



Published in final edited form as:

Angew Chem Int Ed Engl. 2022 April 25; 61(18): e202116999. doi:10.1002/anie.202116999.

Synthesis and Screening of a DNA-Encoded Library of Non-Peptidic Macrocycles

Eric Koesema^b [Dr.], Animesh Roy^b [Dr.], Nicholas G. Paciaroni^b [Dr.], Carlos Coito^b, Michal Tokmina-Roszyk^a [Dr.], Thomas Kodadek^{a,b} [Prof. Dr]

^[a]Department of Chemistry, Scripps Florida, 130 Scripps Way, Jupiter, FL 33458.

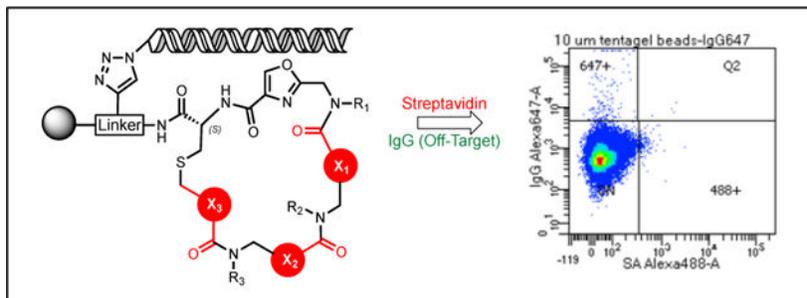
^[b]Deluge Biotechnologies, 6671 W. Indiantown Rd., Suite 50-325, Jupiter, FL 33458.

Abstract

There is considerable interest in the development of libraries of non-peptidic macrocycles as a source of ligands for difficult targets. We report here the solid-phase synthesis of a DNA-encoded library of several hundred thousand thioether-linked macrocycles. The library was designed to be highly diverse with respect to backbone scaffold diversity and to minimize the number of amide N-H bonds, which compromise cell permeability. The utility of the library as a source of protein ligands is demonstrated through the isolation of compounds that bind SA, a model target, with high affinity.

Graphical Abstract

The discovery of ligands for the “undruggable proteome” is likely to require the development of new chemical matter with a greater “molecular wingspan” than traditional Lipinski-compliant small molecules. A DNA-encoded library (DEL) of non-peptidic thioether macrocycles has been constructed and screened for high affinity protein ligands to a model protein, SA.



Keywords

DNA-encoded library; macrocycle; high-throughput screening; solid-phase synthesis; combinatorial chemistry

kodadek@scripps.edu .

Conflict of interest

T.K. is a major shareholder in Deluge Biotechnologies.

Supporting information for this article is given via a link at the end of the document.

Introduction

It is generally acknowledged that molecules with some characteristics outside the “rule of 5”,^[1] particularly molecular weight, will be required to engage difficult protein targets efficiently. Such target proteins lack deep binding pockets suitable for interactions with more traditional drug-like molecules. Engagement of these protein surfaces requires larger, conformationally stable molecules capable of occupying multiple shallow pockets. To that end, macrocycles are of particular interest.^[2] Powerful biological methods, such as phage display^[3] and ribosome display^[4] exist for the creation of large libraries of macrocyclic peptides, which have been shown to be an excellent source of ligands for many different proteins. However, with a few notable exceptions,^[5] macrocyclic peptides exhibit poor cell permeability,^[6] spurring interest in the development of large libraries of non-peptidic macrocycles that would be better able to access intracellular targets. DNA-encoded libraries (DELs)^[7] are of particular interest since these can rival the size of phage display libraries. To date, only a few macrocyclic DELs have been reported. The first, created by Liu and co-workers using DNA-templated chemistry, employed a Wittig reaction to close the ring.^[8] The most recent version of this approach has produced a library of 256,000 macrocycles from which a 40 nM ligand for insulin degrading enzyme was identified.^[9] More recently, Gillingham and colleagues reported a seven cycle DEL comprised of approximately 1.4 million macrocycles in which the ring was closed via amide bond formation.^[10] Ligands with μM dissociation constants were mined from this library. Both of these DELs were created by solution-phase chemistry.

In this study, we report the development of efficient chemistry for the synthesis of DELs of thioether macrocycles by solid-phase split and pool synthesis.^[11] The construction of these libraries on TentaGel resin^[12] allows them to be screened by simply mixing the beads with a fluorescently labeled protein and isolating those that retain the labeled target using a fluorescence activated cell sorter (FACS).^[13] This protocol, in turn, facilitates the use of targets and off targets labeled with different colored dyes, allowing ligands with a high level of selectivity for the target protein to be differentiated from those that do not.^[14] We demonstrate the utility of this system by isolating macrocyclic ligands for SA.

Results and Discussion.

Establishment of Efficient, DNA-Compatible Macrocyclization Conditions.

We sought an efficient strategy for the on-resin macrocyclization of peptoid-inspired conformationally-constrained oligomers (PICCOs)^[15] that would be compatible with DNA encoding technology. A common strategy used for the creation of macrocyclic peptides is thioether formation,^[3b, 16] which often relies on a cysteine residue possessing a protected thiol that is then deprotected to initiate ring closure. However, removal of some thiol protecting groups requires treatment with strong acid, which can be detrimental to the integrity of the DNA encoding tags.^[17] Therefore, we focused on the S-trimethoxyphenol (STmp) protecting group, which is removed efficiently under mild reducing conditions that are likely DNA-compatible.

A linker designed to facilitate ionization by mass spectrometry (Linker-2) was generated on 90 μm TentaGel-RAM resin (Fig. S1). The linker was then acylated with Fmoc-Cys(STmp)-OH. After removal of the Fmoc group, the amine was acylated with 2-(chloromethyl)oxazole-4-carboxylic acid to generate Compound 1a (Fig. 1A). To build the test PICCO, the beads were subjected to 3 rounds of amination with 3-methoxybutylamine and subsequent acylation with 2-(chloromethyl)oxazole-4-carboxylic acid to generate Compound 2 (Fig. 1A). After synthesis of the linear precursor was complete, the STmp group was removed, and the beads were incubated at 37°C for 18 hours to allow for thioether formation. Macrocyclization was then assessed by liquid chromatography/mass spectrometry (LC/MS) as well as matrix assisted laser desorption ionization (MALDI). The molecular ion expected for the macrocyclic product was evident, while the peak expected from the linear precursor was not detected (Figures 1B).

