

Cite this: *Chem. Sci.*, 2020, **11**, 7858

All publication charges for this article have been paid for by the Royal Society of Chemistry

Macrocycle synthesis strategy based on step-wise “adding and reacting” three components enables screening of large combinatorial libraries†

Ganesh K. Mothukuri,^a Sangram S. Kale,^a Carl L. Stenbratt,^a Alessandro Zorzi,^a Jonathan Vesin,^b Julien Bortoli Chapalay,^b Kaycie Deyle,^a Gerardo Turcatti,^b Laura Cendron,^c Alessandro Angelini^{de} and Christian Heinis^{de}*

Macrocycles provide an attractive modality for drug development, but generating ligands for new targets is hampered by the limited availability of large macrocycle libraries. We have established a solution-phase macrocycle synthesis strategy in which three building blocks are coupled sequentially in efficient alkylation reactions that eliminate the need for product purification. We demonstrate the power of the approach by combinatorially reacting 15 bromoacetamide-activated tripeptides, 42 amines, and 6 bis-electrophile cyclization linkers to generate a 3780-compound library with minimal effort. Screening against thrombin yielded a potent and selective inhibitor ($K_i = 4.2 \pm 0.8$ nM) that efficiently blocked blood coagulation in human plasma. Structure–activity relationship and X-ray crystallography analysis revealed that two of the three building blocks acted synergistically and underscored the importance of combinatorial screening in macrocycle development. The three-component library synthesis approach is general and offers a promising avenue to generate macrocycle ligands to other targets.

Received 3rd April 2020

Accepted 25th June 2020

DOI: 10.1039/d0sc01944e

rsc.li/chemical-science

Introduction

Macrocycles, molecules comprised of a ring of 12 or more atoms, have the ability to bind difficult targets considered ‘undruggable’ by traditional small molecules therapeutics.^{1,2} Many important drugs belong to this structural class, such as the immunosuppressant cyclosporine or the antibiotic vancomycin.³ Innovations for designing macrocycles based on protein epitopes^{4,5} or screening large libraries of cyclic peptides by display techniques^{6,7} have enabled the development of peptide-based macrocycles to virtually any target and have led to a number of cyclic peptides that are currently evaluated in clinical studies, most of them targeting extracellular proteins. An important goal that remains a great challenge is the development of macrocycles that are orally available and/or membrane permeable. The generation of such molecules is

ambitious, as they must be small (ideally <1 kDa) and have a small polar surface area to passively cross cell membranes while still maintaining a high affinity for challenging targets.^{8,9} Strategies based on diversity-oriented synthesis allow for generating macrocycles of high structural diversity, but making libraries comprising very large numbers of compounds with such techniques remains work intensive.^{10,11} Solid phase-based strategies were used to synthesize larger macrocycle libraries with smaller effort.^{12–16} Techniques for encoding macrocycles by DNA-templated chemistry^{17,18} or split-and-mix DNA-encoding principles^{19,20} are promising but their application to macrocycles is not trivial as side products can accumulate in the sequential build-up of macrocycles.

Seeking for a high-throughput strategy to generate and screen large libraries of sub-kilodalton macrocycles in micro-well plates, we recently presented an approach based on the combinatorial mixing of m short peptides with n cyclization linkers to generate $m \times n$ macrocycles.²¹ A huge limitation of this technique is that each macrocycle consisted of only two building blocks, meaning that large numbers of m short peptides and n cyclization reagents needed to be prepared to generate large libraries of $m \times n$ macrocycles. The ability to add an additional building block while retaining the efficiency and purity would be enormously beneficial, as with 100 different molecules per building block, one would generate one million ($m \times n \times o$) macrocycles with three building blocks *versus* only 10 000 when two are combined. However, a challenge was to find a chemical strategy to form macrocycles based on three

^aInstitute of Chemical Sciences and Engineering, School of Basic Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland. E-mail: christian.heinis@epfl.ch

^bBiomolecular Screening Facility, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

^cDepartment of Biology, University of Padova, 35131 Padova, Italy

^dDepartment of Molecular Sciences and Nanosystems, Ca' Foscari University of Venice, Via Torino 155, Venezia Mestre, Venice 30172, Italy

^eEuropean Centre for Living Technology (ECLT), Ca' Bottacin, Dorsoduro 3911, Calle Crosera, Venice 30124, Italy

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d0sc01944e

building blocks that can be mixed combinatorially and that form as main products the desired macrocycles.

