How to cite: *Angew. Chem. Int. Ed.* **2024**, *63*, e202400350
doi.org/10.1002/anie.202400350

Macrocycles Hot Paper

Large Libraries of Structurally Diverse Macrocycles Suitable for Membrane Permeation

*Alexander L. Nielsen, Zsolt Bognar[+], Ganesh K. Mothukuri[+], Anne Zarda[+], Mischa Schüttel, Manuel L. Merz, Xinjian Ji, Edward J. Will, Monica Chinellato, Christian R. O. Bartling, Kristian Strømgaard, Laura Cendron, Alessandro Angelini, and Christian Heinis**

Abstract: Macrocycles offer an attractive format for drug development due to their good binding properties and potential to cross cell membranes. To efficiently identify macrocyclic ligands for new targets, methods for the synthesis and screening of large combinatorial libraries of small cyclic peptides were developed, many of them using thiol groups for efficient peptide macrocyclization. However, a weakness of these libraries is that invariant thiol-containing building blocks such as cysteine are used, resulting in a region that does not contribute to library diversity but increases molecule size. Herein, we synthesized a series of structurally diverse thiol-containing elements and used them for the combinatorial synthesis of a 2,688-member library of small, structurally diverse peptidic macrocycles with unprecedented skeletal complexity. We then used this library to discover potent thrombin and plasma kallikrein inhibitors, some also demonstrating favorable membrane permeability. X-ray structure analysis of macrocycle-target complexes showed that the size and shape of the newly developed thiol elements are key for binding. The strategy and library format presented in this work significantly enhance structural diversity by allowing combinatorial modifications to a previously invariant region of peptide macrocycles, which may be broadly applied in the development of membrane permeable therapeutics.

Introduction

Traditional drug discovery methods have focused predominantly on small molecules and monoclonal antibodies; however, small molecules often are not able to engage with challenging biological targets such as proteins with flat and featureless protein surfaces, and antibodies are not membrane permeable, limiting the applicability of both classes in certain therapeutic areas. A molecule format that could potentially fill this gap are the macrocycles.[1] These ring-shaped structures have a high conformational stability that allows strong and specific binding to challenging targets. Their slightly larger size compared to classical small molecules also allows for more molecular contact with the target, increasing the overall interaction surface area.[2] Despite their size that typically falls outside the range of conventional druglikeness models such as Lipinski's rule of five (Ro5),[3] some macrocycles are membrane permeable and thus can reach intracellular targets, and some can even be applied orally.[4] Overall, their favorable binding properties and ability to cross membrane barriers offer enormous opportunities for macrocycles in drug development, such as for generating inhibitors of the numerous intracellular protein-protein interactions that have been difficult to target with classical modalities.

Most existing macrocyclic drugs are based on natural products or derivatives thereof,[1,5,6] though for the majority of disease targets, natural ligands do not exist and must therefore be developed de novo using synthetic approaches. However, this discovery process is currently

[*] A. L. Nielsen, Z. Bognar,[+] G. K. Mothukuri,[+] A. Zarda,[+] M. Schüttel, M. L. Merz, X. Ji, E. J. Will, C. Heinis
Institute of Chemical Sciences and Engineering, School of Basic Sciences, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland
E-mail: christian.heinis@epfl.ch

M. Chinellato, L. Cendron
Department of Biology, University of Padova, 35131 Padova, Italy

C. R. O. Bartling, K. Strømgaard
Center for Biopharmaceuticals and Department of Drug Design and Pharmacology, Faculty of Health and Medical Sciences, University of Copenhagen, Jagtvej 162, DK-2100 Copenhagen, Denmark

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

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

[+] These authors contributed equally to this work

hindered by a lack of techniques for developing and screening libraries of permeable macrocyclic ligands to new targets. Cyclic peptide libraries screened by phage or mRNA display usually yield peptides longer than ten amino acids that typically are not membrane permeable, though an exception to this was recently reported being a heavily *N*-methylated cyclic peptide ligand of KRAS that is cell active.[7,8] DNA-encoding technologies have been applied to generate and screen macrocycles libraries and are promising, but the synthesis of good quality libraries is not trivial due to the large number of coupling and deprotection steps in the presence of DNA.[9] In recent years, collections of purified macrocyclic compounds were produced and are offered for classical high-throughput screening but such libraries are rather small. The generation of larger collections comprising ten-thousands of macrocyclic compounds is hindered by typically low-yielding cyclization reactions that necessitate chromatographic purifications of individual macrocycles prior to screening. Consequently, creating large numbers of macrocycles is extremely time-consuming and resource-intensive. Overall, the difficulty of generating these libraries has drastically limited the potential applicability of macrocyclic compounds as therapeutics.

To address these existing issues with macrocyclic drug development, methods for generating and screening large libraries comprising tens of thousands of small peptide-based macrocycles were developed.[10–12] The approaches are based on screening crude products without purification, which is made possible by high-throughput solid-phase peptide synthesis (SPPS) of short random peptides that can be efficiently cyclized via two thiol end groups[13,14] and combinatorial late-stage diversification of the peptides using efficient and selective reactions.[11,12,15] However, a major limitation of current macrocycle libraries is that they have a structurally identical region that is required for macrocyclization (colored groups in Figure 1a and 1b), which compromises the structural diversity of the library and impairs the chances of identifying good binders. Not surprisingly, ligands identified from such libraries were all binding via the variable region and the constant groups are pointing away from the targets.[10–12] The same structural limitation is also found in disulfide- or thioether-cyclized peptide collections produced by cyclizing peptides via thiol groups, where the random amino acids are flanked by two constant cysteines.[16–19] In macrocycle libraries produced most recently in our lab, we cyclized peptides via the two thiol building blocks, 3-mercaptopropionic acid (Mpa; at the N-terminal end) and 2-mercaptoethylamine (Mea; also named cysteamine).[11,12] They contribute 89 and 76 Da to the molecular mass, respectively, which is substantial considering that this comprises 33 % of the molecule weight target of <500 Da for the best chances of membrane permeability. Hence, to increase the structural diversity of our macrocyclic libraries while remaining within the realm of likely permeable compounds, it is essential to diversify one or both of these constant elements.

Accordingly, in this work, we developed a synthetic strategy to efficiently produce short peptides with altered C-terminal thiol-containing elements (Figure 1a and 1b) to contribute to the development of macrocyclic ligands with a high probability of being membrane-permeable. Adding six structurally different C-terminal thiol-containing elements increases the skeletal diversity of the generated libraries 7-fold compared to existing libraries, all without increasing the size of the macrocycles that could negatively affect membrane permeability. Several of the developed thiol-containing elements lack an H-bond donor, reducing the overall polarity of the macrocycles for even increased chances of membrane permeability. Synthesis and screening of structurally diverse libraries of 2,688 macrocycles containing varied C-terminal thiols yielded nanomolar inhibitors of thrombin and plasma kallikrein, of which several exhibited high cellular permeabilities. The herein-developed approach for synthesizing small and structurally diverse macrocycles suitable for membrane permeation is broadly applicable and may offer a solution for developing therapeutics that can be orally applied and/or reach intracellular disease targets.