To explore this macrocyclization strategy more broadly and to assess the DNA-compatibility of the reaction conditions, a DNA-encoded, 27-molecule “mini-library” was generated. These compounds follow the general design depicted in Figure 2 and represent a diverse set of PICCO scaffolds (Fig S2A). As above, Cys-STmp-OH and 2-(chloromethyl)oxazole-4-carboxylic acid were added to 90 μm TentaGel RAM resin displaying Linker-2. As an attachment point for DNA barcode ligation, azido-PEG headpiece DNA (HDNA)^[12, 18] was added to the linker by copper-catalyzed click chemistry.^[19] The beads were then aliquoted into 27 wells of a 96-well filter plate. Parallel synthesis was then carried out, specifically three cycles of amine side-chain/backbone unit (Fig. S3A) addition, deprotection of the cysteine, and cyclization. After each round of amination/acylation, an encoding DNA was enzymatically ligated to build the barcode.^[12] Following cyclization, the DNA barcodes were amplified by PCR, purified, and sequenced. From the 27 test compounds, 22 showed PCR products of the expected size (Fig. S3B). Of these, 20/22 provided the expected DNA sequences (Table 1). Additionally, MALDI analysis revealed the expected masses for 23 of the 27 macrocyclic compounds (Table 1, Fig. S3C). In the four cases where this was not the case, no dominant peak was seen in the MALDI-MS, indicating some failure early in the synthesis rather than simply a failure to cyclize. These data suggest that thioether formation is broadly suitable for PICCO macrocyclization and that the Cys-STmp deprotection conditions are DNA-compatible.

Library Design and Synthesis.

With these preliminary quality control experiments completed satisfactorily, a one-bead, one-compound (OBOC) DEL of approximately 580,000 compounds (Fig. 2A) was constructed on 10 μm TentaGel resin displaying Linker-1 (Fig S1), modified with HDNA, Cys(STmp), and 2-(chloromethyl)oxazole-4-carboxylic acid (Compound 1). This was followed by three cycles of split and pool amination and acylation. After each round of amination/acylation, an encoding DNA was ligated onto the encoding chains, the beads pooled, then redistributed into a new filter plate. Finally, the cysteine was deprotected.

Because 6–8 different carboxylic acid building blocks were employed at the three “main chain” diversity positions, the library has exceptional macrocyclic scaffold diversity. Moreover, a “null” row was included at the second PICCO position (X2 in Fig. 2), where

no carboxylic acid was added to the bead. This results in the omission of the X2 and N-R3 elements shown in Fig. 2A (for these beads only) and generates a library that includes both “2.5-mer” and “3.5-mer” macrocyclic compounds (Fig S4). Twelve different amines were employed at each diversity position.

On-Resin Macrocyclization Assay.

To assess the level of macrocyclization in the library, a FACS-based macrocyclization assay recently developed in our laboratory was employed.^[20] Briefly, several hours after cysteine deprotection, allowing sufficient time for thioether formation, a thiol-reactive fluorescent dye, Monobromobimane (mBBR), is added to the beads. Any linear starting material remaining at this time will be stained by mBBR, but the thioether product will not. The fluorescence intensity on each bead is then measured by FACS, reflecting the amount of uncyclized material on each bead. When this assay was carried out on an aliquot of beads from the library synthesis, the results shown in Fig. 2C were obtained. The library beads (purple peak) displayed only a low level of fluorescence, barely greater than that observed when beads displaying a methionine (thioether) unit were stained with mBBR (red peak, “cyclic control”). In contrast, control beads displaying cysteine (free thiol) displayed a dramatically higher level of fluorescence after staining with mBBR (green peak). These data indicate that the vast majority of library compounds efficiently formed thioether macrocycles.

FACS-Based Screening.

The DEL shown in Fig. 2 was used in a two-color, FACS-based screen to identify macrocycles that selectively bind to a target protein, Streptavidin (SA), over an unrelated off-target, human IgG (Fig. 3). Ten copies of the library (almost six million beads) were incubated with differentially labeled SA (the target) and human IgG (the off-target) as well as a large excess of unlabeled, diverse competitor proteins, then FACS was used to isolate highly fluorescent beads. To minimize the possibility that the fluorophores would affect the outcome, the screen was done in duplicate, but with the labels swapped. Specifically, one screen employed AlexaFluor 647 (A647)-conjugated SA (SA-A647) and AlexaFluor 488 (A488)-labelled human IgG, while the second screen used A488-conjugated SA and A647-labelled human IgG.

The DNA barcodes on the beads displaying a high level of fluorescence in the channel represented by the SA dye were then amplified by PCR and deep sequenced. Because bead screening has a significant false positive rate, we focused on compounds that appeared in the hit pool on at least three different beads (the encoding tag includes a bead-specific barcode^[14]). We have found that these so-called “redundant hits” are usually bona fide target protein ligands while “singletons” are usually false positives.^[21] The redundant hits that appeared in both of the screening experiments were then decoded to generate a list of putative ligands for SA (Table S1). Of these 31 hits, the vast majority (28/31) were the smaller, “2.5-mer” macrocycles, where the X2 and N-R3 elements (Fig. 2A) are absent, indicating that the binding site on SA prefers more compact ligands. This is not surprising since SA is a biotin-binding protein. All 28 of these macrocycles include a pyridine side chain at the second amine position, which is almost always adjacent to a thiazole (24

out of 28 compounds), suggesting that this grouping is important for binding (Fig. 3D). Indeed, the other four hits contained a closely related oxazole at this position (Fig. 3D). Almost half of the screening hits contain the difluorobenzylamine-derived side chain at position N-R1 and there is always an aromatic ring at this position. Finally, at position X3, 25/28 hits have a 1,3-substituted aromatic ring, suggesting that the influence of this spacing and regiochemistry on the conformation of the macrocycles is important for high affinity binding.

Validation of Screening Hits.

