambiguous assignment of group orientations for all the four protein complexes present in the asymmetric unit. No classical secondary structure elements and no non-covalent intramolecular interactions are found in the macrocycle. The molecule appears to adopt a chair-like conformation that fits well the shape of the catalytic pocket.

Interactions between human α-thrombin and M1

The M1 macrocycle fits well into the cleft formed by the active site and the surrounding substrate pockets covering a protein surface of 400.5 $Å^2$. The macrocycles' conformations and interactions are equivalent in the four active sites of the four-thrombin molecules present in the asymmetric unit. A large portion of interactions of M1 with human α-thrombin are mediated by the 5-chlorothiophene-2-carboxamide functional group that accommodates in the primary specificity S1 pocket. This group is trapped in the pocket by a hydrogen bond with the main chain of Gly219 (M1 N7 with Gly219 O) and a molecule of H_2O that bridges the main chain oxygen O9 of M1 with the main chain nitrogen of Gly193 (Glu193 N, Supplementary Table 3)

and main chain nitrogen of Ser195 (Ser195 N, Supplementary Table 3). 5-chlorothiophene-2-carboxamide is further involved in a network of polar contacts with the main chain of the nearby Cys191 (M1 O9 with Cys191 O), Glu192 (M1 O9 with Gly193 N), Gly216 (M1 N7 with Gly216 O and M1 S15 with Gly216 N), Trp215 (M1 S15 with Trp215 N) and the side chain of Cys220 (M1 N7 with Cys220 S; Supplementary Table 3). Interestingly, the chlorine atom 5-chlorothiophene-2-carboxamide functional group points toward the bottom of the S1 pocket where it forms a halogen-aromatic π interaction (4.0 Å) with the aromatic ring of Tyr228 (Fig. 3f). The main chain nitrogen N4 and oxygen O17 of M1 form hydrogen bonds with the main chain oxygen of Gly216 (Gly216 O) and nitrogen of Gly216 (Gly216 N), respectively (Supplementary Table 3). Additionally, the main chain nitrogen N4 and oxygen O17 of M1 can form polar contacts with the main chain nitrogen of Gly219 (Gly219 N) and oxygen of Gly216 (Gly216 O), respectively (Supplementary Table 3). Similarly, the main chain nitrogen N18 of M1 can form two polar contacts with the side chain carboxylic group of Glu192 (Glu2192 OE1 and OE2; Supplementary Table 3). Finally, a molecule of H_2O bridges the main chain nitrogen N27 of M1 with the main chain oxygen of Glu97A (Glu97A O, Supplementary Table 3). Importantly, the binding of M1 to human α-thrombin is mediated by multiple hydrophobic interactions with main and side chains of adjacent enzyme residues (Supplementary Table 3). The macrocycle backbone (C20-C24), including the disulfide bridge S21-S22, lays towards the hydrophobic cage shaped by the side chains of residues His57, Tyr60A, Trp60D (proximal S2 pocket) and Leu99 (distal S3 pocket). The valine side chain (C28-C31) bends the other side of the ring toward the hydrophobic pocket formed by Ile174 and Trp215. Finally, the C35 – C40 phenyl ring run on top of a thrombin loop (Gly216 – Cys220).

Supplementary Tables

Supplementary Table 1. Macrocycle libraries. Size, diversity and physicochemical properties of the macrocycle libraries. *) Backbones: different macrocyclic skeletons, ignoring peripheral groups; backbones having the same atoms but different conformational constrains imposed by peripheral groups (e.g. N-methylation, cyclic side chain, etc.) or bonds are considered as different backbones.

	Library 1	**Library 2**
Library size		
Number macrocycles	4,608	19,968
Structural diversity		
Number scaffolds	384	192
Different backbones*	144	96
Chemical diversity		
Total number of building blocks	41	116
Different amino acids	27	10
Different carboxylic acids	12	104
Flanking thiol groups	2	2
Physicochemical properties		
Molecular weight (average)	624	684
cLogP (average)	0.8	2.1
Polar surface area (average)	228 $Å^2$	211 $Å^2$
Number H-bond donors (average)	5.23	4.31

Supplementary Table 2. Statistics on X-ray structure data collection and refinement. A single crystal was used to collect all diffraction data. Highest-resolution shell statistics are shown within brackets.

Data collection *	
Wavelength (Å)	0.9795
Space group	$P2_1$
Cell parameters	
a, *b*, *c* (Å); α, β, γ (°)	56.25, 100.57, 108.90; 90, 90.11, 90
Resolution (Å)	73.88 – 2.27 (2.35 – 2.27)
Observations	104136 (8954)
Unique	52658 (4711)
Multiplicity	2.0 (1.9)
R_{merge}	0.0411 (0.2527)
R_{pim}	0.021 (0.210)
$\langle I / \sigma(I) \rangle$	16.69 (3.35)
$CC_{1/2}$	0.998 (0.794)
Completeness (%)	93.85 (84.19)
Wilson B-factor	24.25
Refinement	
No. reflections (Used for R_{free} calculation)	52644 (4712)
R_{work} / R_{free}	0.189 / 0.243
Number non-hydrogen atoms	9791
protein (chains A, B, C, D, E, F, H, L)	9229
ligands (M1)	164
solvent	398
Geometry	
RMSD values	
bond lengths (Å)	0.013
bond angles (°)	1.90
Ramachandran plot (%)	
most favored	95.00
additionally allowed	4.91
outliers	0.09
Average B-factor	28.61

Supplementary Table 3. Atoms of the macrocycle of M1 forming hydrophobic interactions with atoms and residues of human α-thrombin (chymotrypsin numbering). Interactions have distances shorter than 4.0 Å and were defined using the software LIGPLOT+ by the web server PROFUNC.

Thrombin atom / residue	M1 atom	Distance (Å)	Interaction
O / Cys191	O9	3.58	PI
N / Glu192	O9	3.61	PI
OE1 / Glu192	N18	3.66	PI
OE2 / Glu192	N18	3.35	PI
N / Trp215	S15	3.73	PI
N / Gly216	S15	3.59	PI
N / Gly216	O17	3.25	HB
O / Gly216	N4	2.92	HB
O / Gly216	N7	3.84	PI
O / Gly216	O17	3.35	PI
N / Gly219	N4	3.89	PI
O / Gly219	N7	3.17	HB
S / Cys220	N7	3.82	PI
N / Gly193 (H_2O)*	O9	2.85 (2.80)	HB
N / Ser195 (H_2O)*	O9	2.85 (3.17)	HB
O / Glu97A (H_2O)*	N27	3.20 (2.80)	HB

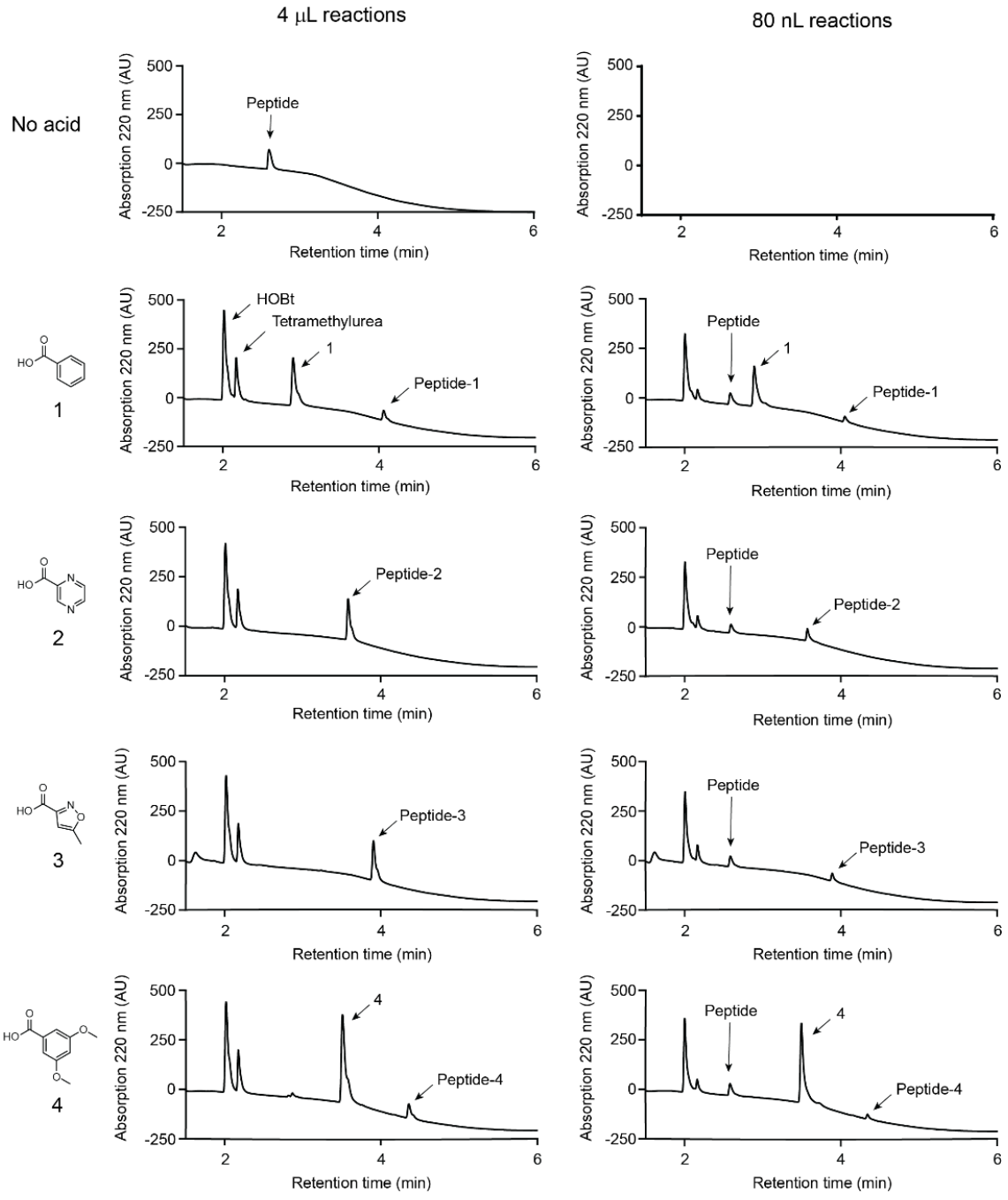
Thrombin atom / residue	M1 atom	Distance (Å)
CD2 / His57	S21	3.84
CE1 / Tyr60A	C24	3.74
CE2 / Tyr60A	C23	3.62
CE2 / Tyr60A	C24	3.86
CZ / Tyr60A	C23	3.65
CZ / Tyr60A	C24	3.34
OH / Tyr60A	C23	3.74

OH / Tyr60A	C24	3.05
CZ2 / Trp60D	C23	3.39
CZ3 / Trp60D	S21	3.90
CH2 / Trp60D	S21	3.84
CH2 / Trp60D	C23	3.35
C / Asn98	C30	3.85
O / Asn98	C30	3.62
CD1 / Leu99	C25	3.85
CG1 / Ile174	C31	3.69
CD1 / Ile174	C31	3.65
CG / Asp189	C12	3.73
OD1 / Asp189	C12	3.28
OD2 / Asp189	C12	3.58
C / Ala190	C11	3.88
O / Ala190	C11	3.51
O / Ala190	C12	3.76
CB / Ala190	C13	3.89
C / Cys191	C8	3.85
C / Cys191	O9	3.48
CA / Glu192	O9	3.73
CD / Glu192	N18	3.82
OE1 / Glu192	C5	3.50
OE1 / Glu192	C6	3.70
OG / Ser195	C19	3.28
CG1 / Val213	CL14	3.67
CG1 / Val213	S15	3.55
O / Ser214	C19	3.31
O / Ser214	C20	3.55
CA / Trp215	S15	3.54