Results and Discussion

To synthesize large libraries of short peptides with altered C-terminal thiol-containing elements, we build on a recently established procedure in which peptides are synthesized by SPPS via a disulfide linker that can be reduced to release the peptides.[13,14] In this strategy, peptides are obtained in high purity without a purification step because we can first deprotect all side-chains on-resin using trifluoroacetic acid (TFA) and then release the washed and deprotected peptides from the resin via a disulfide reduction with 1,4-butanedithiol (BDT). BDT is volatile and can be removed under reduced pressure by rotary vacuum concentration (RVC) to afford only the remaining dithiol peptides in good crude purity (typically 80 % or higher). The C-terminal peptide building block used for this synthesis so far was the aminothiol group **1** (Mea; Figure 1a) that was conjugated to thiol-functionalized resin in a disulfide exchange reaction using pyridyldithioethylamine.[11,13,14] To generate peptides having different thiol-containing groups at the C-terminus (Figure 1b), our initial plan was to use analogous pyridylthiol-activated aminothiols, but we soon realized that their synthetic access was restricted due to the lack of commercial availability of structurally diverse aminothiols or high costs of the material. As an alternative approach, we conjugated different aminothiols onto thiol resin using aminothiols with phenylsulfone as leaving group that is equally suited for disulfide exchange reactions.[20–22] Such building blocks can easily be accessed by reacting thiosulfonates with halogenated *N*-Boc protected amines that are commercially available and cheap (Figure 2a). Importantly, the thiosulfonate precursors were separated from the excess thiosulfonate by extraction without the need of chromatographic purification, which enormously facili-

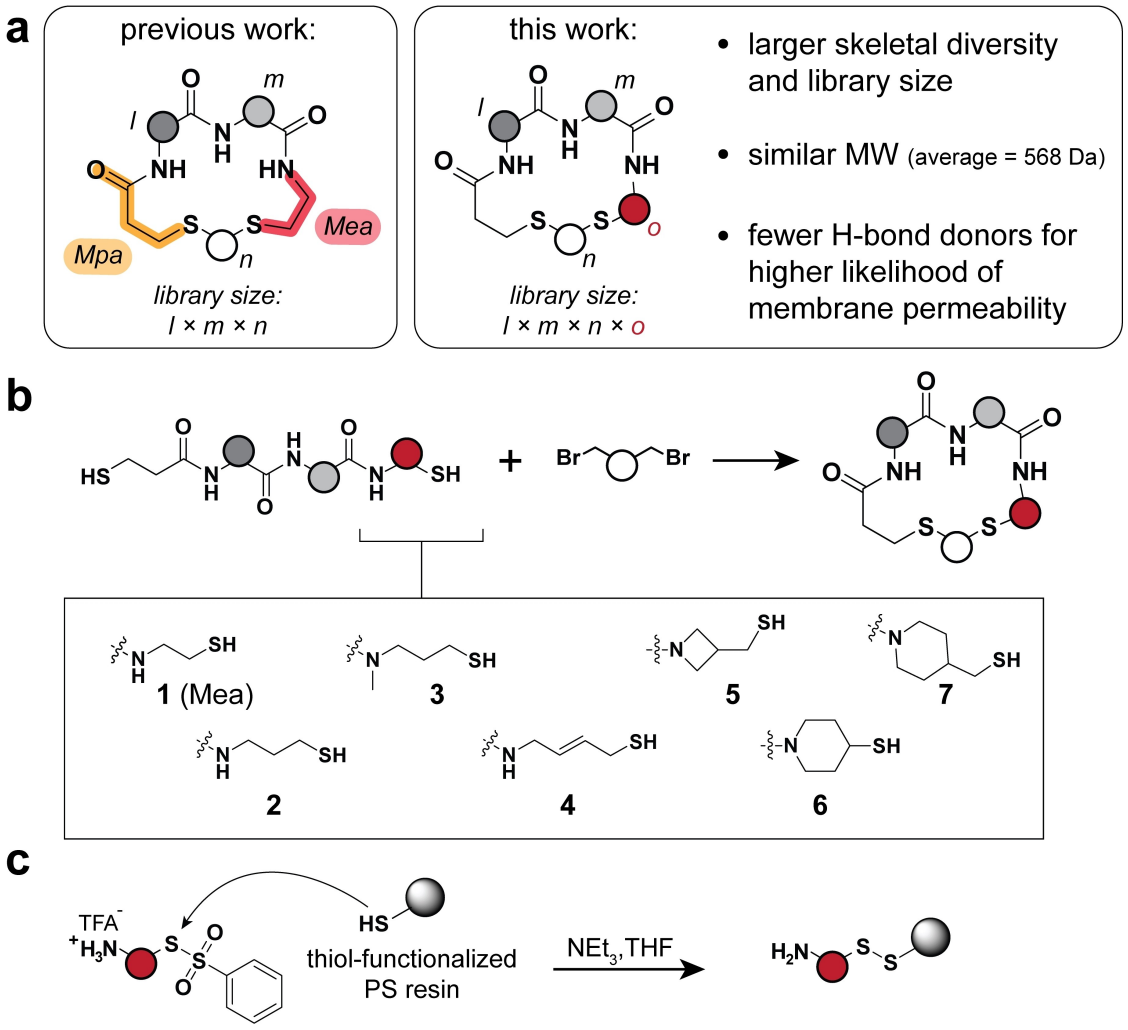


Figure 1. Macrocycle format and synthesis strategy. a) Macrocycle structures and building blocks that can be varied combinatorially to generate large libraries. Gray circles indicate amino acids, Mea and Mpa are the thiol-containing building blocks, and white circles are cyclization linkers. The red circle indicates the herein-developed varied C-terminal aminothiol element. In the chemical structures, standard amide bonds are shown for simplicity although some of them are *N*-methylated. b) Schematic representation of the thiol-thiol cyclization strategy. Chemical structures of diverse aminothiols are shown. c) Schematic representation of the linkage of aminothiols to resin via a disulfide bridge using thiosulfonate building blocks. PS = polystyrene.

tated the preparation of such reagents. Through nucleophilic substitution with sodium benzenethionosulfonate in DMF at 80 °C, we prepared six new *N*-Boc thiosulfonate building blocks on a gram-scale, ready for testing under library-synthesis conditions (Supporting Figure S1; see Supporting Information for the synthesis).

To immobilize the aminothiols **2**–**7** onto the solid support, we first performed Boc deprotection of the *N*-Boc thiosulfonate building blocks and then conjugated the crude mixtures to thiol-resin via a disulfide exchange reaction (Figure 1c and Supporting Scheme S1). To test the new aminothiol-functionalized resins, we synthesized on them the model peptide Mpa-Trp-Ala by SPPS and liberated the resulting dithiol peptides using the reducing agent BDT (Figure 2b). HPLC-MS analysis showed purities ranging between 85 and 98 % (Figure 2c and Supporting Figure S2), indicating the new aminothiol building blocks generate sufficiently pure crude products for the preparation and screening of macrocycle libraries.

To test whether target-specific macrocycles could be identified using the new diversification elements, we designed a library of 2,688 macrocycles that would be generated by cyclizing 384 dithiol peptides with seven bis-electrophilic linker reagents (Figure 3a). The 384 dithiol peptides were designed to each contain: one aminothiol derivative (**1**–**7**), an amino acid that binds the S1 pocket (chosen from three motifs known from literature to improve targeting to trypsin-like serine proteases),[11] a random amino acid (from 27 diverse α, β, γ, and *N*-methylated amino acids), and Mpa as the N-terminal thiol element (Figure 3a and 3b; see Supporting Table S1 for list of synthesized peptides). Primarily due to the variation of the aminothiol building blocks, this library contains a high number of diverse macrocycle backbones (2,107) that