To determine if the redundant screening hits are indeed bona fide SA ligands, 22 of them were re-synthesized by solid-phase parallel synthesis on both 10 μm and 160 μm beads without DNA tags. The 160 μm beads were treated with benzyl bromide after allowing enough time for macrocyclization to go to completion. This modifies any free thiols, while also improving MS ionization, thus making it easier to spot unreacted starting material in the mass spectrometer. The compounds were then released from the beads by treatment with trifluoroacetic acid (TFA) and analyzed by MALDI mass spectrometry. The MALDI data (Fig. S12) indicate the presence of the correct macrocyclic compound for each resynthesized hit, with only two samples showing the presence of linear compound (in addition to the macrocycle) (Table S2). Notably, this treatment also modified the pyridine nitrogen in each of the compounds synthesized. These data confirm that the chemistry proceeded as designed and indicate further that these compounds cyclize readily.

The 10 μm beads displaying resynthesized hits were individually incubated with 100 nM SA-A647 and analyzed for retention of the protein using a flow cytometer. As a negative control, beads displaying only the linker were also tested. The results are shown in Fig. 4A. Nearly all of the beads displaying the various screening hits evinced a much higher brightness than the linker only control. It has been shown that the amount of fluorescently labeled protein captured by a compound displayed on 10 μm TentaGel beads at a given concentration of the target correlates directly with the K_D of the ligand-protein complex.^[13] Thus, the FACS analysis suggests that two of the macrocyclic compounds, DBT-1 and DBT-2 (Fig. 4B), which are nearly identical, differing by only one atom (oxazole vs. thiazole rings at position X1), have the highest affinity for SA, while the remaining compounds bind with slightly lower affinities. These data were used to rank-order the putative hits and to select compounds for further analysis (Fig. 4B).

An on-resin titration experiment was done. DBT-1 and DBT-2, as well as two hits that indicated lower binding affinities, DBT-3 and DBT-4, were titrated with labeled SA. Robust, above-background binding of the labeled SA is observed even at 1 nM protein concentration for each of ligands (Fig. 4C).

SA is a native tetramer, so avidity effects likely contribute to the high affinity binding to bead-displayed ligands. To evaluate the intrinsic affinities in solution, fluorescein conjugates of each hit were synthesized, released from the resin and purified by HPLC (Fig. S6). As a negative control, one of the compounds present in the library that did not show up as a hit (DBT-8, Fig. S6) was also included. Titration of these compounds with unlabeled SA was monitored by an increase in fluorescence polarization (FP). Each of the resynthesized

hits bound to SA with affinities in the high-nanomolar to low-micromolar range (Figs. 4D). Furthermore, the compounds demonstrated selectivity for SA, as no binding was detected when they were exposed to unrelated proteins (Fig. S7). Of note, the relative binding affinities determined from these data correlate nicely with apparent affinities suggested by the on-resin FACS data. Thus, the bead-based FACS method for rank-ordering the putative hits provides a relatively accurate representation of true binding affinities. The ease of hit validation using this methodology is notable. Dozens of compounds can be re-synthesized on a small scale, using exactly the same chemistry employed to create the library, and tested for on-resin binding by FACS in just a few days at little expense.

Nature of the macrocycle-SA interaction.

As discussed above, all the hits identified in this screen include an oxazole or thiazole followed by an aminomethyl pyridine (Fig. 3), strongly suggesting that these residues are pivotal for SA binding. To test this, a series of DBT-1 analogues were synthesized on 10 μm resin and exposed to 100 nM SA-A647. As shown in Fig. 5, substitution of the pyridine ring with a phenyl group essentially abolished binding of the macrocycle to SA. Substitution of the thiazole ring in the main chain with an isoxazole unit also had catastrophic consequences for binding. This was also the case when a 1,4-substituted phenyl ring replaced the thiazole. These data show clearly that the highly conserved elements in the hit pool are critical to high affinity SA binding. More generally, this experiment demonstrates the ease with which preliminary structure activity relationships using parallel solid-phase synthesis of derivatives followed by FACS-based analysis of protein binding.

It is often the case that the linear analogues of macrocyclic protein ligands bind more weakly to the target due to increased conformational flexibility. To test this, beads displaying DBT-1 or DBT-2 or their linear analogues (see Fig. S8) were incubated with 10 nM SA-A647 and assayed for binding. Both macrocycles exhibited stronger binding to SA than their linear analogs (Fig. 6A), although the differences were modest. To address the same question in solution, fluorescein-tagged macrocyclic and linear molecules (Fig. S5) were titrated with SA and binding was again followed by an increase in FP (Fig. 6B). In each case, the macrocyclic compounds showed stronger binding affinities than their linear counterparts, but again the difference was relatively modest (4- to 6-fold).

Lastly, when DBT-1-displaying beads were incubated with SA-A647 in the presence of excess biotin, binding was abolished, arguing that the macrocycle recognizes the biotin-binding pocket of SA (Fig. S9).

PICCO macrocycles are cell permeable.

Having established the ability to synthesize and successfully screen DELs of PICCO macrocycles, we evaluated the cell permeability of a representative group of structures containing oxazole and thiazole units in the cycle (Fig. 7A and S13) using the chloroalkane penetration assay (CAPA).^[22] This involves incubation of the chloroalkane-tagged molecule of interest with mammalian cells that express the HaloTag protein (HTP). If the molecule is cell permeable, then covalent reaction with intracellular HTP blocks labeling of this protein when the cells are subsequently treated with a chloroalkane-linked fluorescent dye. The

results of several such experiments done at different macrocycle concentrations, are shown in Fig. 7B. For comparison, the permeability of a Lipinski-compliant tryptophan-containing small molecule (W-ct) was also measured. The data show that all of the macrocycles are quite cell permeable, with CP_{50} values in the low μM range. Quantitatively, KMR-7 and KMR-8 were about 10-fold less permeable than W-ct, while KMR-9 was about 37 times less permeable, perhaps reflecting the presence of the polar pyridine side chain. Given the molecular masses of KMR-7, -8, and -9 (754, 802 and 805 Daltons, respectively), this an encouraging result and suggests that DELs of PICCOs can serve as an interesting source of probe molecules targeting intracellular proteins. A more comprehensive analysis of PICCO cell permeability will be presented in a separate study (Tokmina-Roszyk, et al., in preparation).