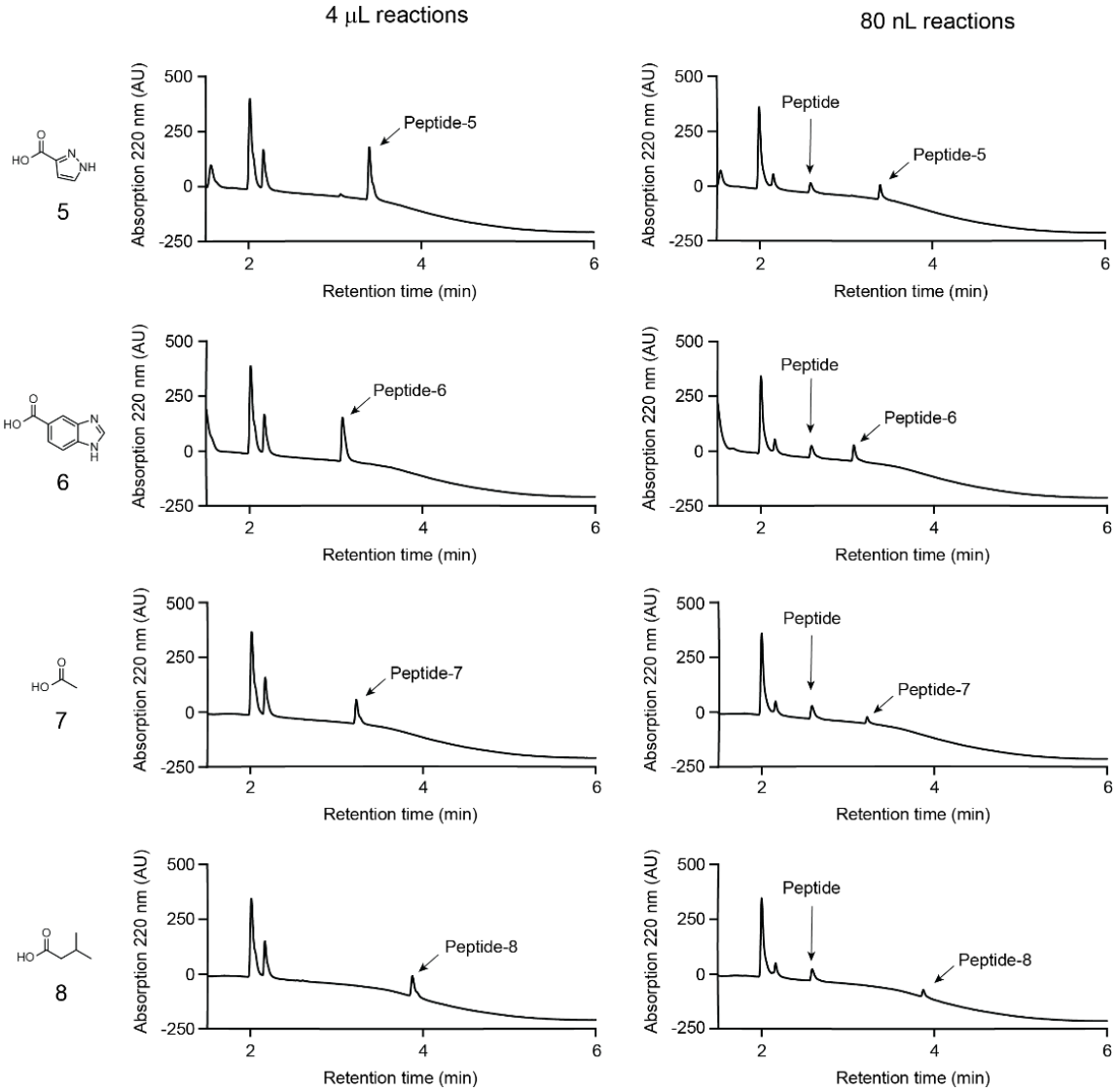
CA / Trp215	O17	3.43
C / Trp215	C13	3.49
C / Trp215	S15	3.47
C / Trp215	O17	3.86
O / Trp215	C13	3.48
O / Trp215	CL14	3.60
CB / Trp215	O17	3.45
CG / Trp215	C30	3.56
CD1 / Trp215	C30	3.45
CD2 / Trp215	C30	3.27
NE1 / Trp215	C30	3.20
CE2 / Trp215	C30	3.05
CE3 / Trp215	N34	3.33
CZ2 / Trp215	C30	3.59
CZ3 / Trp215	N34	3.53
N / Gly216	C10	3.46
N / Gly216	C11	3.58
N / Gly216	C12	3.73
N / Gly216	C13	3.66
CA / Gly216	C11	3.76
CA / Gly216	C12	3.81
O / Gly216	C1	3.41
O / Gly216	C2	3.58
O / Gly216	C36	3.83
O / Gly216	C40	3.19
CA / Glu217	C36	3.79
CA / Glu217	C37	3.80
N / Gly219	C1	3.46
N / Gly219	C2	3.79

N / Gly219	C38	3.86
N / Gly219	C39	3.33
O / Gly219	C11	3.14
CA / Gly226	C12	3.86
CA / Gly226	CL14	3.47
C / Gly226	CL14	3.74
N / Phe227	CL14	3.50
O / Phe227	CL14	3.42
CE1 / Tyr228	CL14	3.73
CZ / Tyr228	CL14	3.60
OH / Tyr228	CL14	3.74

