Linearizable Macrocyclic Peptide Libraries for Affinity Selection-Mass Spectrometry

Michael A. Lee[1†], Joseph S. Brown[1†], Charlotte E. Farquhar[1], Andrei Loas[1], Bradley L. Pentelute[1-4*]

[1] Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States
[2] The Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, United States
[3] Center for Environmental Health Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States
[4] Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, United States
*Email: blp@mit.edu
[†]Equal contribution

1 Abstract

Macrocyclic peptides are attractive for therapeutic development but have been limited in their application to combinatorial library selection from synthetic libraries. Here, we establish a synthetic approach based on split-pool chemistry to produce 100-million membered macrocyclic libraries containing natural and non-natural amino acids. Near-quantitative intramolecular disulfide formation is facilitated by rapid (<10 minute) oxidation by iodine to prepare macrocyclic synthetic libraries in solution. Treatment with dithiothreitol post-affinity selection enables near-quantitative reduction of the library members, rendering the linear analogs amenable to standard tandem mass spectrometry sequencing. We demonstrate the utility of these libraries to discover novel macrocyclic binders to cadherin-2 and the anti-hemagglutinin antibody clone 12ca5. The lead cadherin-binding peptide (**CBP**) is endowed with nanomolar binding affinity measured by biolayer interferometry (BLI, apparent dissociation constant K_D = 53 nM). Structure-activity relationship (SAR) studies including alanine and D-amino acid scans reveal the amino acids responsible for driving affinity (hot-spots) and the positions tolerating mutagenesis (cold-spots). Informed by SAR data, two new macrocyclic libraries are designed to derivatize these positions with a variety of abiotic amino acids based on the hot- and cold-spots. Following affinity selection and experimental validation by BLI, 10 high-affinity ligands out of 10 discovered were identified from the library that derivatized the CBP cold-spots, while zero of the two peptide ligands discovered from the hot-spot library were high-affinity binders. Of these noncanonical CBPs (**NCBP**), **NCBP-4** demonstrates improved affinity to cadherin-2 (K_D = 29 nM). Overall, we expect that this work will pioneer the use of large-scale macrocyclic libraries to further catalyze therapeutic peptide discovery and development.

2 Introduction

Macrocyclic peptides show significant therapeutic promise with advantages over small molecules to disrupt protein-protein interactions and over proteins to cross biological membranes barriers.[1–4] Specifically, macrocyclization can impart several potential benefits to linear precursors, including increased proteolytic stability, cell permeability, and oral bioavailability.[5,6] Proteases often engage and degrade peptides in extended β-strand conformations.[7,8] Macrocyclization can offer proteolytic resistance by limiting conformational accessibility of the peptide backbone to the enzyme active site, and enable the use of specific engineerable scaffolds (e.g., stapled α-helices).[9–12] Cyclization is central to the currently applied design principles to achieve passive cell permeability, in addition to strategies that modulate molecular weight, polar surface area, hydrogen bond interactions, and shape.[13–16] Combining proteolytic stability and passive permeability can impart oral bioavailability for peptide-based drug candidates, which can further be improved by pharmaceutical formulation.[17,18] For these reasons, macrocyclic libraries are preferred for screening with peptide ligand discovery platforms. In addition, the direct identification of macrocyclic peptide binders from these selections streamlines subsequent development by alleviating the need to optimize suitable cyclization sites. Lastly, the conformational constraint imparted by macrocyclization may improve discovery rates of ligands from libraries against challenging targets.[19–21]

Genetically-encoded discovery platforms generally access macrocyclic peptide libraries while focusing on high diversity (>10^8 members),[22–28] while synthetic libraries can access the non-natural chemical space at generally lower diversity (<10^8 members).[29–33] While more stable macrocyclization linkages are preferred (e.g., thioether or alkyl chain),[24,34] the disulfide linkage is suitable at the ligand discovery stage, and does not require any chemical modification or treatment that could compromise genetic amplification in some platforms.[28,35] The disulfide linkage has been used to create macrocyclic libraries for over two decades in phage display discovery platforms,[36–38] and is commonly encountered in clinically-approved drugs.[5,39] Synthetic libraries generally leverage the broader use of non-natural or abiotic functionalities, which have frequently appeared critical to the success of clinical peptide drug candidates including inhibitors to the interleukin-23 receptor (IL-23R), mouse double minute 2 (MDM2), β-catenin, and proprotein convertase subtilisin/kexin type 9 (PCSK9).[17,29,40–42] Because they cannot be genetically-encoded or amplified, synthetic libraries are screened directly[30] as in affinity selection decoded by mass spectrometry (AS-MS).[43,44] With a key exception of DNA-encoded libraries,[45] state-of-the-art synthetic macrocyclic libraries generally number below tens of thousands of individual compounds.[30]

The complexity of decoding macrocyclic peptide sequences in mass spectrometry is a historic limitation for the use of synthetic macrocyclic libraries in affinity selection discovery platforms. Experimental approaches for decoding macrocyclic libraries include computational processing of mass spectra[46–49] and chemically-triggered linearization.[50–55] Computational approaches process primary, secondary, and various tertiary mass spectra of cyclic peptide fragments, and have exceled where database matching is possible.[46,49] For *de novo* sequencing, the complexity of enumerating virtual spectra dramatically increases as the number of monomers and library size increases, and has only been demonstrated up to ~1,000-membered libraries.[47,56] Chemically-triggered linearization adds a synthetic step that must be near-quantitative and high-yielding to enable bottom-up sequencing of non-cyclic peptides, which has been demonstrated at very high diversities.[57] However, most chemical linearization treatments are harsh and/or rely on the inclusion of non-standard chemical functional groups at fixed positions, limiting library composition.[50–55] Moreover, these approaches have yet to be demonstrated on high-diversity libraries (~10^8 members or more).

We demonstrate here ligand discovery from high-diversity libraries (10^8 members) utilizing AS-MS against anti-hemagglutinin antibody clone 12ca5 (hereinafter abbreviated as 12ca5) and cadherin-2. Cyclization by disulfide bond formation is accomplished using aqueous iodine. We verify the integrity of the library utilizing Ellman's assay and size-exclusion chromatography (SEC) to confirm near-quantitative intramolecular macrocyclization of synthetic combinatorial libraries. Linearization is accomplished by mild reduction with heat

and 1,3-dithiothreitol (DTT) and confirmed by Ellman's assay to enable standard tandem MS sequencing. We apply these new macrocyclic high-diversity libraries containing natural and non-natural (or noncanonical) amino acids in an MS-based affinity selection platform for *de novo* peptide ligand discovery.

We demonstrate successful discovery of nanomolar ligands against 12ca5 and the ectodomain of cadherin-2. The 12ca5 protein binds peptides containing the sequence D**DY(A/S).[58,59] While 12ca5 has been used to benchmark linear AS-MS libraries,[33,60] we utilize it here to benchmark and additionally validate the use of the new high-diversity macrocyclic libraries. Cadherin-2 was considered as a second target because of the potential impacts for chemical biology that an affinity reagent could provide, ranging from basic cell adhesion, to neural synapses formation,[61] to the construction of intercalated discs of mammalian heart,[62] as well as potential drug delivery due to its relative tissue selectivity in the brain and heart.[63] These critical roles in biology are generally facilitated by homodimerization in domain 1 and 2.[61,64,65] Thus, we sought to discover ligands that bind to cadherin-2 domains 4 and 5 as they may not interfere with caherin-2 function. Outside of domain 1 and 2, there are no ligands to cadherin-2 to our knowledge.[64,65] Lastly, we demonstrate the incorporation of non-natural amino acids for second-generation ligand discovery in libraries designed with input gained by structure-activity relationship (SAR) data gathered on the initially discovered cadherin-binding peptide (**CBP**). Taken together, the successful discovery of macrocyclic ligands to both targets from AS-MS demonstrates the potential deployment of ultra-large synthetically-prepared macrocyclic libraries for peptide ligand discovery and development.

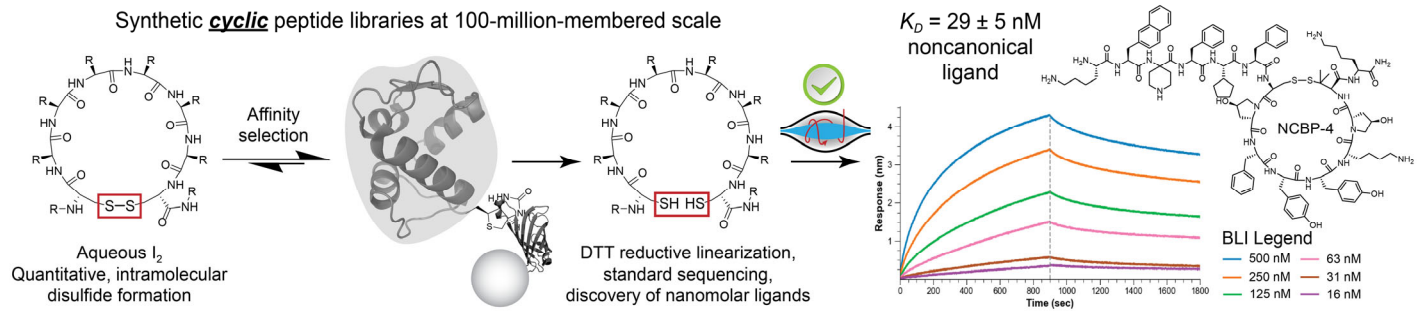


Figure 1. Disulfide linkages allows for high-number diversity libraries compatible with decoding by tandem MS/MS. Libraries of macrocyclic peptides are prepared for affinity selection by oxidation of cysteine analogs using aqueous iodine, providing a near-quantitative conversion to intramolecular macrocyclic peptides. After affinity selection, peptides can be quickly linearized using dithiothreitol (DTT) for sequencing using MS/MS.