Conclusion

In summary, efficient methods for the synthesis and screening of high quality, OBOC DELs of non-peptidic macrocycles has been established. Initial cell permeability data suggests that, as expected, these molecules are quite cell permeable. In addition, a post-screening workflow has been established that rapidly provides useful structure-activity relationship data. We anticipate that the DEL shown in Fig. 2 and related libraries will serve as a useful source of probe molecules for both intracellular and extracellular proteins. We are currently deploying this DEL and related libraries against a variety of challenging targets. These efforts will be reported in due course.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This research was funded by an SBIR grant for the National Institutes of Health (GM130164) to E.K. and a Transformational RO1 grant (GM 133041) to T.K. We thank Dr. Ofelia Utset for administrative oversight of all aspects of the project. All bead sorting and flow cytometry experiments were performed at the Scripps Florida Flow Cytometry Core. Deep sequencing was performed at Scripps Research Florida Genomics Core. We thank the staff of these cores for their technical advice and assistance.

References

- [1]. a) Lipinski CA, *J Pharmacol Toxicol Methods* 2000, 44, 235–249; [PubMed: 11274893] b) Lipinski CA, Lombardo F, Dominy BW, Feeney PJ, *Adv. Drug Deliv. Rev* 1997, 23, 3–25.
- [2]. a) Villar EA, Beglov D, Chennamadhavuni S, Porco JA Jr., Kozakov D, Vajda S, Whitty A, *Nature Chem Biol* 2014, 10, 723–731; [PubMed: 25038790] b) Whitty A, Viarengo LA, Zhong M, *Organic & Biomolecular Chemistry* 2017, 15, 7729–7735; [PubMed: 28876025] c) Giordanetto F, Kihlberg J, *Journal of Medicinal Chemistry* 2014, 57, 278–295; [PubMed: 24044773] d) Heinis C, *Nature Chem Biol* 2014, 10, 696–698. [PubMed: 25038789]
- [3]. a) Smith GP, Petrenko VA, *Chem. Rev* 1997, 97, 391–410; [PubMed: 11848876] b) Heinis C, Winter G, *Curr Opin Chem Biol* 2015, 26, 89–98. [PubMed: 25768886]
- [4]. a) Hanes J, Schaffitzel C, Knapplik A, Plückthun A, *Nature Biotechnol.* 2000, 18, 1287–1292; [PubMed: 11101809] b) Ohuchi M, Murakami H, Suga H, *Curr Opin Chem Biol* 2007, 11, 537–542; [PubMed: 17884697] c) Wilson DS, Keefe AD, Szostak JW, *Proc. Natl. Acad. Sci. USA* 2001, 98, 3750–3755; [PubMed: 11274392] d) Keefe AD, Wilson DS, Seelig B, Szostak JW, *Protein Expr Purif* 2001, 23, 440–446. [PubMed: 11722181]

- [5]. a) Ahlback CL, Lexa KW, Bockus AT, Chen V, Crews P, Jacobson MP, Lokey RS, *Future Med Chem* 2015, 7, 2121–2130; b) Bockus AT, McEwan CM, Lokey RS, *Curr. Top. Med. Chem* 2013, 13, 821–836. [PubMed: 23578026]
- [6]. a) Vinogradov AA, Yin Y, Suga H, *Journal of the American Chemical Society* 2019, 141, 4167–4181; [PubMed: 30768253] b) Rezai T, Bock JE, Zhou MV, Kalyanaraman C, Lokey RS, Jacobson MP, *J. Amer. Chem. Soc* 2006, 128, 14073–14080; [PubMed: 17061890] c) Rezai T, Yu B, Millhauser GL, Jacobson MP, Lokey RS, *J. Amer. Chem. Soc* 2006, 128, 2510–2511. [PubMed: 16492015]
- [7]. a) Clark MA, *Curr Opin Chem Biol* 2010, 14, 396–403; [PubMed: 20346729] b) Neri D, Lerner RA, *Annu Rev Biochem* 2018, 87, 479–502; [PubMed: 29328784] c) Kleiner RE, Dumelin CE, Liu DR, *Chem Soc Rev* 2011, 40, 5707–5717; [PubMed: 21674077] d) Gartner ZJ, Kanan MW, Liu DR, *J Am Chem Soc* 2002, 124, 10304–10306. [PubMed: 12197733]
- [8]. Kleiner RE, Dumelin CE, Tiu GC, Sakurai K, Liu DR, *J Am Chem Soc* 2010, 132, 11779–11791. [PubMed: 20681606]
- [9]. Usanov DL, Chan AI, Maianti JP, Liu DR, *Nat Chem* 2018, 10, 704–714. [PubMed: 29610462]
- [10]. Stress CJ, Sauter B, Schneider LA, Sharpe T, Gillingham D, *Angewandte Chemie International Edition* 2019, 58, 9570–9574. [PubMed: 30938482]
- [11]. Lam KS, Salmon SE, Hersh EM, Hruby VJ, Kazmierski WM, Knapp RJ, *Nature* 1991, 354, 82–84. [PubMed: 1944576]
- [12]. MacConnell AB, McEnaney PJ, Cavett VJ, Paegel BM, *ACS Comb Sci* 2015, 17, 518–534. [PubMed: 26290177]
- [13]. Erharuyi O, Simanski S, McEnaney PJ, Kodadek T, *Bioorg Med Chem Lett* 2018, 28, 2773–2778. [PubMed: 29395976]
- [14]. Mendes K, Malone ML, Ndungu JM, Suponitsky-Kroyter I, Cavett VJ, McEnaney PJ, MacConnell AB, Doran TM, Ronacher K, Stanley K, Utset O, Walzl G, Paegel BM, Kodadek T, *ACS Chem. Biol* 2017, 19, 234–243.
- [15]. Kodadek T, McEnaney PJ, *Chem Commun (Camb)* 2016, 52, 6038–6059. [PubMed: 26996593]
- [16]. a) Kawakami T, Ohta A, Ohuchi M, Ashigai H, Murakami H, Suga H, *Nature Chem Biol* 2009, 5, 888–890; [PubMed: 19915537] b) Fadzen CM, Wolfe JM, Cho C-F, Chiocca EA, Lawler SE, Pentelute BL, *Journal of the American Chemical Society* 2017, 139, 15628–15631. [PubMed: 28992407]
- [17]. Malone ML, Paegel BM, *ACS Comb Sci* 2016, 18, 182–187. [PubMed: 26971959]
- [18]. Clark MA, Acharya RA, Arico-Muendel CC, Belyanskaya SL, Benjamin DR, Carlson NR, Centrella PA, Chiu CH, Creaser SP, Cuzzo JW, Davie CP, Ding Y, Franklin GJ, Franzen KD, Gefter ML, Hale SP, Hansen NJ, Israel DI, Jiang J, Kavarana MJ, Kelley MS, Kollmann CS, Li F, Lind K, Mataruse S, Medeiros PF, Messer JA, Myers P, O’Keefe H, Oliff MC, Rise CE, Satz AL, Skinner SR, Svendsen JL, Tang L, van Vloten K, Wagner RW, Yao G, Zhao B, Morgan BA, *Nature Chem Biol* 2009, 5, 647–654. [PubMed: 19648931]
- [19]. Kolb HC, Sharpless KB, *Drug Discovery Today* 2003, 8, 1128–1137. [PubMed: 14678739]
- [20]. Roy A, Koesema E, Kodadek T, *Angew Chem Int Ed Engl* 2021, 60, 11983–11990. [PubMed: 33682283]
- [21]. Doran TM, Gao Y, Mendes K, Dean S, Simanski S, Kodadek T, *ACS Comb. Sci* 2014, 16, 259–270. [PubMed: 24749624]
- [22]. Peraro L, Zou Z, Makwana KM, Cummings AE, Ball HL, Yu H, Lin YS, Levine B, Kritzer JA, *J Am Chem Soc* 2017, 139, 7792–7802. [PubMed: 28414223]