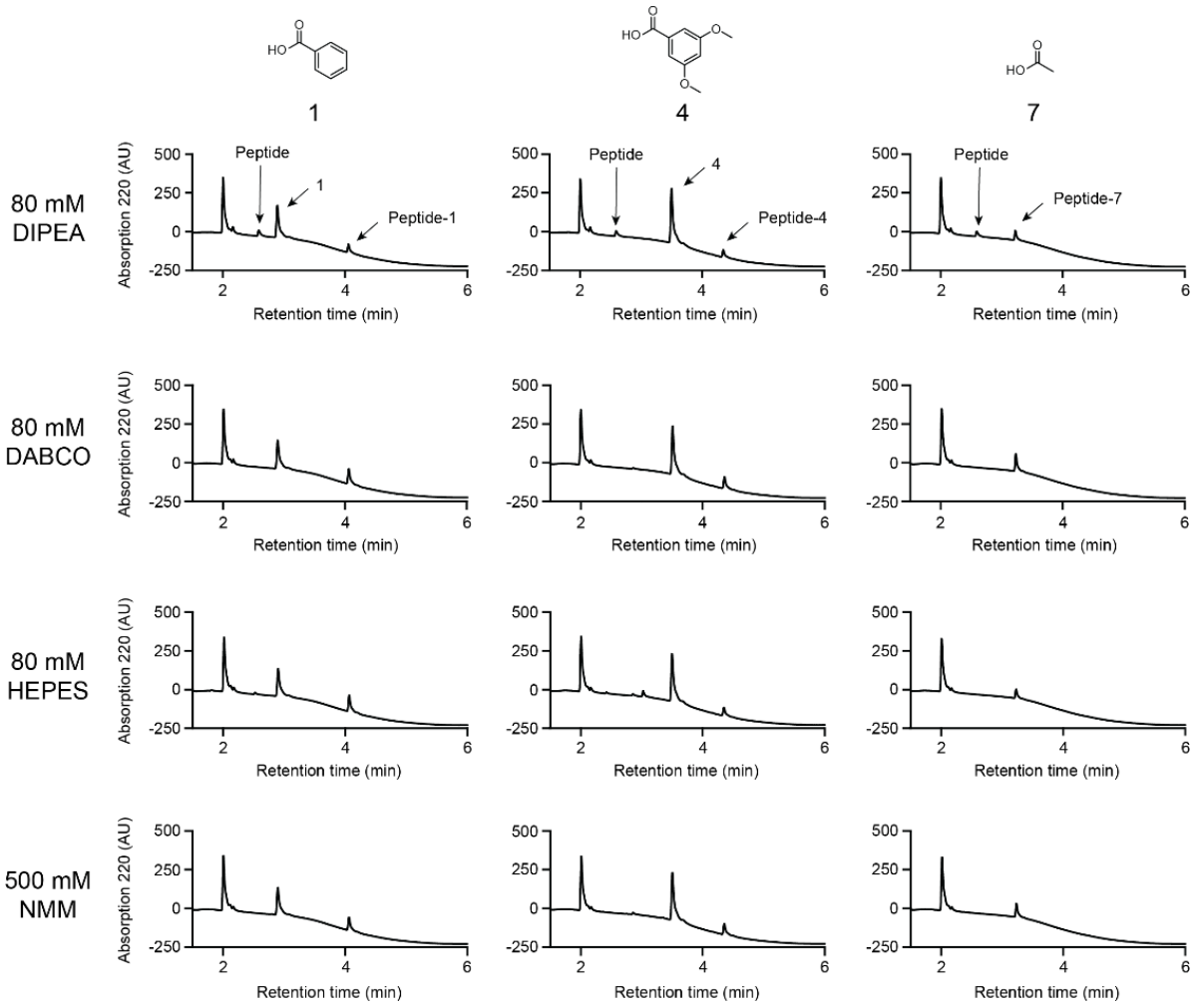
Supplementary Figures



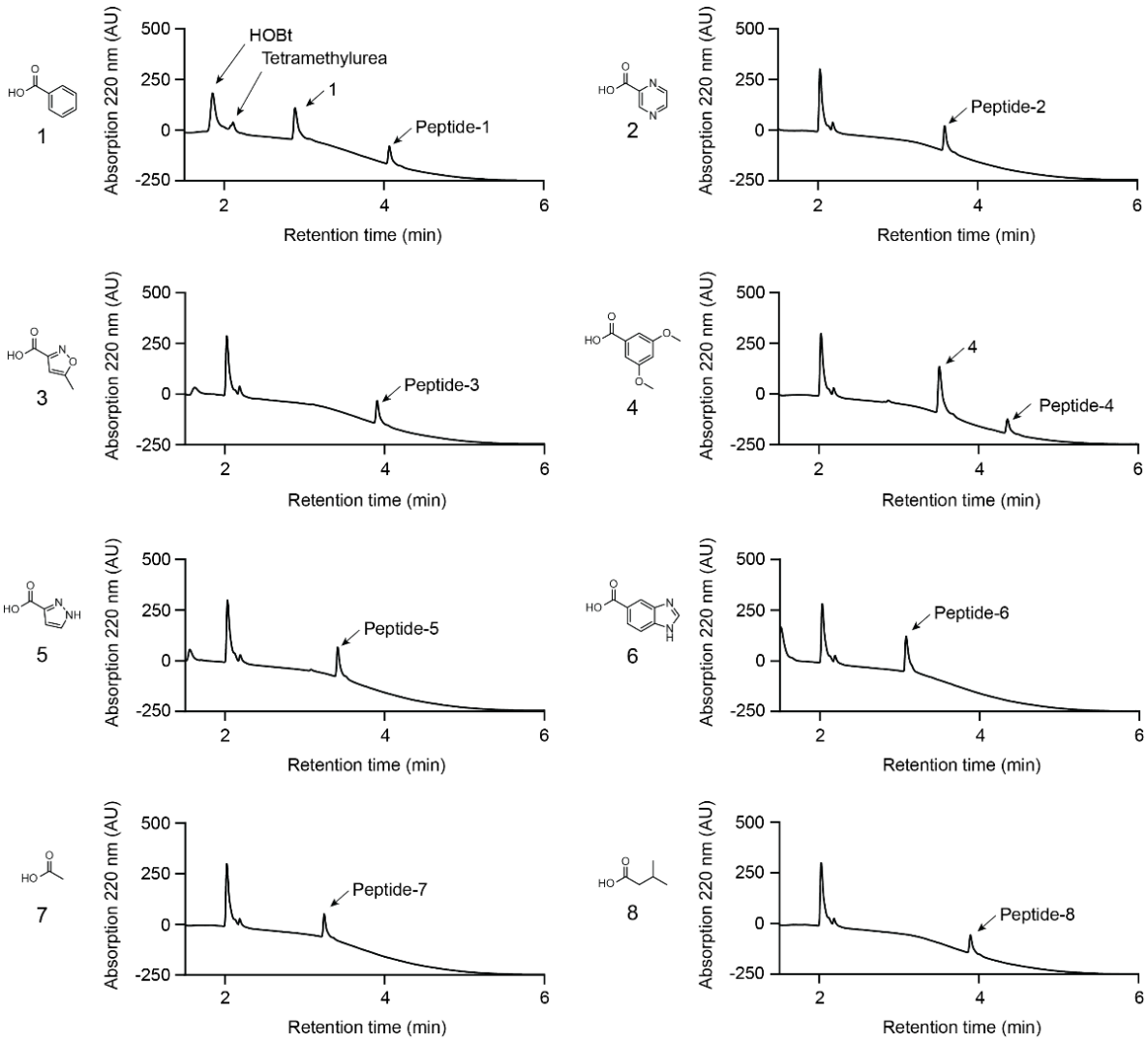
Supplementary Figure 1. Comparison of acylation reactions by pipetting (4 μL volume) and acoustic transfer (80 nL volume) using DIPEA as base. Samples of the reactions were analyzed by LC-MS using a RP column and a 0-60% MeCN/H_2O gradient over 5 minutes. Smaller reaction volumes resulted in significantly higher conversions for all acids when DIPEA was used as the base.



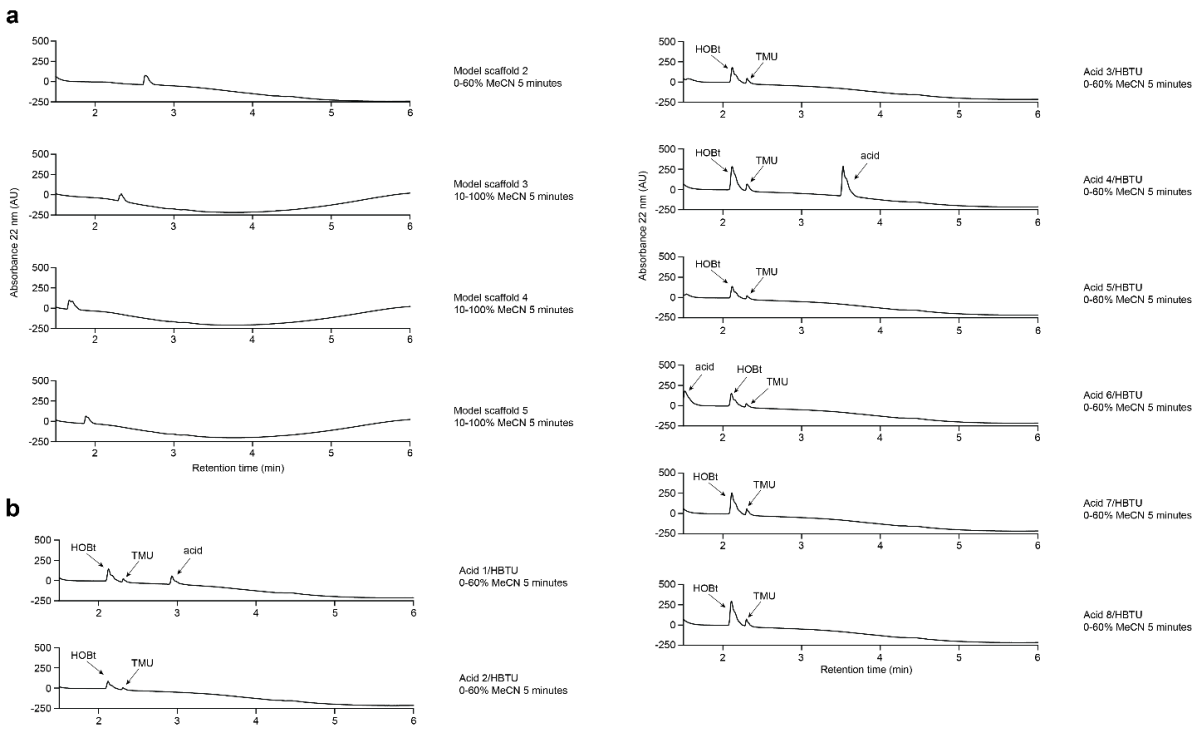
Supplementary Figure 1. Continued



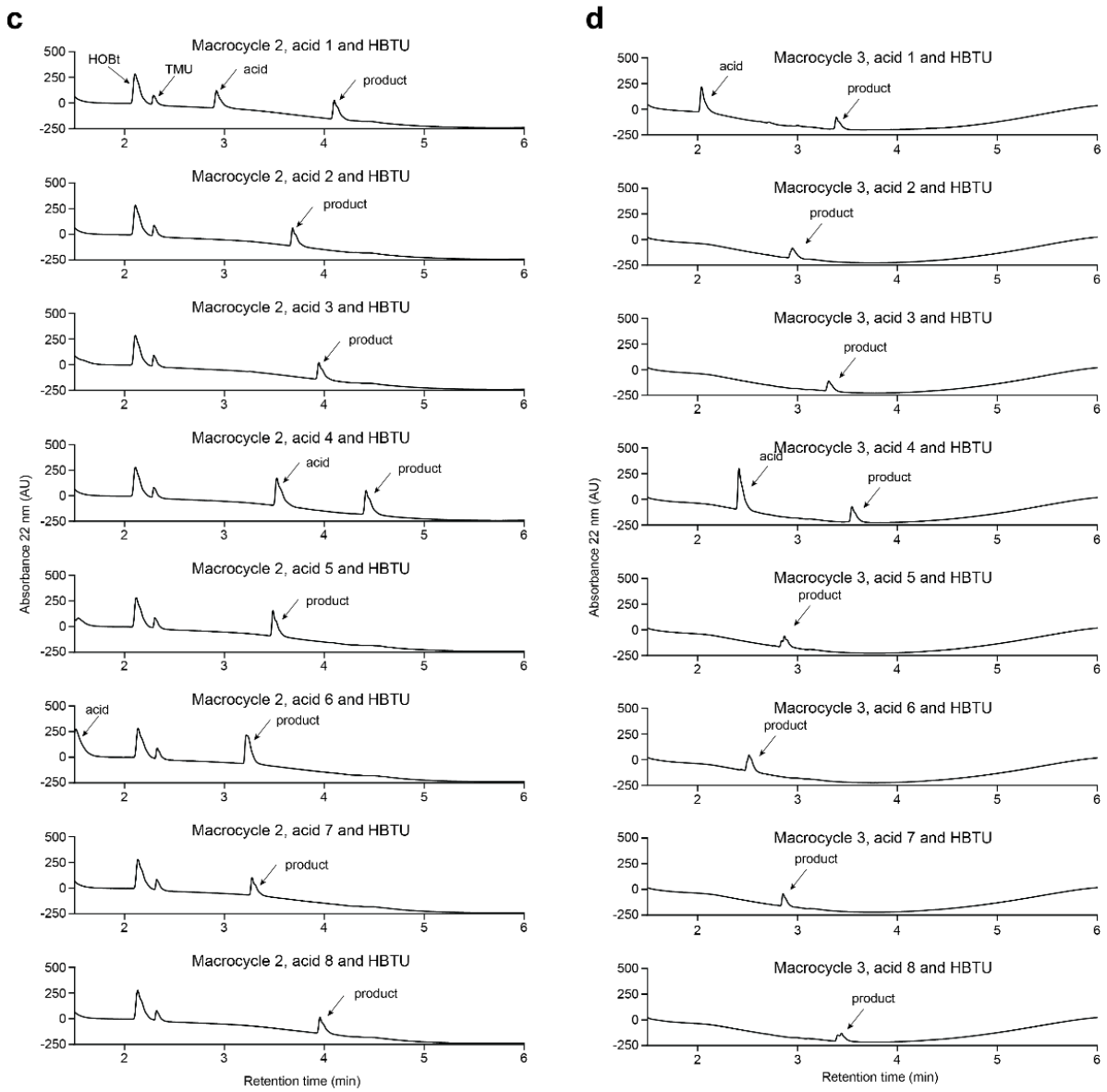
Supplementary Figure 2. Acylation reactions in 80 nL volumes using acoustic dispensing and different bases. Two non-volatile bases (DABCO and HEPES sodium salt) at 80 mM concentration, and a volatile base (NMM) at 500 mM concentration were tested as alternatives to 80 mM DIPEA for the acylation of a model scaffold by three carboxylic acids. Reactions were analyzed by LC-MS using a RP column and a 0-60% MeCN/H_2O gradient over 5 minutes. While reactions with DIPEA did not go to completion, all other bases resulted in quantitative conversion to product.



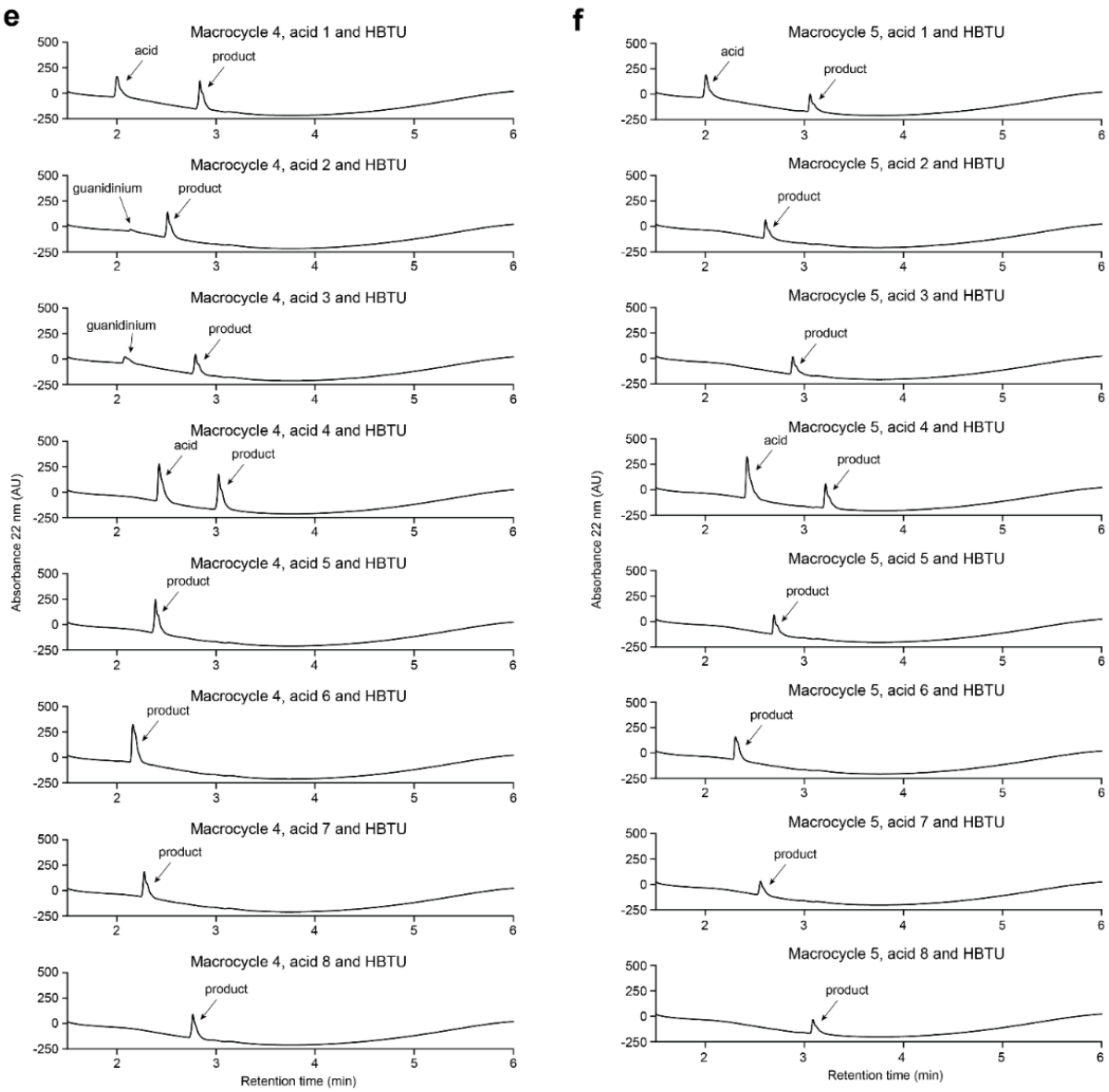
Supplementary Figure 3. Comparison of acylation reactions in 80 nL volumes using acoustic dispending and DABCO as base. The model scaffold 1 was reacted with carboxylic acids **1** – **8** in 80 nL volumes using DABCO as base (80 mM). Reactions were analyzed by LC-MS using a RP column and a 0-60% MeCN/H_2O gradient over 5 minutes.