Herein, we have developed a strategy for macrocycle synthesis that addresses both of these concerns simultaneously—being based on three building blocks and yielding macrocycles of reasonably high purity. This strategy, shown in Fig. 1a, consists of m short peptides functionalized at the N-terminus with bromoacetamide that react with n primary amines added in a combinatorial fashion; the products are subsequently macrocyclized by the combinatorial addition of o bis-electrophile linkers. The first reaction step is based on the sub-monomer strategy used for the generation of peptoid libraries.^{22–24} The second step is based on a thiol-to-amine cyclization reaction reported recently.^{21,25} Important in this strategy was that the three consecutive coupling reactions were efficient and clean and did not interfere with each other to yield macrocycles as main products with minimal side products. This again omitted product purification and enabled the generation and screening of large libraries.

Results and discussion

We tested the feasibility of the “adding and reacting” step-wise macrocycle assembly strategy using (i) the bromoacetamide (BrAc)-functionalized peptide BrAc-Gly-Cys(SO₃H)-R (R = Gly-Trp-NH₂ appendix for better UV absorption), (ii) benzylamine as a primary amine, and (iii) 1,3-bis(bromomethyl)pyridine (**1**) as a cyclization linker, analyzing each step by RP-HPLC and mass spectrometry (Fig. 1b). For the first reaction, we mixed equal volumes (10 μ l) of BrAc-Gly-Cys(SO₃H)-R in H₂O and a 16-fold molar excess of benzylamine in DMSO and analyzed the reaction after incubation for two hours at 37 °C. At final concentrations of 5 mM peptide and 80 mM amine, essentially all of the peptide was alkylated for an excellent first-step yield. The cysteine protecting group was then removed by mixing 2 μ l of the *N*-alkylated peptide (still in the reaction mix containing excess amine; 5 mM peptide stock in 50% H₂O and 50% DMSO, 104 μ M final concentration) with 94 μ l of aqueous buffer (60 mM NH₄HCO₃ buffer, pH 8) containing TCEP (425 μ M, 4-