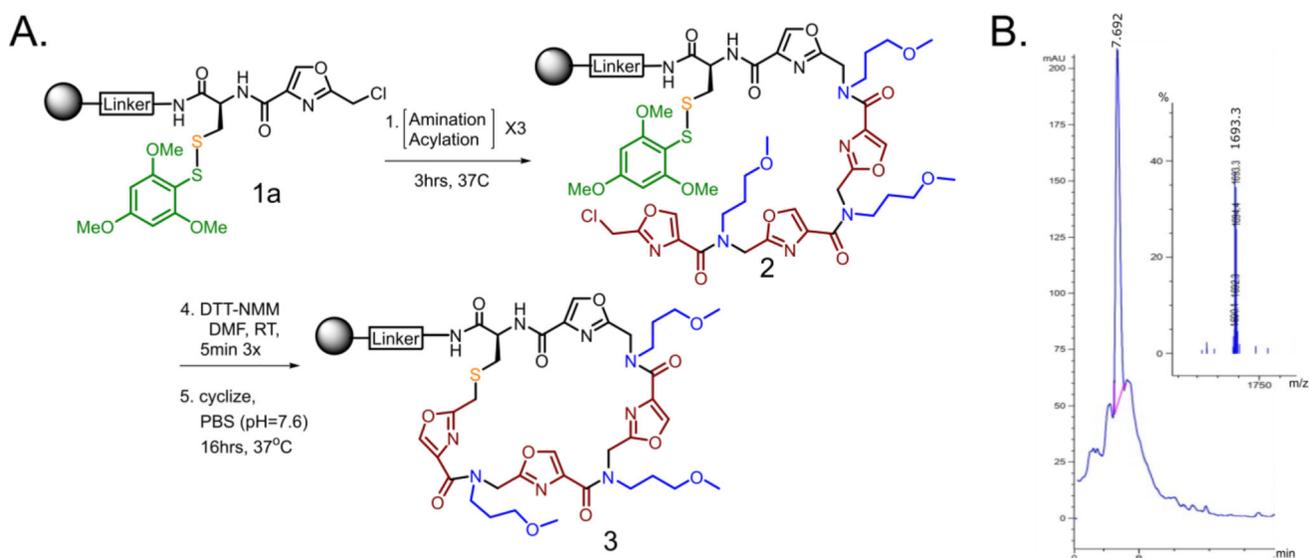


Figure 1: Assessment of PICCO macrocyclization. (A) Beads displaying a linker capped with Cys-STmp and an oxazole (Compound 1a) were subjected to three rounds of amination and acylation to generate a linear PICCO (Compound 2). Cysteine deprotection allows thioether formation, generating the macrocycle (Compound 3). (B) LC/MS analysis demonstrates complete cyclization of the linear precursor. See supporting information (Fig. S2) for complete characterization.