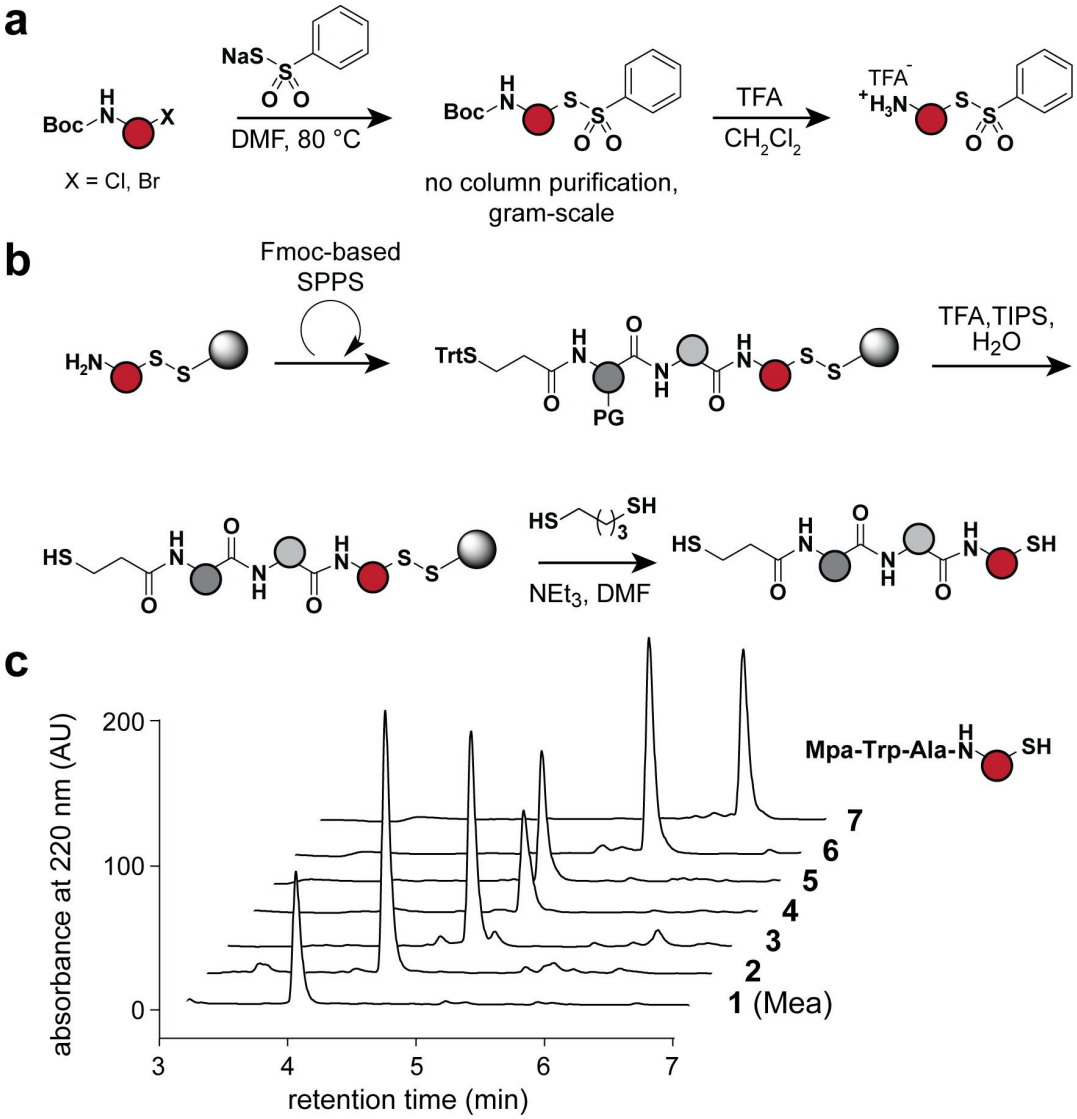


Figure 2. Preparation of aminothiol building blocks. a) Strategy for the "no purification" synthesis of aminothiol compounds activated with a phenylsulfone leaving group. b) SPPS of dithiol peptides using resin carrying a disulfide-linked aminothiol group. c) Stacked HPLC chromatograms of the quality of crude model peptides (Mpa-Trp-Ala-aminothiol X) synthesized with aminothiols **1**–**7**.

are counted by considering only the macrocyclic ring structure and ignoring side chains. This makes the skeletal diversity of the library >4-fold larger than that of the most diverse libraries developed previously (144 and 432 backbones).[10,12] We computationally predicted the physicochemical properties of the generated macrocycles and found that the average molecular weight (568 Da) and other key properties such as polar surface area are well inside the range where macrocyclic compounds are typically able to cross membranes. Overall, this suggested that hits from this library have a good chance of being cell-permeable (Figure 3c and Supporting Table S2), so we proceeded with the library synthesis.

We synthesized the 384 dithiol peptides by automated SPPS in 4×96-well plates using the seven different aminothiols and BDT for peptide release (see workflow in Supporting Figure S3). We determined the peptide yields using Ellman's reagent, which showed an even concentration distribution of compounds prepared with the different aminothiol building blocks (24.9 ± 10.7 mM average concentration; Supporting Figure S4) and a yield of ~60 % based on resin loading. An analysis of 20 randomly chosen peptides by HPLC-MS showed a greater than 80 % average purity, with most impurities stemming from incomplete couplings with challenging Fmoc amino acids during SPPS. This shows that the structurally diverse C-terminal thiol elements are fully compatible with the established SPPS library synthesis, deprotection and reductive release method. More specifically, the results also meant that the disulfide bridges formed by the new aminothiols were sufficiently stable for Fmoc deprotection and on-resin side chain deprotection by TFA, and enough accessible for efficient reductive cleavage with BDT.