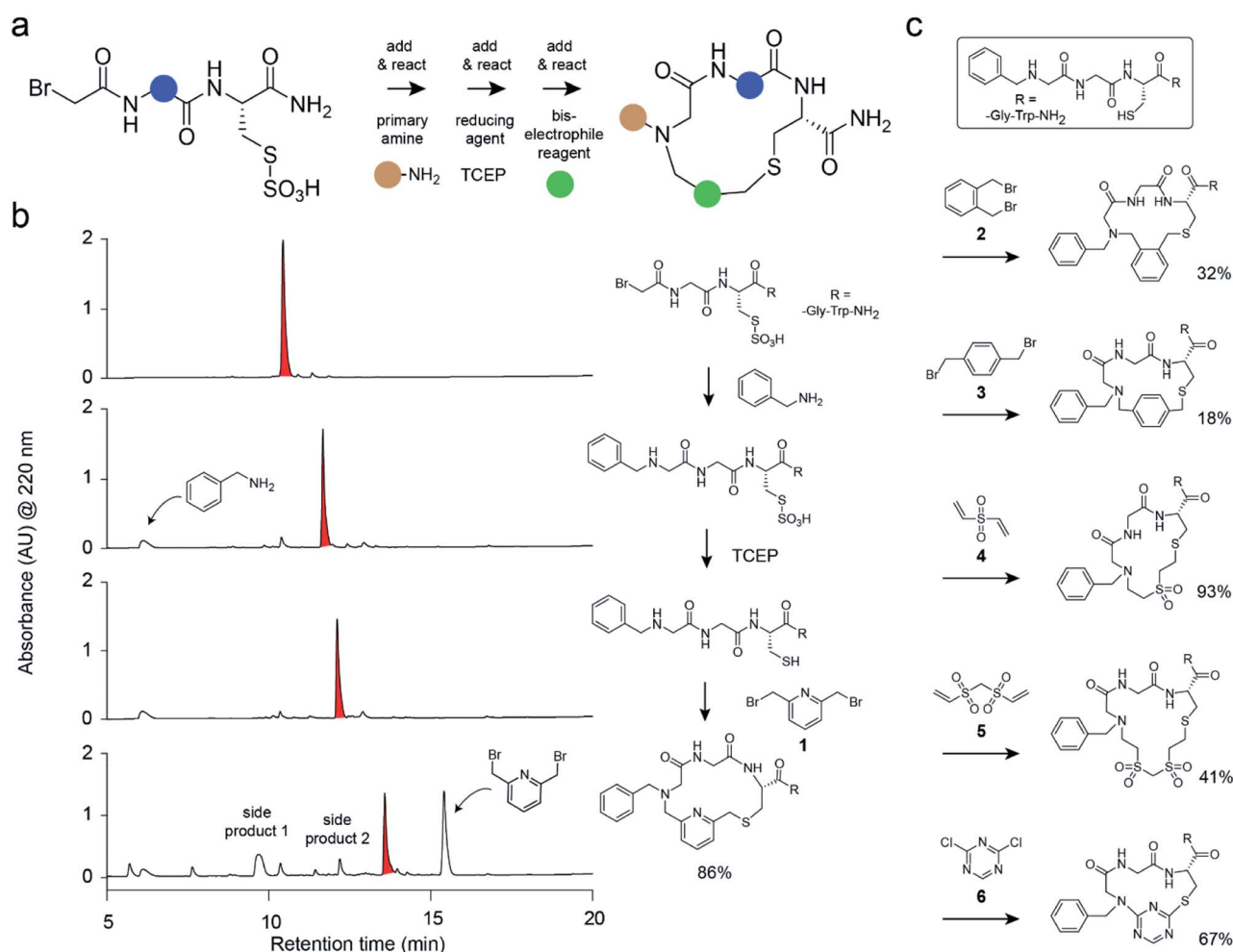


Fig. 1 Combinatorial macrocycle synthesis strategy. (a) Outline of the three-component “add & react” macrocycle synthesis strategy. The colored balls indicate the variable parts in the three components. (b) Reaction steps and intermediate products shown for a macrocycle based on the peptide component BrAc-Gly-Cys(SO₃H)-Gly-Trp-NH₂, the amine component benzylamine, and the cyclization linker **1**. Analytical RP-HPLC chromatograms are shown for the reactions. At none of the indicated reaction steps were the products purified. The peaks of the desired products are colored in red. (c) Cyclization of the peptide-amine conjugate benzyl-NH-Gly-Gly-Cys-Gly-Trp-NH₂ with different bis-electrophile linker reagents. Yields were calculated based on the area of HPLC peaks and are indicated.