3 Results and Discussion

Due to its demonstrated utility in genetically-encoded libraries, disulfide-induced macrocyclization of peptides has become a routine approach with a variety of existing methods to facilitate the oxidization step. Several different methods exist to form the disulfide linkage on single peptides, including oxidation using dimethyl sulfoxide,[66] a gentle stream of air, or aqueous iodine. Ideally, the macrocyclization step can be introduced during standard peptide library synthesis without incurring production delays or yield losses. The isolation of peptides in DMSO-containing solutions could be challenging as the solvent cannot be easily evaporated or lyophilized and solid-phase extraction could incur sample loss due to the DMSO content without further aqueous dilution.[67] In comparison, oxidation utilizing iodine presents itself as a rapid method compatible with mixtures of aqueous or organic solvents, and can even facilitate the formation of disulfide bonds on resin during solid-phase peptide synthesis.[43,68] However, longer reaction times on resin promote iodine-based side reactions, therefore rapid in-solution oxidation is preferred (< 15 min).[69]

Iodine facilitated formation of macrocyclic peptide libraries at 200-million membered scale. We synthesized macrocyclic libraries by split-and-pool solid-phase peptide synthesis using mono-sized 20 µm resin (8.33 g of resin, 2.00 mmol scale total), with each bead providing ~1 pmol of peptide. Two billion-membered libraries were prepared with the designs of **C**X_{12}**C**K and X_6**C**X_6**C**K, where X = all canonical amino acids except Cys, to control disulfide formation, and Ile because it is isobaric in mass with Leu (18 amino acids) and **C** = cysteine, homocysteine, and penicillamine (Figure 3A). The libraries were split into five separate 200-million-membered aliquots and cleaved from the solid phase resin using a cleavage cocktail. After ether trituration and lyophilization, peptide libraries were cyclized in 5% acetonitrile in water (with 0.1% trifluoroacetic acid) at ~2 mg/mL (~1 mM) by dropwise addition of ~1 eq. iodine in methanol until a yellow-brown color persisted. After 5-10 minutes at room temperature in the dark, the reaction was quenched with aqueous ascorbic acid to provide a colorless solution again (3.5 eq.). These libraries were then characterized to verify the efficiency of the oxidation and linearization reactions as well as their structure (intramolecular vs intermolecular disulfide formation).

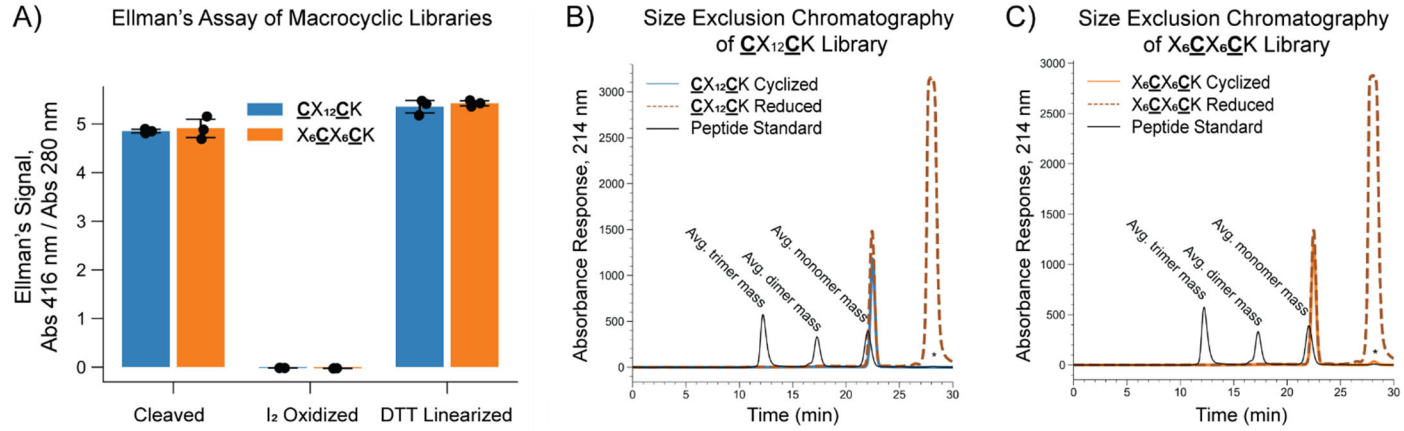


Figure 2. Characterization of macrocyclic libraries based on size and thiol concentration showed near-quantitative formation of intramolecular disulfide bonds. (A) Ellman's assay showed expected changes in total thiol concentration of the peptides directly after cleavage from the solid phase resin, after oxidation by dropwise addition of ~1 eq. of 60 mM iodine in methanol to facilitate disulfide formation (room temperature, 5-10 minutes in the dark, subsequent quench with 3.5 eq. aqueous ascorbic acid), and after reduction using DTT (50 mg/mL, ~1000 eq at 60 ºC for 15 minutes). Free thiol was quantitatively consumed during the oxidation process and was restored after linearization to concentrations comparable to those determined directly after cleavage. (B,C) Size exclusion chromatograms of absorbance at 214 nm of two macrocyclic libraries compared to molecular weight standards corresponding to the average mass of monomeric, dimeric, and trimeric species. Library samples were ran using the cyclized form (later used in affinity selection experiments) and the DTT linearized form, demonstrating the formation of intramolecular disulfide bonds. Peaks marked with an asterisk (*) were residual elements from the sample buffer. **C** = cysteine, homocysteine, and penicillamine.

Iodine-promoted cyclization was highly efficient and provided near-quantitative oxidization to disulfide by thiol quantification using Ellman's assay. We quantitated thiol oxidation by performing an Ellman's assay, normalized by the absorbance of the library at 280 nm (Figure 2A). The thiol content of the library was quantified by Ellman's reagent after cleavage, cold ether trituration, and solid-phase extraction (SPE), to remove any remaining reducing scavengers. Upon aqueous resuspension of the library, a strong thiol signal was observed.

This signal was eliminated completely by the treatment of the library with iodine, ascorbic acid quench, SPE purification, and aqueous resuspension, consistent with the near-quantitative formation of disulfide bonds.

With macrocyclic libraries in hand, a MS-friendly protocol to reduce and linearize the peptides was devised to enable standard tandem sequencing approaches. Both DL-dithiothreitol (DTT) and tris(2-carboxyethyl)phosphine (TCEP) were considered for disulfide reduction. Compared to TCEP, DTT is smaller and more hydrophilic and less likely to be retained on reverse phase columns. Because of the concern for its retention on column, TCEP was utilized in a bead-immobilized form, whereas DTT was directly added to each sample just before mass spectrometry, reducing handling steps and potential sample loss. Due to its high solubility, DTT was utilized at 50 mg/mL, (~1000 eq) at 60 ºC for 15 minutes, whereas immobilized TCEP was utilized at 20 eq at room temperature for 25 minutes per manufacturer protocol.[70] For DTT-treated samples, an SPE purification was performed to remove excess DTT reagent, whereas samples treated with immobilized TCEP were isolated by centrifugation. While TCEP only provided incomplete reduction (40% for X_6**C**X_6**C**K and 85% for **C**X_{12}**C**K), DTT provided near-quantitative disulfide reduction (~100% of original Ellman's signal across all libraries, Figure S1). Additionally, the reduction efficiency by DTT was found to be similar at pH 3 and pH 8 (Figure S1). The direct reduction at pH 3 was performed to mimic a prepared sample in 0.1% formic acid in water, which could then directly be injected in the mass spectrometer for tandem sequencing. Overall, these data support the near-quantitative formation and reduction of disulfide bonds.

Size exclusion chromatography (SEC) confirmed disulfide bonds correspond to formation of intramolecular species, producing an almost-exclusively monomeric macrocyclic peptide library. Utilizing SEC to separate the library by its apparent molecular weight, we assessed if the iodine-facilitated disulfide bond formation was intramolecular or intermolecular. Library samples were injected on a SuperDex® 30 10/300 GL column, which can distinguish the molecular weight range of 100 to 7000 Da, to analyze the presence of monomeric, dimeric, and oligomeric peptides induced by iodine oxidation. Aliquots of 25 µg of library were injected after oxidation by iodine and purification by SPE, as well as in the reduced form after treatment with DTT. A custom low-molecular weight peptide standard was also prepared using peptides that correspond to the average molecular weights of monomeric, dimeric, and trimeric species along with the exclusion limit of the column; the components of the custom standard are given in Table S1. As shown in Figures 2B and 2C, only monomeric library species were observed when compared to the peptide standard, indicating that the disulfide bonds formed were intramolecular rather than intermolecular. This result, in tandem with the data from the Ellman's assay, asserted a nearly quantitative conversion of peptide thiols to intramolecular disulfide bonds and confirmed this technique provides a facile method for preparing high-diversity macrocyclic peptide libraries.