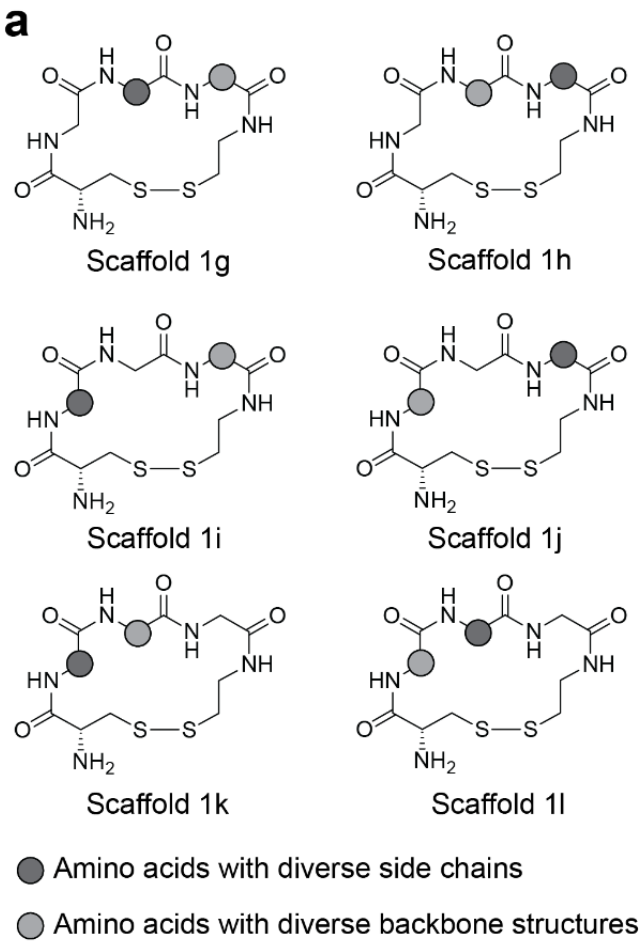
Supplementary Figure 4. Acylation of model scaffolds 2 to 5 in 80 nL volumes and acoustic dispensing. a LC-MS analysis of model scaffolds. Solvent B gradient: 0-60% MeCN, 5 minutes for model scaffold 2; 10-100% MeCN, 5 minutes for model scaffolds 3-5. **b** LC-MS analysis of acids. Solvent B gradient: 0-60% MeCN, 5 minutes. TMU = tetramethyl urea. **c-f** HPLC analysis of acylation reactions for model scaffolds 2 (**c**), 3 (**d**), 4 (**e**) and 5 (**f**). Solvent B gradient: 0-60% MeCN, 5 minutes for model scaffold 2; 10-100% MeCN, 5 minutes for model scaffolds 3-5.