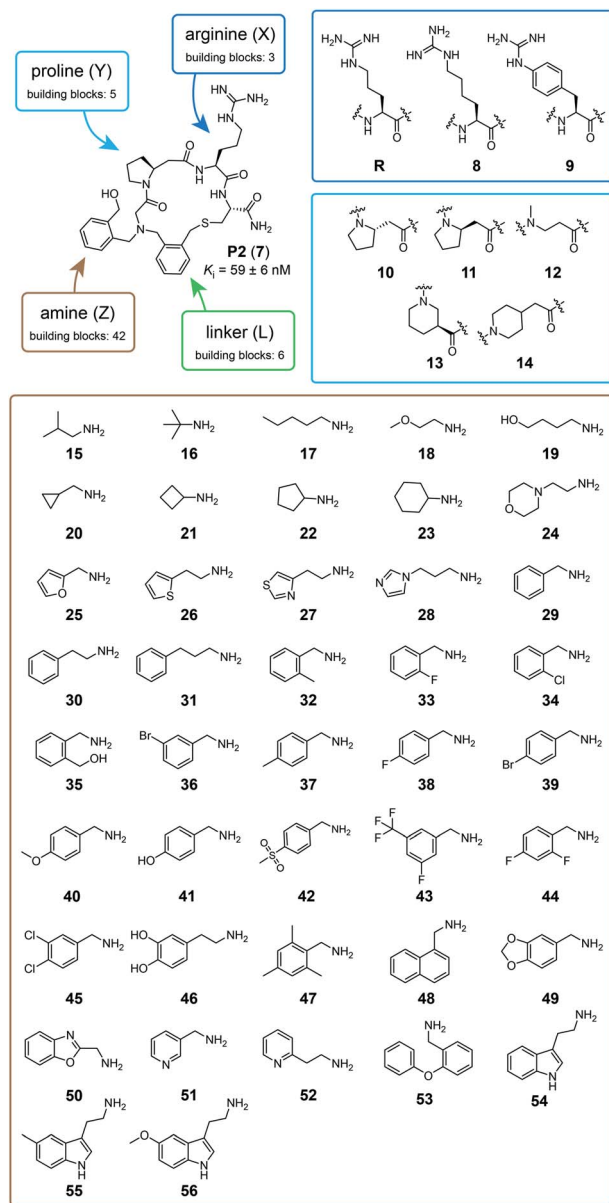


Fig. 2 Macrocyclization library design and building blocks. The chemical structure of the previously developed macrocycle P2 (7) in the top left corner and the three components that are combinatorially varied in the library are indicated in blue (peptide), brown (amine), and green (linker). The eight amino acids used to synthesize the 15 peptide components are framed in dark and light blue. The 42 amine components are shown in a brown frame. The six linker components are shown in Fig. 1.

fold molar excess over peptide) and incubation for one hour at 37 °C. The *N*-alkylated peptide was subsequently cyclized by adding 4 μ l of 20 mM bis-electrophile reagent 1 in DMSO to the 96 μ l of the deprotection reaction mixture containing unprotected peptide, for final concentrations of 100 μ M peptide and 800 μ M cyclization reagent. The desired macrocyclic product was formed with 86% overall yield, as determined by quantifying the area of the analytical HPLC peaks. Small quantities of side products were identified as peptide–1-TCEP conjugate

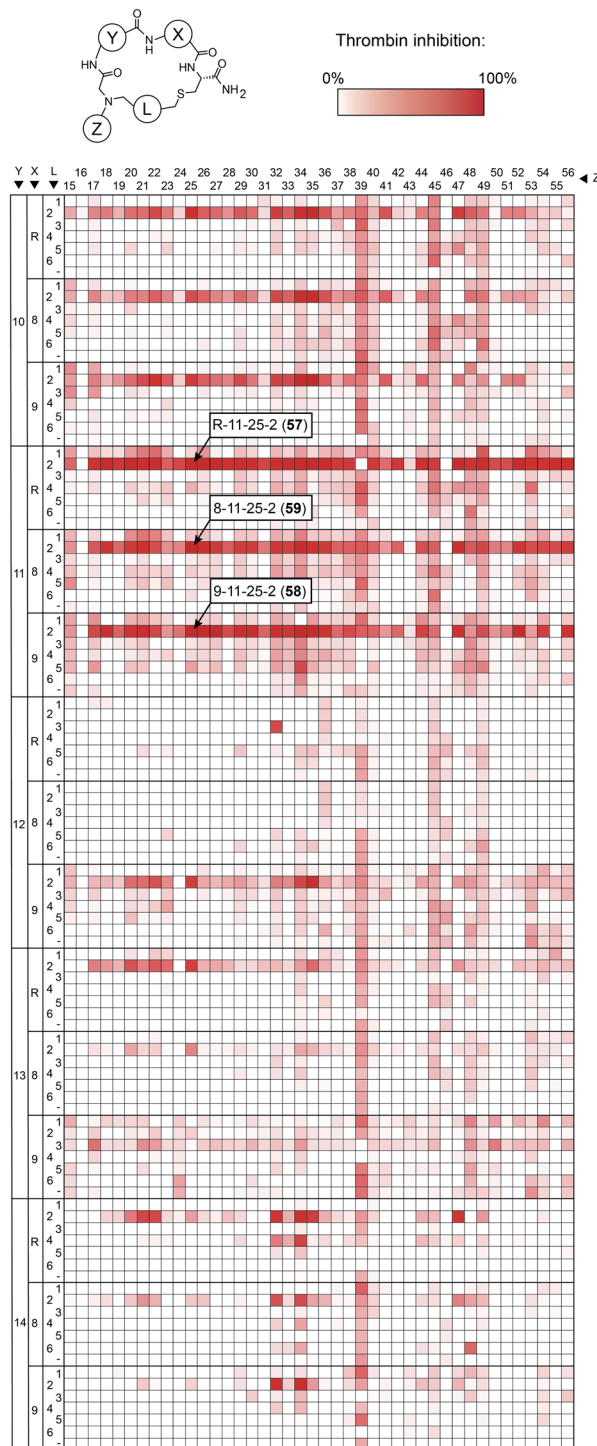


Fig. 3 Thrombin inhibition screen with 3780 macrocycles. Crude macrocycle reactions (13 μ M macrocycle in case of 100% alkylation and cyclization yields) were incubated with thrombin, and the residual protease activity was measured with a fluorogenic substrate. The three macrocycle reactions showing the strongest inhibition are indicated.