The high-diversity macrocyclic libraries verified to be compatible with standard tandem MS/MS sequencing protocols and DTT facilitated the expected recovery of high-confidence peptide sequencing. Small aliquots of library (~1000 beads equivalent to ~1000 sequences) were taken at various points in the affinity selection workflow from cyclization to linearization, including directly after cleavage from resin, after cyclization with iodine, and after linearization with DTT. About 8 µg of peptide library was purified via a C_{18} STAGE tip[71] from three each steps in the protocol: linear from cleavage, iodine-macrocyclized, and DTT-linearized. As expected, the library samples taken directly after cleavage and after linearization showed high sequencing confidence, described by *de novo* sequence IDs using PEAKS Studio 8.5 with an assigned local confidence (ALC) greater than 85% and an absolute mass error <5 ppm (see Figure 3B and Figure S2).[72] Conversely, the macrocyclized library sample showed poor sequencing confidence, consistent with unproductive fragmentation caused by the disulfide macrocycle. Lastly, the linearization by DTT in the mass spectrometry sample provided a significant recovery of the peptides discovered in standard tandem sequencing methods (Figure 3B).

Analysis of the MS/MS sequencing data demonstrated a balanced distribution of amino acid monomers throughout the library, as well as the incorporation of non-canonical cysteine analogs. Histograms showing the monomer distribution among high-confidence sequence assignments are given in Figure 3C. Although penicillamine and methionine are isobaric, the fixed positions of penicillamine allowed for effective filtering of

sequences to prevent inaccurate assignments. Interestingly, the X_6**C**X_6**C**K library design sequenced with higher confidence overall compared to the **C**X_{12}**C**K library design, suggesting a benefit of the intermediate cysteine position during fragmentation events in sequence assignment. Overall, these results corroborate the successful split-and-pool synthesis, oxidation, reduction, and tandem sequencing decoding of the macrocyclic libraries.

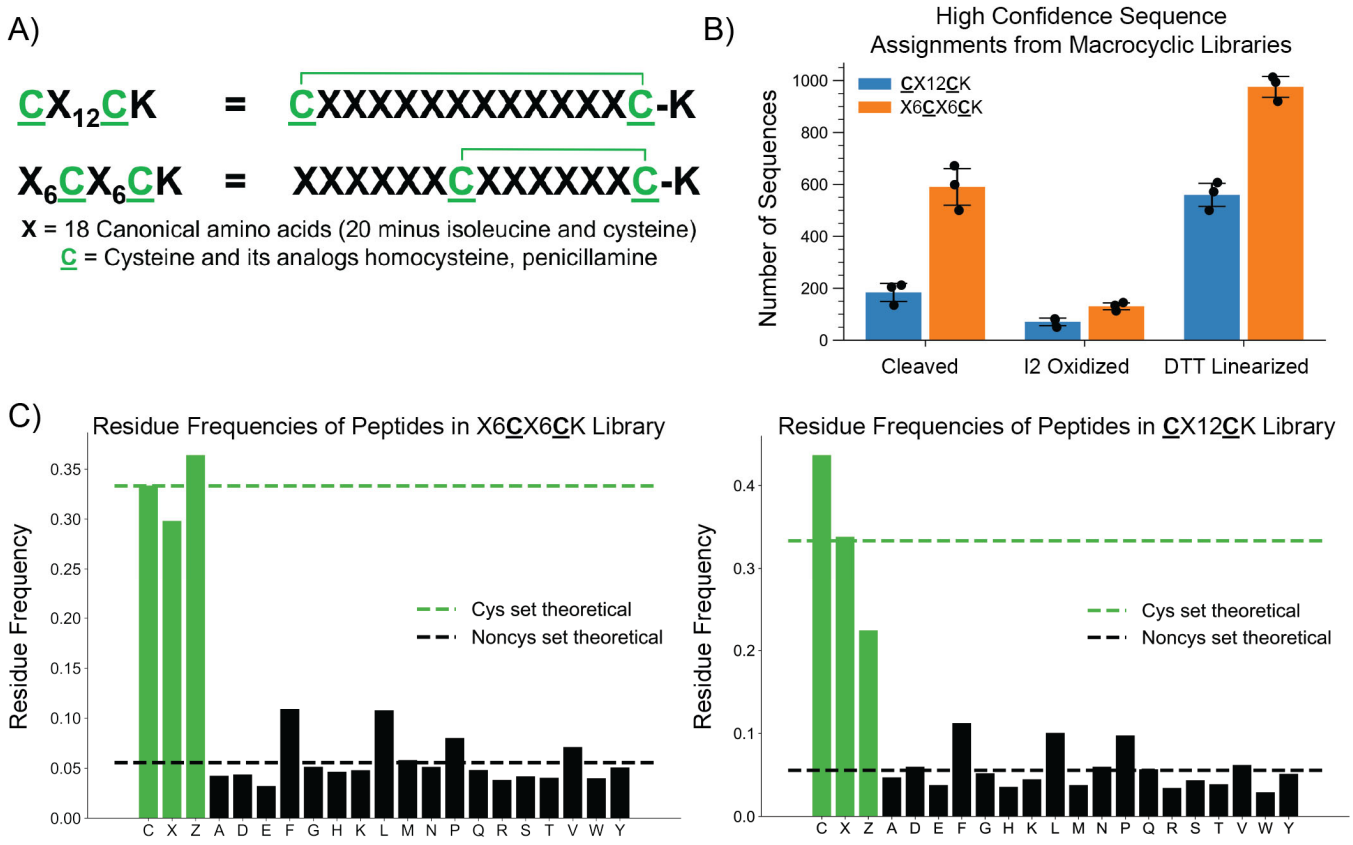


Figure 3. Characterization of the macrocyclic peptide libraries by tandem MS/MS sequencing shows successful split-and-pool synthesis of 15-residue libraries. (A) The macrocyclic peptide libraries were synthesized according to two designs, with a large 12-member macrocycle or a smaller 6-member macrocycle. Additionally, cysteine analogs including homocysteine and penicillamine were used to increase the diversity around the resulting disulfide linkage. (B) High-confidence sequence assignments of ~1000-member library samples directly after cleavage, after oxidation by I_2, and after subsequence reduction using DTT show a loss and gain of sequencing capabilities with the macrocyclization process as expected (n = 3). High-confidence sequences were determined as having a calculated average local confidences (ALC) in PEAKS Studio 8.5 for the *de novo* sequence assignment by PEAKS Studio 8.5 greater than 85% and an absolute mass error <5 ppm. (C) Normalized residue frequencies assignments (as a fraction) show a balanced incorporation of amino acids in the variable positions (black) and cysteine analog positions (green, X is homocysteine, Z is penicillamine).