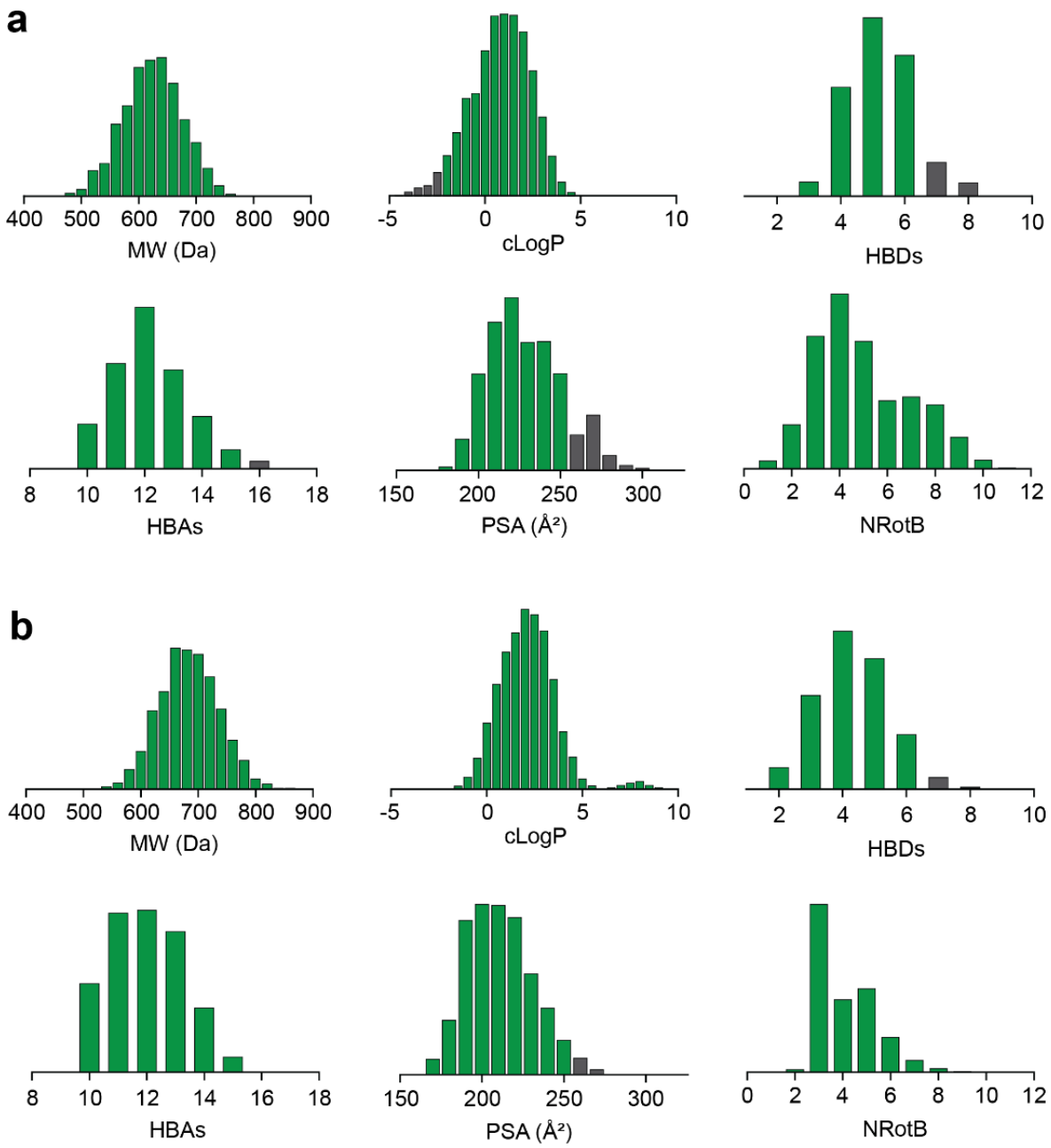
Supplementary Figure 4. Continued



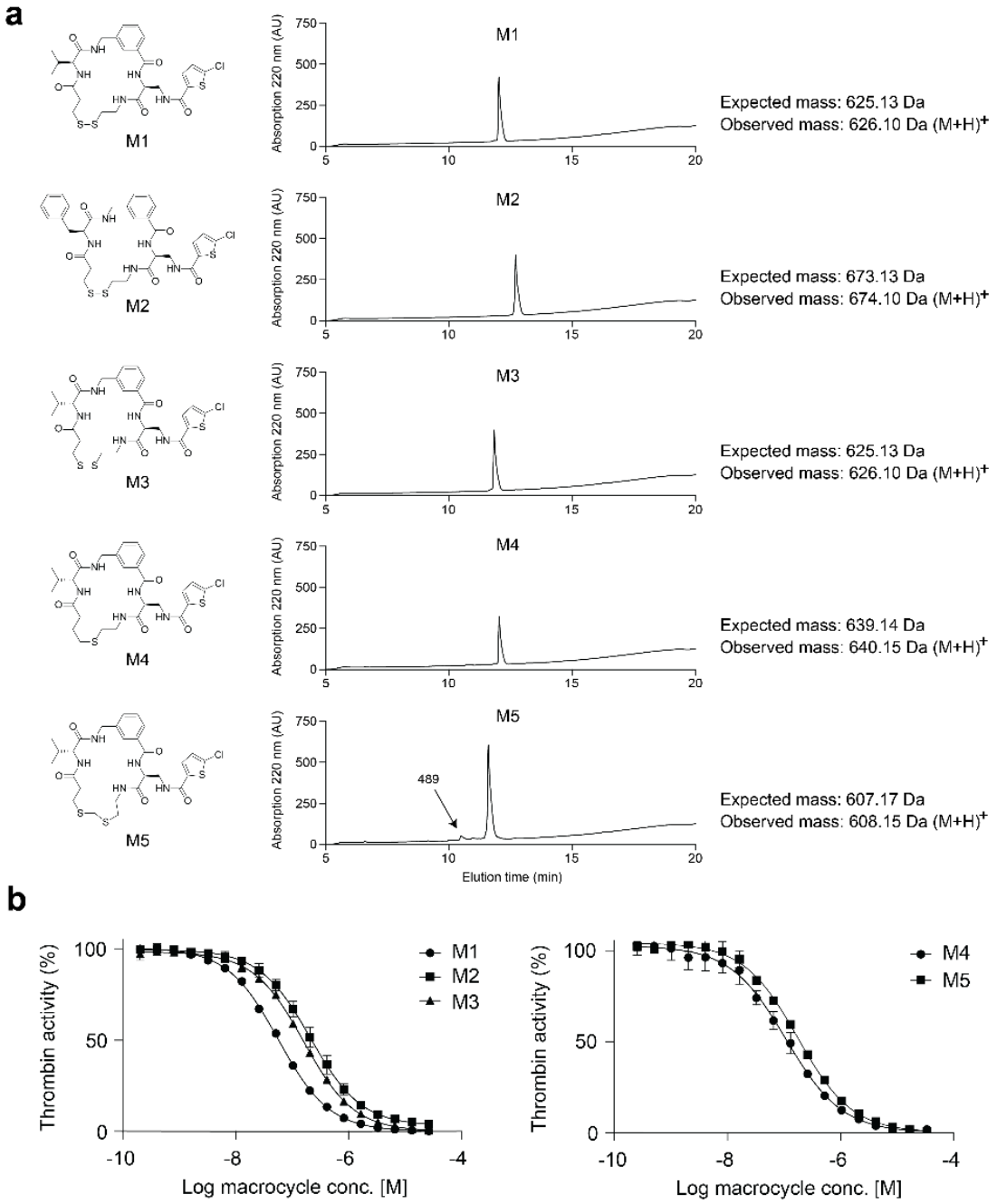
Supplementary Figure 4. Continued



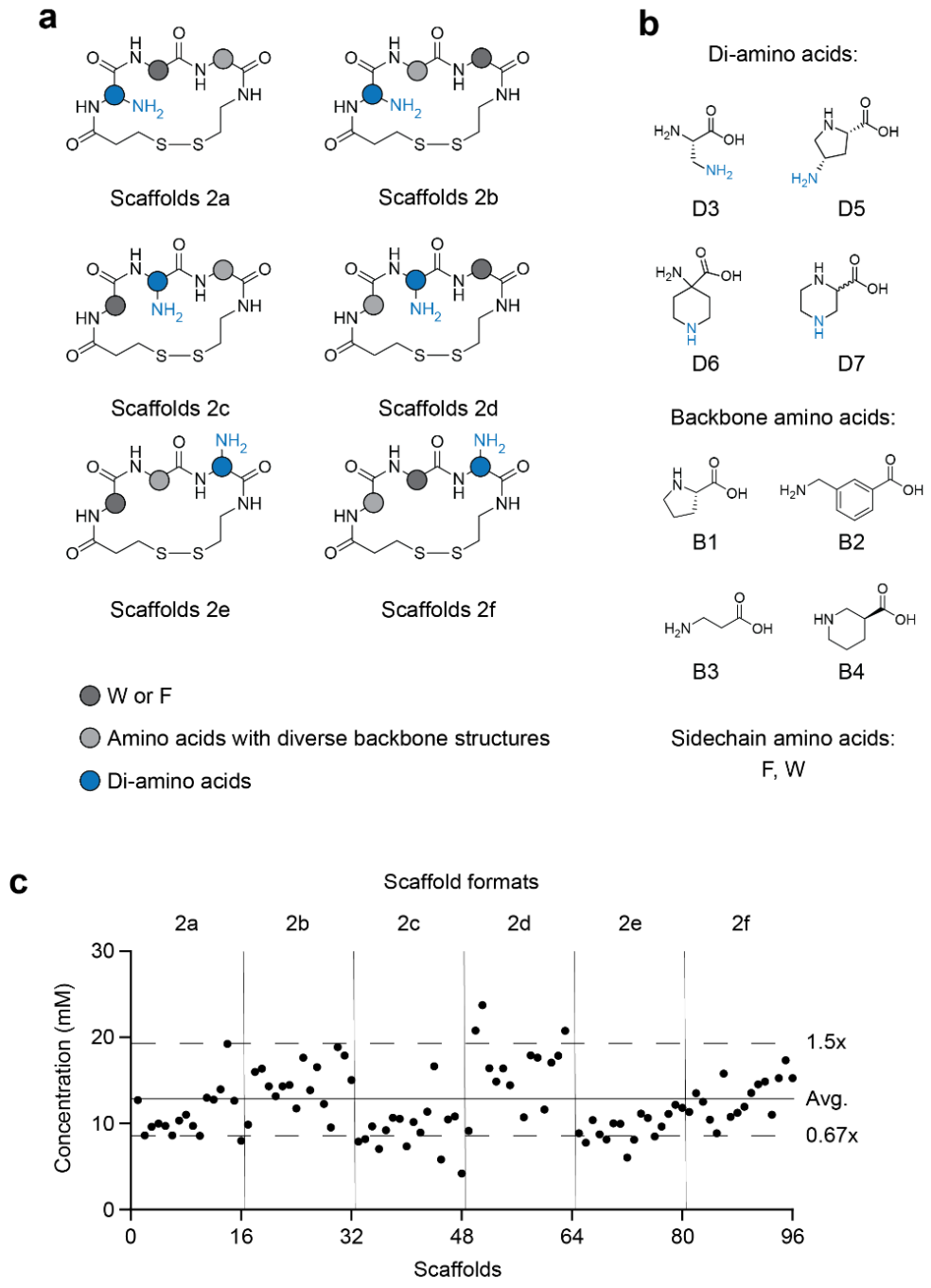
Supplementary Figure 5. Preparation of cyclic peptide scaffolds containing an N-terminal amino group. a Six different scaffold formats. **b** Amino acids used for scaffold synthesis.



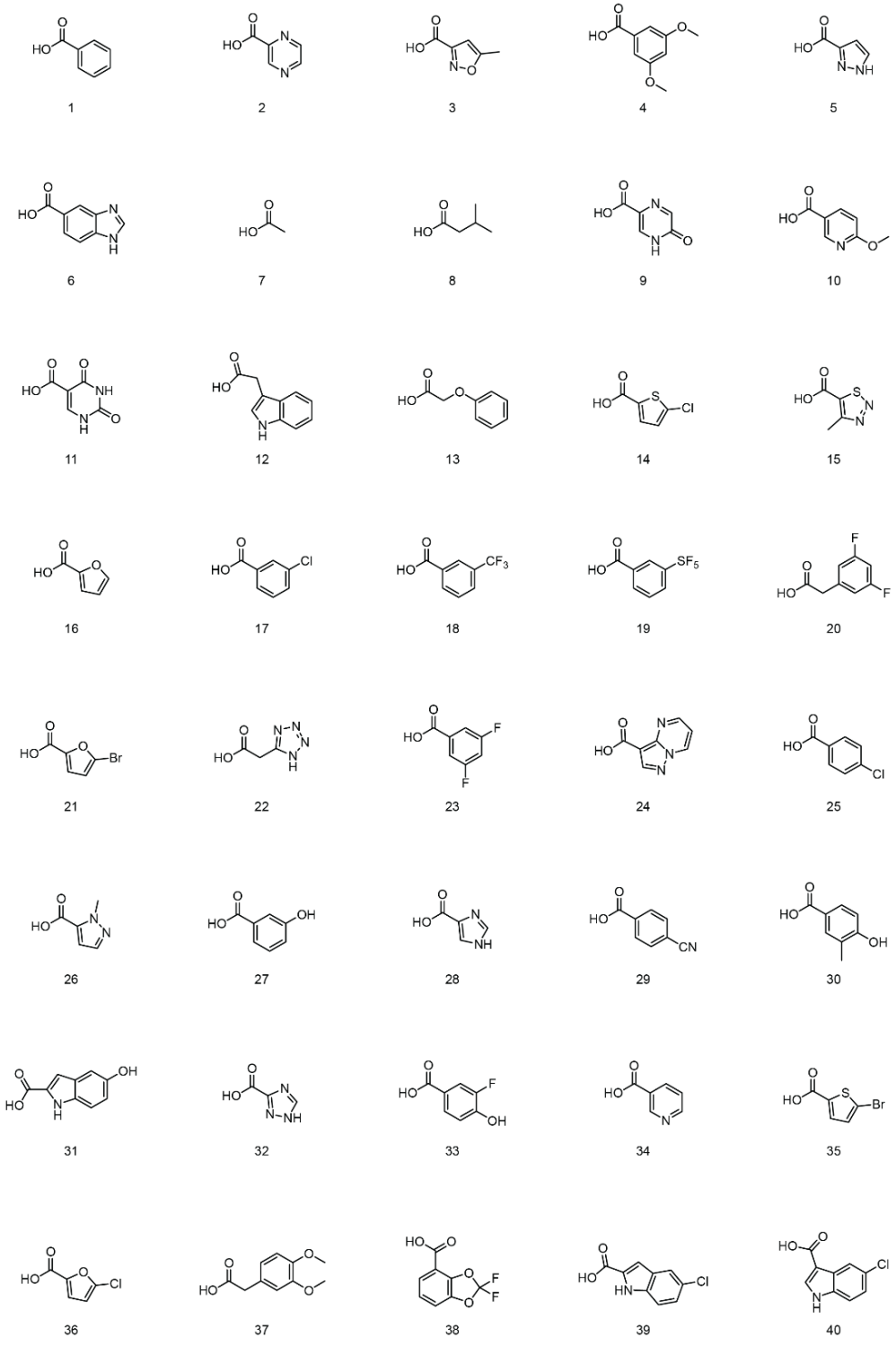
Supplementary Figure 6. Physicochemical properties of Library 1 (thrombin screen) and Library 2 (MDM2 screen). Properties were calculated using DataWarrior software. Regions compliant with Kihlberg's rules for permeability (ref) are colored green. The majority of both libraries fall in a space that is predicted to be cell permeable. MW = molecular weight, cLogP = calculated n-octanol/water partition coefficient, HBD = hydrogen bond donors, HBA = hydrogen bond acceptors, PSA = polar surface area, NRotB = number of rotatable bonds. **a** Library 1 (for thrombin screen). **b** Library 2 (for MDM2 screen).