(Fig. S1†). The thiol-to-amine macrocyclization of the same *N*-alkylated peptide with the bis-electrophile reagents 2 to 6 gave macrocycle yields between 18 and 93% (based on HPLC peaks; Fig. 1c and S2†), with the low yielding reaction coming from the most constrained linker.



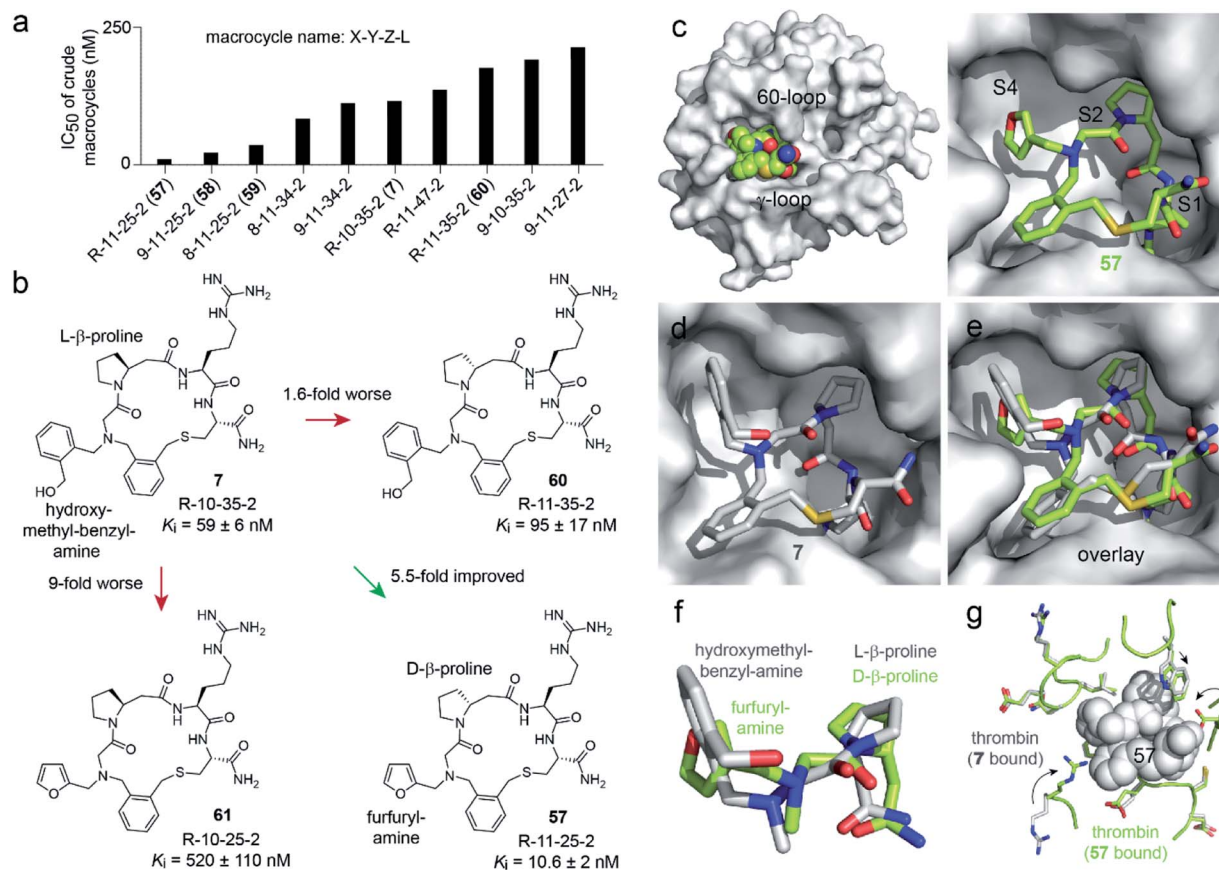


Fig. 4 Structure–activity relationship. (a) The thrombin inhibition assay was repeated for the top 48 hits at nine 2-fold dilutions of the crude macrocyclization reactions. The 10 hits with the lowest IC₅₀ values are shown in panel (a), and the IC₅₀ values of all 48 reactions are shown in Fig. S6†. (b) Macrocycle variants of 7 containing only one of the two building block modifications present in the hit compound 57. (c) X-ray structure of 57 bound to human thrombin at a resolution of 2.32 Å (PDB 6T7H). (d) Previously solved X-ray structure of 7 bound to human thrombin (PDB 6GWE).²¹ (e) Overlay of 7 (grey) and 57 (green) with the thrombin structure of 6T7H shown. (f) Zoom in of the amine substituents and β-homoprolines shown in panel (e). (g) Comparison of the thrombin backbone and side chains of 6T7H (green) and 6GWE (grey) at the macrocycle-binding region. Large conformational changes of side chains are indicated by arrows.