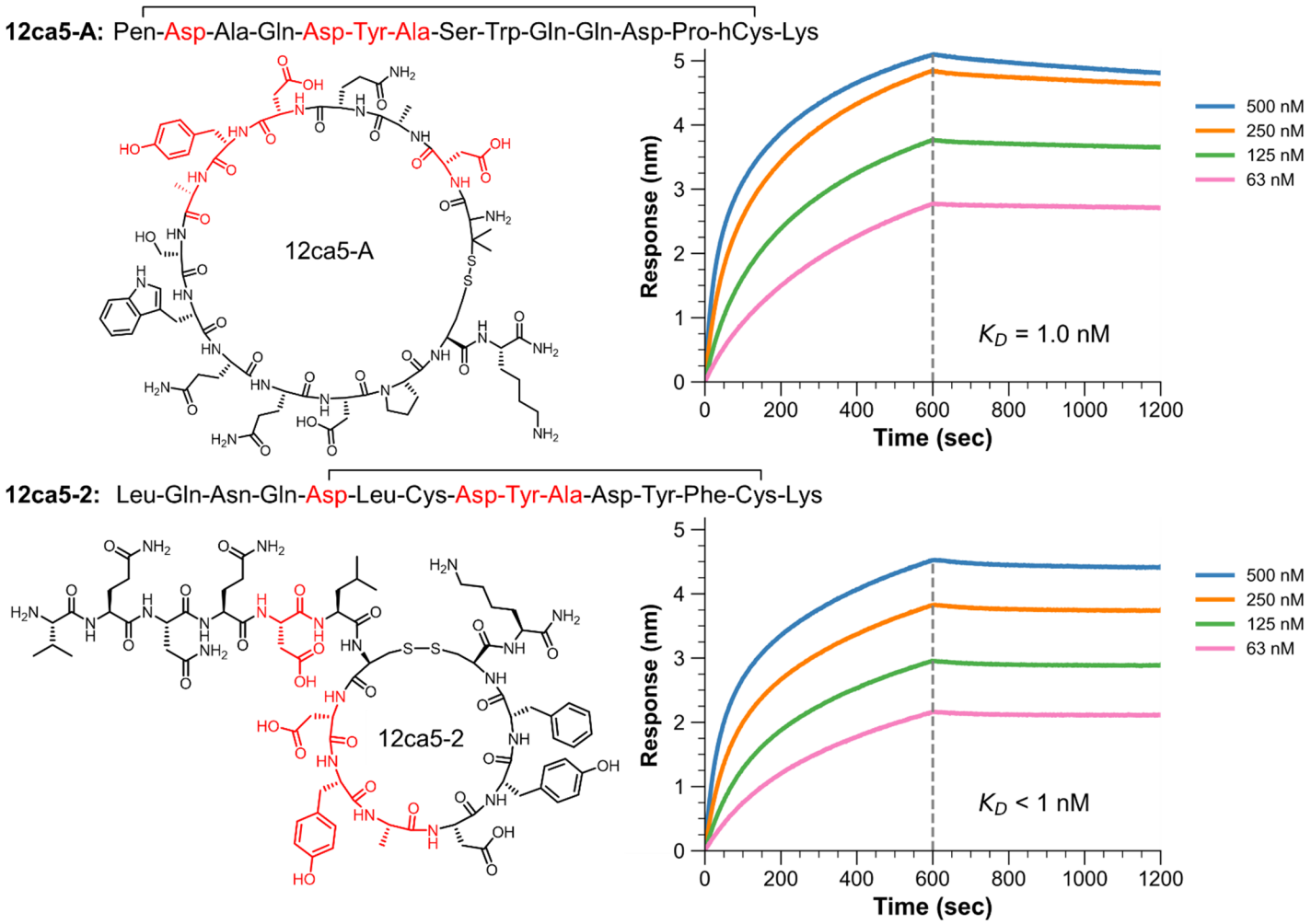


Figure 4. High-affinity macrocyclic peptide ligands to 12ca5 were enriched and identified via affinity selection-mass spectrometry (AS-MS), and the binding affinity was confirmed and measured using biolayer interferometry (BLI). All peptides were prepared with Lys(Biotin)-aminohexanoic acid (Ahx) attached to the N-terminus of the shown sequence and immobilized onto the BLI tip. The BLI tip was then dipped into solutions containing varying concentrations of cadherin-2 to record the concentration-dependent association and dissociation events. The characteristic 12ca5-binding motif D**DY(A/S) is highlighted in red and appeared exclusively at a single position in the X_6**C**X_6**C**K library (7 discovered peptides). In comparison, only one motif-containing peptide was discovered from the **C**X_{12}**C**K library.

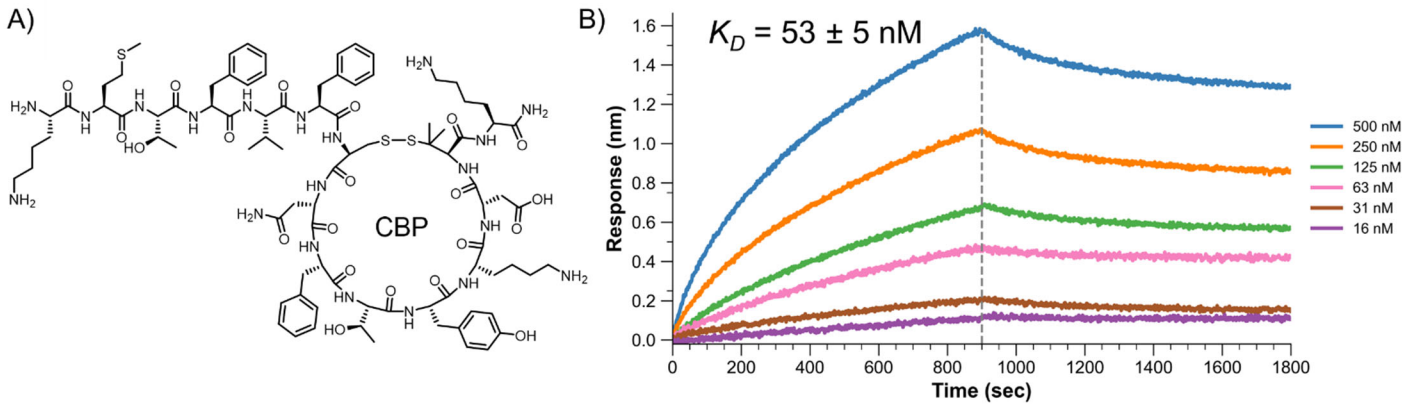
Macrocyclic low-nanomolar peptide ligands were discovered by affinity selection-mass spectrometry (AS-MS) performed against 12ca5 as a model protein target. The anti-hemagglutinin protein 12ca5 binds peptides containing the sequence D**DY(A/S) and has been used to benchmark AS-MS libraries.[33,58–60] Seven high affinity peptide ligands were pulled down from the X_6**C**X_6**C**K library design, while only one peptide was from the **C**X_{12}**C**K library (see Table S2 for all identified sequences). A select number of these sequences were synthesized and validated (see Figure S3) for their binding affinity using biolayer interferometry (BLI). All identified binders exhibited apparent dissociation constants (K_D) in the single-digit nanomolar to estimated high picomolar range, nearing the lower limit of detection for the instrument (see Figures 4 and S4). The binding motif was present in the same position or frameshift in all sequences found from the X_6**C**X_6**C**K library. Specifically, the cysteine analog in the middle of the library design was located inside of the 12ca5 motif at the

third position (i.e., D*ΨDY(A/S), where Ψ was discovered to be cysteine, homocysteine, or penicillamine). This outcome could potentially be explained by macrocyclization mimicking the geometry of the proline present at the same position within the natural substrate (HA peptide, YPYDVPDYA), leading to a stronger affinity. Moreover, this trend also suggests the X_6**C**X_6**C**K library design is more amenable to enrich high affinity binders against 12ca5 relative to the **C**X_{12}**C**K library.

Nanomolar affinity binders to cadherin-2 (CDH2) were discovered by AS-MS with the macrocyclic libraries. Due to the critical roles of CDH2 in adhesion in neural and cardiac junctions, the discovery of ligands outside of the protein homodimerization site could be of importance toward its study without affecting biological function. The homodimerization of cadherin-2 is largely driven by molecular interactions involving domains 1 and 2 of the ectodomain.[61,62,64,65] Based on these considerations, a fusion protein construct comprised of domains 4 and 5 of the CDH2 ectodomain (residues 498 to 724, Uniprot P15116) was used as the target for AS-MS selections. The fusion protein was cloned and expressed in mammalian cells, purified using SEC and Anti-Protein-C (clone HPC4) affinity tag purification, and verified by analytical SEC, reducing and non-reducing SDS-PAGE gels, as well as Western blotting (Figure S5). To our knowledge, no peptide or small molecular ligands have been reported to these CDH2 domains outside the homodimerization site.[64,65]

Nanomolar peptide ligands were discovered by AS-MS with both X_6**C**X_6**C**K and **C**X_{12}**C**K libraries against the CDH2[498-724] fusion protein. Only one peptide with high sequencing confidence was enriched against CDH2 (KMTFLFCNFTYKDZK, called cadherin-binding peptide, **CBP**, where Z is penicillamine disulfide bonded to C). **CBP** was synthesized and tested for its binding affinity to CDH2 by BLI, yielding a 53 nM K_D value (Figure 5). To verify sequence binding specificity, a scramble sequence that preserved the size of the macrocycle was synthesized and shown to have negligible binding response to CDH2 by BLI (Figure S6). The linearized form of **CBP** was also tested and demonstrated a greatly-reduced binding response (0.2 nm versus 1.6 nm response for the macrocycle), with an affinity of K_D of 150 nM (Figures S7 and S8). The more rigid structure of the macrocyclic **CBP** thus appears to favor higher affinity.

Structure-activity relationships (SAR) were delineated to characterize **CBP** using single residue replacement studies (alanine and D-amino acid scans) and truncations. In all SAR studies the Cys7-Pen14 disulfide bond was maintained to provide a consistent macrocycle structure to optimize from. First, an alanine scan was performed by synthesis of 13 variants featuring individual alanine mutations, which were assayed by BLI against CDH2 (Table S3, Figure S9). This alanine scan revealed multiple residues to be important for binding (hot-spots) including Lys4, Phe6, Lys12, and Lys15, due to the complete ablation of binding to CDH2 observed in BLI assays (Figure S10). Other residues including Met2, Thr3, Leu5, Thr10, Tyr11, and Asp13 had no effect on binding (cold-spots), suggesting they are not drivers of **CBP** binding to CDH2. Second, a truncation study focused at shortening **CBP** from the N-terminus, producing five additional peptides for BLI testing (Table S4 and Figure S11). BLI assays with these peptides confirmed the impact of the N-terminal residues on binding affinity, especially Phe6 as well as Leu5, while Met2 and Thr3 contributed minimally to CDH2 binding (Figure S12). Third, a D-amino acid scan of **CBP** was performed by iteratively replacing L-amino acids with D-amino acids to determine the impact of stereochemistry on the ligand interactions.[73,74] Notable hot-spots identified from the D-amino acid scan were Phe6 and Tyr11, further reinforcing the importance of the aromatic residues for binding (see Table S5 and Figures S13-S14). In summary, these initial SAR studies outline the hot-spot residues that appear to drive the high affinity binding of CBP to CDH2, including Lys4, Phe6, Lys12, and Lys15, while Met2, Thr3, Leu5, Thr10, and Asp13 are designated as cold-spots with a minimal effect on binding.