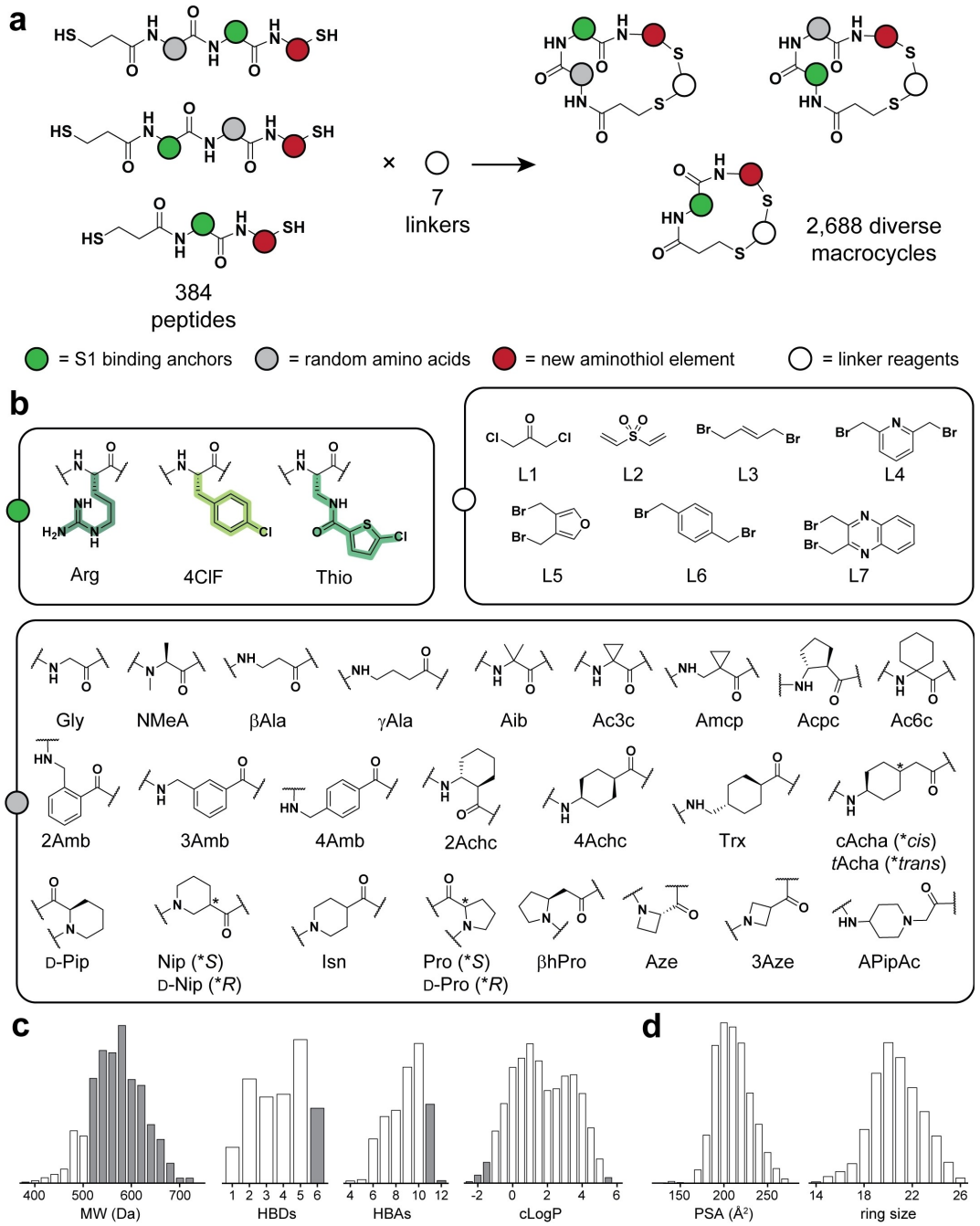


Figure 3. Macrocycle library designed for the inhibition of trypsin-like serine proteases. a) Format of linear peptides before and after cyclization by bis-electrophilic reagents. In the chemical structures, standard amide bonds are shown for simplicity although some of them are *N*-methylated. b) Amino acids and linkers used for library synthesis. The colored dots indicate the four groups of building blocks. c) Histograms of selected physicochemical properties of the macrocyclic library calculated using DataWarrior.[31] Marked in white is the area in accordance with the Ro5.[3] d) Additional library properties not belonging to the Ro5.

Next, we cyclized the dithiol peptides using bis-electrophilic linker reagents. We applied seven linkers having different lengths and shapes to impose different conformations onto the peptide backbones.[10] To limit side products stemming from oxidation such as disulfide cyclized or oligomerized peptide observed in previous work, we introduced a new procedure in which we freshly reduced the thiol groups by incubation with BDT, removed the reducing agent by RVC, and added the linker reagents to the fully reduced and dried peptide (Figure 4a). We performed the cyclization reactions at a 40 nmol scale, distributing the peptides into 384-well polypropylene source plates by acoustic liquid transfer. HPLC-MS analysis of two random peptides cyclized with the seven different linkers showed nearly quantitative cyclization with all linkers (Supporting Figure S5). Thus, we were able to minimize disulfide oxidation

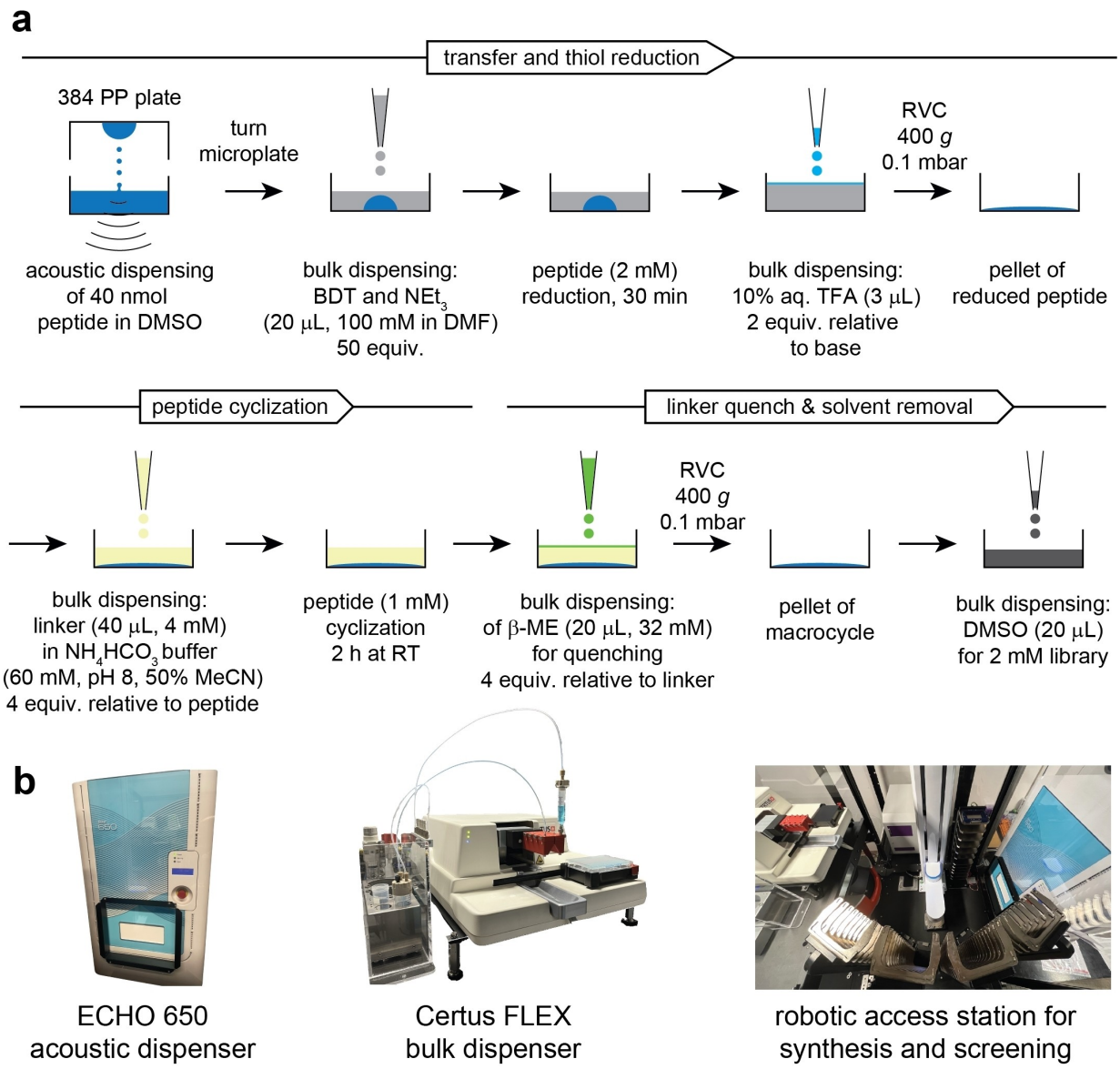


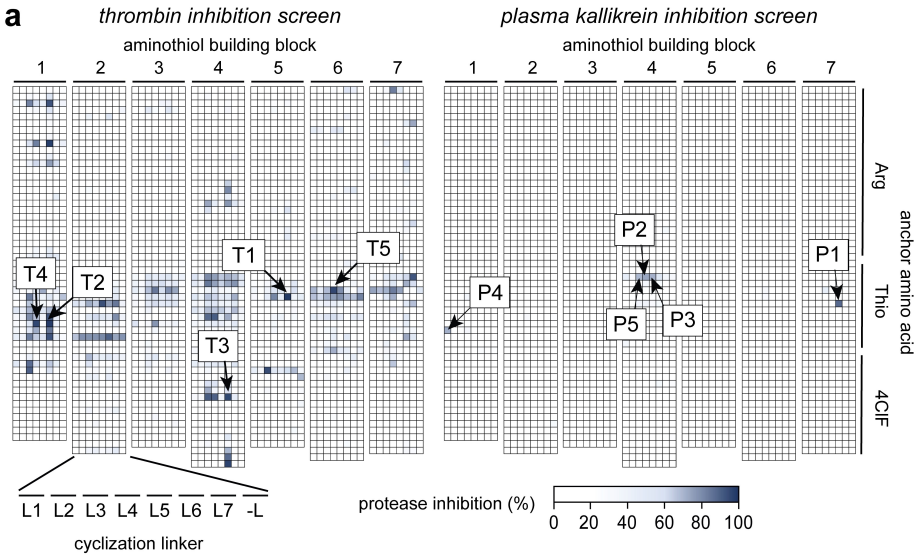
Figure 4. Workflow for the combinatorial cyclization of dithiol peptides in microtiter plates. a) Workflow for the reduction, cyclization, and quenching of the linker reagent in 384-well plates. b) Robotic instrumentation used for the full automation and tip-less peptide library cyclization and activity screening.