We next applied the strategy to synthesize a combinatorial library of 3780 macrocycles, which we screened against the coagulation protease thrombin, a validated target of anti-coagulation therapy.²⁶ For this library, we sought to improve the performance of the macrocyclic thrombin inhibitor P2 (7; K_i = 59 ± 6 nM) that was identified previously using the $m \times n$ library discussed above (Fig. 2).²¹ Therefore, the included building blocks were variants or homologs of the structural features of 7, as shown in Fig. 2.

To make the short peptide unit, we synthesized 15 bromoacetamide-functionalized peptides of the form BrAc-Y-X-Cys(SO₃H)-NH₂ (Fig. S3†) on solid phase and purified them by HPLC (Fig. S4†). Next, we combinatorially reacted the 15 peptides with 42 amines and cyclized the products with 6 bis-electrophile reagents, as described above for the model peptide. We analyzed 24 of the reactions by LC-MS, which showed that most of them contained the desired macrocycle as the main product, with the exception of the reactions involving linkers 3 and 6, which yielded only small macrocycle quantities or even no macrocycle (Fig. S5†). We screened the 3780 reactions containing the crude mixture of macrocycles against

human thrombin using a fluorogenic protease substrate (Fig. 3). The final concentration of macrocycle in the assay was 13 μM in case of 100% yield for all synthesis steps. Fig. 3 indicates the extent of thrombin inhibition, with red indicating 100% inhibition and white indicating 0% inhibition.

Peptides containing D-β-homoproline and cyclized with linker 2 showed the strongest thrombin inhibition, with no real preference for the arginine analogs. The cyclization linker 2 is the same one as in the parental macrocycle 7, confirming that this linker is most suited, likely due to the necessity for maintaining a similar backbone size as the original compound.

In contrast, the amino acid 11 is D-β-homoproline, which has an inverted stereocenter from the L-β-homoproline (10) in the parent compound 7. The characterization of the top hits is described in detail in the ESI Results.† The best inhibitor, the macrocycle 58 (K_i = 4.2 ± 0.8 nM) was 14-fold better than the parental macrocycle 7.

We next assessed which of the two substitutions, L-β-homoproline (10) to D-β-homoproline (11), or hydroxymethyl-benzyl-amine (35) to furfurylamine (25) contributed more to the activity improvement, and how much (Fig. 4b and S8†). Both



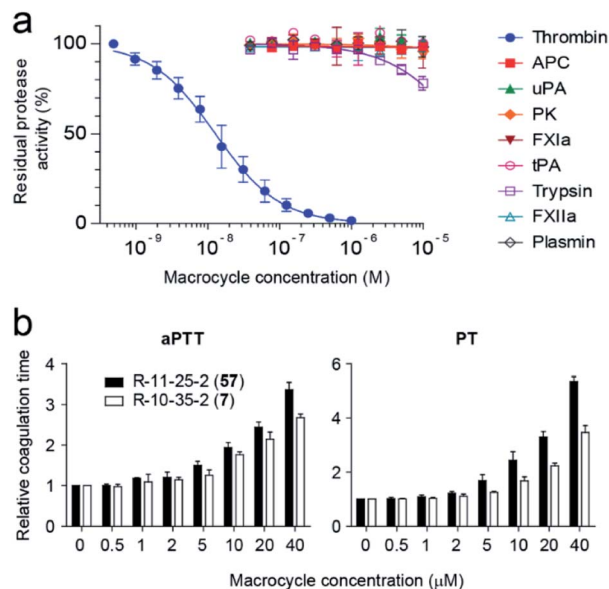


Fig. 5 Characterization of macrocycle hits. (a) Specificity profiling of macrocycle 57. (b) Prolongation of the coagulation time of human plasma in the presence of different concentrations of macrocycle 7 and 57. Coagulation was triggered to activate the intrinsic coagulation pathway (aPTT) and the extrinsic coagulation pathway (PT).

individual modifications made the K_i worse, the first was 1.6-fold and the second was 9-fold weaker than the parent. This result was surprising given that the two combined building blocks, present in 57, increased the activity 5.5-fold. It suggested that the two structural changes to the parent compound 7 acted synergistically, and that macrocycle 57 would not have been identified if the two building blocks had not been altered in a combinatorial fashion. This in turn validated the strategy of synthesizing and screening many macrocycles and testing many building block combinations.