C)

CBP residue number →	1	2	3	4	5	6	8	9	10	11	12	13	15
SAR Study ↓	Lys	Met	Thr	Phe	Leu	Phe	Asn	Phe	Thr	Tyr	Lys	Asp	Lys
Alanine scan	250	70	60	NB	80	NB	100	120	50	70	NB	50	NB
Truncation from N-terminus		60	70	100	190	NB							
D-amino acid scan	50	70	140	60	50	3000	60	150	60	NB	60	50	NB
Mutated in Hot-spot Library													
Mutated in Cold-spot Library													

Fold Change (BLI K_D)
None, ~ 50 nM
Minor, >2-fold increase
Major, >50-fold increase

Figure 5. Macrocyclic peptide libraries enabled discovery of a 53 nM peptide ligand to a portion of the ectodomain of cadherin-2. (A) Structure of CBP. (B) BLI experiment reports the affinity of CBP to CDH2 with K_D = 53 nM binding affinity. All peptides were prepared with Lys(Biotin)-Ahx attached to the N-terminus in addition to the sequence shown and immobilized onto the BLI tip. The BLI tip was then dipped into solutions containing varying concentrations of CDH2 to record the concentration-dependent association and dissociation events. (C) Summary of experimental (alanine scan, D-amino acid scan, and N-terminal truncation study) SAR data. "NB" denotes non-binding. This SAR information was used to inform the designation of CBP "hot-spots" and "cold-spots," which do or do not drive high affinity binding, respectively.

SAR data informed the design of two focused libraries based on **CBP**: one to derivatize the high-affinity hot-spot residues, and the other to derivatize the cold-spot residues non-essential for binding. A common approach to design noncanonical libraries is to diversify the hot-spots.[75,76] However, peptide development and optimization often also considers the cold-spot residues that do not drive high affinity binding to the target. These cold-spot residues can be non-intuitively critical to improving binding affinity, solubility, or proteolytic stability.[42,43] Thus, we chose to compare the results from both maturation strategies. The set of non-canonical amino acids for incorporation in both the hot- and cold-spot focused libraries were selected based on the consensus data provided by the docking, alanine scan, D-amino acid scan, and truncation studies (Figure 6A and 6B). These libraries were synthesized and subjected to validation by SEC as shown in Figure S15 to demonstrate the lack of apparent oligomerization after disulfide bond formation.

No high-affinity ligands were discovered from the hot-spot focused library by AS-MS, while the cold-spot library provided ten new high-affinity noncanonical macrocyclic cadherin-2 peptide binders (**NCBPs**). From the hot-spot library, only two candidates were identified with high sequencing confidence, which featured multiple mutations from the original **CBP** sequence and shared replacements including Thr10Msn, Tyr11Dph, and Lys15Arg. However, the two hot-spot candidates (**NCBP-1** and **NCBP-2**) were synthesized and tested by BLI, revealing that they were non-binders to CDH2 under these conditions (Figures S16-S17). From the cold-spot library, ten noncanonical putative binders with high sequencing confidence were discovered and synthesized (Table S6, Figure S16). All discovered sequences from the cold-spot library (**NCBP-3 to NCBP-12**) were high-affinity binders to CDH2 in the BLI experiments, with determined K_D values between 20 and 50 nM (Figure S17).

The significant improvement in binding affinities observed by BLI upon derivatization of the cold-spot residues support this strategy as an efficient avenue for further optimization.

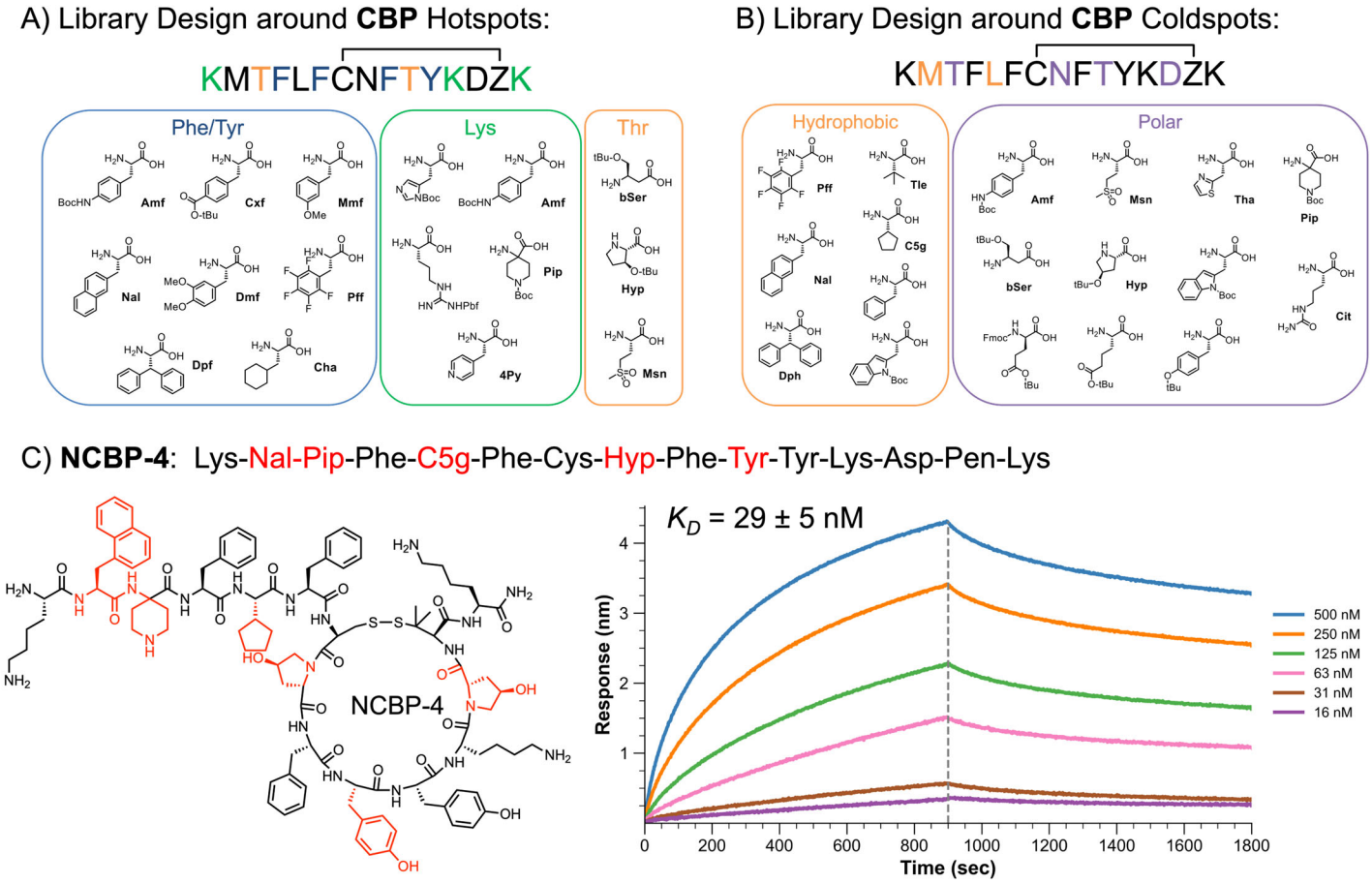


Figure 6. Single-peptide SAR information informs combinatorial library design and affinity maturation with noncanonical amino acids. (A and B) From the single peptide SAR studies summarized in Figure 5C, two libraries were designed to perform affinity maturation. The first library focused on minimally derivatizing the hot-spots by matching the original natural amino acids properties (e.g., hydrophobicity or positive charge). The second library focused around diversifying the cold-spots to examine the possibility that the cold-spots could be mutated to improve the overall binding of the CBP peptide, either by pre-arranging the conformation of the peptide or facilitating new binding interactions with CDH2. Both libraries were prepared using split-pool synthesis, except the entirety of the theoretical sequence space was sampled by the library due to its smaller focused design. Specifically, the number of beads used in split-pool synthesis approximately matched the theoretical diversity: Hot-spot library total number of beads: 2.5 x 10^6 with theoretical sequence space diversity: 2.7 x 10^6 and cold-spot library total number of beads: 7.0 x 10^5 with theoretical sequence space diversity: 7.2 x 10^5. (C) Sequence and structure of **NCBP-4** discovered from affinity selection and its BLI binding response, which exhibited high-affinity binding (K_D = 29 ± 5 nM). Nal = 3-(2-naphthyl)-L-alanine, Pip = 4-aminopiperidine-4-carboxylic acid, C5g = cyclopentylglycine, and Hyp = L-trans-4-hydroxyproline.

The **NCBP-4** noncanonical binder exhibits nanomolar binding affinity to CDH2 (K_D = 29 ± 5 nM). The results from AS-MS of the cold-spot macrocyclic library show most cold-spot amino acids were replaced in the identified sequences (Table S6). In all candidates, Met2 was replaced by a hydrophobic amino acid, with 3-(2-

naphthyl)-L-alanine (Nal) appearing in six of the ten. Similarly, Leu5 was replaced by Phe, Nal, cyclopentylglycine (C5g), or pentafluoro-L-phenylalanine (PFf). For the polar subset of amino acids in the cold-spot library, the replacements made to **CBP** were more mixed. Asn8 and Thr10 were replaced with a diverse set of amino acids, possibly indicating their lack of contribution to the binding interaction. Thr3 was replaced by cationic 4-aminopiperidine-4-carboxylic acid (Pip) and polar L-β-Homoserine (bSer). And lastly, Asp13 demonstrated a preferred replacement to trans-4-hydroxyproline (Hyp), appearing in five of ten candidates. With these results in mind, **NCBP-4** was chosen for detailed investigation by BLI as it featured a consensus of amino acid replacements including Met2Nph, Thr3Pip, Leu5C5g, and Asp13Hyp (Figure S18). **NCBP-4** exhibited clear concentration-dependent binding to CDH2 and a resulting binding affinity of K_D = 29 ± 5 nM determined by BLI (Figure 6). Moreover, **NCBP-4** demonstrated a stronger response illustrated by a higher BLI signal (~4.3 nm), more than double the response seen with **CBP** (~1.6 nm, Figure S19).