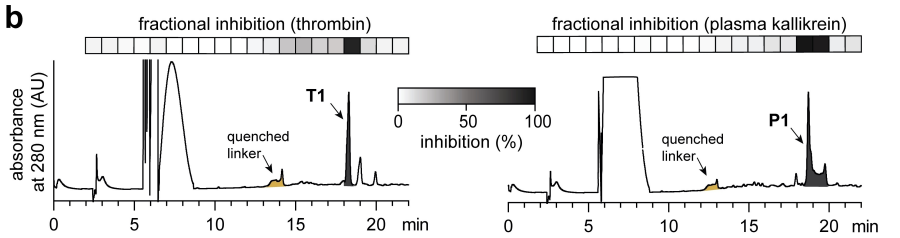
side products by developing a procedure to dispense bis-electrophiles to reduced and dried rather than dissolved peptides. Our workflow generated macrocyclic compounds at a larger scale than previously (40 nmol vs. 160 pmol),[11] producing material for more than 100 screens and thus allowing screening of the same library against multiple targets, as shown herein. While we synthesized the macrocyclic molecules in 384-well plates in this work, we have preliminary data showing that the same procedures are applicable in 1,536-well plates, allowing us to further increase the library size and throughput by 4-fold with an equivalent effort. As final quality check, we analyzed by LC-MS 28 randomly picked library members wherein we ensured that all seven aminothiol building blocks and cyclization linkers were represented. While some reactions showed major side products such as truncated peptide or non-identified species, and some products were difficult to detect due to low UV absorbance of the chemical structures, the desired products could be detected in 27 of the 28 samples and for 16 of them, the desired macrocycle was the main product (Supporting Figure S5).

We screened the crude library against five trypsin-like serine proteases, namely thrombin, kallikrein 5 (KLK5), human coagulation factor XI (FXI), human coagulation factor XII (FXII), and plasma kallikrein (PK) using a final concentration of 10 µM (thrombin) or 20 µM (other proteases) crude compounds. To identify hit compounds based on activity and not just binding, we measured the ability of the macrocycles to inhibit enzymatic cleavage of fluorogenic 7-amino-4-methylcoumarin (AMC) substrates. The entire

screen was fully automated using a robotic access station (Figure 4b) so that the high-throughput screen for all targets, comprising a total of >15,000 assay wells (1,536-well plates), was performed in only a few hours. Pleasingly, the screens against two of the targets, thrombin and PK, showed many hits with high activities (Figure 5a). Given the much stronger inhibition seen for these two targets than for KLK5, FXI and FXII (Supporting Figure S6) we focused in our analysis and the characterization on hits form the thrombin and PK screens. Structure-activity analysis showed that all aminothiol building blocks generated hits in the thrombin screen (Supporting Figure S7), wherein the hits for PK were found only for macrocycles based on aminothiols **1**, **4**, and **7**, underscoring the importance of skeletal diversity in libraries.

We next analyzed whether hits of the crude macrocycle screens were based on the anticipated macrocycles or potential side-products that can occasionally give rise to activity in crude product screens.[10] HPLC-MS analysis showed that the predominant peptide was the desired





c

macrocycle	% inhibition in screen	K_i (nM)	sequence
T1	95.3	19.5 ± 3.2	Mpa Thio Nip **5** (L6)
T2	94.6	10.9 ± 5.6	Mpa *t*Acha Thio **1** (L6)
T3	92.5	123 ± 47	Mpa 4ClF Nip **4** (L6)
T4	92.5	26 ± 1	Mpa *t*Acha Thio **1** (L4)
T5	92.3	81 ± 5	Mpa Thio βAla **6** (L4)

macrocycle	% inhibition in screen	K_i (μM)	sequence
P1	85	0.51 ± 0.04	Mpa Thio Pro **7** (L6)
P2	72	6.5 ± 0.5	Mpa Thio Gly **4** (L4)
P3	70	12.3 ± 1.1	Mpa Thio Gly **4** (L5)
P4	70	13.6 ± 2.6	Mpa Thio APipAc **1** (L1)
P5	68	3.8 ± 0.2	Mpa Thio Gly **4** (L3)

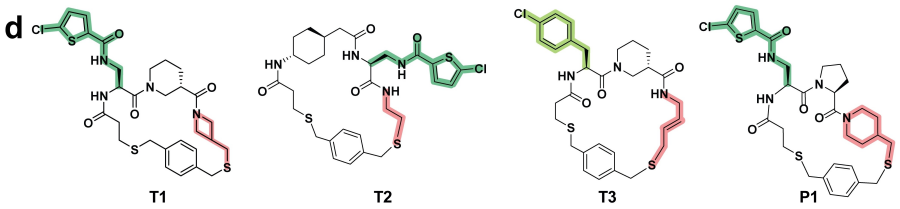


Figure 5. Macrocyclic library screening. a) Heatmap showing residual protease inhibition of thrombin (10 μM macrocycle) and PK (20 μM macrocycle). A non-linear color scale was chosen to emphasize compounds with higher activity (dark blue) and hide inhibition below 30% (all white) as shown by scale bar. Arrows indicate and label macrocycles that were purified and showed the highest protease inhibitory activity. b) Identification of active species in crude cyclization reaction of **T1** and **P1**. Fractional inhibition shows the extent of protease inhibition found for indicated HPLC fractions, each one corresponding to the indicated 1-minute elution times. c) Activities of purified macrocycles. Average values and SDs are based on four independent experiments (**T1**–**T3**), or two (**T4**) or four (**T5, P1**–**P5**) replicate measurements performed in one experiment. d) Chemical structures of selected hits. The group binding to the S1 pocket is highlighted in green and the aminothiol element in red.

macrocycle in all crude DMSO stocks of the top seven thrombin hits (**T1**–**T7**) and top five PK hits (**P1**–**P5**) (Supporting Figure S8). To confirm that the inhibitory effect stemmed from the macrocyclic species as opposed to the presence of excess quenched linkers, side products, or peptide aggregates, we repeated the cyclization of **T1** and **P1** at 40 nmol scales, fractionated the crude cyclization mixture using preparative reversed-phase (RP) HPLC, and measured the activity of all collected fractions. In both cases, strong inhibition was observed only for fractions containing the desired macrocycle (Figure 5b). To determine IC_{50} and K_i values, we re-synthesized all 12 hits on a larger scale (25 μmol) and purified them by preparative RP-HPLC (Figure 5c and 5d, Supporting Figures S9–S12). Of the seven thrombin hits that showed 92 % or greater inhibition in the crude screen, five (**T1**–**T5**) were nanomolar binders, and the one with the highest affinity (**T2**) had a measured K_i of 10.9 ± 5.6 nM (Figure 5c). **T2** was around 4-fold more potent than inhibitors identified previously by screening of other random libraries (best $K_i = 42$ nM),[10,11] suggesting that the higher structural diversity directly translated into the isolation of better inhibitors. For PK, we confirmed the inhibition for all top five hits (**P1**–**P5**) and determined the best inhibitor to be **P1** ($K_i = 0.51 \pm 0.04$ μM). For PK, previous screening of other synthetic libraries performed in our lab had not yielded any binders (not published), so the sub-micromolar inhibitor was likely found as a result of the larger structural diversity in this screen. Testing the five thrombin inhibitors in the PK activity assay showed that they are all highly specific for thrombin with a selectivity of over 1,000-fold. In contrast, the five PK inhibitors blocked thrombin to a similar extent, with the exception of **P1** that showed more than 50-fold selectivity for PK over thrombin (Supporting Figure S13).