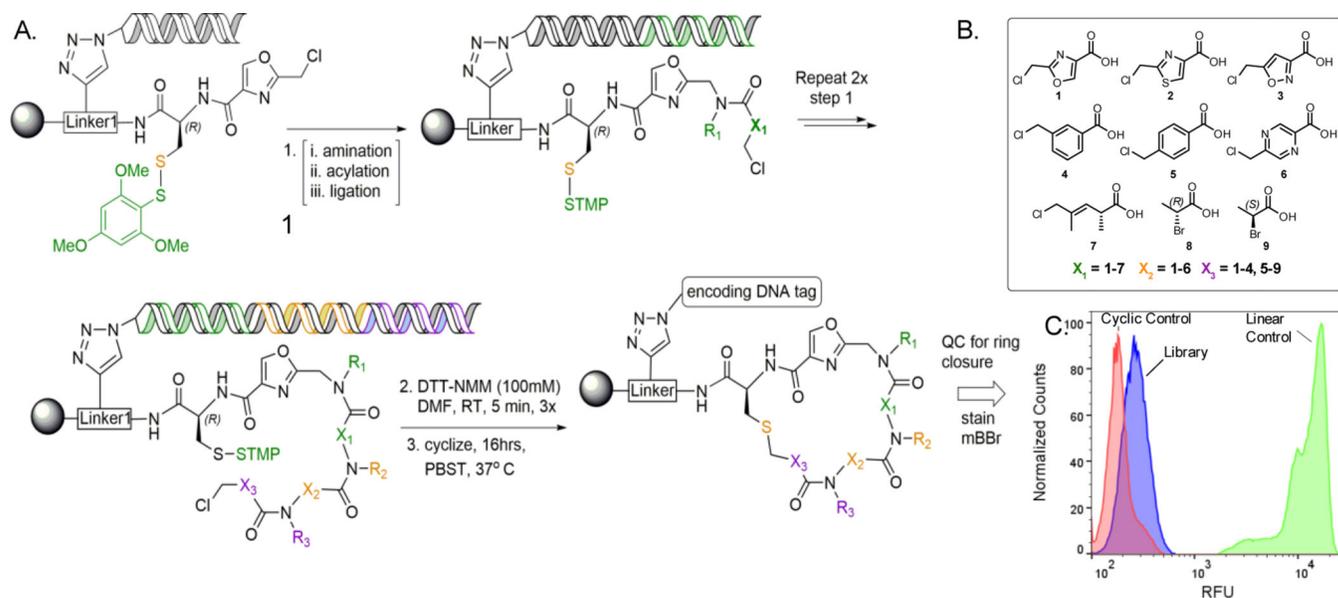
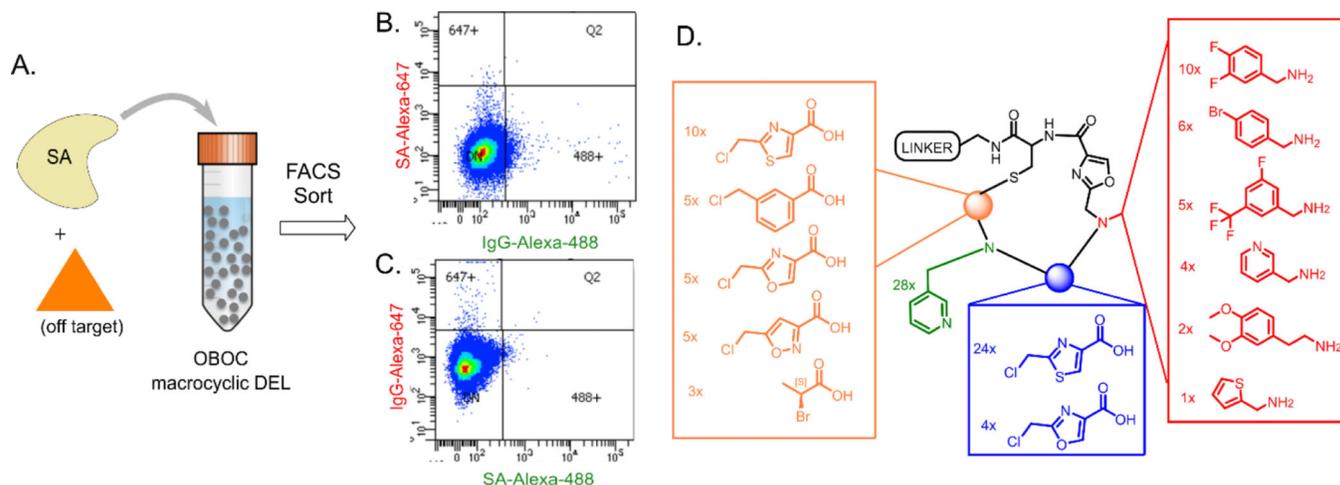


Figure 2: General library design. (A) The library was synthesized by solid-phase split-and-pool synthesis on 10 μ m resin possessing a linker (Linker-1) modified with HDNA, Cys(STmp), and an oxazole (compound 1). The linear precursors were formed after 3 rounds of amination/acylation/enzymatic ligation using the backbone units shown (B). The STmp group was then removed to allow thioether formation. (C) After macrocyclization, aliquots (about 10,000 beads) were stained with a thiol-reactive fluorescent dye (mBBr) to confirm that the vast majority of the library completed cyclization. Control compounds are shown in Fig. S5.

**Figure 3.**

Summary of screening experiments. (A) Library beads were incubated with fluorescently-labeled Streptavidin (SA) and an orthogonally labelled off-target, then sorted by FACS. (B,C) FACS plots showing data from the SA-A647/IgG-A488 (B) and IgG-A647/ SA-A488 (C) replicate screens. Beads that shifted into the 647+ or 488+ gates were collected, and the encoding DNAs deep sequenced for structure determination. (D) "Bottom-up" analysis of 28 "2.5-mer" SA hits, indicating the frequency of each chemical unit at each position. This illustrates a conserved pyridine side-chain (green) that is almost exclusively adjacent to a thiazole backbone unit (blue).

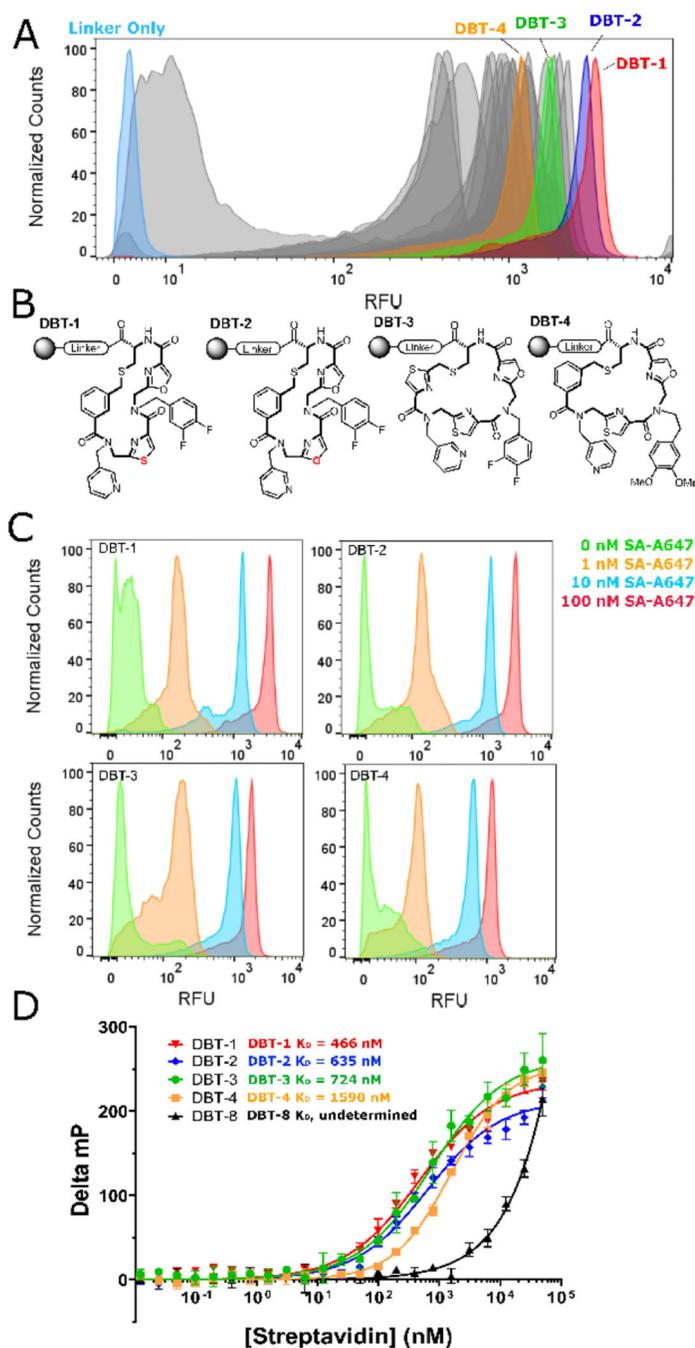


Figure 4: Validation of screening hits. (A) FACS histogram showing the relative fluorescence of beads displaying resynthesized hits after incubation with 100 nM SA-A647. Highlighted samples were analyzed further. (B) Structures of the selected hits. (C) Titration of beads displaying DBT compounds with the indicated concentration of SA-A647, followed by FACS analysis. FACS histograms are shown. (D) Fluorescence polarization data for the selected DBT compounds.