To understand why the two building blocks α - β -homoproline (11) and furfurylamine (25) acted synergistically, we crystallized macrocycle 57 in complex with human thrombin and solved the structure at a 2.32 Å resolution (PDB 6T7H; Fig. 4c and S9†). The macrocycle bound to the same sub-sites of thrombin as the parental macrocycle 7 (PDB 6GWE; Fig. 4d),²¹ but the backbones followed rather different trajectories (RMSD = 0.70; Fig. 4e and f). Most likely, the two β -homoproline stereoisomers in 7 and 57 imposed a different conformation onto the backbones, which altered the positioning of the *N*-substituents. As a result, the furfuryl substituent binds optimally to the S4 pocket only if the macrocycle backbone contains α - β -homoproline, and conversely, the hydroxymethyl-benzyl substituent only binds well in the presence of the α - β -homoproline (a detailed structure discussion is provided in the ESI†).

Inhibition studies with macrocycle 57 and a panel of nine human proteases, which are structurally and functionally most similar to thrombin, revealed a high target selectivity (Fig. 5a). We also assessed the ability of macrocycle 57 to block coagulation in human blood by measuring the activated partial thromboplastin time (aPTT) and prothrombin time (PT). aPTT

and PT measure the time until coagulation upon initiation of the intrinsic and extrinsic pathways, respectively, and results are reported in terms of the concentration of the tested compound required for a 2-fold increase in coagulation time ($EC_{2\times}$). Macrocycle 57 prolonged aPTT and PT in human plasma with 1.4- and 1.9-fold better $EC_{2\times}$ values than its parent 7 (Fig. 5b).

In a final experiment, we tested if the “add & react” macrocycle synthesis strategy works with crude instead of HPLC-purified peptides, as crude peptides can be produced with a smaller effort and thus in larger numbers. The alkylation and macrocyclization reactions worked nearly as good as with purified peptides (Fig. S10†). We estimate that the approach can be used to generate and screen libraries of ten- to hundred-thousands of macrocycles in an academic setting and probably much larger libraries with an industrial infrastructure.

Conclusions

In summary, we present a new strategy for efficiently synthesizing combinatorial libraries of macrocycles based on three components that are sequentially added and reacted. Importantly, all chemical ligation reactions are efficient and selective so that the main products are the macrocycles, allowing for screening without purification. It is this omission of a purification step that enables the screening of large numbers of macrocycles, as demonstrated in this work by a screen of 3780 compounds with a minimal synthetic effort. Importantly, we show that the best inhibitors found from this library are based on two key synergistic modifications that do not improve the affinity if introduced individually. This observation underscores the importance of screening large libraries of macrocycles in which multiple building blocks are combinatorially varied. The three-component macrocycle synthesis approach is general and offers a promising avenue to generate macrocycle ligands to other targets.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by the Swiss National Science Foundation (Grant 157842 and the NCCR Chemical Biology). The authors would like to thank Diamond Light Source for beam time (proposal MX21741) and the staff of beamline I03 for assistance with crystal testing and data collection.

Notes and references

- 1 E. M. Driggers, S. P. Hale, J. Lee and N. K. Terrett, *Nat. Rev. Drug Discovery*, 2008, 7, 608–624.
- 2 E. Valeur, S. M. Guéret, H. Adihou, R. Gopalakrishnan, M. Lemurell, H. Waldmann, T. N. Grossmann and A. T. Plowright, *Angew. Chem., Int. Ed.*, 2017, 56, 10294–10323.