To determine the importance of the size and shape of the newly introduced aminothiol elements in the binding interaction, we used X-ray crystallography to analyze the macrocycles bound to thrombin. We obtained diffraction-quality crystals of thrombin co-crystallized with compounds **T1** (PDB 8ASF, 2.57 Å resolution) and **T3** (PDB 8ASE, 2.55 Å resolution). From the structures, we see that the macrocycles perfectly occupy the enzyme's active site, with the S1-binding chlorothiophene and 4-chlorophenylalanine groups pointing directly into the S1 pocket (Figure 6a and 6b; see Supporting Results and Supporting Figures S14–S16 for additional discussion of the structures). An overlay of the two inhibitors showed an overall similar binding orientation, though the positions of the α-carbons of the S1-binding groups differ greatly due to the different lengths but identical binding modes of these groups (Figure 6c). The different shapes and dimensions of the aminothiol groups seem to thus accommodate the structural differences of the S1-binding groups. Taken together, the structural results showed that the larger structural diversity in the backbones of the macrocycles enabled the discovery of perfectly fitting inhibitors.

Finally, to measure the membrane permeability of these compounds, we first determined the passive permeability of the thrombin inhibitors **T1**–**T5** using the parallel artificial membrane permeability assay (PAMPA).[23] Two macrocycles, **T1** and **T3**, displayed an effective permeability coefficient (P_e) of greater than 1×10^{-6} cm/s (Figure 7a,

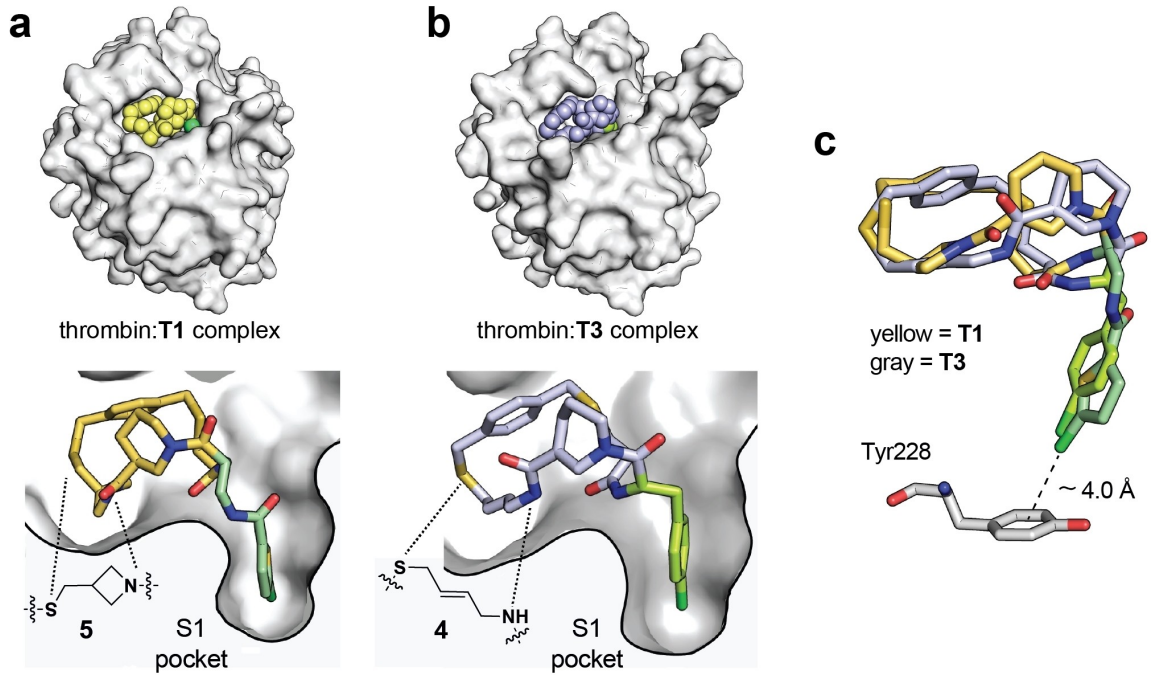


Figure 6. X-ray structures of thrombin:macrocycle complexes. a,b) Surface view of thrombin (grey) with bound **T1** (a) or **T3** (b). The lower parts of the panels show the macrocycles as sticks, wherein the S1 binding anchor is highlighted in green. The regions of the aminothiol building blocks are indicated. c) Superposition of **T1** and **T3** in the active site and highlighting of the halogen-π interaction with Tyr228.

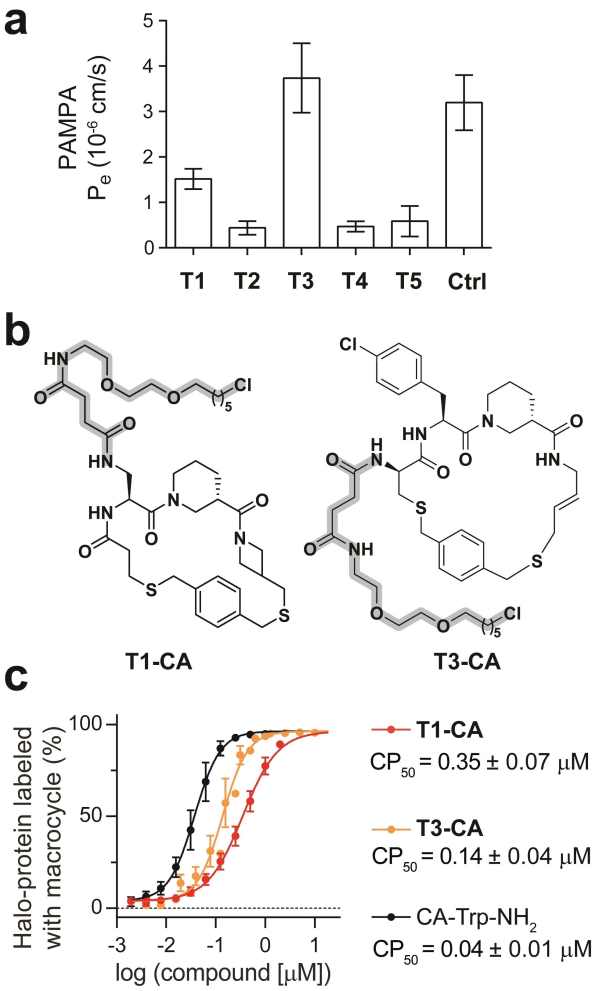


Figure 7. Macrocycle membrane permeability. a) PAMPA permeability after 12 h incubation at 50 μM concentration. Ctrl = previously reported membrane permeable macrocycle **15** (see Supporting Figure S17 for structure) was used as positive control. Mean values and SDs of three replicates done in the same PAMPA plate are shown. b) Structures of macrocycles **T1** and **T3** containing a chloroalkane (CA) tag (gray shaded) for quantification of compound that reaches the cell cytosol by the CAPA assay. c) CAPA results after exposure of cells to macrocycles for 4 hrs. The percent of HaloTag occupied by macrocycle-CA was quantified by reaction of free HaloTag with fluorescent TAMRA-CA and flow cytometry. Mean CP_{50} values ± SD from four experiments performed on different days. CA-Trp-NH_2 was used as positive control.