A specific mutant of **NCBP-4** was constructed with serine mutated at the hot-spots to specifically examine if the cold-spot residues were effective in creating new binding interactions with CDH2 (sequence: Ser-Nal-Pip-Ser-C5g-Ser-Cys-Hyp-Ser-Tyr-Ser-Ser-Hyp-Pen-Ser-NH2). Interestingly, this mutant retained moderate binding to CDH2 with an apparent K_D of ~ 150 nM (Figure S20). This serine-substituted mutant was also tested against 12ca5 to assess non-specific binding. The mutant showed no affinity towards 12ca5 (Figure S19). This result suggests the amino acid replacements made in the cold-spots of **CBP** to deliver the **NCBP** compound series resulted in peptide ligands that maintain target specificity despite the presence of additional hydrophobic residues, which are commonly associated with non-specific interactions.[77,78]

4 Conclusions

We established a protocol for the split-and-pool synthesis of high-diversity macrocyclic peptide libraries and demonstrated their use in the discovery of nanomolar ligands against two protein targets of different structure. Formation of the macrocycle is performed using a simple disulfide bond, which is rapidly installed in aqueous solution using iodine. Quantification of free thiol content via Ellman's assay in both the cyclized and linearized forms confirmed the complete conversion of thiols to disulfides in the library, while SEC revealed that the disulfide bonds were formed exclusively intramolecularly without oligomerization. While the disulfide bond is not the most robust linkage for cyclization, there are several approaches available to replace it when needed to improve stability toward therapeutic development.[11,79,80]

Affinity selection was able to identify protein-specific ligands from these synthetic macrocyclic libraries for both a model protein (12ca5) and a novel target, cadherin-2, which participates in basic biological adhesion of cells in neural and cardiometabolic function. Both the canonical **CBP** and non-canonical **NCBP-4** peptides demonstrated concentration-dependent binding to CDH2, binding specificity, and high affinity with K_D values of 50 and 29 nM, respectively. After examining the SAR of **CBP**, several hot-spot residues were revealed to be critical for binding to CDH2, featuring several hydrophobic and cationic residues.

From the SAR studies, two additional macrocyclic libraries containing a diverse set of noncanonical amino acids were synthesized focusing on the affinity-driving hot-spots and the non-essential cold-spots, respectively. The subsequent affinity selection experiments investigated the hypotheses of whether the hot-spots can be further refined or if the cold-spots can become meaningful contributors to the binding affinity upon maturation with focused library designs. Overall, AS-MS utilizing the hot-spot CDH2-focused library did not provide any binders or improvement to the original **CBP** peptide. However, AS-MS experiments utilizing the cold-spot library were able to provide several candidates individually validated to be high-affinity binders. Of these, **NCBP-4** was examined more closely for its high affinity (K_D ~29 nM), specificity, and specific side-chain contributions demonstrated by the amino acids that replaced the original CBP cold-spot residues.

Utilizing noncanonical amino acids in combinatorically prepared macrocyclic libraries, we demonstrated the rapid affinity maturation of **CBP**. This process was successful through the replacement of cold-spot residues

with noncanonical monomers. Overall, due to the improvements that macrocyclization often offer over linear peptide scaffolds, we expect this work to be fundamental to the impactful deployment of macrocyclic synthetic libraries for the advancement of peptide therapeutic discovery and development.

5 Data Availability Statement

The data supporting the findings of this study are available in the main text and Supporting Information.

6 Author Contributions

All authors designed research. Michael A. Lee, Joseph S. Brown and Charlotte E. Farquhar performed experimental work. All authors performed data analysis. Michael A. Lee and Joseph S. Brown wrote the manuscript with input from all authors. Bradley L. Pentelute and Andrei Loas supervised the work.

7 Acknowledgements

Funding for this work was provided by Novo Nordisk A/S. We thank Dr. Thomas E. Nielsen and Dr. Uli Stilz for their helpful discussion in support of our work. Joseph S. Brown acknowledges support from Pharmaceutical Research and Manufacturers of America (PhRMA) Foundation through the Postdoctoral Fellowship in Drug Discovery. Michael A. Lee acknowledges support from the MIT Dean of Science Fellowship.

8 Conflict of Interest Statement

The authors declare the following competing interests: Bradley L. Pentelute is a co-founder and/or member of the scientific advisory board of several companies focusing on the development of protein and peptide therapeutics.

9 Methods

9.1 Rapid oxidation of peptide thiols for intramolecular cyclization using iodine

After cleavage and lyophilization, each split-pool prepared peptide library (e.g., 10 mg, 4.2 μmol, average molecular weight ~ 2400 g/mol) was resuspended at 2 mg/mL in 5% AcN in Water (0.1% TFA) and treated with 10 μL portions of freshly prepared 60 mM I_2 in MeOH until the solution remained yellow. For the example 10 mg scale, approximately 70 μL I_2 stock solution total was used, resulting in ~1 eq of I_2 with respect to the library. The iodine-treated library was incubated for 10 minutes in the dark at rt, upon which 15 μL of freshly prepared 1 M ascorbic acid was added (~3.6 eq of ascorbic acid with respect to the library example). This solution was immediately loaded onto a pre-equilibrated SPE column, SPE-purified to remove any remaining iodine and ascorbic acid, and lyophilized. The lyophilized powder was then resuspended at approximately 0.1 mM and its thiol concentration as quantified by Ellman's assay.

9.2 Reductive linearization of cyclized library

Peptide libraries were resuspended at 0.62 mg/mL to mimic the maximum concentration possibly isolated at the end of AS-MS, due to the maximum capacity of the STAGE tip (8 μg for a double plug, using 13 μL). As described in the Main Text, reduction after resuspension at pH 3 from 0.1% formic acid in ultrapure water was successful using 1,4-DL-dithiothreitol (DTT, Chem-Impex Cat: 00127). The reduction was also tested in 200 mM sodium phosphate, 5 mM EDTA, pH 8 with DTT, and immobilized tris(2-carboxyethyl)phosphine (TCEP, Thermo Fisher Scientific, 77712). DTT was freshly prepared in a stock solution of 500 mg/mL and added to samples to

provide 50 mg/mL final, 1000 eq and incubated at 60 °C for 15 minutes. Samples were then SPE-purified, lyophilized, and Ellman's quantified upon resuspension. Immobilized TCEP beads (8 mM stock suspension) were washed three times before use with the assay buffer using centrifugation at 1000 rcf for 1 minutes. Treatment of the library peptides with immobilized TCEP used 20 eq for 25 minutes at room temperature rocking on a nutating mixer. The supernatant was isolated from centrifugation, lyophilized, and Ellman's' quantified.

9.3 Sequencing validation of reduced and oxidized libraries

A small portion of library resin was measured and made into a 1 mg/mL stock solution in DMF. Several aliquots of ~1000 beads (for 20 µm resin, this will be about ~10 µL of 1 mg/mL stock) were taken, centrifuged, and aspirated of DMF. Each aliquot was cleaved from the solid phase support using 60 µL of 95% trifluoroacetic acid, 2.5% water and 2.5% triisopropylsilane at 60 °C for 15 minutes. Half of the liquid was then evaporated under a gentle stream of N_2, followed by dilution to a total volume of 240 µL using water (0.1% TFA). A third of the solution was aliquoted to represent the library before cyclization. The remaining solution was cyclized according to the Section 3.5. The cyclized peptide library and aliquoted linear library were both prepared for nLC/MS-MS analysis according to Section 3.4. The dried library samples were then reconstituted in water (0.1% TFA) at a concentration of 100 pg/µL/peptide (for example, prepare 8 µg of an aliquot of 1000 peptides in 80 µL). Half of the cyclized peptide sample was then linearized using DTT as described in Section 3.6. All three types of samples, peptide post-cleavage, cyclized, and cyclized then reduced, were subjected to nLC-MS/MS analysis as described in Section 3.11.

9.4 Ellman's thiol quantification assay

The thiol concentration of suspended peptides was completed using Ellman's reagent (Millipore-Sigma, 5,5′-dithiobis(2-nitrobenzoic acid), D8130, ≥98%, BioReagent) using the following conditions. Ellman's stock solution was prepared at 10.0 mM and assay buffer was 1x PBS at pH 8.0 with 1 mM of EDTA. Nonsterile Greiner 96-well polystyrene plates (Millipore-Sigma, M2936) were used. Using the assay buffer to have a final 200 µL well volume, 3.6 µL of Ellman's stock solution was combined with the peptide solution to give a final peptide intended concentration of 0.1 mM, which was determined to be within the linear regime in which signal could be observed from a standard curve constructed using cysteine (Millipore-Sigma, C7352, ≥98%, BioReagent). After combining, all materials were incubated for 7 minutes, and then read at 416 nm using a TECAN Spark Plate Reader. The concentration of the library was inferred by measuring its absorbance at 280 nm (NanoQuant Plate) and was used to normalize the Ellman's thiol concentration to account for slight variations in the intended resuspended concentrations.