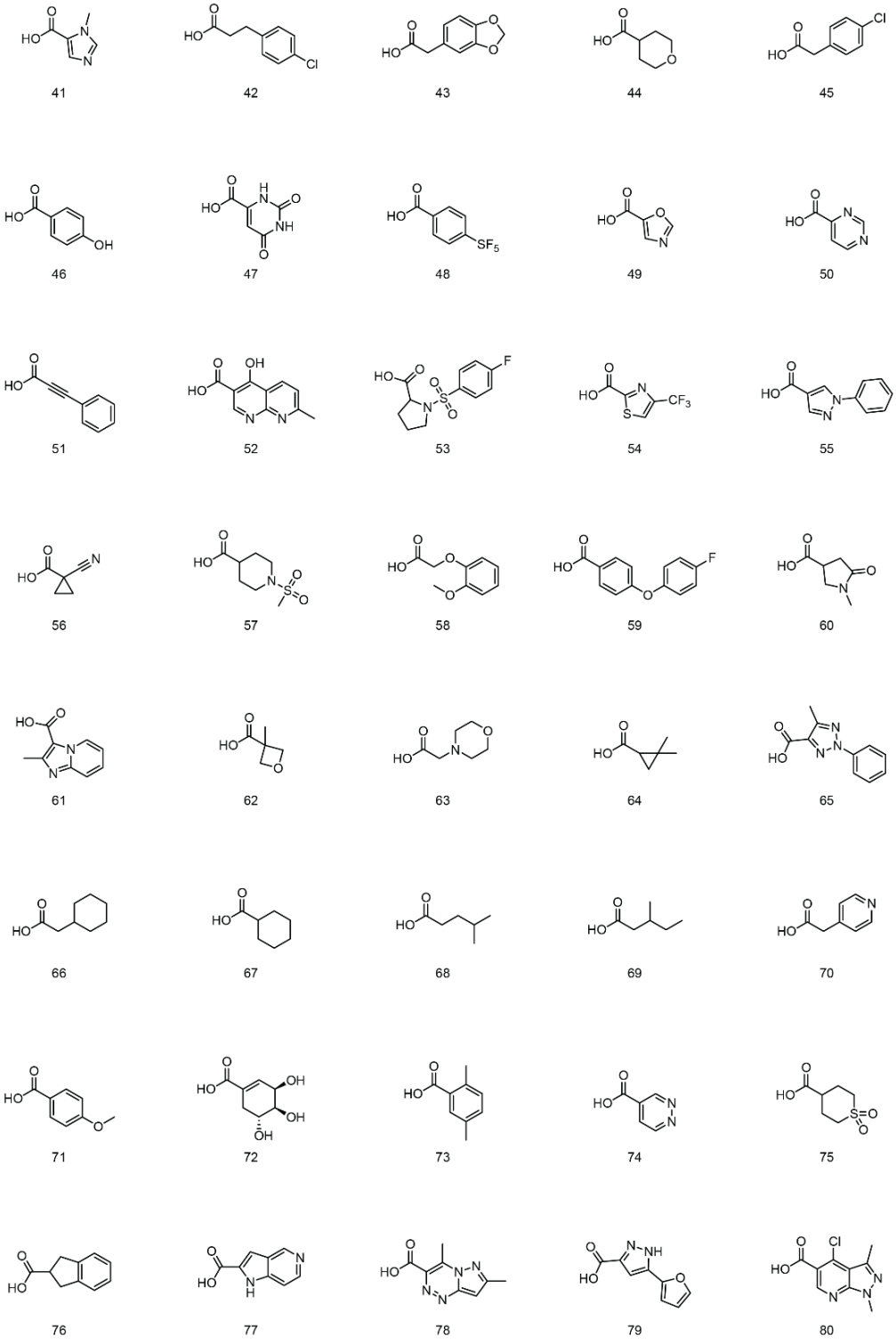
Supplementary Figure 7. Thrombin inhibitors M1 to M5. a Chemical structures and analytical HPLC chromatograms obtained using a 0-100% MeCN/H_2O gradient over 15 minutes. **b** For all macrocycles, an 18-point, two-fold serial dilution was performed in 50 μL volumes. Thrombin was added (50 μL, 2 nM final conc.), followed 10 minutes later by fluorogenic substrate Z-Gly-Gly-Arg-AMC (50 μL, 50 μM final conc.). The increase in fluorescence was measured over 30 minutes. Residual thrombin activity was determined by dividing the slope of fluorescence intensity over time for each well by the slope of control wells without macrocycle. Mean values and SDs are indicated for three independent measurement.



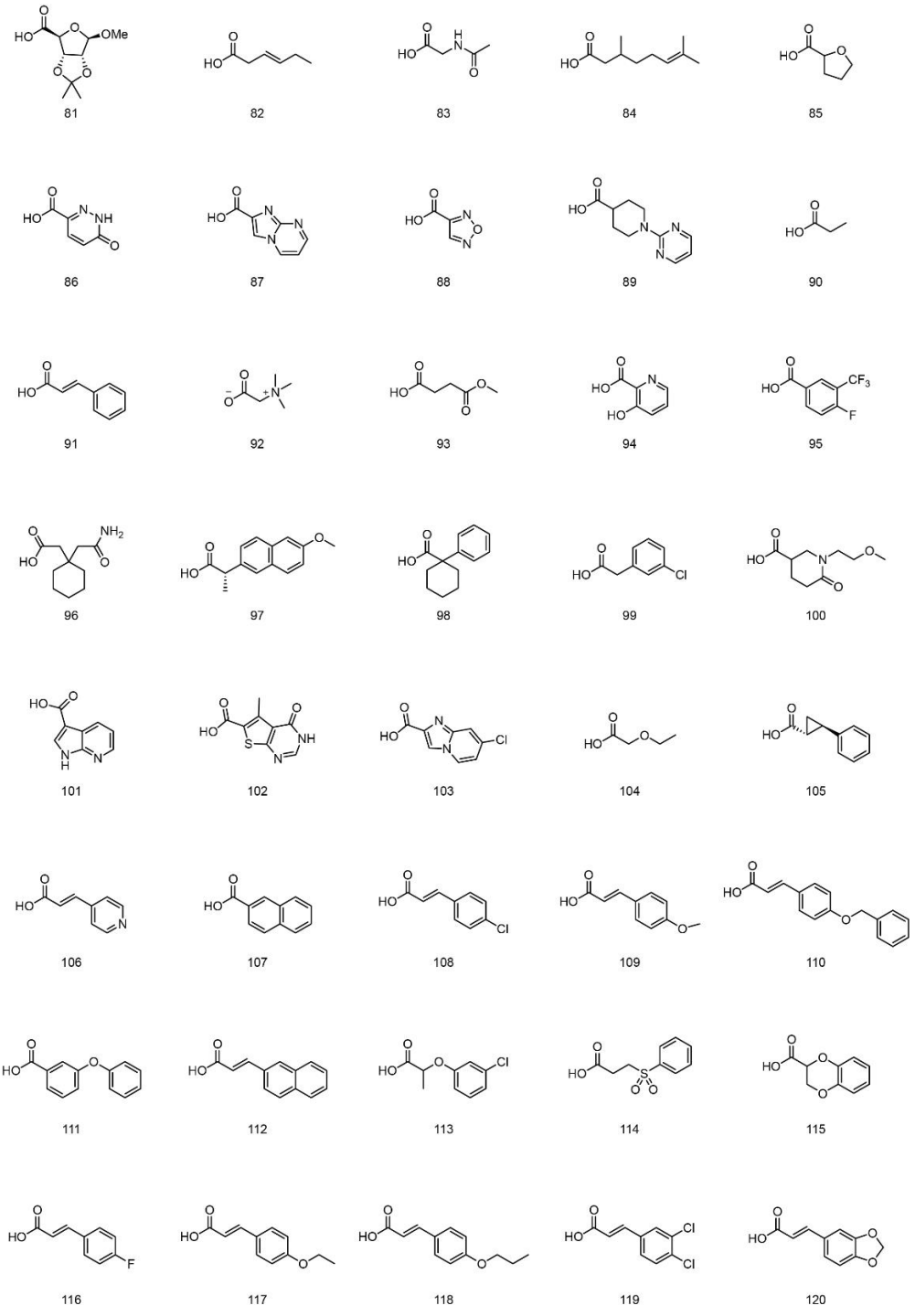
Supplementary Figure 8. Scaffold synthesized for Library 2 (MDM2 screen). a Format of scaffolds in Library 2. **b** Amino acids used for the scaffold library synthesis. All combinations of four di-amino acids, four backbone amino acids, two sidechain amino acids, and six sub-library formats, were synthesized. **c** Yields of tryptophan-containing scaffolds after cyclative release. The average concentration of scaffold was 12.9 mM as determined by nanodrop absorbance. The average purity measured by LC-MS was around 90%.



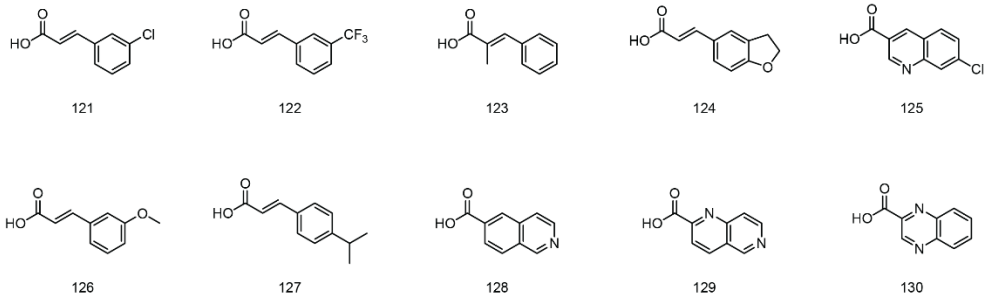
Supplementary Figure 9. Carboxylic acids. Overview of carboxylic acids used to acylate peripheral amines in cyclic peptide scaffolds.



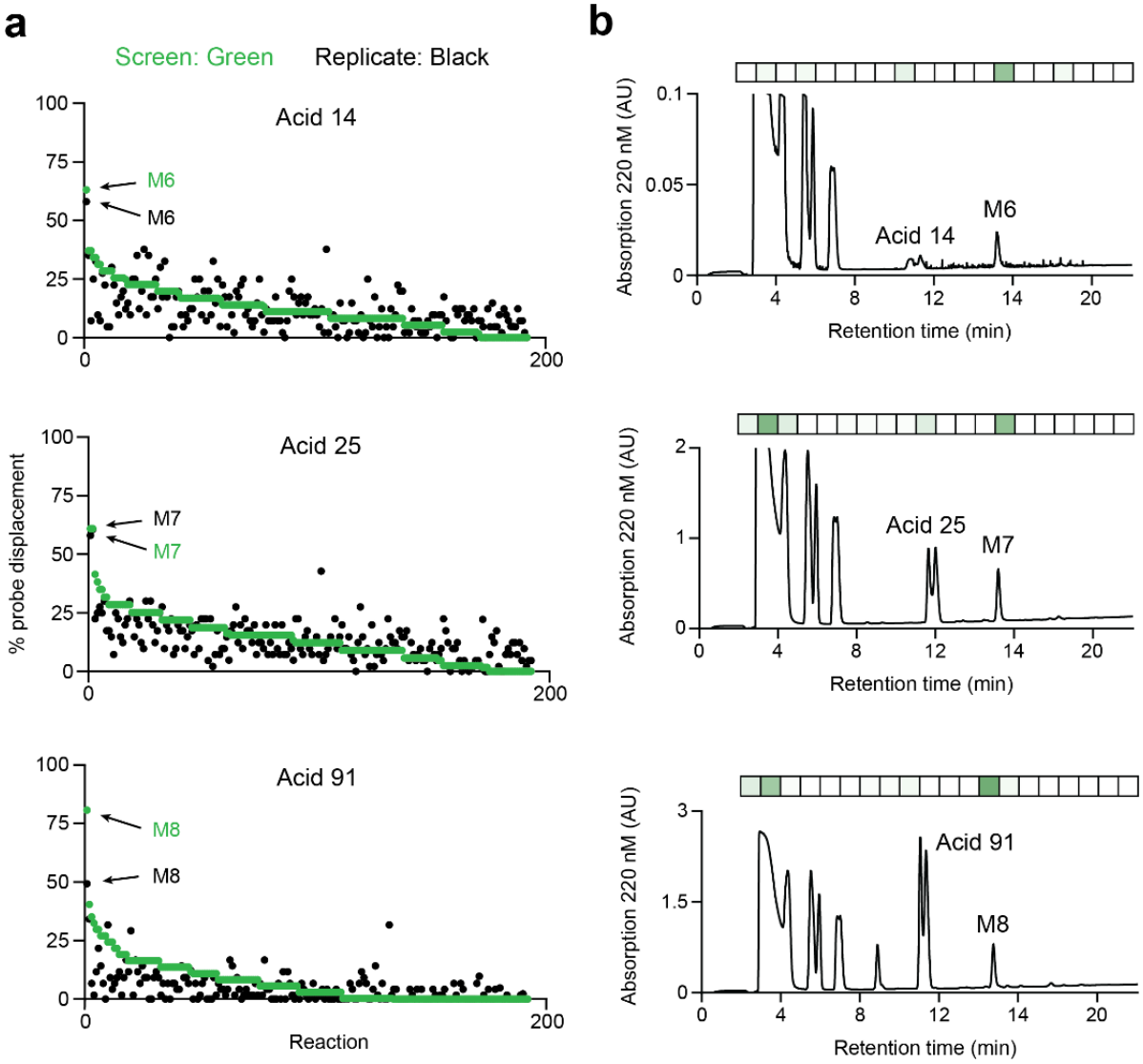
Supplementary Figure 9. Continued



Supplementary Figure 9. Continued



Supplementary Figure 9. Continued



Supplementary Figure 10. Repetition of MDM2 screen with acids that showed strong reporter peptide displacement in screen, and identification of active species in reactions of M6 to M8. a Data of initial screen is shown in green and data of replicate is shown in black. Macrocycles are ordered according to their activity measured in the initial screen. **b** Chromatographic separation acylation reaction yielding M6, M7 and M8 and analysis of fractions for MDM2 binding species.