Supporting Table S3), which is in the range of the macrocycle-based drug cyclosporin that has a good membrane permeability ($P_e = 2\times10^{-6}$ cm/s).[24] To evaluate if **T1** and **T3** were also able to penetrate the plasma membranes of live cells, we applied the chloroalkane penetration assay (CAPA).[25,26] In this assay, compounds bearing a chloroalkane (CA) group are added to cultured cells that express HaloTag[27] protein in the cytosol and CA-tagged compounds that enter the cell react covalently with HaloTag. The %-HaloTag occupied with compound-CA, being proportional to the amount that crossed the membrane, is quantified with a cell-permeable fluorescent chase and flow cytometric analysis. To prepare for the assay, we therefore synthesized macrocycle analogs bearing a CA group by either replacing the chlorothiophene group in **T1** with an amino acid having CA as side chain to obtain the comparably sized macrocycle **T1-CA**, or by appending the CA to the N-terminus of **T3** to generate **T3-CA**, for the latter molecule risking that the larger size might reduce the permeability compared to the non-tagged macrocycle (Figure 7b and Supporting Figures S18–S20). After incubating cells for four hours with different concentrations of **T1-CA** and **T3-CA** (2-fold dilutions) we observed sub-micromolar half-maximal cellular penetration (CP_{50}) values of 0.35 ± 0.07 μM and 0.14 ± 0.04 μM, respectively (Figure 7b), which are found in CAPA for passively permeable compounds such as the highly permeable CA-Trp-NH_2 used as a control (0.04 ± 0.01 μM). This result effectively validated the strategy of diversifying a previously constant element in macrocycle libraries, which yielded more potent inhibitors that are also membrane-permeable. Notably, some of the identified binders showed good membrane permeability in artificial membranes (PAMPA) and live cells because the size of the macrocycles was maintained small, close to 500 Da.

Conclusion

In this work, we developed a strategy for efficiently synthesizing large libraries of peptide macrocycles that have a more than 4-fold higher structural diversity than previous libraries, important for identifying binders to targets, and yet are sufficiently small sized to achieve good membrane permeability. We achieved this by diversifying elements of macrocyclic peptides that are typically kept constant in library formation such as the cysteines or cysteamines used for cyclizing peptides via thioether or disulfide linkages. This diversification became possible via the development of an efficient synthesis pathway for generating structurally diverse, phenylsulfone-activated aminothiol building blocks that could be incorporated into peptides by SPPS. We tested our strategy in a proof-of-concept study, making a library of 2,688 compounds that culminated in the discovery of nanomolar inhibitors of two proteases, including some that displayed good membrane permeability.

The high structural diversity directly translated into the isolation of better binders from random libraries, with the best thrombin inhibitor ($K_i = 10.9$ nM) identified being 4-fold more potent than those isolated previously from other libraries (best $K_i = 42$ nM),[10] and the library is the first one in which we found inhibitors for plasma kallikrein. The importance of the expanded structural diversity is supported by the X-ray structural analysis of two of the macrocycles bound to thrombin, which showed that the specific size and shape of the aminothiol groups are needed to optimally accommodate the neighboring macrocycle building blocks for forming optimal contacts with the thrombin's active site. Importantly, several of the identified binders showed good membrane permeability in artificial membranes (PAMPA) and in live cells (CAPA), which was likely a result of keeping the size of the macrocycles rather small, close to 500 Da. Such a result would likely not have been possible if

the library diversity had been increased through adding an additional amino acid position for diversification due to the increase in size and the addition of an extra amide bond, the latter typically much reducing the membrane permeability due to its polar nature.

While we synthesized a rather small library of only 2,688 macrocycles in this study, much larger libraries can be generated with the applied methods. We recently modified a commercial 96-well plate peptide synthesizer to make peptides in 384-well plates.[28] The instrument holds four 384-well reactor plates allowing us to synthesize 1,536 peptides in one run and in less than three days. As mentioned above, we also have preliminary data that shows that the cyclization reactions and procedures established herein for 384-well plates are applicable to 1,536-well plates, allowing the synthesis of ten-thousands or even over hundred-thousand macrocyclic compounds with the described expanded skeletal diversity. Such larger libraries may be beneficial for developing macrocycle-based ligands to more challenging targets than the herein-applied proteases, such as protein-protein interactions.

We had used in this work seven different aminothiol building blocks, but this number may be grown further by synthesizing additional ones as for example aminothiol groups containing side chains. In addition, the *N*-terminal thiol-containing building block Mpa, which was kept constant in this work, may be diversified in the future too. Building blocks analogous to Mpa, may be prepared by reacting halogenated carboxylic acids with trityl mercaptan. Together with a set of diverse aminothiols, even greater macrocycles diversities for generating membrane permeable ligands could be generated. Finally, the macrocycle backbones may additionally be varied by introducing building blocks that add ester or thioamide bonds, that are also suited to improve membrane permeability.[29,30] We are optimistic that the herein developed new approach for synthesizing small and structurally highly diverse, membrane permeable macrocycles will be broadly applicable to many proteins and may offer a solution to develop therapeutics to some of the most challenging intracellular disease targets.

Abbreviations

AMC	7-amino-4-methyl coumarin
BDT	1,4-butanedithiol
CA	chloroalkane
CAPA	chloroalkane penetration assay
Mea	2-mercaptoethylamine
P_e	effective permeability coefficient
FXI	factor XI
FXII	factor XII
CP_{50}	half-maximal cellular penetration
HBD	hydrogen bond donors
HTS	high-throughput screening
KLK5	tissue kallikrein 5
Mpa	3-mercaptopropionic acid
PAMPA	parallel artificial membrane permeability assay
PK	plasma kallikrein
PP	polypropylene
PS	polystyrene
RVC	rotary vacuum concentration
RP	reversed-phase
Ro5	rule of five
SPPS	solid-phase peptide synthesis
β-ME	β-mercaptoethanol

Author Contributions

A.L.N, G.K.M and C.H. conceptualized the study; G.K.M innovated the synthetic strategy for synthesizing aminothiol reagents; Z.B. established the synthesis of aminothiols and testing in SPPS; A.L.N., Z.B., A.Z., and M.S. performed the chemical synthesis; A.L.N. A.Z., M.M., X.J. and E.J.W. conducted the biochemical experiments; M.C., L.C. and A.A. generated the co-crystals and solved the X-ray structures; C.R.O.B and K.S. performed the CAPA studies, A.L.N. and C.H. wrote the original draft of the manuscript, and all authors reviewed the final version.

Acknowledgements

We thank Peter M. F. Panzar, Elena Covato, Bo Fu, and Violaine Chirat for technical assistance and building block synthesis. Prof. Joshua Kritzer is acknowledged for providing the HaloTag-expressing HeLa cell line. Kaycie Butler and Mark D. Nolan are recognized for proofreading the manuscript. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 101020521-TARGET. This work was also supported by the Swiss National Science Foundation (grant 192368) and the BactiVax European Training Network (ETN), funded under the H2020-MSCA-ITN-2019 call (grant agreement ID: 860325). Open Access funding provided by École Polytechnique Fédérale de Lausanne.

Conflict of Interest

The authors declare the following competing financial interest(s): A.L.N., Z.B., G.K.M., M.S., M.M., and C.H. are inventors of a patent that covers aspects of this work. C.H. is a co-founder of Orbis Medicines. M.S. is an employee of Orbis Medicines. The remaining authors declare no competing interests.

Data Availability Statement

Supporting Information accompanies this paper and includes: Supporting results, Figures, and tables as well as experimental procedures, compound characterization data, and copies of HPLC chromatograms and NMR spectra. Raw data of results shown in figures is provided in an Excel

file. X-ray diffraction data, coordinates, and structure factors for the X-ray crystal structures are deposited with the PDB under the accession numbers 8ASF and 8ASE.