Sample preparation to measure thiol signal from cleavage and SPE of library (reduced library): Approximately 50 mg of peptide library (peptide + resin) was globally deprotected and cleaved from resin with 95% (v/v) TFA, 2.5% (v/v) water, and 2.5% (v/v) triisopropylsilane, for 15 minutes at 60 ºC (~20 mL/g of resin). Precipitated peptide was triturated (3 x 100 mL/g resin) with cold diethyl ether, resuspended in 5% acetonitrile in water (0.1% TFA) and solid-phase extracted. After lyophilization, this sample was resuspended in assay buffer at 0.1 mM and measured for its thiol concentration by Ellman's assay.

9.5 Size exclusion chromatography (SEC) of libraries

Size exclusion chromatography (SEC) was performed using an Agilent 1260 Infinity II LC System with a Superdex 30 Increase 10/300 GL column (10 x 300 mm, 9 µm particle size from Cytiva Life Sciences, separation MW range 100 to 7000 Da). 25 µg of library was injected in 200 µL of total solution. Cyclized peptide samples were aliquoted from the main stock prepared according to Sections 3.5 and 3.3. Linearized samples were prepared by adding 1000 eq from a 50 mg/mL DTT stock solution and heating the sample to 60 °C for 10 minutes before dilution to 200 µL using 1x PBS. Column conditions: isocratic 1x PBS for 1.5 column volumes

at 0.8 mL/min. Buffer blanks were prepared for both cyclized and linearized samples and were subtracted from the library samples. A custom mass standard was prepared by adding 10 µg of a mixture of peptides corresponding to the following molecular weights: 1807, 3750, 5312, 8305 for average monomer mass, average dimer mass, average trimer mass, and the exclusion limit of the SEC column.

10 References

1. Cunningham, A. D., Qvit, N. & Mochly-Rosen, D. Peptides and peptidomimetics as regulators of protein–protein interactions. *Curr Opin Struct Biol* **44**, 59–66 (2017).
2. Wells, J. A. & McClendon, C. L. Reaching for high-hanging fruit in drug discovery at protein–protein interfaces. *Nature* **450**, 1001–1009 (2007).
3. Henninot, A., Collins, J. C. & Nuss, J. M. The Current State of Peptide Drug Discovery: Back to the Future? *J Med Chem* **61**, 1382–1414 (2018).
4. Muttenthaler, M., King, G. F., Adams, D. J. & Alewood, P. F. Trends in peptide drug discovery. *Nat Rev Drug Discov* **20**, 309–325 (2021).
5. Góngora-Benítez, M., Tulla-Puche, J. & Albericio, F. Multifaceted roles of disulfide bonds. peptides as therapeutics. *Chem Rev* **114**, 901–926 (2014).
6. Vinogradov, A. A., Yin, Y. & Suga, H. Macrocyclic Peptides as Drug Candidates: Recent Progress and Remaining Challenges. *J Am Chem Soc* **141**, 4167–4181 (2019).
7. Tyndall, J. D. A., Nall, T. & Fairlie, D. P. Proteases universally recognize beta strands in their active sites. *Chem Rev* **105**, 973–999 (2005).
8. Fairlie, D. P. *et al.* Conformational selection of inhibitors and substrates by proteolytic enzymes: Implications for drug design and polypeptide processing. *J Med Chem* **43**, 1271–1281 (2000).
9. Wang, D. *et al.* Enhanced Metabolic Stability and Protein-Binding Properties of Artificial α Helices Derived from a Hydrogen-Bond Surrogate: Application to Bcl-xL. *Angewandte Chemie International Edition* **44**, 6525–6529 (2005).
10. Spokoyny, A. M. *et al.* A perfluoroaryl-cysteine SNAr chemistry approach to unprotected peptide stapling. *J Am Chem Soc* **135**, 5946–5949 (2013).
11. Tugyi, R., Mezö, G., Fellinger, E., Andreu, D. & Hudecz, F. The effect of cyclization on the enzymatic degradation of herpes simplex virus glycoprotein D derived epitope peptide. *Journal of Peptide Science* **11**, 642–649 (2005).
12. Walensky, L. D. *et al.* Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix. *Science (1979)* **305**, 1466–1470 (2004).
13. Naylor, M. R., Bockus, A. T., Blanco, M. J. & Lokey, R. S. Cyclic peptide natural products chart the frontier of oral bioavailability in the pursuit of undruggable targets. *Curr Opin Chem Biol* **38**, 141–147 (2017).
14. Pye, C. R. *et al.* Nonclassical Size Dependence of Permeation Defines Bounds for Passive Adsorption of Large Drug Molecules. *J Med Chem* **60**, 1665–1672 (2017).
15. Bhardwaj, G. *et al.* Accurate de novo design of membrane-traversing macrocycles. *Cell* **185**, 3520-3532.e26 (2022).
16. Mizuno-Kaneko, M. *et al.* Molecular Design of Cyclic Peptides with Cell Membrane Permeability and Development of MDMX-p53 Inhibitor. *ACS Med Chem Lett* (2023) doi:10.1021/ACSMEDCHEMLETT.3C00102.
17. Iskandar, S. E. & Bowers, A. A. mRNA Display Reaches for the Clinic with New PCSK9 Inhibitor. *ACS Med Chem Lett* (2022) doi:10.1021/ACSMEDCHEMLETT.2C00319.
18. Bourne, G. *et al.* Oral peptide inhibitors of interleukin-23 receptor and their use to treat inflammatory bowel diseases. vol. 2 (2017).

19. Benhamou, R. I. *et al.* Macrocyclization of a Ligand Targeting a Toxic RNA Dramatically Improves Potency. *ChemBioChem* **21**, 3229–3233 (2020).
20. Hacker, D. E. *et al.* Direct, Competitive Comparison of Linear, Monocyclic, and Bicyclic Libraries Using mRNA Display. *ACS Comb Sci* **22**, 306–310 (2020).
21. Gao, Y. & Kodadek, T. Direct comparison of linear and macrocyclic compound libraries as a source of protein ligands. *ACS Comb Sci* **17**, 190–195 (2015).
22. Passioura, T., Katoh, T., Goto, Y. & Suga, H. Selection-Based Discovery of Druglike Macrocyclic Peptides. *https://doi.org/10.1146/annurev-biochem-060713-035456* **83**, 727–752 (2014).
23. Tavassoli, A. SICLOPPS cyclic peptide libraries in drug discovery. *Curr Opin Chem Biol* **38**, 30–35 (2017).
24. Goto, Y. & Suga, H. The RaPID Platform for the Discovery of Pseudo-Natural Macrocyclic Peptides. *Acc Chem Res* **54**, 3604–3617 (2021).
25. Deyle, K., Kong, X. D. & Heinis, C. Phage Selection of Cyclic Peptides for Application in Research and Drug Development. *Acc Chem Res* **50**, 1866–1874 (2017).
26. Bacon, K. *et al.* Isolation of Chemically Cyclized Peptide Binders Using Yeast Surface Display. *ACS Comb Sci* **22**, 519–532 (2020).
27. Wong, J. Y. K. *et al.* Genetically-encoded discovery of proteolytically stable bicyclic inhibitors for morphogen NODAL. *Chem Sci* **12**, 9694–9703 (2021).
28. Owens, A. E., Iannuzzelli, J. A., Gu, Y. & Fasan, R. MOrPH-PhD: An Integrated Phage Display Platform for the Discovery of Functional Genetically Encoded Peptide Macrocycles. *ACS Cent Sci* **6**, 368–381 (2020).
29. Biron, É., Vézina-Dawod, S. & Bédard, F. Synthetic Strategies for Macrocyclic Peptides. *Practical Medicinal Chemistry with Macrocycles* 205–241 (2017) doi:10.1002/9781119092599.CH9.
30. Habeshian, S. *et al.* Synthesis and direct assay of large macrocycle diversities by combinatorial late-stage modification at picomole scale. *Nature Communications 2022 13:1* **13**, 1–14 (2022).
31. Pomplun, S. *et al.* De Novo Discovery of High-Affinity Peptide Binders for the SARS-CoV-2 Spike Protein. *ACS Cent Sci* **7**, 156–163 (2021).
32. Zhang, G. *et al.* Rapid de novo discovery of peptidomimetic affinity reagents for human angiotensin converting enzyme 2. *Commun Chem* **5**, 1–10 (2022).
33. Quartararo, A. J. *et al.* Ultra-large chemical libraries for the discovery of high-affinity peptide binders. *Nat Commun* **11**, 3183 (2020).
34. White, C. J. & Yudin, A. K. Contemporary strategies for peptide macrocyclization. *Nature Chemistry 2011 3:7* **3**, 509–524 (2011).
35. Ekanayake, A. I. *et al.* Genetically Encoded Fragment-Based Discovery from Phage-Displayed Macrocyclic Libraries with Genetically Encoded Unnatural Pharmacophores. *J Am Chem Soc* **143**, 5497–5507 (2021).
36. Wrighton, N. C. *et al.* Small Peptides as Potent Mimetics of the Protein Hormone Erythropoietin. *Science (1979)* **273**, 458–463 (1996).
37. Fairbrother, W. J. *et al.* Novel peptides selected to bind vascular endothelial growth factor target the receptor-binding site. *Biochemistry* **37**, 17754–17764 (1998).
38. DeLano, W. L., Ultsch, M. H., De Vos, A. M. & Wells, J. A. Convergent solutions to binding at a protein-protein interface. *Science (1979)* **287**, 1279–1283 (2000).
39. Zhang, H. & Chen, S. Cyclic peptide drugs approved in the last two decades (2001–2021). *RSC Chem Biol* **3**, 18–31 (2022).
40. Mortensen, K. T., Osberger, T. J., King, T. A., Sore, H. F. & Spring, D. R. Strategies for the Diversity-Oriented Synthesis of Macrocycles. *Chem Rev* **119**, 10288–10317 (2019).
41. Sayago, C. *et al.* Deciphering Binding Interactions of IL-23R with HDX-MS: Mapping Protein and Macrocyclic Dodecapeptide Ligands. *ACS Med Chem Lett* **9**, 912–916 (2018).