4.8 References

(1) Scott, D. E.; Bayly, A. R.; Abell, C.; Skidmore, J. Small Molecules, Big Targets: Drug Discovery Faces the Protein–Protein Interaction Challenge. *Nat. Rev. Drug Discov.* **2016**, *15* (8), 533–550. https://doi.org/10.1038/nrd.2016.29.

(2) Dang, C. V.; Reddy, E. P.; Shokat, K. M.; Soucek, L. Drugging the “undruggable” Cancer Targets. *Nat. Rev. Cancer* **2017**, *17* (8), 502–508. https://doi.org/10.1038/nrc.2017.36.

(3) Driggers, E. M.; Hale, S. P.; Lee, J.; Terrett, N. K. The Exploration of Macrocycles for Drug Discovery - An Underexploited Structural Class. *Nat. Rev. Drug Discov.* **2008**, *7* (7), 608–624. https://doi.org/10.1038/nrd2590.

(4) Villar, E. A.; Beglov, D.; Chennamadhavuni, S.; Porco, J. A.; Kozakov, D.; Vajda, S.; Whitty, A. How Proteins Bind Macrocycles. *Nat. Chem. Biol.* **2014**, *10* (9), 723–731. https://doi.org/10.1038/nchembio.1584.

(5) Giordanetto, F.; Kihlberg, J. Macrocyclic Drugs and Clinical Candidates: What Can Medicinal Chemists Learn from Their Properties? *J. Med. Chem.* **2014**, *57* (2), 278–295. https://doi.org/10.1021/jm400887j.

(6) Mortensen, K. T.; Osberger, T. J.; King, T. A.; Sore, H. F.; Spring, D. R. Strategies for the Diversity-Oriented Synthesis of Macrocycles. *Chem. Rev.* **2019**, *119* (17), 10288–10317. https://doi.org/10.1021/acs.chemrev.9b00084.

(7) Peacock, H.; Suga, H. Discovery of De Novo Macrocyclic Peptides by Messenger RNA Display. *Trends Pharmacol. Sci.* **2021**, *42* (5), 385–397. https://doi.org/10.1016/j.tips.2021.02.004.

(8) Deyle, K.; Kong, X.-D.; Heinis, C. Phage Selection of Cyclic Peptides for Application in Research and Drug Development. *Acc. Chem. Res.* **2017**, *50* (8), 1866–1874. https://doi.org/10.1021/acs.accounts.7b00184.

(9) Stress, C. J.; Sauter, B.; Schneider, L. A.; Sharpe, T.; Gillingham, D. A DNA-Encoded Chemical Library Incorporating Elements of Natural Macrocycles. *Angew. Chemie - Int. Ed.* **2019**, *58* (28), 9570–9574. https://doi.org/10.1002/anie.201902513.

(10) Li, Y.; De Luca, R.; Cazzamalli, S.; Pretto, F.; Bajic, D.; Scheuermann, J.; Neri, D. Versatile Protein Recognition by the Encoded Display of Multiple Chemical Elements on a Constant Macrocyclic Scaffold. *Nat. Chem.* **2018**, *10* (4), 441–448. https://doi.org/10.1038/s41557-018-0017-8.

(11) Gartner, Z. J. DNA-Templated Organic Synthesis and Selection of a Library of Macrocycles. *Science* **2004**, *305* (5690), 1601–1605. https://doi.org/10.1126/science.1102629.

(12) Ermert, P. Design, Properties and Recent Application of Macrocycles in Medicinal Chemistry. *Chim. Int. J. Chem.* **2017**, *71* (10), 678–702. https://doi.org/10.2533/chimia.2017.678.

(13) Gao, K.; Shaabani, S.; Xu, R.; Zarganes-Tzitzikas, T.; Gao, L.; Ahmadianmoghaddam, M.; Groves, M. R.; Dömling, A. Nanoscale, Automated, High Throughput Synthesis and Screening for the Accelerated Discovery of Protein Modifiers. *RSC Med. Chem.* **2021**, *12* (5), 809–818. https://doi.org/10.1039/d1md00087j.

(14) Sangouard, G.; Zorzi, A.; Wu, Y.; Ehret, E.; Schüttel, M.; Kale, S.; Diaz Perlas, C.; Vesin, J.; Bortoli Chapalay, J.; Turcatti, G.; Heinis, C. Picomole-scale Synthesis and Screening of Macrocyclic Compound Libraries by Acoustic Liquid Transfer. *Angew. Chemie Int. Ed.* **2021**, anie.202107815. https://doi.org/10.1002/anie.202107815.

(15) Mothukuri, G. K.; Kale, S. S.; Stenbratt, C. L.; Zorzi, A.; Vesin, J.; Bortoli Chapalay, J.; Deyle, K.; Turcatti, G.; Cendron, L.; Angelini, A.; Heinis, C. Macrocycle Synthesis Strategy Based on Step-Wise “Adding and Reacting” Three Components Enables Screening of Large Combinatorial Libraries. *Chem. Sci.* **2020**, *11* (30), 7858–7863. https://doi.org/10.1039/d0sc01944e.

(16) Kale, S. S.; Bergeron-Brlek, M.; Wu, Y.; Kumar, M. G.; Pham, M. V.; Bortoli, J.; Vesin, J.; Kong, X. D.; Franco Machado, J.; Deyle, K.; Gonschorek, P.; Turcatti, G.; Cendron, L.; Angelini, A.; Heinis, C. Thiol-to-Amine Cyclization Reaction Enables Screening of Large Libraries of Macrocyclic Compounds and the Generation of Sub-Kilodalton Ligands. *Sci. Adv.* **2019**, *5* (8), eaaw2851. https://doi.org/10.1126/sciadv.aaw2851.

(17) Fitzgerald, P. R.; Paegel, B. M. DNA-Encoded Chemistry: Drug Discovery from a Few Good Reactions. *Chem. Rev.* **2021**, *121* (12), 7155–7177. https://doi.org/10.1021/acs.chemrev.0c00789.

(18) Bentley, M. R.; Ilyichova, O. V.; Wang, G.; Williams, M. L.; Sharma, G.; Alwan, W. S.; Whitehouse, R. L.; Mohanty, B.; Scammells, P. J.; Heras, B.; Martin, J. L.; Totsika, M.; Capuano, B.; Doak, B. C.; Scanlon, M. J. Rapid Elaboration of Fragments into Leads by X-Ray Crystallographic Screening of Parallel Chemical Libraries (REFiLX). *J. Med. Chem.* **2020**, *63* (13), 6863–6875. https://doi.org/10.1021/acs.jmedchem.0c00111.

(19) Brik, A.; Lin, Y. C.; Elder, J.; Wong, C. H. A Quick Diversity-Oriented Amide-Forming Reaction to Optimize p-Subsite Residues of HIV Protease Inhibitors. *Chem. Biol.* **2002**, *9* (8), 891–896. https://doi.org/10.1016/S1074-5521(02)00184-9.

(20) Habeshian, S.; Sable, G.; Schuttel, M.; Merz, M.; Heinis, C. A Cyclative Release Strategy to Obtain Pure Cyclic Peptides Directly from Solid Phase. submitted.

(21) Doak, B. C.; Over, B.; Giordanetto, F.; Kihlberg, J. Oral Druggable Space beyond the Rule of 5: Insights from Drugs and Clinical Candidates. *Chem. Biol.* **2014**, *21* (9), 1115–1142. https://doi.org/10.1016/j.chembiol.2014.08.013.

(22) Matsson, P.; Doak, B. C.; Over, B.; Kihlberg, J. Cell Permeability beyond the Rule of 5. *Adv. Drug Deliv. Rev.* **2016**, *101*, 42–61. https://doi.org/10.1016/j.addr.2016.03.013.

(23) Stangier, J.; Clemens, A. Pharmacology, Pharmacokinetics, and Pharmacodynamics of Dabigatran Etexilate, an Oral Direct Thrombin Inhibitor. *Clin. Appl. Thromb.* **2009,** *15*, 9S-16S. https://doi.org/10.1177/1076029609343004.

5 Conclusions

Macrocyclic compounds show great promise as therapeutics for difficult targets but it is due to challenges involved in their development, most noteworthy the synthetic difficulties associated with generating large libraries needed for high-throughput screening, that they are not a highly utilized modality in drug discovery. It is for this reason that I sought to develop new methods for the high-throughput synthesis of small macrocycles. A first method developed based on solid phase peptide synthesis and cyclative release afforded pure macrocycles in bio-compatible solvent (DMSO), with only a tertiary amine base as a byproduct. Combinatorial modification of these macrocycles with diverse carboxylic acids using acoustic dispensing provided access to tens of thousands of macrocycles on picomole scale, which I utilized to find nanomolar inhibitors for two proof-of-concept targets. Together, the new methods represent significant improvements in the throughput of macrocycle synthesis. The specific conclusions from each project are discussed below.

5.1. A cyclative release strategy to obtain pure cyclic peptides directly from solid phase

In the developed high-throughput cyclic peptide synthesis strategy, peptides are synthesized using SPPS on an amine that is linked to the solid phase through a disulfide bond, and contain a thiol at the N-terminus. Treatment with TFA removes the protecting groups from amino acids without cleaving the peptide, thus allowing for facile removal of all deprotection byproducts. Subsequent treatment with an amine base results in an intramolecular disulfide exchange, and produces pure cyclic peptides in DMSO. Use of a volatile amine base allows for its removal if desired. I had used the method to synthesize a 96-member library, which I screened against thrombin to determine compatibility with biochemical assays. No undesired interference was observed, and a weak inhibitor was found. Next, a 342-member peptide/peptoid library was synthesized to demonstrate the facile increase in scale, compatibility with reactions other than acylation, and to assess the physicochemical properties.

Thus I was able to reach my goal of establishing a method for making macrocycles that did not require purification and thus allowed production of the macrocycles at high throughput. Compared to previously reported methods, my cyclative release represents a large improvement in terms of throughput and library purity: the macrocycles are obtained in good

yield, high purity, and with almost no byproducts. The method has become the most commonly used method of macrocycle synthesis in our laboratory.

Despite the ease of synthesis and the utility of the cyclative release strategy, some drawbacks exist. The first is that rather short peptides may not all be releases efficiently. While peptides of the format Mpa-Xaa-Xaa-Xaa-Mea undergo cleavage in high yield, and with minimal dimer formation, smaller peptides of the format Mpa-Xaa-Xaa-Mea suffered lower recovery and higher conversion to dimeric species. This issue is especially prevalent when amino acids that would result in a rigid, strained macrocycle are used. The second drawback to the method is that it results in macrocycles containing a disulfide bond. While the disulfide bond is the reason why I can synthesize so many macrocycles, it remains the largest metabolic liability if these compounds are to be used as drugs. The disulfide bond is likely to be readily reduced by glutathione in the cytosol of cells. The limitation of the reducible disulfide bridge may be addressed in future by reducing the cyclic peptides and cyclizing them using bis-electrophile linker reagents prior to screening. Alternative, the disulfide bond of a macrocycle identified in a screen may be replaced by appropriate isosteres. Nevertheless, the synthesis and screening method represents a powerful hit-finding technology.