Keywords: cell permeability · cyclization · high-throughput screening · macrocycle · protease inhibitor

[1] E. M. Driggers, S. P. Hale, J. Lee, N. K. Terrett, *Nat. Rev. Drug Discovery* **2008**, *7*, 608–624.
[2] D. Garcia Jimenez, V. Poongavanam, J. Kihlberg, *J. Med. Chem.* **2023**, *66*, 5377–5396.
[3] C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, *Adv. Drug Delivery Rev.* **2001**, *46*, 3–26.
[4] B. Over, P. Matsson, C. Tyrchan, P. Artursson, B. C. Doak, M. A. Foley, C. Hilgendorf, S. E. Johnston, M. D. Lee, R. J. Lewis, P. McCarren, G. Muncipinto, U. Norinder, M. W. D. Perry, J. R. Duvall, J. Kihlberg, *Nat. Chem. Biol.* **2016**, *12*, 1065–1074.
[5] M. R. Naylor, A. T. Bockus, M.-J. Blanco, R. S. Lokey, *Curr. Opin. Chem. Biol.* **2017**, *38*, 141–147.
[6] X. Ji, A. L. Nielsen, C. Heinis, *Angew. Chem. Int. Ed.* **2023**, *63*, e202308251, DOI 10.1002/anie.202308251.
[7] A. Ohta, M. Tanada, S. Shinohara, Y. Morita, K. Nakano, Y. Yamagishi, R. Takano, S. Kariyuki, T. Iida, A. Matsuo, K. Ozeki, T. Emura, Y. Sakurai, K. Takano, A. Higashida, M. Kojima, T. Muraoka, R. Takeyama, T. Kato, K. Kimura, K. Ogawa, K. Ohara, S. Tanaka, Y. Kikuchi, N. Hisada, R. Hayashi, Y. Nishimura, K. Nomura, T. Tachibana, M. Irie, H. Kawada, T. Torizawa, N. Murao, T. Kotake, M. Tanaka, S. Ishikawa, T. Miyake, M. Tamiya, M. Arai, A. Chiyoda, S. Akai, H. Sase, S. Kuramoto, T. Ito, T. Shiraishi, T. Kojima, H. Iikura, *J. Am. Chem. Soc.* **2023**, *145*, 24035–24051.
[8] M. Tanada, M. Tamiya, A. Matsuo, A. Chiyoda, K. Takano, T. Ito, M. Irie, T. Kotake, R. Takeyama, H. Kawada, R. Hayashi, S. Ishikawa, K. Nomura, N. Furuichi, Y. Morita, M. Kage, S. Hashimoto, K. Nii, H. Sase, K. Ohara, A. Ohta, S. Kuramoto, Y. Nishimura, H. Iikura, T. Shiraishi, *J. Am. Chem. Soc.* **2023**, *145*, 16610–16620.
[9] L. Plais, J. Scheuermann, *RSC Chem Biol* **2022**, *3*, 7–17.
[10] S. S. Kale, M. Bergeron-Brlek, Y. Wu, M. G. Kumar, M. V. Pham, J. Bortoli, J. Vesin, X.-D. Kong, J. F. Machado, K. Deyle, P. Gonschorek, G. Turcatti, L. Cendron, A. Angelini, C. Heinis, *Sci. Adv.* **2019**, *5*, eaaw2851.
[11] S. Habeshian, M. L. Merz, G. Sangouard, G. K. Mothukuri, M. Schüttel, Z. Bognár, C. Díaz-Perlas, J. Vesin, J. Bortoli Chapalay, G. Turcatti, L. Cendron, A. Angelini, C. Heinis, *Nat. Commun.* **2022**, *13*, 3823.
[12] M. L. Merz, S. Habeshian, B. Li, J.-A. G. L. David, A. L. Nielsen, X. Ji, K. Il Khwildy, M. M. Duany Benitez, P. Phothirath, C. Heinis, *Nat. Chem. Biol.* **2024**, *20*, 624–635 .
[13] S. Habeshian, G. A. Sable, M. Schüttel, M. L. Merz, C. Heinis, *ACS Chem. Biol.* **2022**, *17*, 181–186.
[14] Z. Bognar, G. K. Mothukuri, A. L. Nielsen, M. L. Merz, P. M. F. Pânzar, C. Heinis, *Org. Biomol. Chem.* **2022**, *20*, 5699–5703.
[15] M. D. Nolan, M. Schüttel, E. M. Scanlan, A. L. Nielsen, *Pept. Sci.* **2024**, *116*, e24310.
[16] D. P. Fairlie, A. Dantas de Araujo, *Pept. Sci.* **2016**, *106*, 843–852.
[17] R. A. Roof, K. Sobczyk-Kojiro, A. J. Turbiak, D. L. Roman, I. D. Pogozheva, L. L. Blazer, R. R. Neubig, H. I. Mosberg, *Chem. Biol. Drug Des.* **2008**, *72*, 111–119.
[18] L. E. J. Smeenk, D. Timmers-Parohi, J. J. Benschop, W. C. Puijk, H. Hiemstra, J. H. van Maarseveen, P. Timmerman, *ChemBioChem* **2015**, *16*, 91–99.
[19] F. F. Faucher, D. Abegg, P. Ipock, A. Adibekian, S. Lovell, M. Bogyo, *Isr. J. Chem.* **2023**, *63*, DOI 10.1002/ijch.202300020.
[20] N. S. Zefirov, N. V. Zyk, E. K. Beloglazkina, A. G. Kutateladze, *Sulfur Rep.* **1993**, *14*, 223–240.
[21] L. Lin, M. E. Olson, T. Sugane, L. D. Turner, M. A. Tararina, A. L. Nielsen, E. K. Kurbanov, S. Pellett, E. A. Johnson, S. M. Cohen, K. N. Allen, K. D. Janda, *J. Med. Chem.* **2020**, *63*, 11100–11120.
[22] E. A. Stura, L. Le Roux, K. Guitot, S. Garcia, S. Bregant, F. Beau, L. Vera, G. Collet, D. Ptchelkine, H. Bakirci, V. Dive, *J. Biol. Chem.* **2012**, *287*, 33607–33614.
[23] M. Kansy, F. Senner, K. Gubernator, *J. Med. Chem.* **1998**, *41*, 1007–1010.
[24] D. Lee, S. Lee, J. Choi, Y.-K. Song, M. J. Kim, D.-S. Shin, M. A. Bae, Y.-C. Kim, C.-J. Park, K.-R. Lee, J.-H. Choi, J. Seo, *J. Med. Chem.* **2021**, *64*, 8272–8286.
[25] L. Peraro, Z. Zou, K. M. Makwana, A. E. Cummings, H. L. Ball, H. Yu, Y.-S. Lin, B. Levine, J. A. Kritzer, *J. Am. Chem. Soc.* **2017**, *139*, 7792–7802.
[26] L. Peraro, K. L. Deprey, M. K. Moser, Z. Zou, H. L. Ball, B. Levine, J. A. Kritzer, *J. Am. Chem. Soc.* **2018**, *140*, 11360–11369.
[27] G. V. Los, L. P. Encell, M. G. McDougall, D. D. Hartzell, N. Karassina, C. Zimprich, M. G. Wood, R. Learish, R. F. Ohana, M. Urh, D. Simpson, J. Mendez, K. Zimmerman, P. Otto, G. Vidugiris, J. Zhu, A. Darzins, D. H. Klaubert, R. F. Bulleit, K. V. Wood, *ACS Chem. Biol.* **2008**, *3*, 373–382.
[28] M. Schüttel, E. Will, G. Sangouard, A. Zarda, S. Habeshian, A. L. Nielsen, C. Heinis, *J. Pept. Sci.* **2024**, *4*, e3555.
[29] P. Ghosh, N. Raj, H. Verma, M. Patel, S. Chakraborti, B. Khatri, C. M. Doreswamy, S. R. Anandakumar, S. Seekallu, M. B. Dinesh, G. Jadhav, P. N. Yadav, J. Chatterjee, *Nat. Commun.* **2023**, *14*, 6050.
[30] Y. Hosono, S. Uchida, M. Shinkai, C. E. Townsend, C. N. Kelly, M. R. Naylor, H.-W. Lee, K. Kanamitsu, M. Ishii, R. Ueki, T. Ueda, K. Takeuchi, M. Sugita, Y. Akiyama, S. R. Lokey, J. Morimoto, S. Sando, *Nat. Commun.* **2023**, *14*, 1416.
[31] T. Sander, J. Freyss, M. Von Korff, C. Rufener, *J. Chem. Inf. Model.* **2015**, *55*, 460–473.

Manuscript received: January 5, 2024
Accepted manuscript online: April 11, 2024
Version of record online: May 24, 2024