42. Guerlavais, V. *et al.* Discovery of Sulanemadlin (ALRN-6924), the First Cell-Permeating, Stabilized α-Helical Peptide in Clinical Development. *J Med Chem* **66**, 9401–9417 (2023).
43. Garrigou, M. *et al.* Accelerated Identification of Cell Active KRAS Inhibitory Macrocyclic Peptides using Mixture Libraries and Automated Ligand Identification System (ALIS) Technology. *J Med Chem* **65**, 8961–8974 (2022).
44. Lim, S. *et al.* Discovery of cell active macrocyclic peptides with on-target inhibition of KRAS signaling. *Chem Sci* **12**, 15975–15987 (2021).
45. Fair, R. J., Walsh, R. T. & Hupp, C. D. The expanding reaction toolkit for DNA-encoded libraries. *Bioorg Med Chem Lett* **51**, 128339 (2021).
46. Behsaz, B. *et al.* De Novo Peptide Sequencing Reveals Many Cyclopeptides in the Human Gut and Other Environments. *Cell Syst* **10**, 99-108.e5 (2020).
47. Townsend, C. *et al.* CycLS: Accurate, whole-library sequencing of cyclic peptides using tandem mass spectrometry. *Bioorg Med Chem* **26**, 1232–1238 (2018).
48. Kavan, D., Kuzma, M., Lemr, K., Schug, K. A. & Havlicek, V. CYCLONE - A utility for de novo sequencing of microbial cyclic peptides. *J Am Soc Mass Spectrom* **24**, 1177–1184 (2013).
49. Mohimani, H. *et al.* Cycloquest: Identification of cyclopeptides via database search of their mass spectra against genome databases. *J Proteome Res* **10**, 4505–4512 (2011).
50. Lee, J. H., Meyer, A. M. & Lim, H. S. A simple strategy for the construction of combinatorial cyclic peptoid libraries. *Chemical Communications* **46**, 8615–8617 (2010).
51. Liang, X., Vézina-Dawod, S., Bédard, F., Porte, K. & Biron, E. One-Pot Photochemical Ring-Opening/Cleavage Approach for the Synthesis and Decoding of Cyclic Peptide Libraries. *Org Lett* **18**, 1174–1177 (2016).
52. Borges, A. *et al.* Facile de Novo Sequencing of Tetrazine-Cyclized Peptides through UV-Induced Ring-Opening and Cleavage from the Solid Phase. *ChemBioChem* **24**, e202200590 (2023).
53. Simpson, L. S. & Kodadek, T. A cleavable scaffold strategy for the synthesis of one-bead one-compound cyclic peptoid libraries that can be sequenced by tandem mass spectrometry. *Tetrahedron Lett* **53**, 2341–2344 (2012).
54. Elashal, H. E., Cohen, R. D., Elashal, H. E. & Raj, M. Oxazolidinone-Mediated Sequence Determination of One-Bead One-Compound Cyclic Peptide Libraries. *Org Lett* **20**, 2374–2377 (2018).
55. Menegatti, S. *et al.* Reversible cyclic peptide libraries for the discovery of affinity ligands. *Anal Chem* **85**, 9229–9237 (2013).
56. Novák, J., Lemr, K., Schug, K. A. & Havlíček, V. CycloBranch: De Novo Sequencing of Nonribosomal Peptides from Accurate Product Ion Mass Spectra. *J Am Soc Mass Spectrom* **26**, 1780–1786 (2015).
57. Tran, N. H. *et al.* Deep learning enables de novo peptide sequencing from data-independent-acquisition mass spectrometry. *Nature Methods 2018 16:1* **16**, 63–66 (2018).
58. Rini, J. M., Schulze-Gahmen, U. & Wilson, I. A. Structural evidence for induced fit as a mechanism for antibody-antigen recognition. *Science (1979)* **255**, 959–965 (1992).
59. Houghten, R. A. *et al.* Generation and use of synthetic peptide combinatorial libraries for basic research and drug discovery. *Nature 1991 354:6348* **354**, 84–86 (1991).
60. Brown, J. S. *et al.* Unsupervised machine learning leads to an abiotic picomolar peptide ligand. *ChemRxiv* (2023) doi:10.26434/CHEMRXIV-2023-TWS4N.
61. Takeichi, M. The cadherin superfamily in neuronal connections and interactions. *Nature Reviews Neuroscience 2006 8:1* **8**, 11–20 (2006).
62. Kostetskii, I. *et al.* Induced Deletion of the N-Cadherin Gene in the Heart Leads to Dissolution of the Intercalated Disc Structure. *Circ Res* **96**, 346–354 (2005).
63. Karlsson, M. *et al.* A single–cell type transcriptomics map of human tissues. *Sci Adv* **7**, (2021).

64. Williams, E. J., Williams, G., Gour, B., Blaschuk, O. & Doherty, P. INP, a novel N-cadherin antagonist targeted to the amino acids that flank the HAV motif. *Mol Cell Neurosci* **15**, 456–464 (2000).
65. Burden-Gulley, S. M. *et al.* Novel peptide mimetic small molecules of the HAV motif in N-cadherin inhibit N-cadherin-mediated neurite outgrowth and cell adhesion. *Peptides (N.Y.)* **30**, 2380–2387 (2009).
66. Tam, J. P., Wu, C. R., Liu, W. & Zhang, J. W. Disulfide Bond Formation in Peptides by Dimethyl Sulfoxide. Scope and Applications. *J Am Chem Soc* **113**, 6657–6662 (1991).
67. Hennion, M. C. Solid-phase extraction: method development, sorbents, and coupling with liquid chromatography. *J Chromatogr A* **856**, 3–54 (1999).
68. Yang, Y., Hansen, L., Fraczek, A., Badalassi, F. & Kjellström, J. DMF-Assisted Iodination Side Reaction during the Preparation of Disulfide Peptides, Its Substrate/Solvent/pH Dependence, and Implications on Disulfide-Peptide Production. *Org Process Res Dev* **25**, 2090–2099 (2021).
69. Yang, Y., Hansen, L., Fraczek, A., Badalassi, F. & Kjellström, J. DMF-Assisted Iodination Side Reaction during the Preparation of Disulfide Peptides, Its Substrate/Solvent/pH Dependence, and Implications on Disulfide-Peptide Production. *Org Process Res Dev* **25**, 2090–2099 (2021).
70. Thermo Scientific (Pierce Biotechnology). *Instructions Immobilized TCEP Disulfide Reducing Gel (Cat 77712) Version 1325.3 MAN0011439*. (2023).
71. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature Protocols 2007 2:8* **2**, 1896–1906 (2007).
72. Vinogradov, A. A. *et al.* Library Design-Facilitated High-Throughput Sequencing of Synthetic Peptide Libraries. *ACS Comb Sci* **19**, 694–701 (2017).
73. Simon, M. D. *et al.* D-Amino Acid Scan of Two Small Proteins. *J Am Chem Soc* **138**, 12099–12111 (2016).
74. Peeters, T. L. *et al.* d-Amino acid and alanine scans of the bioactive portion of porcine motilin. *Peptides (N.Y.)* **13**, 1103–1107 (1992).
75. Touti, F., Gates, Z. P., Bandyopdhyay, A. & Lautrette, G. In-solution enrichment identifies peptide inhibitors of protein – protein interactions protein-protein interactions. *Nat Chem Biol* 318–339 (2019) doi:10.1038/s41589-019-0245-2.
76. Ye, X. *et al.* Binary combinatorial scanning reveals potent poly-alanine-substituted inhibitors of protein-protein interactions. *Communications Chemistry 2022 5:1* **5**, 1–10 (2022).
77. Frutiger, A. *et al.* Nonspecific Binding - Fundamental Concepts and Consequences for Biosensing Applications. *Chem Rev* **121**, 8095–8160 (2021).
78. Ng, S. *et al.* De-risking Drug Discovery of Intracellular Targeting Peptides: Screening Strategies to Eliminate False-Positive Hits. *ACS Med Chem Lett* acsmedchemlett.0c00022 (2020) doi:10.1021/acsmedchemlett.0c00022.
79. Dias, R. L. A. *et al.* Protein ligand design: From phage display to synthetic protein epitope mimetics in human antibody Fc-binding peptidomimetics. *J Am Chem Soc* **128**, 2726–2732 (2006).
80. Kourra, C. M. B. K. & Cramer, N. Converting disulfide bridges in native peptides to stable methylene thioacetals. *Chem Sci* **7**, 7007–7012 (2016).