- 3 F. Giordanetto and J. Kihlberg, *J. Med. Chem.*, 2014, **57**, 278–295.
- 4 J. A. Robinson, S. DeMarco, F. Gombert, K. Moehle and D. Obrecht, *Drug Discovery Today*, 2008, **13**, 944–951.
- 5 M. Pelay-Gimeno, A. Glas, O. Koch and T. N. Grossmann, *Angew. Chem., Int. Ed.*, 2015, **54**, 8896–8927.
- 6 K. Deyle, X.-D. Kong and C. Heinis, *Acc. Chem. Res.*, 2017, **50**, 1866–1874.
- 7 Y. Huang, M. M. Wiedmann and H. Suga, *Chem. Rev.*, 2019, **119**, 10360–10391.
- 8 E. A. Villar, D. Beglov, S. Chennamadhavuni, J. A. Porco, D. Kozakov, S. Vajda and A. Whitty, *Nat. Chem. Biol.*, 2014, **10**, 723–731.
- 9 M. R. Naylor, A. T. Bockus, M. J. Blanco and R. S. Lokey, *Curr. Opin. Chem. Biol.*, 2017, **38**, 141–147.
- 10 P. Ermert, *CHIMIA International Journal for Chemistry*, 2017, **71**, 678–702.
- 11 K. T. Mortensen, T. J. Osberger, T. A. King, H. F. Sore and D. R. Spring, *Chem. Rev.*, 2019, 9b00084.
- 12 M. E. Fitzgerald, C. A. Mulrooney, J. R. Duvall, J. Wei, B.-C. Suh, L. B. Akella, A. Vrcic and L. A. Marcaurelle, *ACS Comb. Sci.*, 2012, **14**, 89–96.
- 13 L. A. Marcaurelle, E. Comer, S. Dandapani, J. R. Duvall, B. Gerard, S. Kesavan, M. D. Lee, H. Liu, J. T. Lowe, J.-C. Marie, C. A. Mulrooney, B. A. Pandya, A. Rowley, T. D. Ryba, B.-C. Suh, J. Wei, D. W. Young, L. B. Akella, N. T. Ross, Y.-L. Zhang, D. M. Fass, S. A. Reis, W.-N. Zhao, S. J. Haggarty, M. Palmer and M. A. Foley, *J. Am. Chem. Soc.*, 2010, **132**, 16962–16976.
- 14 J. Morimoto and T. Kodadek, *Mol. Biosyst.*, 2015, **11**, 2770–2779.
- 15 A. P. Treder, J. L. Hickey, M.-C. J. Tremblay, S. Zaretsky, C. C. G. Scully, J. Mancuso, A. Doucet, A. K. Yudin and E. Marsault, *Chem.–Eur. J.*, 2015, **21**, 9249–9255.
- 16 B. K. W. Chung, C. J. White and A. K. Yudin, *Nat. Protoc.*, 2017, **12**, 1277–1287.
- 17 Z. J. Gartner, B. N. Tse, R. Grubina, J. B. Doyon, T. M. Snyder and D. R. Liu, *Science*, 2004, **305**, 1601–1605.
- 18 D. L. Usanov, A. I. Chan, J. P. Maianti and D. R. Liu, *Nat. Chem.*, 2018, **10**, 704–714.
- 19 Y. Li, R. de Luca, S. Cazzamalli, F. Pretto, D. Bajic, J. Scheuermann and D. Neri, *Nat. Chem.*, 2018, **10**, 1–8.
- 20 C. J. Stress, B. Sauter, L. A. Schneider, T. Sharpe and D. Gillingham, *Angew. Chem., Int. Ed.*, 2019, **58**, 9570–9574.
- 21 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 and C. Heinis, *Sci. Adv.*, 2019, **5**, eaaw2851.
- 22 R. N. Zuckermann, J. M. Kerr, S. B. H. Kent and W. H. Moos, *J. Am. Chem. Soc.*, 1992, **114**, 10646–10647.
- 23 P. G. Alluri, M. M. Reddy, K. Bachhawat-Sikder, H. J. Olivos and T. Kodadek, *J. Am. Chem. Soc.*, 2003, **125**, 13995–14004.
- 24 A. M. Webster and S. L. Cobb, *Chem.–Eur. J.*, 2018, **24**, 7560–7573.
- 25 N. Assem, D. J. Ferreira, D. W. Wolan and P. E. Dawson, *Angew. Chem., Int. Ed.*, 2015, **54**, 8665–8668.
- 26 K. A. Arsenault, J. Hirsh, R. P. Whitlock and J. W. Eikelboom, *Nat. Rev. Cardiol.*, 2012, **9**, 402–414.