5.2. Synthesis and screening of large macrocycle libraries by late stage modification at picomole scale

In the second project, I had developed a method for combinatorially diversifying macrocycles with the goal of making tens of thousands of macrocyclic compounds. Using a simple, robust amide bond forming reaction, I was able to couple over 100 carboxylic acids to structurally diverse macrocyclic scaffolds bearing peripheral amine groups. The reaction was capable of being performed in nanoliter volumes using an acoustic dispensing instrument, and was applicable to peptidic and non-peptidic macrocycles. To determine the compatibility with biochemical assays, as well as the effect of a large diversity increase, two proof-of-concept screens were performed. For both screens, macrocycle backbones bearing amine handles were generated using the cyclative release strategy developed in the first project.

In the first screen, roughly 4,000 macrocycles were synthesized and tested for thrombin inhibition, and a double-digit nanomolar inhibitor was found. Several other, weaker inhibitors were identified which contained similar macrocycle structures, which provided convincing SAR data to suggest that the hits were genuine. While one carboxylic acid was crucial for activity, only a small fraction of the macrocycles containing it showed any activity, demonstrating the

importance of the combination of macrocycle backbone and acid. Thrombin bound by the best inhibitor was analyzed by X-ray crystallography, which provided a final confirmation of target engagement.

A second proof of concept screen was performed against MDM2, which engages in a PPI with P53 and is thus more relevant to our interests. Roughly 20,000 macrocycles were synthesized and screened using a FP competition assay. Several interesting series of hit compounds were identified, but upon subsequent hit validation experiments, only one series demonstrated that the desired macrocycles were responsible for the observed activity. Direct binding of the best fluorescein labeled inhibitor showed a K_D of 380 nM. To further demonstrate the power of the method and to optimize the activity of the inhibitor, an additional 23 carboxylic acids were coupled to the active scaffold, and potency was increased to 31 nM.

With this method, I was able to achieve my ultimate goal of combinatorial reaction of *m* macrocycle scaffolds with *n* chemical building blocks to obtain large, diverse macrocycle libraries. Combining a robust diversification reaction with the cyclative release strategy, I have made a method where compounds do not need any purification at any step during synthesis, and can be screened directly as the final crude mixtures without interference from side products. With minimal time and effort, I was able to synthesize and screen 20,000 macrocycles, which already exceeds the size of most commercial macrocycle screening collections. This represents a massive improvement in the throughput of macrocycle synthesis, and a significant contribution to the field. In a secondary application, the method is very useful for the rapid optimization of hits. Tailored libraries can be readily generated in order to obtain SAR information and optimize potency and physicochemical properties in a combinatorial, rather than iterative, fashion. The combined cyclative release and peripheral diversification strategies are currently being used in our laboratory for screening targets of therapeutic interest.

In the future, the platform could be improved to increase throughput and diversity of the macrocycles. Further automation will help to increase the total size of libraries that are synthesized and screened. Currently, our plate reader is not capable of the rapid measurement that would be required for libraries of greater than 100,000 macrocycles. In addition, all plate sealing, de-sealing, and centrifugation steps need to be performed manually, which limits the throughput. If these functions, along with a higher-throughput plate reader, were integrated into an automated platform with robotics, vast numbers of macrocycles, exceeding a million, could be synthesized and screened. Our laboratory is currently in the process of establishing such an automated platform. The scope of diversification reactions could also be increased in the future. Any robust reaction that is tolerant of many functional groups, and is compatible

with biochemical or cellular assays could be used. Importantly, reactions beyond amide bond formation would also improve the physicochemical properties of the library compounds by minimizing polar surface area and HBDs. In the research presented here, the diversification reactions were applied to disulfide cyclized macrocycles, which as previously mentioned, could suffer from unfavorable reduction in vivo. However, as I demonstrated in proof-of-concept reactions, the diversification can be applied to macrocycles of any type. Thus there is the potential for combining the method with new macrocycle backbone generation technologies for making even more attractive compounds.

CV

Sevan Habeshian

Rue du Simplon 32A, 1020 Renens CH
+41 079 624 36 83
Sevan.Habeshian@Gmail.com

Education

2017 – 2021* PhD in chemical biology, **EPFL** (Lausanne VD, Switzerland)

2006 – 2010 MSc in chemistry, **Brandeis University** (Waltham MA, USA)

2006 – 2010 BA in chemistry, **Brandeis University** (Waltham MA, USA)

Professional experience

2017 – Present **Doctoral student**
EPFL Laboratory of Therapeutic Proteins and Peptides Lausanne, CH
Advisor: Professor Christian Heinis

- Developed a high throughput synthetic method to make small, cell permeable macrocycles with high purity.
- Developed an acoustic droplet ejection-compatible reaction to diversify macrocycles on picomole scale.
- Screened a total of 25,000 macrocycles against two proof-of-concept targets, and identified inhibitors with low nanomolar affinity for both.
- Optimized potency and physicochemical properties of a macrocyclic PPI inhibitor

2016 – 2017 **Principal Research Associate – Discovery Chemistry Group**
X-Chem Pharmaceuticals, Waltham MA, USA
Promoted to join a partner-facing project team

- Worked as part of a project team to provide pharma/biotech partners with optimized compounds for their targets of interest.
- Analyzed DNA-encoded library selection data to identify hits for several protein target classes, including kinases, PPIs, and epigenetic proteins.
- Independently designed compounds for resynthesis off-DNA.
- Aided in hit-to-lead optimization for an internal program.

2010 – 2016 **Senior Research Associate – Library Synthesis *Group***
X-Chem Pharmaceuticals, Waltham MA, USA
First member hired into the Library Synthesis Group

- Synthesized DNA-encoded small-molecule libraries utilizing chemical and enzymatic ligation techniques, and a variety of chemical reactions; personally responsible for the synthesis of over 80 billion compounds.
- Developed novel on-DNA reactions and designed libraries that have produced lead compounds.
- Managed a collection of greater than 20,000 chemical building blocks.
- Mentored and trained new hires in library synthesis and analysis.

2008-2010 **Research Assistant**
Laboratory of Prof. Isaac Krauss, Brandeis University Waltham MA, USA
Performed senior/Master's thesis research, assisted other research projects

- Synthesized, purified, and characterized oligosaccharide derivatives. Performed total synthesis of an oligomannan azide important for the development of HIV vaccines
- Contributed to the synthesis of bromophycolide natural products by synthesizing multi gram-scale quantities of an intermediate product.

Publications

(1) **Habeshian, S.**; Merz, M. L.; Sangouard, G. ; Mothukuri, G. ; Schüttel, M. ; Vesin, J. ; Chapalay, J. B. ; Turcatti, G. ; Cendron, L. ; Angelini, A. ; Heinis, C. Generation and screening of large macrocycle diversities by late-stage modification at picomole scale. **2021**, *Manuscript in preparation.*

(2) **Habeshian, S.**; Sable, G. A.; Schüttel, M.; Merz, M. L.; Heinis, C. A cyclative release strategy to obtain pure cyclic peptides directly from solid phase. **2021**, *Submitted*

(3) Guilinger, J. P.; Archna, A.; Augustin, M.; Bergmann, A.; Centrella, P. A.; Clark, M. A.; Cuozzo, J. W.; Däther, M.; Guié, M. A.; **Habeshian, S.**; Kiefersauer, R.; Krapp, S.; Lammens, A.; Lercher, L.; Liu, J.; Liu, Y.; Maskos, K.; Mrosek, M.; Pflügler, K.; Siegert, M.; Thomson, H. A.; Tian, X.; Zhang, Y.; Konz Makino, D. L.; Keefe, A. D. Novel Irreversible Covalent BTK Inhibitors Discovered Using DNA-Encoded Chemistry. *Bioorganic Med. Chem*. **2021**, *42*, 116223.

(4) Brown, D. G.; Brown, G. A.; Centrella, P.; Certel, K.; Cooke, R. M.; Cuozzo, J. W.; Dekker, N.; Dumelin, C. E.; Ferguson, A.; Fiez-Vandal, C.; Geschwindner, S.; Guié, M. A.; **Habeshian, S.**; Keefe, A. D.; Schlenker, O.; Sigel, E. A.; Snijder, A.; Soutter, H. T.; Sundström, L.; Troast, D. M.; Wiggin, G.; Zhang, J.; Zhang, Y.; Clark, M. A. Agonists and Antagonists of Protease-Activated Receptor 2 Discovered within a DNA-Encoded Chemical Library Using Mutational Stabilization of the Target. *SLAS Discov.* **2018**, *23 (5)*, 429–436.

(5) Cuozzo, J. W.; Centrella, P. A.; Gikunju, D.; **Habeshian, S.**; Hupp, C. D.; Keefe, A. D.; Sigel, E. A.; Soutter, H. H.; Thomson, H. A.; Zhang, Y.; Clark, M. A. Discovery of a Potent BTK Inhibitor with a Novel Binding Mode by Using Parallel Selections with a DNA-Encoded Chemical Library. *ChemBioChem* **2017**, *18 (9)*, 864–871.

(6) Johannes, J. W.; Bates, S.; Beigie, C.; Belmonte, M. A.; Breen, J.; Cao, S.; Centrella, P. A.; Clark, M. A.; Cuozzo, J. W.; Dumelin, C. E.; Ferguson, A. D.; **Habeshian, S.**; Hargreaves, D.; Joubran, C.; Kazmirski, S.; Keefe, A. D.; Lamb, M. L.; Lan, H.; Li, Y.; Ma, H.; Mlynarski, S.; Packer, M. J.; Rawlins, P. B.; Robbins, D. W.; Shen, H.; Sigel, E. A.; Soutter, H. H.; Su, N.;

Troast, D. M.; Wang, H.; Wickson, K. F.; Wu, C.; Zhang, Y.; Zhao, Q.; Zheng, X.; Hird, A. W. Structure Based Design of Non-Natural Peptidic Macrocyclic Mcl-1 Inhibitors. *ACS Med. Chem. Lett*. **2017**, *8 (2)*, 239–244.

(7) Soutter, H. H.; Centrella, P.; Clark, M. A.; Cuozzo, J. W.; Dumelin, C. E.; Guie, M. A.; **Habeshian, S.**; Keefe, A. D.; Kennedy, K. M.; Sigel, E. A.; Troast, D. M.; Zhang, Y.; Ferguson, A. D.; Davies, G.; Stead, E. R.; Breed, J.; Madhavapeddi, P.; Read, J. A. Discovery of Cofactor-Specific, Bactericidal Mycobacterium Tuberculosis InhA Inhibitors Using DNA-Encoded Library Technology. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113 (49)*, E7880–E7889.

(8) Litovchick, A.; Dumelin, C. E.; **Habeshian, S.**; Gikunju, D.; Guié, M. A.; Centrella, P.; Zhang, Y.; Sigel, E. A.; Cuozzo, J. W.; Keefe, A. D.; Clark, M. A. Encoded Library Synthesis Using Chemical Ligation and the Discovery of SEH Inhibitors from a 334-Million Member Library. *Sci. Rep.* **2015**, *5*.

(9) MacPherson, I. S.; Temme, J. S.; **Habeshian, S.**; Felczak, K.; Pankiewicz, K.; Hedstrom, L.; Krauss, I. J. Multivalent Glycocluster Design through Directed Evolution. *Angew. Chemie - Int. Ed.* **2011**, *50 (47)*, 11238–11242.