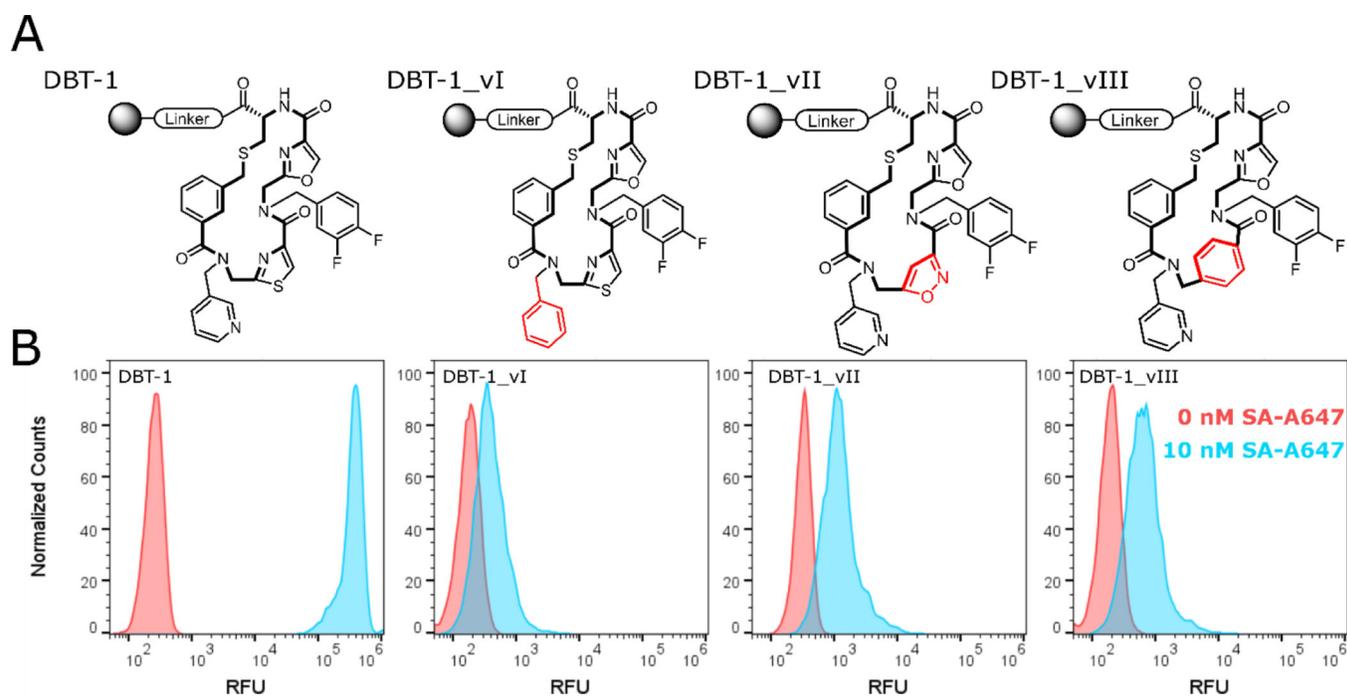


Figure 5: Structure-activity relationship. (A) Structures of DBT-1 and three analogs tested. (B) FACS histograms showing the level of fluorescent SA binding to the corresponding bead-displayed analogs. The structural alterations relative to DBT-1 are highlighted in red.

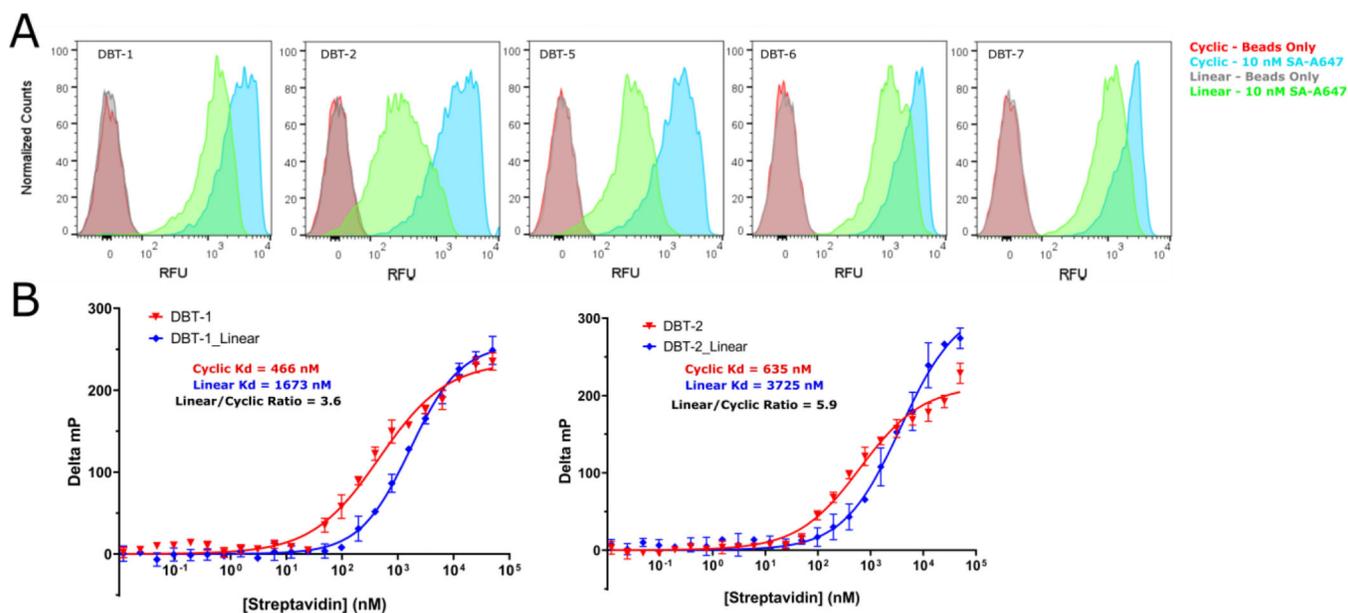


Figure 6:

Comparison of macrocyclic hits with corresponding linear analogues. (A) A set of macrocyclic and linear compounds was resynthesized on-resin, incubated with 10 nM SA-A647, then assayed by FACS for SA binding. Higher RFU suggests a stronger binding affinity for the macrocyclic compounds. (B) Fluorescence polarization data indicating the binding affinities of two purified macrocyclic compounds and their linear analogues.

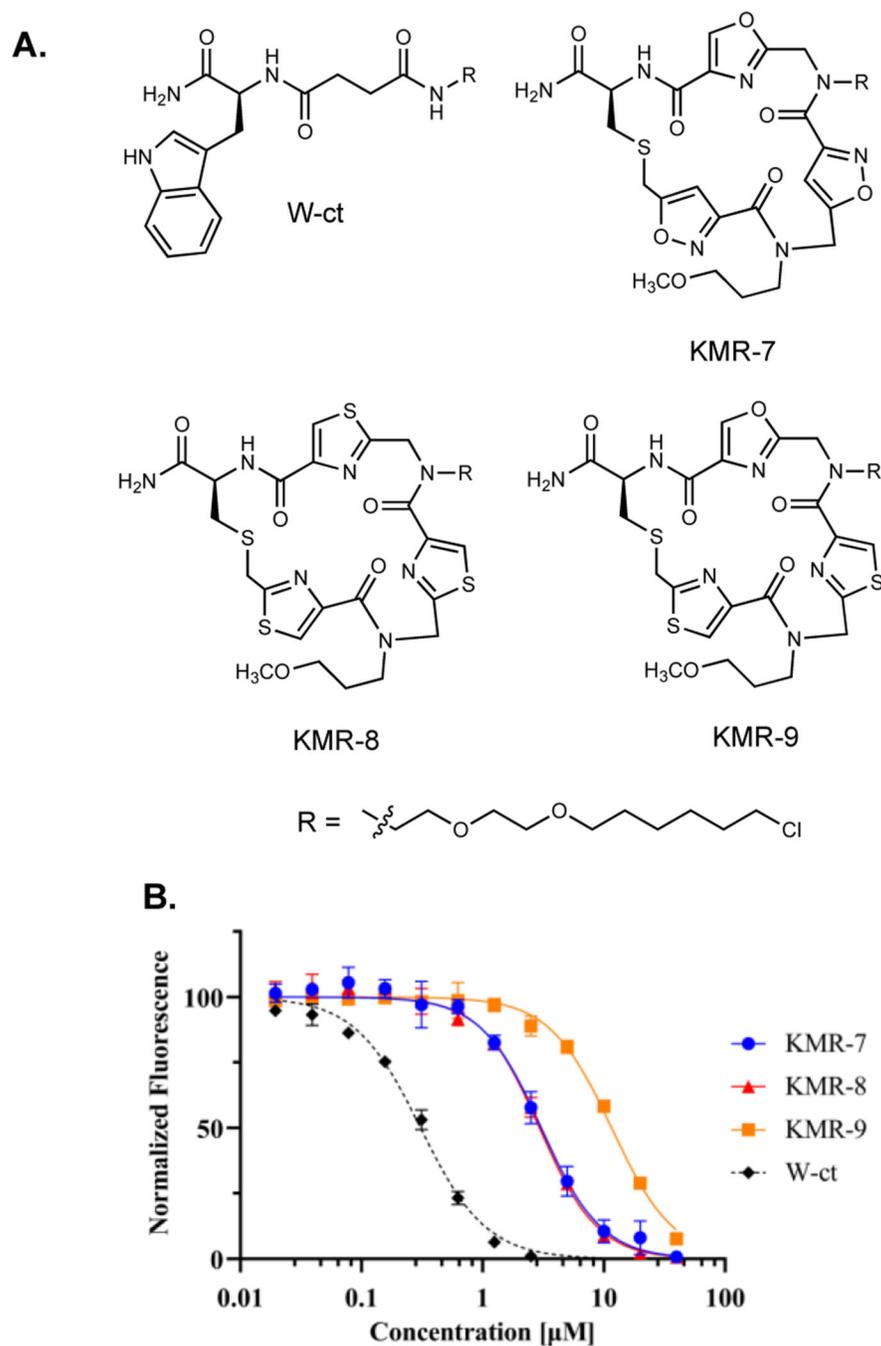


Figure 7: Cell permeability of select PICCO macrocycles. (a) Structures of the molecules analyzed. (B) Relative level of fluorescence of cells treated with the compounds indicated followed by chloroalkane-tagged fluor (see SI for details).

Table 1 –

A set of 27 DNA-encoded PICCO macrocycles was synthesized in individual wells of a filter plate, then assayed by DNA sequencing and MALDI. Verified sequences and masses are highlighted in yellow.

Cmpd	DNA Barcode	Expected Mass
1	130124011501260117012801	1815
2	130124011501260117012802	1824
3	130124011501260117012803	1762
4	130124011504260417042804	1782
5	130124011504260417042805	1771
6	130124011504260417042806	1709
7	130124011507260717072807	1711
8	130124011507260717072808	1722
9	130124011507260717072809	1658
10	130224021502260217022801	1723
11	130224021502260217022802	1737
12	130224021502260217022803	1675
13	130224021505260517052804	1658
14	130224021505260517052805	1667
15	130224021505260517052806	1605
16	130224021508260817082807	1700
17	130224021508260817082808	1709
18	130224021508260817082809	1646
19	130324031503260317032801	1618
20	130324031503260317032802	1627
21	130324031503260317032803	1565
22	130324031506260617062804	1660
23	130324031506260617062805	1669
24	130324031506260617062806	1607
25	130324031509260917092807	1669
26	130324031509260917092808	1678
27	130324031509260917092809	1616